CODEN: JASMAN 4966

THE JOURNAL of the Acoustical Society of America

Vol. 103, No. 5, Pt. 1 May 1998

SOUNDINGS SECTION

ACOUSTICAL NEWS—USA 2241

USA Meetings Calendar 2243

ACOUSTICAL STANDARDS NEWS 2248

Standards Meetings Calendar 2248

OBITUARIES 2255

BOOK REVIEWS 2256

REVIEWS OF ACOUSTICAL PATENTS 2257

SELECTED RESEARCH ARTICLES [10]

The force between two parallel rigid plates due to the radiation pressure of broadband noise: An acoustic Casimir effect — Andrés Larraza, Christopher D. Holmes, Robert T. Susbilla, Bruce Denardo — 2267

The effect of compression ratio and release time on the categorical rating of sound quality — Arlene C. Neuman, Matthew H. Bakke, Carol Mackersie, Sharon Hellman, Harry Levitt — 2273

GENERAL LINEAR ACOUSTICS [20]

Lamb waves as thickness vibrations superimposed on a membrane carrier wave — J. D. Achenbach — 2283

Aperture synthesis of time-limited X waves and analysis of their propagation characteristics — Argyrios A. Chatzipetros, Amr M. Shaarawi, Ioannis M. Besieris, Mohamed A. Abdel-Rahman — 2287

Wave propagation in a wavy fiber–epoxy composite material: Theory and experiment — Kwang Yul Kim, Wei Zou, Wolfgang Sachse — 2296

Sound transmission through a periodic cascade with application to drill pipes — Niels J. C. Lous, Sjoerd W. Rienstra, Ivo J. B. F. Adan — 2302

NONLINEAR ACOUSTICS, MACROSONICS [25]

A finite element algorithm for the study of nonlinear standing waves — L. Elvira-Segura, E. Riera-Franco de Sarabia — 2312

The acoustic radiation force in lossless fluids in Eulerian and Lagrangian coordinates — K. Beissner — 2321

(Continued)

CONTENTS—*Continued from preceding page*

UNDERWATER SOUND [30]

Measurement and localization of interface wave reflections from a buried target	Eric Smith, Preston S. Wilson, Fred W. Bacon, Jason F. Manning, John A. Behrens, Thomas G. Muir	2333
Scattering of interface waves from pointlike obstacles	Eric Smith	2344
A fast integral equation approach to acoustic scattering from three-dimensional objects using a natural basis set	D. T. Moroney, P. J. Cullen	2354
Acoustical properties of end-initiated explosive line charges	William J. Marshall	2365
Nonlinear acoustic method for bubble density measurements in water	A. M. Sutin, S. W. Yoon, E. J. Kim, I. N. Didenkulov	2377

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

Radiation impedance of resonant ultrasound spectroscopy modes in fused silica	Hong Zhang, R. S. Sorbello, Carsten Hucho, Joseph Herro, Jeffrey R. Feller, D. E. Beck, Moises Levy, D. Isaak, J. D. Carnes, O. Anderson	2385
A new approximation method for thermoacoustic calculations	Richard Raspet, James Brewster, Henry E. Bass	2395
Focusing and steering through absorbing and aberrating layers: Application to ultrasonic propagation through the skull	Mickaël Tanter, Jean-Louis Thomas, Mathias Fink	2403
Line-focusing of ultrasonic *SV* wave by electromagnetic acoustic transducer	Hirotsugu Ogi, Masahiko Hirao, Toshihiro Ohtani	2411
Lamb wave assessment of fatigue and thermal damage in composites	Michael D. Seale, Barry T. Smith, W. H. Prosser	2416

TRANSDUCTION [38]

A finite element analysis of an interferometric optical fiber hydrophone	Jong-in Im, Yong-rae Roh	2425
Ultrasonic three-dimensional probe modeling in anisotropic solids	A. Jonas Niklasson	2432

STRUCTURAL ACOUSTICS AND VIBRATION [40]

Modal excitation and control of membrane waves on cylindrical shells	Joseph E. Bondaryk	2443
Sound radiation by a simply supported unbaffled plate	B. Laulagnet	2451

ACOUSTICAL MEASUREMENTS AND INSTRUMENTATION [58]

Measurement of the reflection matrix of a discontinuity in a duct	Mohamed Akoum, Jean-Michel Ville	2463

ACOUSTIC SIGNAL PROCESSING [60]

Known source detection predictions for higher order correlators	Lisa A. Pflug, George E. Ioup, Juliette W. Ioup	2469
Reference-beam detection for scanning laser acoustic microscopy	Moisés Cywiak, Cristina Solano, Glen Wade, S. Isakson	2478
An adaptive noise reduction stethoscope for auscultation in high noise environments	Samir B. Patel, Thomas F. Callahan, Matthew G. Callahan, James T. Jones, George P. Graber, Kirk S. Foster, Kenneth Glifort, George R. Wodicka	2483
Projection data error estimation and correction for high-resolution tomographic acoustic microscopy	Richard Y. Chiao, Davis Kent, Hua Lee	2492

(Continued)

CONTENTS—*Continued from preceding page*

PHYSIOLOGICAL ACOUSTICS [64]

Spectral-ripple representation of steady-state vowels in primary auditory cortex — Huib Versnel, Shihab A. Shamma — 2502

PSYCHOLOGICAL ACOUSTICS [66]

Within-channel gap detection using dissimilar markers in cochlear implant listeners — Monita Chatterjee, Qian-Jie Fu, Robert V. Shannon — 2515

Analysis of the performance of a model-based optimal auditory signal processor — Lisa C. Gresham, Leslie M. Collins — 2520

Beneficial effects of notched noise on intensity discrimination in the region of the ''severe departure'' — Christopher J. Plack — 2530

Auditory filter nonlinearity at 2 kHz in normal hearing listeners — Stuart Rosen, Richard J. Baker, Angela Darling — 2539

Modeling interaural-delay sensitivity to frequency modulation at high frequencies — Kourosh Saberi — 2551

Forward masked excitation patterns in multielectrode electrical stimulation — Monita Chatterjee, Robert V. Shannon — 2565

The masking-level difference in low-noise noise — Joseph W. Hall III, John H. Grose, William Morris Hartmann — 2573

The influence of stimulus envelope and fine structure on the binaural masking level difference — David A. Eddins, Laura E. Barber — 2578

Masker fluctuation and the masking-level difference — John H. Grose, Joseph W. Hall III — 2590

Spontaneous classification of complex tones at high and ultrasonic frequencies in the bat, *Megaderma lyra* — Annemarie Preisler, Sabine Schmidt — 2595

The pitch of a mistuned harmonic: Evidence for a template model — Jian-Yu Lin, William M. Hartmann — 2608

Monaural phase discrimination by macaque monkeys: Use of multiple cues — David B. Moody, Colleen G. Le Prell, Andrew J. Niemiec — 2618

Lateralization and binaural discrimination of patients with pontine lesions — Vered Aharonson, Miriam Furst, Robert A. Levine, Michael Chaigrecht, Amos D. Korczyn — 2624

Spectro-temporal factors in two-dimensional human sound localization — Paul M. Hofman, A. John Van Opstal — 2634

SPEECH PRODUCTION [70]

Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers — Michael Blomgren, Yang Chen, Manwa L. Ng, Harvey R. Gilbert — 2649

Automatic parametrization of differentiated glottal flow: Comparing methods by means of synthetic flow pulses — Helmer Strik — 2659

SPEECH PERCEPTION [71]

''How to milk a coat:'' The effects of semantic and acoustic information on phoneme categorization — Susan Borsky, Betty Tuller, Lewis P. Shapiro — 2670

Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration — Ken W. Grant, Brian E. Walden, Philip F. Seitz — 2677

SPEECH PROCESSING AND COMMUNICATION SYSTEMS [72]

An analysis of the effect of basilar membrane nonlinearities on noise suppression — Jayadev Billa, Amro El-Jaroudi — 2691

MUSIC AND MUSICAL INSTRUMENTS [75]

Jet-wave amplification in organ pipes — Shigeru Yoshikawa — 2706

(Continued)

CONTENTS—*Continued from preceding page*

BIOACOUSTICS [80]

Acoustic anisotropy in bovine cancellous bone	A. Hosokawa, T. Otani	2718

LETTERS TO THE EDITOR

Comments on ''Biot model of sound propagation in water-saturated sand'' [J. Acoust. Soc. Am. 97, 199–214 (1995)] [30]	R. D. Stoll	2723
Response to: ''Comments on 'Biot model of sound propagation in water-saturated sand' '' [J. Acoust. Soc. Am. 103, 2723–2725 (1998)] [30]	N. P. Chotiros	2726
Ultrasonic surface waves above rectangular-groove gratings [35]	Luc Kelders, Jean F. Allard, Walter Lauriks	2730
Comments on ''A re-examination of risk estimates from the NIOSH Occupational Noise, Hearing Survey'' [J. Acoust. Soc. Am. 101, 950–963 (1997)] [50]	Robert A. Dobie	2734
Response to ''Comments on 'A re-examination of risk estimates from the NIOSH Occupational Noise, Hearing Survey' '' [J. Acoust. Soc. Am. 103, 2734 (1998)] [50]	Mary M. Prince, Leslie T. Stayner, Randall J. Smith, Stephen J. Gilbert	2736
Species differences of distortion-product otoacoustic emissions: Comment on ''Interpretation of distortion product otoacoustic emission measurements. I. Two stimulus tones'' [J. Acoust. Soc. Am. 102, 413–429 (1997)] [64]	M. L. Whitehead	2740

CUMULATIVE AUTHOR INDEX

2743

NOTES CONCERNING ARTICLE ABSTRACTS

1. The number following the abstract copyright notice is a Publisher Item Identifier (PII) code that provides a unique and concise identification of each individual published document. This PII number should be included in all document delivery requests for copies of the article.
2. PACS numbers are for subject classification and indexing. See June and December issues for detailed listing of acoustical classes and subclasses.
3. The initials in brackets following the PACS numbers are the initials of the JASA Associate Editor who accepted the paper for publication.

Document Delivery: Copies of journal articles can be ordered from the new Articles in Physics online document delivery service (URL: http://www.aip.org/articles.html).

CONTENTS

page

Technical Program Summary (Westin) A8

Schedule of Technical Session Starting Times (Westin) A9

Technical Program Summary (Sheraton) A10

Schedule of Technical Session Starting Times (Sheraton) A11

Map of Seattle A12

Map of Meeting Rooms (Westin) A14

Map of Meeting Rooms (Sheraton) A16

Calendar—Technical Program A18

Calendar—Other Events A32

Meeting Information A34

Guidelines for Presentations A42

Dates of Future Meetings A43

Technical Sessions (1a_), Monday Morning 2747

Technical Sessions (1p_), Monday Afternoon 2783

Technical Sessions (2a_), Tuesday Morning 2823

Technical Sessions (2p_), Tuesday Afternoon 2860

Technical Sessions (3a_), Wednesday Morning 2904

Technical Sessions (3p_), Wednesday Afternoon 2938

ASA Medals and Awards Plenary Session, Wednesday Afternoon 2953

R. Bruce Lindsay Award Encomium 2953

Honorary Fellowship Award Encomium 2953

Helmholtz–Rayleigh Interdisciplinary Silver Medal Award Encomium 2953

Gold Medal Award Encomium 2953

Technical Sessions (4a_), Thursday Morning 2954

Technical Sessions (4p_), Thursday Afternoon 2993

Technical Sessions (5a_), Friday Morning 3032

Technical Sessions (5p_), Friday Afternoon 3063

Author Index to Abstracts 3097

Regional Chapters 3108

Application Forms 3110

Sustaining Members 3116

Index to Advertisers 3120

SOUNDINGS

This front section of the *Journal* includes acoustical news, views, reviews, and general tutorial or selected research artcles chosen for wide acoustical interest and written for broad acoustical readership.

ACOUSTICAL NEWS—USA

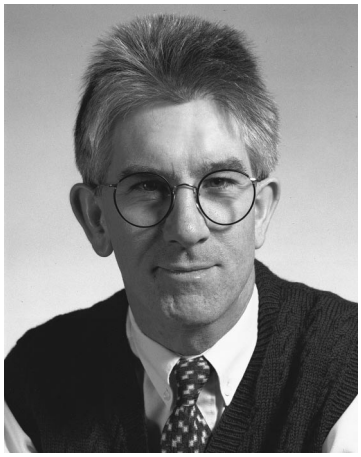
Elaine Moran
Acoustical Society of America, 500 Sunnyside Boulevard, Woodbury, New York 11797

Editor's Note: *Deadline dates for news items and notices are 2 months prior to publication.*

New Fellows of the Acoustical Society of America

Armand Dancer—For research on effects of intense sound on hearing.

John P. Seiler—For contributions to acoustical measurements and standards development.



Quentin Summerfield—For research in auditory grouping and speech perception.

Harvey Sussman—For contributions to the understanding of speech production.

The 134th meeting of the Acoustical Society of America held in San Diego, California

The 134th meeting of the Acoustical Society of America was held 1–5 December 1997 at the Town and Country Hotel in San Diego, California. This was the sixth time the Society has met in this city, the previous meetings being held in 1952, 1969, 1976, 1983, and 1990.

The meeting drew a total of 996 registrants. There were 175 nonmembers registered of whom over 33% joined the Society at the time of registration. Additionally, 189 students were in attendance.

Attesting to the international ties of our organization, 112 of the registrants (that is, about 11%) were from outside North America, the United States, Canada, and Mexico, which accounted for 854, 26, and 4, respectively. There were 31 registrants from Japan, 9 each from Sweden and France, 8 each from Germany and Russia, 7 from the United Kingdom, 6 from Denmark, 5 each from Brazil and Italy, 4 from the Netherlands, 3 each from Norway and Spain, and Australia, 2 each from Finland, Korea and Singapore, and 1 each from Austria, China, India, Israel, and Slovakia.

A total of 679 papers organized into 77 sessions covered the areas of interest of all 12 Technical Committees and 1 Technical Group. The meeting also included 21 different meetings dealing with standards, indicating a healthy level of activity in this area of our endeavors.

The local meeting committee arranged technical tours to the Hubbs Sea-World Research Institute and to FLIP, the 355-foot-long manned spar buoy developed and operated by the Marine Physical Laboratory of Scripps Institute of Oceanography.

The tutorial lecture series was continued at the San Diego meeting.

 0001-4966/98/103(5)/2241/4/$10.00

FIG. 1. President Lawrence Crum (l) and Executive Director Charles Schmid present the Distinguished Service Citation to Alice H. Suter.

FIG. 2. Robert. W. Young (c), new Honorary Fellow, receives congratulations from President Lawrence Crum and Vice President Ilene J. Busch-Vishniac (l), Evelyn Young and Robert S. Gales (r).

FIG. 3. Herman Medwin (c), recipient of the Silver Medal in Acoustical Oceanography, receives congratulations from President Lawrence Crum (l), Eileen Medwin and Clarence Clay (r).

FIG. 4. Robert E. Apfel (c), recipient of the Silver Medal in Physical Acoustics, is congratulated by Vice President Ilene Busch-Vishniac and President Lawrence Crum (l) and Nancy Apfel (r).

FIG. 5. Patricia K. Kuhl (c), recipient of the Silver Medal in Speech Communication, is congratulated by Vice President Ilene Busch-Vishniac, President Lawrence Crum (l) and Kenneth Stevens (r).

FIG. 6. Russell Johnson (c), recipient of the Wallace Clement Sabine Medal, with ASA President Lawrence Crum (l) and William Cavanaugh (r).

FIG. 7. President Lawrence Crum and Vice President Ilene Busch-Vishniac (r) present the Science Writing Award for Professionals in Acoustics to John R. Potter, Michael J. Buckingham, and Chad L. Epifanio (l–r).

FIG. 8. President Lawrence Crum (l) and Vice President Ilene Busch-Vishniac (r) present certificates to the newly elected Fellows of the Acoustical Society of America.

Whitlow W. L. Au of the Hawaii Institute of Marine Biology lectured on ''The Dolphin Echolocation System'' to an audience of about 100, attesting to the broad interest of the subject.

The plenary session included the presentation of awards, announcement of newly elected Fellows of the Society, and recognition of the meeting organizers. The President presented eight Society awards (see Figs. 1 through 7) and announced the election of 24 new Fellows. The Science Writing Award in Acoustics for Journalists was awarded to Richard Wolkomir for his article in the February 1996 issue of *Smithsonian* titled ''Decibel by Decibel, Reducing the Din to a Very Dull Roar.'' Mr. Wolkomir was unable to attend the meeting. The Science Writing Award for Professionals in Acoustics was presented to Michael J. Buckingham, John R. Potter, and Chad L. Epifanio for their article in the February 1996 issue of *Scientific American* titled ''Seeing Underwater with Background Noise.'' The Distinguished Service Citation was presented to Alice H. Suter of Alice H. Suter Associates, Ashland, Oregon, ''for increasing public awareness of acoustics and for dedicated service to the Society as editor of *Echoes*.'' Honorary Fellowship was conferred upon Robert W. Young, Consultant in Acoustics, San Diego, California, ''for contributions encompassing many areas of acoustics, to international standards and acoustical terminology, and for sustained and devoted service to the Society.'' The Silver Medal in Acoustical Oceanography was presented to Herman Medwin of Ocean Acoustics Associates, Pebble Beach, California, ''for contributions to the understanding of acoustical scattering, absorption and ambient noise, particularly in relation to the acoustics of bubbles in the sea.'' The Silver Medal in Physical Acoustics was presented to Robert E. Apfel of Yale University, New Haven, Connecticut, ''for contributions to the understanding of acoustic cavitation, acoustic radiation pressure, and the bioeffects of medical ultrasound.'' The Silver Medal in Speech Communication was presented to Patricia K. Kuhl of the University of Washington, Seattle, ''for contributions to the understanding of the innate and learned aspects of speech perception and production.'' The Wallace Clement Sabine Medal was presented to Russell Johnson of Artec Consultants, New York, New York ''for contributions to the understanding of the acoustics of performance spaces and the design of concert halls, theaters and opera houses throughout the world.'' Election of the following persons to Fellow grade was announced: Robert J. Bernhard, Nicholas P. Chotiros, Walter L. Clearwaters, Diana Deutsch, Gary W. Elko, Albert S. Feng, Shimshon Frankenthal, Andrew S. Harris, Yun-Fan Hwang, Akira Ishmaru, C. Scott Johnson, Gerald R. Kidd, Marjorie R. Leek, Brenda L. Lonsbury-Martin, Nicholas Makris, Matthew A. Nobile, Daniel R. Raichel, Richard J. Salvi, Klaus Scherer, Astrid Schmidt-Nielsen, Guido F. Smoorenburg, Arnold Tubis, Richard M. Warren, and George Winzer (see Fig. 8).

The President expressed the Society's thanks to the Local Committee for the excellent execution of the meeting, which clearly required meticulous planning. He introduced the Chair of the Meeting, Mauro Pierucci, who acknowledged the contributions of the members of his committee including: Michael J. Buckingham and Ann E. Bowles, Technical Program Cochairs; Grant B. Deane and Joel R. Bock, Audio-Visual; Paul A. Baxley, Socials and Receptions; Charles F. Greenlaw, Exhibits; Robert J. Vent, Facilities; Frederick H. Fisher, Public Relations; Mary Pierucci, Penny Buckingham, Janet Bock and Donna Baxley, Hospitality; D. Vance Holliday, Paper Copying Service; Jo Griffith and Gail Smith, Registration; Duncan E. McGhee, Signs; Francis X. Byrnes, Technical Tours, Linda Sax Crossland, Visitors' Program.

The President also extended thanks to the members of the Technical Program Organizing Committee: Michael J. Buckingham and Ann E. Bowles, Technical Program Cochairs; Darrell R. Jackson, Christian P. de Moustier, and Gerald R. D'Spain, Acoustical Oceanography; Ann E. Bowles and David A. Helweg, Animal Bioacoustics; David Lubman and Neil Shaw, Architectural Acoustics; Lawrence A. Crum, Bioresponse to Vibration/Biomedical Ultrasound; Scott D. Sommerfeldt, Education in Acoustics; Steven Baker, Engineering Acoustics; R. Dean Ayres, Musical Acoustics; Louis A. Sutherland, Noise; Andres Larraza and Thomas L. Hofler, Physical Acoustics; Robert V. Shannon, Psychological and Physiological Acoustics; David I. Havelock and Gary R. Wilson, Signal Processing in Acoustics; Terrance M. Nearey, Edward Auer, and Paul Iverson, Speech Communication; Courtney B. Burroughs and Scott D. Sommerfeldt, Structural Acoustics and Vibration; Newell O. Booth and Michael D. Collins, Underwater Acoustics; Gail Smith and Graziella Bruni, Coordinators.

LAWRENCE A. CRUM
President 1997–1998

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.

1998

13–15 May	28th Annual Symposium, Ultrasonic Industry Assn., King of Prussia, PA [Ultrasonic Industry Assn., Inc., P.O. Box 1420, Cherry Hill, NJ 08034-0054; Tel.: 609-424-8998; Fax: 609-424-9248].
4–7 June	7th Symposium on Cochlear Implants in Children, Iowa City, IA [Center for Conferences and Institutes, The University of Iowa, 249 Iowa Memorial Union, Iowa City, IA 52242-1317; Tel.: 800-551-9029; Fax: 319-335-3533].
20–26 June	135th meeting of the Acoustical Society of America/16th International Congress on Acoustics, Seattle, WA [ASA, 500 Sunnyside Blvd., Woodbury, NY 11797, Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org, WWW: http://asa.aip.org].
26 Jun.–1 Jul.	International Symposium on Musical Acoustics, ISMA 98, Leavenworth, WA [Maurits Hudig, Catgut Acoustical Society, 112 Essex Ave., Montclair, NJ 07042, Fax: 201-744-9197; E-mail: catgutas@msn.com, WWW: www.boystown.org/isma98].
7–12 July	Vienna and the Clarinet, Ohio State Univ., Columbus, OH [Keith Koons, Music Dept., Univ. of Central Florida, P.O. Box 161354, Orlando, FL 32816-1354, Tel.: 407-823-5116; E-mail: kkons@pegasus.cc.ucf.edu].

12–15 July Transportation Research Board A1F04 Summer Meeting on Transportation Related Noise and Vibration, Tallahassee, FL [Win Lindeman, Florida Department of Transportation, 605 Suwannee St., M.S. 37, Tallahassee, FL 32399-0450, Tel.: 850-488-2914; Fax: 850-922-7217; E-mail: win.lindeman@dot.state.fl.us].

9–14 Aug. International Acoustic Emission Conference, Hawaii [Karyn S. Downs, Lockheed Martin Astronautics, P.O. Box 179, M.S. DC3005, Denver, CO 80201; Tel.: 303-977-1769; Fax: 303-971-7698; E-mail: karyn.s.downs.@lmco.com].

13–17 Sept. American Academy of Otolaryngology—Head and Neck Surgery, San Francisco, CA [American Academy of Otolaryngology—Head and Neck Surgery, One Prince St., Alexandria, VA 22314; Tel.: 703-836-4444; Fax: 703-683-5100].

18–19 Sept. 6th Annual Conference on Management of the Tinnitus Patient, Iowa City, IA [Richard Tyler, Univ. of Iowa, Dept. of Otolaryngology—Head & Neck Surgery, 200 Hawkins Dr., C21GH, Iowa City, IA 52242; Tel.: 319-356-2471; Fax: 319-353-6739; E-mail: rich-tyler@uiowa.edu].

12–16 Oct. 136th meeting of the Acoustical Society of America, Norfolk, VA [ASA, 500 Sunnyside Blvd., Woodbury, NY 11797, Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org, WWW: http://asa.aip.org].

15–19 March Joint meeting: 137th meeting of the Acoustical Society of America/Forum Acusticum [Acoustical Society of America, 500 Sunnyside Blvd., Woodbury, NY 11797, Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; WWW: asa.aip.org].

27–30 June ASME Mechanics and Materials Conference, Blacksburg, VA [Mrs. Norma Guynn, Dept. of Engineering Science and Mechanics, Virginia Tech, Blacksburg, VA 24061-0219; Fax: 540-231-4574; E-mail: nguynn@vt.edu; WWW: http://www.esm.vt.edu/mmconf/]. Deadline for receipt of abstracts: 15 January 1999.

ACOUSTICAL STANDARDS NEWS

Avril Brenig, Standards Manager
Standards Secretariat, Acoustical Society of America, 120 Wall Street, 32nd Floor, New York, New York 10005-3993 [Tel.: (212) 248-0373; Fax: (212) 248-0146; e-mail: asastds@aip.org]

George S. K. Wong
Acoustical Standards, Institute for National Measurement Standards, National Research Council, Ottawa, Ontario K1A 0R6, Canada [Tel.: (613) 993-6159; Fax: (613) 990-8765; e-mail: george.wong@nrc.ca]

American National Standards (ANSI Standards) developed by Accredited Standards Committees S1, S2, S3, and S12 in the areas of acoustics, mechanical vibration and shock, bioacoustics, and noise, respectively, are published by the Acoustical Society of America (ASA) through the American Institute of Physics (AIP). In addition to these standards, a Catalog of Acoustical Standards—ASA Catalog 17-1997—is available. For a copy, contact Avril Brenig.

Comments are welcomed on all material in Acoustical Standards News.

Standards Meetings Calendar—National

The next meetings of the Accredited Standards Committees for national standards development, the review of International Standards, and planning of future meetings are scheduled to be held during the joint 135th ASA/16th ICA Meetings at the Westin Hotel and the Seattle Sheraton Hilton Hotel, Seattle, WA, 20–28 June 1998. Meetings of the Accredited Standards Committees are open to the public for participation. Organizations and individuals with expertise that may enhance the performance of the acoustical standards effort are encourage to participate.

• ***Monday, 22 June 1998, 7:00 p.m***.
ASACOS Steering Committee

• ***Tuesday, 23 June 1998, 7:30 a.m.***
ASA Committee on Standards (ASACOS). Meeting of the committee that directs the Acoustical Society Standards Program

• ***Wednesday, 24 June 1998, 9:00 a.m.***
Accredited Standards Committee S2 on Mechanical Vibration and Shock (also Technical Advisory Group for ISO/TC 108 Mechanical Vibration and Shock as well as ISO/TC 108/SC1, Balancing, Including Balancing Machines; ISO/TC 108/SC2 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/ SC5 Condition Monitoring and Diagnostics of Machines, and ISO/TC108/SC 6 Vibration and Shock Generating Equipment).

• ***Thursday, 25 June 1998, 8:00 a.m.***
Standards Committee Plenary Group
[To discuss national items relevant to all S Committees (S1, S2, S3, and S12), plus a review of the international standardization activities under ISO and IEC for which the S Committees form the respective U.S. Technical Advisory Groups (TAGS)]

• ***Thursday, 25 June 1998, 9:30 a.m.***
Accredited Standards Committee S12 on Noise (also Technical Advisory Group for ISO/TC 43/SC 1—Noise and ISO/TC 94/SC 12—Hearing Protection)

• ***Thursday, 25 June 1998, 2.00 p.m.***
Accredited Standards Committee S3 on Bioacoustics (also Technical Advisory Group for ISO/TC 43 Acoustics, IEC/TC 29—Electroacoustics, and ISO/TC 108/SC 4—Human Exposure to Mechanical Vibration and Shock)

• ***Thursday, 25 June 1998, 3.30 p.m.***
Accredited Standards Committee S1 on Acoustics (also Technical Advisory Group for ISO/TC 43—Acoustics and IEC/TC 29—Electroacoustics)

American National Standards Institute (ANSI) Calendar

18 May 1998
• ANSI Seminar: For the Good of the People, Washington, DC

19 May 1998
• Organizational Member Council (OMC) Meeting, Washington, DC

20 May 1998
• ANSI/International Advisory Committee Forum (AIF), Washington, DC
• (AIF)/International Advisory Committee Forum, Washington, DC

21 May 1998
• (IAC)/International Advisory Committee, Washington, DC

1 June 1998
• ISO Technical management Board—(Day 1 of 2), Geneva, Switzerland
• ANSI Seminar: For the Good of the People, Chicago, IL

2 June 1998
• ISO Technical management Board—(Day 2 of 2), Geneva, Switzerland
• ANSI Seminar: The American Way, Chicago, IL

3 June 1998
• ANSI Seminar**:** Participate Effectively, Chicago, IL

4 June 1998
• ISO Council—(Day 1 of 2), Geneva, Switzerland

5 June 1998
• ISO Council—(Day 2 of 2), Geneva, Switzerland

10 June 1998
• Executive Standards Council (ExSC) Meeting—(Day 1 of 2), Quincy, MA

11 June 1998
• Executive Standards Council (ExSC) Meeting—(Day 2 of 2), Quincy, MA

23 July 1998
• ANSI Board of Directors Executive Committee Meeting, New York, NY

19 August 1998
• ANSI/International Advisory Committee Forum, New York, NY
• (AIF)/International Advisory Committee Forum, New York, NY

26 August 1998
• (IAC)/International Advisory Committee, New York, NY

9 September 1998
• Executive Standards Council (ExSC) Meeting—(Day 1 of 2), New York, NY

10 September 1998
• Executive Standards Council (ExSC) Meeting—(Day 2 of 2), New York, NY

14 September 1998
• ISO Technical Management Board—(Day 1 of 2), Geneva, Switzerland

15 September 1998
• ISO Technical Management Board—(Day 2 of 2), Geneva, Switzerland

16 September 1998
• ISO General Assembly—(Day 1 of 2), Geneva, Switzerland

 0001-4966/98/103(5)/2248/6/$10.00

17 September 1998
• ISO General Assembly—(Day 2 of 2), Geneva, Switzerland

18 September 1998
• ISO Council, Geneva, Switzerland

21 September 1998
• ANSI Seminar: The American Way, Washington, DC

22 September 1998
• Organizational Member Council (OMC) Meeting, Washington, DC
• ANSI Seminar: American Access to the Regional Standardization Process, Washington, DC
• ANSI Seminar: Author Artisan and Statesman, Washington, DC

24 September 1998
• ANSI Board of Directors Meeting—(Day 1 of 2)
• ANSI Seminar: For the Good of the People, Washington, DC

25 September 1998
• ANSI Board of Directors Meeting—(Day 2 of 2)
• ANSI Seminar: Strengthening the U.S. Voice in International Standard

12 October 1998
• International Electrotechnical Commission (IEC) 62nd General Meeting, Houston, TX

27 October 1998
• Information Systems Standards Board (ISSB) Meeting, New York, NY

28 October 1998
• Information Systems Standards Panel (IISP) Meeting—(Day 1 of 2)

29 October 1998
• Information Systems Standards Panel (IISP) Meeting—(Day 2 of 2)

Standards News from the United States

(Partially derived from *ANSI Reporter*, and *ANSI Standards Action*, with appreciation)

Call for comment on proposed American National Standards and American National Standards proposed for Reaffirment or Withdrawal

This section lists proposed American National Standards that solicit public comments and review. The final dates for offering comments, listed in parentheses, are for information only. Copies of these documents are available from ANSI.

SHOCK AND VIBRATION (30 September 1997)

BSR S2.2-1959, Methods for the Calibration of Shock and Vibration Pickups (**reaffirmation of** ANSI S2.2-1959 (R1990))

Selects five methods as the standard methods for the calibration of acceleration, velocity, and displacement pickups. They are described in Section 4 of this standard. It is impracticable to calibrate all pickups by one standard method over the entire frequency and amplitude range of vibrations and shocks to be measured by the pickups. Several methods are accordingly described. Each method is limited to a range of frequency and amplitude, and to the weight of pickup that can be calibrated. The limitations may include, in addition, other variables such as volume of the pickup and temperature of operation.

BSR S2.4-1976, Method for Specifying the Characteristics of Auxiliary Analog Equipment for Shock and Vibration Measurements (**reaffirmation of** ANSI S2.4-1976 (R1990))

Applies to the auxiliary equipment used between a shock or vibration transducer and the final indicator, recorder, or signal processor. This document presents a standard format for indicating pertinent characteristics, but does not in any respect become a standard on the performance of the equipment. Since this standard was prepared to cover a wide variety of equipment in considerable detail, not all items will be pertinent to a specific piece of equipment. Also, it is not the intent of this standard to establish an ironclad rule as to which of the characteristics should be included, although in many cases important characteristics are emphasized.

BSR S2.5-1962, Recommendations for Specifying the Performance of Vibration Machines (**reaffirmation of** ANSI S2.5-1962 (R1990))

Examines three types of vibration machines that are used for exciting vibrations in equipment or structures, in order to provide assurance that such equipment or structures are able to withstand the conditions of vibration they may encounter in later service. Shakers, such as are used for sifting, mixing, or unloading operations, and machines designed specifically for fatigue testing are excluded. Three types of vibration machines are considered: the mechanical direct-drive: the mechanical reaction type, which includes the vibration exciter and vibration machine as separate items; and the electrodynamic. Other types, involving hydraulic excitation, resonant systems, piezoelectric application for instrument calibrations, and special vibration machines for testing packages come under the general scope of this standard: however, the exploratory nature of this work does not make a complete coverage of this field appear feasible at this time.

BSR S2.7-1982, Balancing Terminology (**reaffirmation of** ANSI S2.7-1982 (R1986))

Contains a collection of terms and definitions relating to balancing procedures and equipment. The terms are grouped into seven major categories covering the subject of (1) machines, (2) rotors, (3) unbalance, (4) balancing, (5) balancing machines and equipment, (6) flexible rotors, and (7) rotating rigid free-bodies (i.e., aerospace vehicles). An alphabetical index follows at the end.

BSR S2.8-1972, Resilient Mountings, Guide for Describing the Characteristics of Resilient Mountings (**reaffirmation of** ANSI S2.8-1972 (R1986))

Outlines, in standardized form, what data should be presented to enable the experienced designer to apply resilient mountings correctly. Also, the standard defines terminology, in a further effort to ease the problem of communication between user and manufacturer. It is beyond the scope of this standard to present characteristics of resilient mountings.

BSR S2.9-1976, Nomenclature for Specifying Damping Properties of Materials (**reaffirmation of** ANSI S2.9-1976 (R1990))

Presents the preferred nomenclature (parameters, symbols, and definitions) for specifying the damping properties of uniform materials and uniform specimens, where ''uniform'' implies homogeneity on a macroscopic scale. For reference purposes, certain nonstandard damping parameters are discussed (Appendix A). Also included is a classification of damping test parameters (Appendix B). The term ''damping'' as used in this standard is restricted to those properties of a material that are related to the process of cyclic energy dissipation during harmonic vibration. This definition specifying excludes: energy transfer devices such as dynamic absorbers; external damping processes such as air-damping or acoustic radiation; and energy dissipation during transient or random vibration.

BSR S2.10-1971, Methods for Analysis and Presentation of Shock and Vibration Data (**reaffirmation of** ANSI S2.10-1971 (R1990))

Covers vibration in the following idealized classes that are described in 2.2§ 1) Periodic vibration; (2) Aperiodic vibration; (3) Random vibration; (4) Transient vibration (including shock). It is assumed that the data are available as time-histories of a variable associated with shock and vibration, for example, acceleration, velocity, force, etc., and that any distortion resulting from the transducer, recording system, etc., has been eliminated. Also, assumed that the mean value of a random vibration is zero. It is recognized that in many instances the vibration does not conform to this classification of two or more classes. Suggestions are given for separating the classes as a necessary step in data analysis.

BSR S2.11-1969, Calibrations and Tests for Electrical Transducers Used for Measuring Shock and Vibration, Selection of (**reaffirmation of** ANSI S2.11-1969 (R1986))

Includes considerations relevant to commonly employed electromechanical shock and vibration measurement transducers, but not to those transducers primarily designed for measurement of acoustic or pressure phenomena. In general, the tests and calibrations listed are pertinent to the transducer portion of integrated measuring systems, but are not intended to be comprehensive in regard to associated circuitry or electronic accessory instruments.

BSR S2.14-1973, Methods for Specifying the Performance of Shock Machines (**reaffirmation of** ANSI S2.14-1973 (R1986))

Details the shock machines employed for diagnostic testing and for demonstration of evaluation of the effect of shock conditions representative of use environment. Several techniques for generating the desired shock motion are discussed. Two basic types of pulse shock machines are considered: those that provide the shock by a controlled reduction of the velocity of the test specimen (declaration or impact), and those that apply the shock

by a controlled increase in the specimen velocity (acceleration or impulse). A vibration machine employed to generate controlled transient vibratory shocks and some pulse shocks is also discussed.

BSR S2.15-1972, Specification for the Design, Construction, and Operation of Class HI (High Impact) Shock-Testing Machine for Lightweight Equipment (**reaffirmation of** ANSI S2.15-1972 (R1986))

Describes the design and construction of a class HI (High-Impact) shock-testing machine. By means of this standardized design, the ability of various types of equipment to withstand shock loadings may be compared. Also outlines recommended test procedures.

BSR S2.17-1980, Machinery Vibration Measurement (**reaffirmation of** ANSI S2.17-1980 (R1986))

Details vibration measurement quantities, equipment, and procedures involved in operating machinery. Calibration of vibration measurement and calibration equipment are discussed in the document. The standard has applications to preventive maintenance programs, equipment selection, and equipment quality.

BSR S2.19-1980, Mechanical Vibration—Balance Quality Requirements of Rigid Rotors—Part 1, Determination of Permissible Residual Unbalance (**reaffirmation of** ANSI S2.19-1989)

Contains recommended balance tolerances for various types of rotors and a detailed explanation of how they are to be understood for selected rotor planes. This document also deals with methods of allocating the recommended permissible residual unbalance to rotors with narrowly spaced, overhung, and/or unsymmetrically located correction planes. This standard is the U.S. counterpart of ISO 1940/1-1986, *Mechanical Vibration Balance Quality Requirements of Rigid Rotors—Part 1: Determination of Permissible Residual Unbalance*. It is a revised version of the prior standard ANSI S2.19-1975, American National Standard for Balance Quality of Rotating Rigid Bodies.

BSR S2.20-1983, Air Blast Characteristics for Single Point Explosions in Air, with a Guide to Evaluation of Atmospheric Propagation and Effects (**reaffirmation of** ANSI S2.20-1983 (R1989))

Provides consensus quantitive definitions of explosion characteristics for a single point explosion in air, along with methodologies for scaling these characteristics for a wide range of yield and ambient air conditions. Factors for use with common solid explosives are also included. Methods are provided for predictions of long range propagation under atmospheric refractive influences. Target damage estimation procedures are provided for use in explosion operational planning and evaluation.

BSR S2.31-1979, Methods for the Experimental Determination of Mechanical Mobility, Part I: Basic Definitions and Transducers (**reaffirmation of** ANSI S2.31-1979 (R1986))

Covers basic concepts and definitions and serves as a guide for the selection, calibration, and evaluation of the transducers and instruments used in mobility measurements. The material in Part I is common to most mobility measurement tasks. This standard is the first part of a series of five standards covering the experimental determination of mechanical mobility of structures by a variety of methods appropriate for different test situations. This document supersedes ANSI S2.6-1963 (R1976).

BSR S2.32-1982, Methods for the Experimental Determination of Mechanical Mobility, Part II: Measurements Using Single-Point Translation Excitation (**reaffirmation of** ANSI S2.32-1982 (R1990))

Comprises the second part of a set of five standards covering the experimental determination of mechanical mobility of structures by a variety of methods appropriate for different test situations. Part I of the set (American National Standard Methods for the Experimental Determination of Mechanical Mobility, Part I: Basic Definitions and Transducers, ANSI S2.31-1979) covers basic concepts and definitions as well as instruments used in mobility measurements. The material in Part I is common to most mobility measurement tasks. The present Part II of the set describes measurements in situations in which single-point translational excitation with an attached vibration exciter is appropriate.

BSR S2.34-1984, Guide to the Experimental Determination of Rotational Mobility Properties and the Complete Mobility Matrix (**reaffirmation of** ANSI S2.34-1984 (R1990))

Determines the mobility properties of structures derived from the complex amplitudes of translational and rotational responses and the complex amplitudes of excitation forces and moments within the audio frequency range.

BSR S2.38-1982, Field Balancing Equipment—Description and Evaluation (**reaffirmation of** ANSI S2.38-1982 (R1990))

Provides the description and evaluation of portable equipment used for field (in-place) balancing of rotating machinery. It tells the equipment manufacturer which performance parameters need to be specified and in what terms, so that the prospective user can assess the applicability of the equipment for his purpose. At the same time, a prospective user is given guidance on how to describe his requirements to the equipment manufacturer.

BSR S2.40-1984, Mechanical Vibration of Rotating and Reciprocating Machinery—Requirements for Instruments for Measuring Vibration Severity (**reaffirmation of** ANSI S2.40-1984 (R1990))

Establishes the requirements of instrumentation for accurately measuring the vibration severity of machinery. Limitations for frequency, sensitivity, amplitude range, calibration, and environmental factors are presented in this standard.

BSR S2.41-1985, Mechanical Vibration of Large Rotating Machines with Speed Range from 10 to 200 rev/s—Measurement and Evaluation of Mechanical Vibration Severity *in situ* (**reaffirmation of** ANSI S2.41-1985 (R1990))

Describes measurement and evaluation of vibration severity of large rotating machinery *in situ* and is the U.S. counterpart of ISO 3945-1977. This standard, which is in complete technical agreement with ISO 3945-1977, is not applicable to reciprocating machinery. The values of vibration severity recommended as limits in this standard are intended to serve as standard values for machines of similar type, when measured in accordance with the procedures described in this standard. The recommended vibration limits may be used for acceptance standards or for monitoring the satisfactory performance of the machine during service operations.

BSR S2.42-1982, Procedures for Balancing Flexible Rotors (**reaffirmation of** ANSI S2.42-1982 (R1990))

Classifies rotors into groups by their balancing requirements as influenced by the rotor's flexural stiffness and unbalance distribution. Certain classes of rotors may be balanced by normal or modified rigid rotor techniques. Other more flexible rotors may require high-speed balancing. The fundamentals of flexible rotor balancing are discussed as well as methods of assessment of final unbalance. Guidance is given in judging the final balance quality; however, this standard is not intended to serve as an acceptance specification for any rotor group. It is offered to provide direction on how to avoid gross deficiencies or unattainable requirements.

BSR S2.43-1984, Criteria for Evaluating Flexible Rotor Balance (**reaffirmation of** ANSI S2.43-1984 (R1990))

Specifies criteria for evaluating flexible rotors in a balancing facility in terms of (a) permissible vibration at specified measuring points (derived from ISO 2372-1974, Mechanical vibration of machines with operating speeds from 10 to 200 rev/s-Basis, for specifying evaluation standards, and ISO 3945-1977, Mechanical vibrations of large rotating machines with speed range from 10 to 200 rev/s—Measurement and evaluation of vibrations severity *in situ*; and (b) permissible residual unbalance in specified correction planes [derived from ASA STD 2-1975 (ANSI 2.19-1974), American National Standard for Mechanical Vibration—Balance Quality Requirements of Rigid Rotors—Part I, Determination of Permissible Residual Unbalance]. This standard is the U.S. counterpart of ISO 5343-1983 and is to be used in conjunction with American National Standard Procedures for Balancing Flexible Rotors, ANSI S2.42-1982 (ASA 46-1982)

BSR S2.45-1983, Electrodynamic Test Equipment for Generating Vibration—Method of Describing Equipment Characteristics (**reaffirmation of** ANSI S2.45-1983 (R1990))

Provides a method for specifying the characteristics of electrodynamic test equipment for generating vibration and serves as a guide to the selection of such equipment. It applies to electrodynamic vibration generators and power amplifiers, both individually and in combination. The standard provides means to assist a prospective user to calculate the performance of equipment provided by two or more manufacturers, even if the vibration generator and the power amplifier are from different manufacturers.

BSR S2.46-1989, Characteristics to be Specified for Seismic Transducers (**reaffirmation of** ANSI S2.46-1989 (R1991))

Specifies the characteristics of a seismic transducer (pickup) that can serve as an adequate description of the capabilities of the pickup to perform a shock or vibration measurement in various environments. It is intended as a guide to instruct in specifying, selecting, or preparing performance description by users. This standard is the national counterpart of ISO 8042-

1988, *Shock and Vibration Measurements—Characteristics to be Specified for Seismic Pick-Ups.*

BSR S2.47-1990, Vibration of Buildings—Guidelines for the Measurement of Vibrations and Evaluation of Their Effects on Buildings (**reaffirmation of** ANSI S2.47-1990)

Provides guidelines for the measurement of building vibrations and evaluation of their effects on buildings. It is intended to establish the basic principles for carrying out vibration measurement and processing data, with regard to evaluating vibration effects on buildings. The evaluation of the effects of building vibration is primarily directed at structural response, and includes appropriate analytical methods where the frequency, duration, and amplitude can be defined. It is the U.S. counterpart of the International Standard, ISO 4866-1990.

BSR S2.48-1993, Servo-Hydraulic Test Equipment for Generating Vibration—Methods of Describing Characteristics (**reaffirmation of** ANSI S2.48-1993)

Provides a method for specifying the characteristics of servo-hydraulic test equipment for generating vibration and serves as a guide to the selection of such equipment. It applies to servo-hydraulic vibration generators and power amplifiers, both individually and in combination. The standard provides means to assist a prospective user to calculate and compare the performance of equipment provided by two or more manufacturers, even if the vibration generator and the power amplifier are from different manufacturers.

BSR S2.58-1983, Auxiliary Tables for Vibration Generators—Methods of Describing Equipment Characteristics (**reaffirmation of** ANSI S2.58-1983 (R1990))

Provides a method for specifying the characteristics of eight types of auxiliary tables for vibration generators. It serves as a guide to the prospective user of auxiliary tables to assist in the objective comparison of the performance of auxiliary tables available from different manufacturers.

BSR S2.60-1987, Balancing Machines—Enclosures and Other Safety Measures (**reaffirmation of** ANSI S2.60-1987)

Describes a balancing machine's potential hazards to its operator and the surrounding workshop area, relates accident probabilities, and enumerators precautionary measures. The major portion of the standard deals with the danger from rotor particles or components separating during rotation, and specifies safety measures in five different protection classes. The effects of fragment mass, shape, material, and velocity are discussed and a standard projectile for penetration resistance testing is established. Finally, a table lists parameters for fragment enclosures for a representative series of general-purpose balancing machines.

BSR S2.61-1989, Guide to the Mechanical Mounting of Accelerometers (**reaffirmation of** ANSI S2.61-1989 (R1991))

Specifies methods for mounting contact accelerometer and delineates the limitations of the recommended methods, such as frequency range of interest, amplitude, and phase measurement accuracy. Characteristics of the mounting arrangements used by specific accelerometers, which should be specified by the manufacturer are established and guidance is provided to the user to optimize performance of a recommended mounting method.

BSR S3.34-1986, Human Exposure to Vibration Transmitted to the Hand, Guide for Measurement and Evaluation of (**reaffirmation of** ANSI S3.34-1986 (R1992))

Specifies the recommended method for the measurement, data analysis, and reporting of human exposure to hand-transmitted vibration. A standard format is established for measurement, data analysis, and reporting of hand-transmitted vibration, periodic or random, in up to three orthogonal axes, in the frequency range from 5.6 Hz to 1400 Hz. Three Appendices, which are not part of the standard, provide information for the assessment of exposure to hand-transmitted vibration, on the latent period for the development of symptoms due to hand-transmitted vibration, and on further publications on the subject.

ACCELEROMETERS

BSR/IEEE 836-1992 (R1997), Recommended Practice for Precision Centrifuge Testing of Linear Accelerometers (**reaffirmation of** ANSI/IEEE 836-1992) (11 Nov. 1997)

Describes the conduct and analysis of precision tests that are to be performed on linear accelerometers using centrifuge techniques.

Final actions on American National Standards

ANSI's Board of Standards Review has taken the final action indicated on the standards listed below. When the approved standards are published, an announcement will be carried in *Standards Action*, in ANSI's Supplements to the Catalog of American National Standards, and in trade publications.

ACOUSTICS

ANSI S1.4-1983 (R1997), Specification for Sound Level Meters (**reaffirmation of** ANSI S1.4-1983 (R1994)): 29 July 1997.

ANSI S1.6-1984 (R1997), Preferred Frequencies, Frequency Levels, and Band Numbers for Acoustical Measurements (**reaffirmation of** ANSI S1.6-1984 (R1994)): 29 July 1997.

ANSI S1.8-1989 (R1997), Reference Quantities for Acoustical Levels (**reaffirmation of** ANSI S1.8-1989): 29 July 1997.

ANSI S1.10-1966 (R1997), Method for the Calibration of Microphones (**reaffirmation of** ANSI S1.10-1966 (R1986)): 7 July 1997.

ANSI S1.12-1967 (R1997), Specifications for Laboratory Standard Microphones (**reaffirmation of** ANSI S1.12- 1967) (R1986)): 29 July 1997.

ANSI S1.22-1992 (R1997), Scales and Sizes for Frequency Characteristics and Polar Diagrams in Acoustics (**reaffirmation of** ANSI S1.22-1992): 29 July 1997.

ANSI S1.25-1991 (R1997), Specification for Personal Noise Dosimeters (**reaffirmation of** ANSI S1.25-1991): 29 July 1997.

ANSI S1.40-1984 (R1997), Specification for Acoustical Calibrators (**reaffirmation of** ANSI S1.40-1984): 29 July 1997.

ANSI S3.14-1997 (R1997), Rating Noise with Respect to Speech Interference (**reaffirmation of** ANSI S3.14-1997 (R1986)): 29 July 1997.

ANSI S12.5-1990 (R1997), Requirements for the Performance and Calibration of Reference Sources (**reaffirmation of** ANSI S12.5-1990): 29 July 1997.

ANSI S12.10-1985 (R1997), Methods for the Measurement and Designation of Noise Emitted by Computer and Business Equipment (**reaffirmation of** ANSI S12.10-1985 (R1990)): 29 July 1997.

ANSI S12.11-1987 (R1997), Method for the Measurement of Noise Emitted by Small Air-Moving Devices (**reaffirmation of** ANSI S12.11-1987 (R1993)): 29 July 1997.

ANSI S12.12-1992 (R1997), Engineering Method for the Determination of Sound Power Levels of Noise Sources Using Sound Intensity (**reaffirmation of** ANSI S12.12-1992): 29 July 1997.

ANSI S12.14-1992 (R1997), Methods for the Field Measurement of the Sound Output of Audible Public Warning Devices Installed at Fixed Locations Outdoors (**reaffirmation of** ANSI S12.14-1992): 29 July 1997.

ANSI S12.15-1992 (R1997), Acoustics—Portable Electric Power Tools, Stationary and Fixed Electric Power Tools, and Gardening Appliances—Measurement of Sound Emitted (**reaffirmation of** ANSI S12.15-1992): 29 July 1997.

ANSI S12.16-1992 (R1997), Guidelines for the Specification of Noise of New Machinery (**reaffirmation of** ANSI S12.16-1992): 29 July 1997.

ANSI S12.30-1990 (R1997), Guidelines for the Use of Sound Power Standards and for the Preparation of Noise Test Codes (**reaffirmation of** ANSI S12.30-1990): 29 July 1997.

ANSI S12.33-1990 (R1997), Engineering Methods for the Determination of Sound Power Levels of Noise Sources in a Special Reverberation Test Room (**reaffirmation of** ANSI S12.33-1990): 29 July 1997.

ANSI S12.34-1988 (R1997), Engineering Methods for the Determination of Sound Power Levels of Noise Sources for Essentially Free-Field Conditions over a Reflecting Plane (**reaffirmation of** ANSI S12.34-1988 (R1993)): 29 July 1997.

ANSI S12.36-1990 (R1997), Survey Methods for the Determination of Sound Power Levels of Noise Sources (**reaffirmation of** ANSI S12.36-1990): 29 July 1997.

BIOACOUSTICS

ANSI S3.4-1990 (R1997), Procedure for the Computation of Loudness of Noise (**reaffirmation of** ANSI S3.4-1980 (R1992)): 29 July 1997.

ANSI S3.13-1987 (R1997), Mechanical Coupler for Measurement of Bone Vibrators (**reaffirmation of** ANSI S3.13-1987 (R1993)): 29 July 1997.

ANSI S3.21-1978 (R1997), Method for Manual Pure-Tone Threshold Audiometry (**reaffirmation of** ANSI S3.21-1978 (R1992)): 29 July 1997.

ANSI S3.35-1985 (R1997), Method of Measurement of Performance Characteristics of Hearing Aids under Simulated *in-situ* Working Conditions (**reaffirmation of** ANSI S3.35-1985 (R1990)): 29 July 1997.

ANSI S3.37-1987 (R1997), Preferred Earhook Nozzle Thread for Postauricular Hearing Aids (**reaffirmation of** ANSI S3.37-1987 (R1992)): 29 July 1997.

ANSI S3.42-1992 (R1997), Testing Hearing Aids with a Broad-Band Noise Signal (**reaffirmation of** ANSI S3.42-1992): 29 July 1997.

LOUDSPEAKERS

ANSI S4.56-1997, Recommended Practice for Professional Audio—Subjective Evaluation of Loudspeakers (new standard): 30 September 1997.

Standards News from Abroad

(Partially derived from ANSI Reporter and ANSI Standards Action, with appreciation)

Newly published ISO and IEC Standards and Technical Reports

This section lists new and revised standards and technical reports recently approved and promulgated by ISO and IEC. These documents together with catalog listing several thousand current ISO and IEC standards are available from ANSI.

ISO Standards

ACOUSTICS (TC 43)

ISO 10847: 1997, Acoustics—*In-situ* determination of insertion loss of outdoor noise barriers of all types.
ISO 10843: 1997, Acoustics—Methods for the description and physical measurement of single impulses or series of impulses.
ISO 11819-1: 1997, Acoustics—Measurement of the influence of road surfaces on traffic noise—Part 1: Statistical Pass-By method.

International documents submitted to the U. S. for vote and/or comment

Some of the documents processed recently by the ASA Standards Secretariat. Dates in parentheses are deadlines for submission of comments and recommendation for vote, and they are for information only.

U.S. TAG	ISO Documents
S3 & S12	**1st ISO/CD 12124**—''Procedures for the Measurement of real-ear Characteristics of hearing aids,'' 4 June 1997.
S12	**ISO/DIS 362**—Acoustics—Measurement of Noise Emitted by Accelerating Road Vehicles—Engineering Method (Affirmative w/Comments), 7 February 1997.
	ISO/DIS 14163—Acoustics—Guidelines for Noise Control for Silencers (Negative w/Comments), 31 January 1997.
	ISO/DIS 1680—Acoustics—Test Code for the Measurement of Airborne Noise Emitted by Rotating Electrical Machinery (Negative w/Comments), 7 February 1997.
	ISO 1996/2:1987 DAM 1—Acoustics—Description and Measurement of Environmental Noise—Part 2: Acquisition of Data Pertinent to Land Use, Amendment 1 (Affirmative w/Comments), 12 March 1997.
	ISO/DIS 3741—Acoustics—Determination of Sound Power Levels of Noise Sources Using Sound Pressure—Precision Methods for Reverberation Rooms (Negative w/Comments), 31 January 1997.
	ISO/DIS 10846-3—Acoustics and Vibration—Laboratory Measurements of Vibro-acoustic transfer Properties of Resilient Elements, Part 3: Dynamic Stiffness of Elastic Supports for Translatory Motion—Indirect Method (Affirmative w/o Comments), 19 June 1997.
	ISO/FDIS 10847—Acoustics—*In situ* determination of Insertion Loss of Outdoor Noise Barriers of All Types (Negative w/Comments), 30 April 1997.
	ISO/FDIS 10846-1—Acoustics and Vibration—Laboratory Measurements of Vibro-acoustic transfer Properties of Resilient Elements, Part 1: Principles and Guidelines, 18 June 1997.
	ISO/FDIS 10846-2—Acoustics and Vibration—Laboratory Measurements of Vibro-acoustic transfer Properties of Resilient Elements, Part 2: Dynamic Stiffness of Elastic Supports for Translatory Motion—Direct Method, 18 June 1997.
	ISO/DTR 11688-2—Acoustics—Recommended Practice for the Design of Low-noise Machinery and Equipment—Part 2: Introduction into Physics of Low-noise Design (Affirmative w/o Comments), 24 January 1997.
	2nd ISO/CD 3740—Determination of Sound Power Levels of Noise Sources—Guidelines for the Use of Basic Standards (Affirmative w/Comments), 13 May 1997.
	4th ISO/CD 3747—Determination of Sound Power Levels of Noise Sources Using Sound Pressure—Engineering Comparison Method for Use in an Approximately Reverberant Field *in situ* (Affirmative w/Comments), 13 May 1997.
	1st ISO/CD 3891-1—''Determination, Measurement and Management of Aircraft Noise—Part 1: Description and Measurement of Aircraft Noise Heard on the Ground'' (Negative w/Comments), 1 July 1997.
	1st ISO/CD 3891-4—''Description, Measurement and Management of Aircraft Noise—Part 4: Airport Noise Management and Land use'' (Negative w/Comments), 1 July 1997.
	2nd ISO/CD 5129—Acoustics—Measurement of Aircraft Interior Sound Pressure Levels in-Flight (Affirmative w/Comments), 22 May 1997.
	ISO/CD 13325—Tyres—Coast-by Test Method for Measuring Tyre Road Sound Emission (ISO/TC 31 N577) (Affirmative w/Comments), 16 April 1997.
	1st ISO/CD 15667—Acoustics—Guidelines for Noise Control by Enclosures and Cabins (Affirmative w/Comments), 19 June 1997.

U.S. TAG	IEC Documents
S1	**IEC/TC29/352/Second CD 1672**—Electroacoustics—Sound Level Meters (Revision of IEC 651:1979 and IEC 804:1995), **Affirmative w/Comments**, 13 February 1997.
	IEC/TC 29/357/CD, First IEC/DAM—Amendment to IEC 651:1979, Sound Level Meters—Electrostatic Compatibility Requirements and Test Procedures, 12 May 1997.
	IEC/TC29/359/CD, First IEC/DAM—Amendment to IEC 804:1985, Sound Level Meters—Electromagnetic and Electrostatic Compatibility Requirements and Test Procedures, 12 May 1997.
	IEC/TC 29/360/CD, First IEC/DAM—Amendment to IEC 1252:1993, Sound Level Meters—Electromagnetic and Electrostatic Compatibility Requirements and Test Procedures, 12 May 1997.
	IEC/TC29/362/CDV 1672—Electroacoustics—Sound Level Meters (Revision of IEC 651:1979 and IEC 804:1995), 12 May 1997.
	IEC/TC29/368CD—IEC 61669, Ed.1: Equipment for

the Measurement of Real-Ear Characteristics of Hearing Aids, 4 June 1997.

S3 **IEC/TC 29/350/FDIS**, Amendment 2 to IEC 118-2, Ed. 2—Hearing Aids—Part 2: Measurements of Frequency Response Using Steady State Broad Band Input Signals (Abstention w/Comments), 24 February 1997.

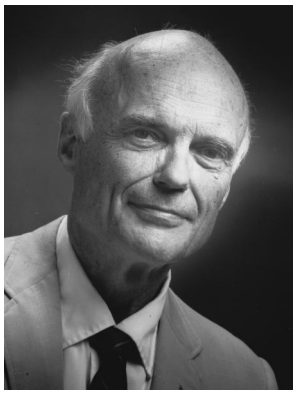
IEC/TC 29/356/Second CD, Amendment 1 to IEC 118-6—Hearing Aids—Part 6: Characteristics of Electrical Input Circuits for Hearing Aids (Abstention w/ Comments), 23 April 1997.

IEC 352/CDV (2nd IEC/CD 1672)—Electroacoustics—Sound Level Meters (Revision of IEC 651:1979 and IEC 804:1985) (Affirmative w/comments), 3 March 1997.

OBITUARIES

This section of the Journal publishes obituaries concerning the death of Fellows of the Society and other acousticians eminent in the world of acoustics. When notified, the Editor-in-Chief solicits a summary of the person's life and contributions from an ASA member thoroughly familiar with the details, if possible. If a promised obituary is never received, a brief obituary notice may be published later.

Ross E. Williams • 1922–1997



Ross E. Williams, a Fellow of the Acoustical Society of America and a valued contributor and leader in the underwater acoustics research community, died 8 November 1997 at the age of 75. Williams was noted for major contributions to the knowledge and applications of underwater sound. He had been an associate director of Hudson Laboratories of Columbia University, and a founder and President of Ocean and Atmospheric Sciences, Inc. (OAS) of Dobbs Ferry, New York.

Born in Carlinville, Illinois, Williams was raised in Scarsdale, New York where he graduated first in his high school class. He graduated in 1943, Summa Cum Laude, in physics and mathematics from Bowdoin College in Brunswick, Maine. He was elected to Phi Beta Kappa and Sigma Xi. He served as an officer in the US Navy aboard the USS YORKTOWN in WWII. He earned a master's degree in physics in 1947 at Columbia University, then worked for Sperry Products, Incorporated. He became a graduate assistant at Columbia in 1949, where he received a Ph.D. in physics in 1955 doing his research on Pi Meson scattering.

After 5 years as a consultant to the U.S. government and industry, Dr. Williams joined the Hudson Laboratories of Columbia University as a Senior Research Associate, and became Associate Director in 1966. In 1968 he was appointed Professor of Applied Science and Engineering at Columbia, and Director of the Ocean Engineering Program, a post he retained until 1974. During that time he published extensively in underwater signal processing and underwater propagation. He served as Technical Director of Project Artemis, a large active sonar research project sponsored by the Office of Naval Research.

In 1974 Williams joined Ocean and Atmospheric Sciences, Inc. where he served at different times as President, Chief Scientist, and Board Chairman, and where he also continued his career as a scientist, specializing in underwater sound research. In addition to more than 30 papers published in scientific journals, Williams authored more than 100 technical reports on sponsored research programs, and served numerous committees for the Navy and for professional societies such as IEEE and the American Society of Photogrammetry.

Throughout his life Williams loved the New Hampshire woods and wildlife, and bought tracts of land in New Hampshire to preserve them as national forest. He also collected and drove antique cars, and was an active lay leader and trustee of Asbury United Methodist Church in Scarsdale, New York. Ross Williams will be fondly remembered by his colleagues and his clients, reflecting his own personal warmth, his broad experience and deep technical understanding. He is survived by his wife, Madeline; his children, Katherine J. Williams and Ross Edward Williams; and his sister, Beverly Whitehead.

JAMES W. JENKINS
JULIUS I. BOWEN

BOOK REVIEWS

James F. Bartram
94 Kane Avenue, Middletown, Rhode Island 02842

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).]*

Mediacoustic—Teaching Acoustics by Computer

Héléne Gröll, Jean Paul Lefevre, *et al.*

01dB, Villeurbane, France.
CD-ROM. Price in the U.S. $699.

This CD-ROM teaching software package represents the collaborative efforts of acoustical experts from Ecole Polytechnique de Lausanne, Liverpool University, E.N.T.P.E., Acouphen, Estudi Acustica, Laboratoire informatique ARIA, and other European institutions. The disk is available in two languages: French and English (with British spelling). With the advent of multi-media in computers, the acoustics teacher is no longer relegated to providing pedagogical examples through lectures or ''chalk talks.'' The range of coverage and material depth permit use of this CD-ROM for introducing basic acoustics to high school students while incorporating sufficient sophistication to serve college level introductory courses. *Mediacoustic*, based on a central core, contains the fundamentals of acoustics plus a number of specialized modules. The central core consists of: 1. Introduction, 2. Basic Physics, 3. Noise & Man, 4. Room Acoustics, and 5. Noise Control. The program layout is fairly self-obvious, with hypertexts providing additional information but the user might get sidetracked occasionally that he or she might have to return to the main text by clicking onto the index. Each module contains sound clips, text, pictures, photographs, and animated cartoons featuring a rather amusing pair of caricatured ears named Audio and Acoussa (who slightly resemble the Potatoheads of television renown). Audio and Acoussa play musical instruments, sing, operate audio gear, and perform other chores in their joint capacity as virtual teaching aides.

An extremely brief historical overview is provided, starting with Vitruvius of the 1st century B.C. who made use of acoustic amphorae to modify the acoustical environment of amphitheaters. Bypassing the discoveries of Galileo, Leonardo, Mersenne, Sauveur, Grimaldi, and others, a sudden leap is made to the eighteenth century and beyond, with descriptions of acoustical contributions by d'Alembert, Helmholtz, Lord Rayleigh, and Sabine.

In the segment on basic physics, a ''slinky'' coil is shown to represent sound transmission; it would be more desirable to more accurately demonstrate the propagation effect by showing rarefactions and condensations in the format of dots representing fluid molecules in motion about their equilibrium points, with their distribution density representing rarefaction (less dense, fewer dots) and condensation (denser, more dots)—this was done in a different context where plane waves are shown being generated by a piston.

Among the commendable features of the disk are the superimposition of pure tones to demonstrate the harmonic content of a signal, the reduction of a signal by 3 dB or 6 dB to simulate the effect of doubling and quadrupling the distance from a source, the difference between pink and white noise, the modification of a signal passing through an octave or one-third octave filter, the effect of inserting noise barriers, simulation of different sound sources (jackhammer, musical instrument, singing voice, lowing cow, overhead jetcraft, etc.).

A high school student would benefit from the introduction to principles of propagation, the attributes of sound, definition of the decibel, and how the ear functions. Some of the more complex equations entailing calculus are discreetly hidden within hypertexts. On the collegiate level, both engineering and architectural students can hear, understand, and appreciate examples of reverberation effects. One particularly effective method of demonstrating pure tones and the concept of octaves and their harmonics, and at the same time, showing the effect of superpositioning, is the separate activation of mono-frequency signals defined by

$$p_1 = p_{\max} \sin(\omega t),$$

$$p_2 = p_{\max} \sin(2\omega t),$$

$$p_3 = p_{\max} \sin(3\omega t)$$

and the composite signal:

$$p = p_{\max}[\sin(\omega t) + \sin(2\omega t) + \sin(3\omega t)].$$

The listener can hear for him/herself the effect of combining the harmonics. The elements of the frequency analyzer are also described but some of the instrumentation shown do not appear to be state-of-the-art. The *2001* theme from Richard Strauss's *Also Sprach Zarathusa* provides a vivid example of the range from ppp (triple pianissimo) to fff (triple fortissimo). The module on hearing contains nearly all of the elements of interest to aspiring audiologists in dealing with frequency range of hearing, the concept of loudness and equal loudness contours, subjective judgement of loudness, differential loudness threshold, and loudness effects of short versus long sounds, and other aspects of psychoacoustics. Audiograms of people with normal hearing and partial deafness are shown and discussed, and there is even a musical excerpt to demonstrate how a person with normal hearing perceives differently from a person who has been sensorineurally affected by noisy environment at work. In the modules on architectural acoustics and noise control, emphasis is properly placed on the importance of the design stage to prevent acoustical debacles.

Other excellent examples of acoustic phenomena abound, which clarify the difference between longitudinal and transverse waves, illustrate the difference in propagation velocities in gases, liquids, and solids, the difference between the velocities of light waves and sound waves (thunder is heard a few seconds after the lightning clap), the effects of reflection, refraction, and diffraction, methods of cutting down on noise, the principle of the Helmholtz resonator, etc. The program material also would provide a suitable review course for practicing engineers and other professionals with appreciable mathematical backgrounds.

This disk may be deemed a meritorious teaching device, but its steep price of $699 will surely preclude its widespread dissemination. But if this program can be mounted on a server, then the cost takes on a more reasonable aspect.

DANIEL R. RAICHEL
Department of Mechanical Engineering
Graduate School and University Center
Steinman Hall, City University of New York
New York, New York 10031

REVIEWS OF ACOUSTICAL PATENTS

Daniel W. Martin
7349 Clough Pike, Cincinnati, Ohio 45244

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.

Reviewers for this issue:

JOHN ERDREICH, *Ostergaard Associates, 100 Executive Drive, West Orange, New Jersey 07052*
HARVEY H. HUBBARD, *325 Charleston Way, Newport News, Virginia 23606*
SAMUEL F. LYBARGER, *101 Oakwood Road, McMurray, Pennsylvania 15317*
D. LLOYD RICE, *11222 Flatiron Drive, Lafayette, Colorado 80026*
CARL J. ROSENBERG, *Acentech Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*
ERIC E. UNGAR, *Acentech Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*
ROBERT C. WAAG, *University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, New York 14642*

5,691,476

43.35.Zc METHOD FOR ULTRASONIC IMAGING AND DEVICE FOR PERFORMING THE METHOD

Eric I. Madaras, assignor to the United States of America
25 November 1997 (Class 73/644); originally filed 7 September 1993

This patent discloses a method for ultrasonic imaging of structures and flaws in a test specimen having an irregular contact surface. The ultrasonic transducer is coupled acoustically to the irregular surface through a number of ultrasonic waveguides that are thin and bendable so that they adapt to variations in the distance between the transducer and different parts of the contact surface. Since the waveguides are of equal length the delay times are also equal, and all parts of the irregular contact surface receive the sound waves in phase. The waveguides can be arranged in the form of an ultrasonic brush. By bevelling the bristle ends, shear mode waves can be produced.—DWM

5,668,865

43.38.Si ECHO CANCELER E-SIDE SPEECH DETECTOR

Donald Lars Duttweiler *et al.*, assignors to Lucent Technologies, Incorporated
16 September 1997 (Class 379/410); filed 26 February 1996

This telephone echo canceler performs speech detection on near-end, far-end, and outgoing signals. If the echo canceler output is greater than the echo signal, and a coarse far-end detector does not detect near-end speech, then a more accurate near-end detector is used to fine-tune the canceler output.—DLR

5,649,020

43.38.Si ELECTRONIC DRIVER FOR AN ELECTROMAGNETIC RESONANT TRANSDUCER

G. W. McClurg *et al.*, assignors to Motorola, Incorporated
15 July 1997 Class (381/151); filed 24 July 1994

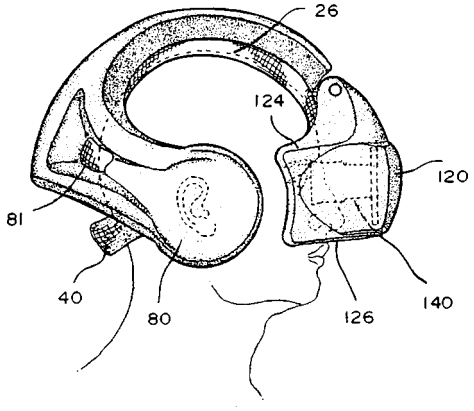
A vibrotactile alerting device is described which can be used to replace motor/offset weight alerting devices in applications such as cellular telephones. Methods of driving an electromagnetic resonant transducer utilizing resonant spring and nonlinear hardening spring resonant armature systems are disclosed. The device may be used for both vibrotactile alerting and voice communication.—JE

5,682,172

43.38.Si HEADSET FOR PRESENTING VIDEO AND AUDIO SIGNALS TO A WEARER

Paul J. Travers and Robert E. Taylor, assignors to Forte Technologies, Incorporated
28 October 1997 (Class 345/8); filed 30 December 1994

A headset is described for physically presenting a visual display and audio signals to the user. The headset has a beam extending over the vertex



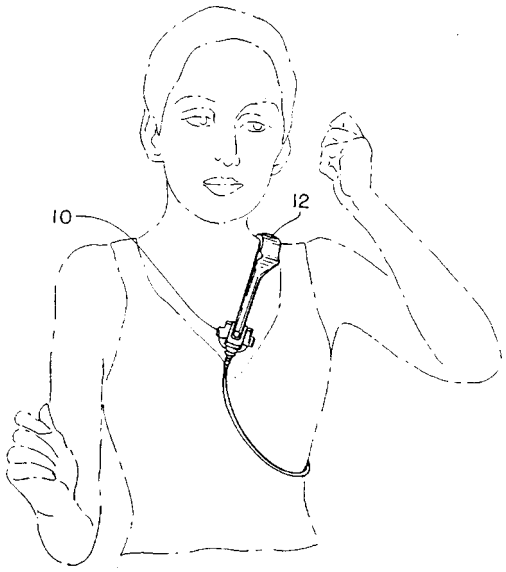
of the head that contacts the brow at one end and has a nape strap at the other end for adjusting the fit.—SFL

5,687,244

43.38.Si BONE CONDUCTION SPEAKER AND MOUNTING SYSTEM

Peter Untersander, assignor to Stanton Magnetics, Incorporated
11 November 1997 (Class 381/151); filed 28 March 1996

The patent shows a bone conduction receiver placed against the ster-



num to produce audible sounds, such as music, as vibrations. A mounting system is shown to hold the bone vibrator in place.—SFL

5,632,045

43.40.Ng ANTIVIBRATION GLOVE

D. D. Chase and D. A. Talonn, assignors to Chase Ergonomics, Incorporated
27 May 1997 (Class 2/161.6); filed 8 May 1995

This is an antivibration glove comprised of one to three layers of viscoelastic material. The damping layers may be of different thickness for different regions of the glove.—JE

5,673,437

43.40.Ng VIBRATION ATTENUATING MEMBER AND METHOD OF MAKING SAME

D. D. Chase and D. A. Talonn, assignors to Chase Ergonomics, Incorporated
7 October 1997 (Class 2/167); filed 6 May 1996

Use of one to three layers of resilient material to reduce hand-transmitted vibration is described. To assure registration of each layer with the others during manufacture, the patent describes a pin registration system.—JE

5,655,980

43.40.Tm VIBRATION DAMPING DEVICE FOR SPORTING IMPLEMENTS

Ahid D. Nashif and Gopichand Koganti, assignors to Roush Anatrol, Incorporated
12 August 1997 (Class 437/520); filed 2 November 1995

This patent pertains to dynamic vibration absorbers that may be attached to a tennis racquet or built into it for the purpose of reducing the shock and vibration transmitted to the user's hand. Two configurations are described: one consists of a mass that is attached to the strings near the racquet's throat; the other consists of a viscoelastically damped cantilever that is built into the handle near its end. The absorber is tuned to the racquet's natural frequency that it is intended to suppress, so that it vibrates essentially 180° out of phase with its attachment point.—EEU

5,669,594

43.40.Tm VIBRATION ISOLATING SYSTEM

David L. Platus and Donald A. Durran, assignors to Minus K Technology
23 September 1997 (Class 267/141); filed 3 May 1995

This patent is a continuation or extension of others relating to ''minus K'' systems. These systems make use of elastic structures that are loaded beyond their critical buckling limit, and that then exhibit a negative spring rate. In such a system a buckled structure is arranged in mechanical parallel with a conventional spring having a positive spring rate that is selected so that the net spring rate of the assembly is near zero. The present patent relates to an assembly that can provide near-zero spring rates for all six rigid-body degrees of freedom.—EEU

5,671,909

43.40.Tm OSCILLATING ROD MEMBER EQUIPPED WITH DYNAMIC DAMPER

Masaaki Hamada and Takashiro Aoi, assignors to Tokai Rubber Industries
30 September 1997 (Class 267/141); filed in Japan 13 May 1994

A rod or a shaft that may vibrate axially, laterally, and/or rotationally is provided with a vibration absorber (tuned damper) that in essence consists of a rubber sleeve with one or more massive rings imbedded in it. By placing the portion of the sleeve that contains a mass ring over a reduced diameter portion of the rod, or by configuring that portion of the sleeve to have an over-sized inside diameter, room is provided for excursions of the ring relative to the rod. The damper is tuned by appropriate selection of the ring masses and the cross-sectional geometry of the rubber sleeve.—EEU

5,676,352

43.40.Tm VIBRATION ABSORBER WITH MOUNT

Thomas Mayer *et al.*, assignors to Daimler-Benz Aerospace AG
14 October 1997 (Class 267/140.4); filed in Germany 1 December 1994

This vibration absorber assembly is intended to be fastened to the inside surface of the rim of a wheel like that used on rail vehicles. It consists of a number of short, flat metal strips with plastic material between them, together with a support element with its outer surface curved so as to fit the curvature of the inside surface of the wheel rim. Several such absorber assemblies are to be mounted around the circumference of the wheel.—EEU

5,678,840

43.40.Tm VIBRATION DAMPING DEVICES FOR SKIS AND OTHER APPLICATIONS

Stepan S. Simonian, Torrance, CA
21 October 1997 (Class 280/602); filed 20 March 1995

Two kinds of damping devices are described in this patent. One is called a ''cantilevered impact/friction damper (CIFD)'' and the other an ''improved constrained layer damper.'' A CIFD consists of a tubular cantilever spring that is filled with a particulate material and carries a mass at its tip, which mass can impact against adjustable limit stops if its excursion is great enough. The constrained layer dampers, which are meant to be attached to the upper surface of a ski, are special cases of spaced constrained viscoelastic layer configurations that have been discussed in the literature and have been employed in other applications.—EEU

5,620,303

43.50.Gf ROTOR SYSTEM HAVING ALTERNATING LENGTH ROTOR BLADES FOR REDUCING BLADE-VORTEX INTERACTION (BVI) NOISE

Robert C. Moffitt and Joseph A. Visintainer, assignors to Sikorsky Aircraft Corporation
15 April 1997 (Class 416/87); filed 11 December 1995

This patent applies to the control of blade–vortex interaction noise which is of particular concern during low forward-speed descending flight of rotorcraft. Blade–vortex interactions are most prevalent for descents during which the downward velocity of the rotor causes it to fly into its own wake. Tip vortices are shed from one rotor blade and interact with a following blade as it rotates through the vortex field. This interaction induces impulsive air loading which in turn creates an acoustic pressure wave that is the source of blade–vortex interaction noise. The use of different length rotor blades causes a spatial separation of shed vortices, a weakening of blade–vortex interactions, and a reduction in associated noise.—HHH

5,663,700

43.50.Gf SOUND DAMPENING SOLENOID

Glenn C. Spence and Dennis A. Maller, assignors to Trombetta Corporation
2 September 1997 (Class 335/257); filed 25 August 1995

This patent applies to push-type, pull-type, and rotary solenoids in order to minimize the noise created by the impact of the moving element against the stop or pole piece. Air pressure in the space between the pole piece and the moving element is regulated by vents to reduce the velocity of the moving element at the time of impact.—HHH

5,670,756

43.50.Gf SILENCER

Akihiko Ohtaka *et al.*, assignors to Honda Gilken Kogyo Kabushiki Kaisha
23 September 1997 (Class 181/256); filed 25 September 1995

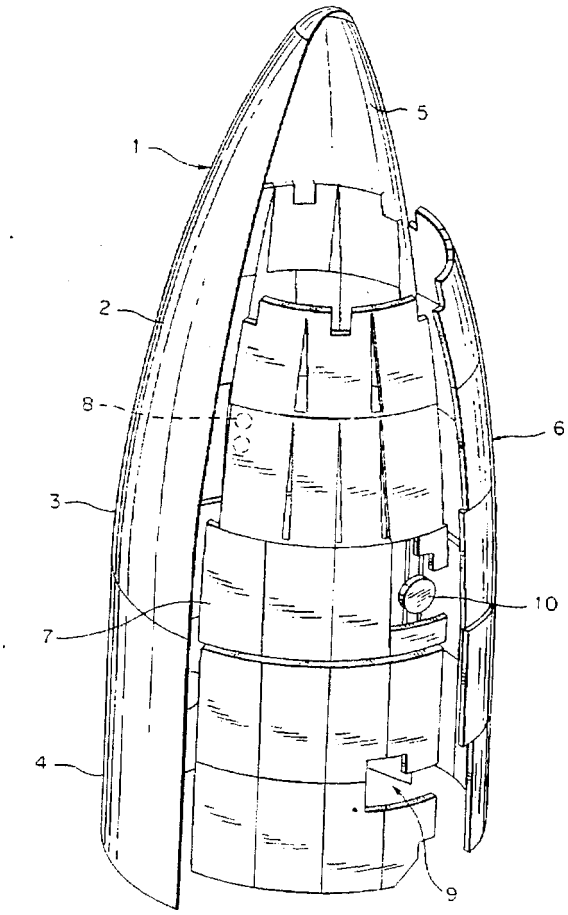
This patent relates to exhaust silencers which have an inner sleeve with pores and heat resistant fibers packed between the inner and outer sleeves, such as those on some automobiles. The fibers are configured with the addition of a heat resistant binder to enhance the press fitting of the fibrous liner.—HHH

5,670,758

43.50.Gf ACOUSTIC PROTECTION ON PAYLOAD FAIRINGS OF EXPENDABLE LAUNCH VEHICLES

Ingo U. Borchers *et al.*, assignors to Oerlikon-Contraves AG
23 September 1997 (Class 181/286); filed in Switzerland 20 April 1995

This patent relates to the reduction of environmental noise in the frequency range below 90 Hz which may cause damage or malfunction in



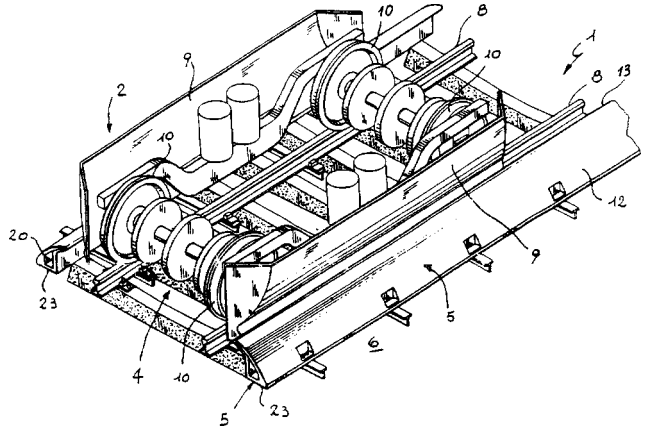
rocket vehicle payloads. Arrays of tuned acoustical absorbers, molded from low surface density foamed plastic, are bonded to the inside surfaces of the external vehicle fairings. Such arrays are noted to be effective during liftoff and the initial ascent phases of flight without large weight penalties.—HHH

5,671,685

43.50.Gf ACOUSTIC BARRIER, PARTICULARLY FOR RAILWAY SUPERSTRUCTURES

Stefano Ughi, Milano, Italy
30 September 1997 (Class 105/452); filed in Italy 28 February 1995

This patent is aimed at the attenuation of wheel-rail noise during the passage of trains. It consists of one member attached to the train in the



vicinity of the bogies, and another located near ground level and attached to the road bed. These two members together form a barrier to the sideways propagation of noise from the bogies.—HHH

5,678,676

43.50.Gf ROLLER SHAFT MOUNTING FOR SOUND AND MOTION CONTROL

Bruce A. Pierson, assignor to HK Systems, Inc.
21 October 1997 (Class 193/37); filed 20 January 1995

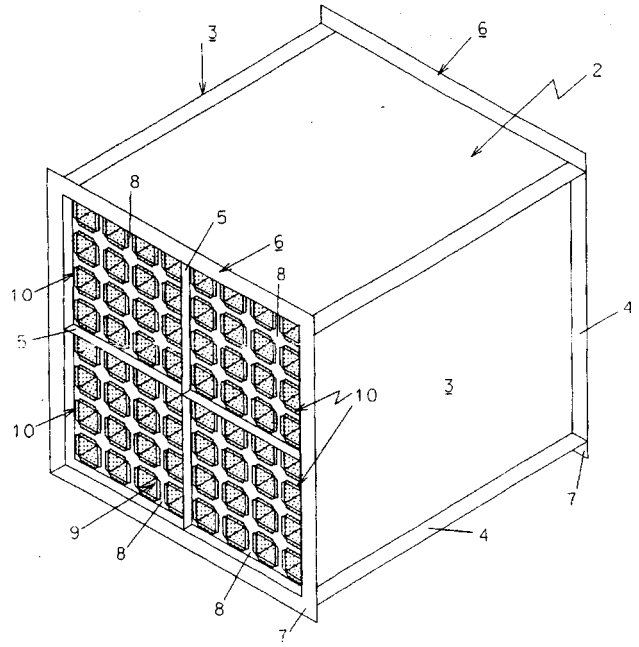
This patent relates to the control of noise and vibration of a roller conveyer by tightening the fit of each roller shaft in the conveyer frame. A tight fit is provided either by tapering each rotor shaft or by providing a tapered cap for each shaft.—HHH

5,679,931

43.50.Gf SOUND ATTENUATING APPARATUS AND METHOD OF FORMING THE SAME

John H. Furse and Robert A. Noreen, assignors to AAF-International
21 October 1997 (Class 181/224); filed 10 July 1995

This patent relates to the reduction of noise in an airflow along with a standardized method for adding passive acoustical treatment. Depending on



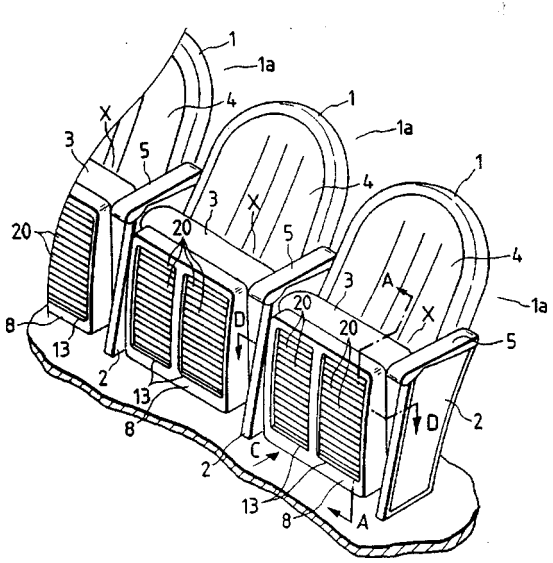
the amount of airflow, the size of duct, and the amount of noise reduction required, treatment can be added systematically.—HHH

5,700,052

43.55.Ev CHAIR FOR AN ACOUSTICALLY DESIGNED BUILDING

Ryokichi Yamazaki *et al.*, assignors to Yamaha Corporation
23 December 1997 (Class 297/217.3); filed in Japan 14 November 1994

The theater-type chair has absorptive material in the seat bottom, covered by slats **20** of a movable louver **13**. When the seat **3** is down (occupied), the louver is closed. When the seat is up (unoccupied), the louver is open, and the sound absorptive infill provides the means to absorb as much sound as the seat plus occupant, thus providing a more stable reverberation time for the hall.—CJR

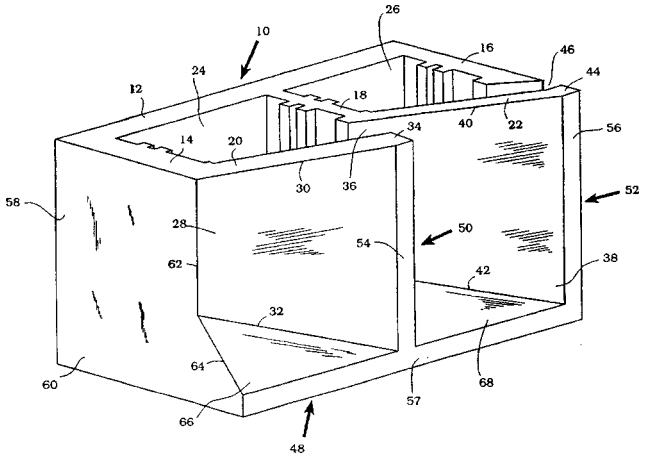


5,700,983

43.55.Ev SOUND ATTENUATING STRUCTURAL BLOCK

Kelly L. VonDross, assignor to Best Block Company
23 December 1997 (Class 181/285): filed 26 August 1996

This structural concrete masonry block has angled faces and open slits that allow sound energy to penetrate the block. With no infill, the block provides a resonant absorber with peak absorption around 200 Hz. With



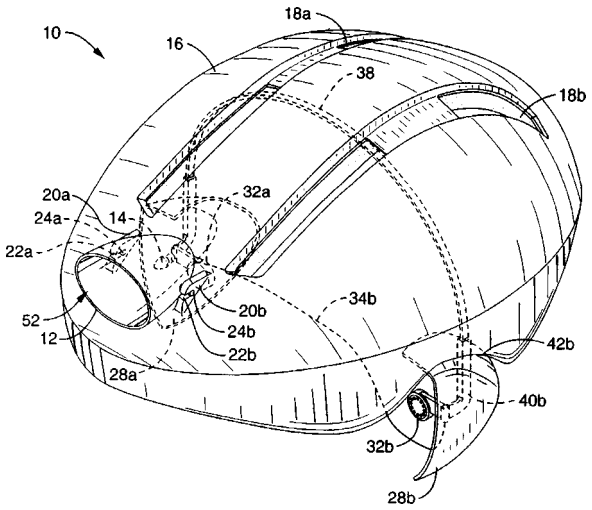
fibrous absorption in the cores, the absorption characteristics cover a broader frequency range, with coefficients of absorption around 0.4 at higher frequencies. The design of the block provides good structural support. Grooves on the sides of the core allow a separator to be slipped into place to divide the cavity so that the rear portion can be filled with concrete to carry additional structural weight.—CJR

5,691,514

43.66.Qp REARWARD SOUND ENHANCING APPARATUS

Timothy J. Landis, assignor to OP-D-OP, Incorporated
25 November 1997 (Class 181/129); filed 16 January 1996

This helmet for cyclists has a cup **12** mounted in a rearward facing direction, containing a microphone **14** for responding to noises behind the cyclist. The microphone signal is amplified and reproduced at the wearer's



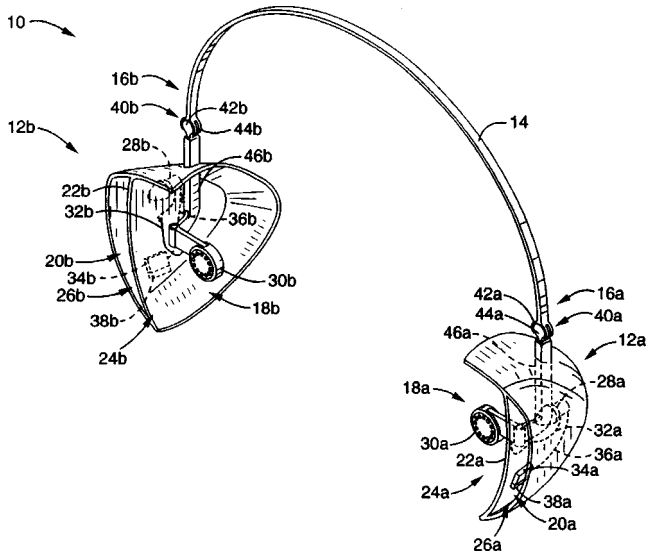
ears by earphones **32a** and **32b** to alert the cyclist aurally to vehicles that are following. Housings **28b** are shaped to reduce wind turbulence noise at the wearer's ears.—DWM

5,691,515

43.66.Qp REARWARD SOUND ENHANCING APPARATUS

Timothy J. Landis, assignor to OP-D-OP, Incorporated
25 November 1997 (Class 181/129); filed 28 June 1996

This patent discloses a headset to be worn by cyclists to alert them to the noise of vehicles that are following them. The patent is closely related to the same inventor's patent 5,691,514 (reviewed above). In this patent microphones **28a** and **28b** are mounted in an outer compartment of ear covers



12a and **12b**, respectively, in place of the helmet-mounted microphone cup in patent 5,691,514. Earphones **30a** and **30b** are in inner compartments. Shields **22a** and **22b** are intended to prevent acoustical feedback between the microphones and earphones.—DWM

5,696,356

43.66.Qp PASSIVE SOUND GATHERING APPARATUS

James P. Dudley *et al.*, assignors to OP-D-OP, Incorporated
9 December 1997 (Class 181/136); filed 7 March 1996

A cyclist helmet is shown in this patent that eliminates the electronic amplification provided in patents 5,691,514 and 5,691,515 (both reviewed above), and substitutes a pair of helmet-mounted sound gathering horns with rearward facing mouths. The throats of the horns deliver the acoustically amplified noise to shells covering the wearer's ears.—DWM

5,682,020

43.66.Ts SEALING OF HEARING AID TO EAR CANAL

Robert J. Oliveira, Maplewood, MN
28 October 1997 (Class 181/130); filed 17 January 1996

The patent shows a flexible sealing device that is particularly useful for in-the-canal hearing aids. It is held in place on the surface of the aid by an adhesive layer on one face. The opposite face, which makes contact with the skin of the ear canal is a layer of retarded recovery polymeric foam, provides an effective acoustic seal.—SFL

5,687,241

43.66.Ts CIRCUIT ARRANGEMENT FOR AUTOMATIC GAIN CONTROL OF HEARING AIDS

Carl Ludvigsen, assignor to Topholm & Watermann APS
11 November 1997 (Class 381/68.4); filed in Germany 1 December 1993

A circuit arrangement is shown for a programmable hearing aid in which the processing circuit contains means for continuous determination of percentile values of the input signal, whereby the gain and/or frequency response of the system are optimized. Block diagrams of the signal processing circuits are given.—SFL

5,687,242

43.66.Ts HEARING AID CONTROLS OPERABLE WITH BATTERY DOOR

Lester Iburg, assignor to Resistance Technology, Incorporated
11 November 1997 (Class 381/69.2); filed 11 August 1995

The patent shows a battery door for a hearing aid that also includes two switching functions. After the battery door is closed, two single-pole switches are available by rotating the battery holder either way using a small knob provided on the battery holder. The switches appear to make contact only when rotating pressure is applied.—SFL

5,573,015

43.66.Vt EXTRUDED EAR PLUG

Colin D. Williams, Charlotte, NC
12 November 1996 (Class 128/864); filed 28 March 1995

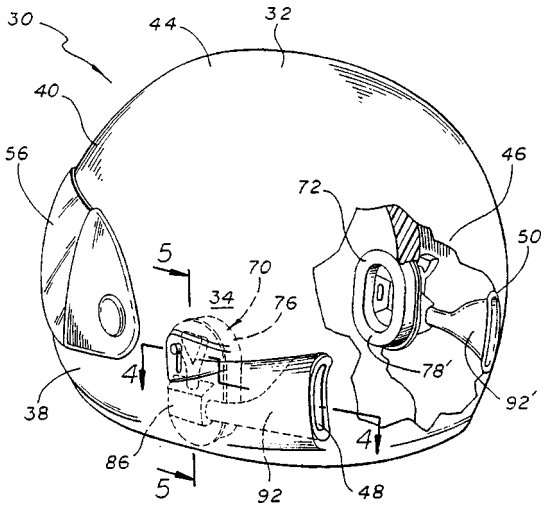
A method for producing an extruded, as opposed to molded, ear plug is described. The ear plug consists of a rigid core surrounded by a foam sheath.—JE

5,632,048

43.66.Vt PROTECTOR HEARING HELMET

D. Mortell *et al.*, assignors to Protector Development
27 May 1997 (Class 2/423); filed 20 September 1995

A helmet is described which incorporates a chamber enclosing each of the ears of a motorcycle or bicycle rider. The chambers include a rearward facing opening to the outside of the helmet and a ventilation loop which can



be opened or closed. No acoustical treatment of the openings is disclosed. The authors say that the helmet reduces wind noise. Nothing is said about protection from intense sounds, nor is the helmet likely to provide such.—JE

5,619,180

43.66.Wv APPARATUS FOR PROVIDING VIBROTACTILE SENSORY SUBSTITUTION OF FORCE FEEDBACK

M. J. Massimino, Houston, TX
8 April 1997 (Class 340/407.1); originally filed 3 November 1993

A vibrotactile stimulus is provided to the operator of a remote manipulator. The vibrotactile stimulus amplitude is proportional to the force applied by the remote manipulator. Sensor location on the manipulator may be replicated by the location of the feedback transducer on the operator or by the frequency of the vibrotactile stimulation.—JE

5,673,703

43.66.Wv APPARATUS FOR AUTOMATED DETERMINATION OF LOW FREQUENCY TACTILE THRESHOLDS

J. H. Fisher *et al.*, assignors to Ztech L.C.
7 October 1997 (Class 128/739); filed 17 February 1995

Several methods exist for the purpose of determining vibrotactile thresholds in humans. Many of these require an analog input signal to the transducer which has been produced by the computer and then converted to an analog signal in a D/A converter internal to the computer. This system uses a frequency-modulated signal from a standard PC to control a motor which drives a tactile stimulator. One stated goal is ''to provide a system and method to accurately and reproducibly determine a patient's threshold to vibrotactile stimuli...''. However, no method is provided for controlling the force of contact between the stimulator and the patient.—JE

5,673,361

43.72.Ar SYSTEM AND METHOD FOR PERFORMING PREDICTIVE SCALING IN COMPUTING LPC SPEECH CODING COEFFICIENTS

Mark A. Ireton, assignor to Advanced Micro Devices, Incorporated
30 September 1997 (Class 395/2.28); filed 13 November 1995

This speech analysis method is an improvement on a previously known method of linear prediction analysis. The FLAT analysis method involves iteratively removing contributions of successive reflection coefficients from the speech power as represented in coefficient matrices. As the matrix element magnitudes are reduced, a loss of resolution occurs in the digital values. This patent adjusts for that loss by scaling the matrix at each iteration.—DLR

5,668,925

43.72.Gy LOW DATA RATE SPEECH ENCODER WITH MIXED EXCITATION

Joseph Harvey Rothweiler *et al.*, assignors to Martin Marietta Corporation
16 September 1997 (Class 704/220); filed 1 June 1995

This vocoder uses line spectral frequencies (LSF) coding, performing simultaneous analyses using different numbers of blocks per frame and choosing the best fitting result. A variable distortion threshold is adjusted to maintain a roughly constant transmit data rate. The pitch coding is improved by a 1-bit jitter code, which controls the addition of random fundamental frequency jitter in the receiver.—DLR

5,671,327

43.72.Gy SPEECH ENCODING APPARATUS UTILIZING STORED CODE DATA

Masami Akamine *et al.*, assignors to Kabushiki Kaisha Toshiba
23 September 1997 (Class 395/2.28); filed in Japan 21 October 1991

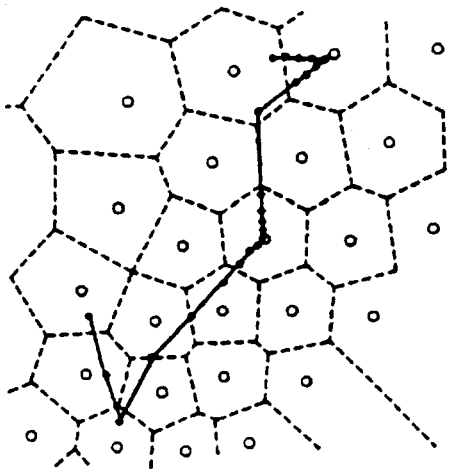
This patent describes a code-excited (CELP) vocoder which uses an adaptive codebook. Linear predictors are computed for frames of input speech and the residual is divided into four subframes for the codebook search. On finding the excitation which best reproduces the target speech subframe, the codebook is modified by merging the target excitation into the winning codebook entry.—DLR

5,675,701

43.72.Gy SPEECH CODING PARAMETER SMOOTHING METHOD

Willem Bastiaan Kleijn and Hans Petter Knagenhjelm, assignors to Lucent Technologies, Incorporated
7 October 1997 (Class 395/2.31); filed 28 April 1995

A vector quantizer speech vocoder uses some type of codebook for encoding one or more aspects of the speech signal. However, the code resulting from each search encodes not the specific target, but a region in a



space of many potential targets. This patent describes a method of processing (''dithering'') the code sequences before transmission so as to minimize the perceptual effect of abrupt transitions across regions of the code space.—DLR

5,675,702

43.72.Gy MULTI-SEGMENT VECTOR QUANTIZER FOR A SPEECH CODER SUITABLE FOR USE IN A RADIOTELEPHONE

Ira A. Gerson *et al.*, assignors to Motorola, Incorporated
7 October 1997 (Class 395/2.32); filed 26 March 1993

The patent describes the construction of a vector-sum-excited (VSELP) vocoder using a fixed point lattice (FLAT) method of computing the linear prediction filters. The FLAT technique allows efficient implementation on a fixed point processor. The method is applicable to either single or multi-segment codebook searches and performs additional scalar quantization to reduce codebook storage requirements.—DLR

5,673,362

43.72.Ja SPEECH SYNTHESIS SYSTEM IN WHICH A PLURALITY OF CLIENTS AND AT LEAST ONE VOICE SYNTHESIZING SERVER ARE CONNECTED TO A LOCAL AREA NETWORK

Tatsuro Matsumoto, assignor to Fujitsu Limited
30 September 1997 (Class 395/2.69); filed in Japan 12 November 1991

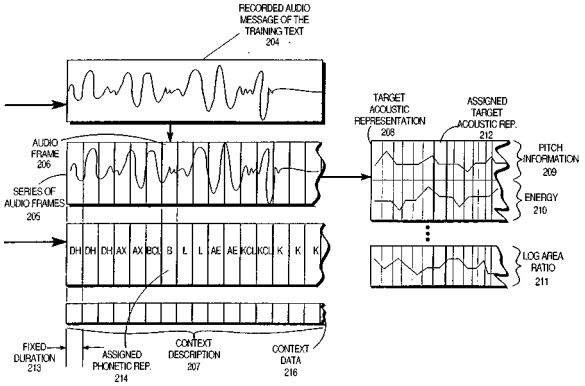
This patent describes the construction of a speech synthesis server suitable for providing text-to-speech audio services to a number of clients, similar to the type of data services provided by a typical local area network file server. Various functional units in the server perform text parsing, text-to-phoneme, phoneme-to-parameter, and parameter-to-waveform conversions, routing partially processed streams among these functions as required.—DLR

5,668,926

43.72.Ja METHOD AND APPARATUS FOR CONVERTING TEXT INTO AUDIBLE SIGNALS USING A NEURAL NETWORK

Orhan Karaali *et al.*, assignors to Motorola, Incorporated
16 September 1997 (Class 704/232); filed 28 April 1994

This method consists of a procedure for training a neural network to generate specific sounds such as speech sounds when a given input is applied to the network. In a training phase, sounds either in the form of



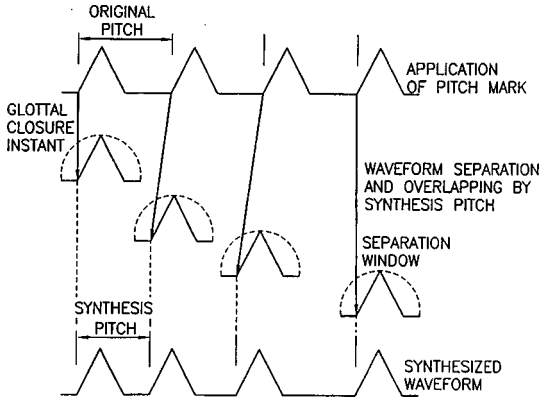
waveform frames or analysis vectors such as log area ratios are associated with phonetic symbols. A phonetic sequence is then sequentially applied to the network, reproducing the acoustic output by use of an appropriate playback or decoding method.—DLR

5,671,330

43.72.Ja SPEECH SYNTHESIS USING GLOTTAL CLOSURE INSTANTS DETERMINED FROM ADAPTIVELY-THRESHOLDED WAVELET TRANSFORMS

Masaharu Sakamoto *et al.*, assignors to International Business Machines Corporation
23 September 1997 (Class 395/2.77); filed in Japan 21 September 1994

This speech synthesis system is based on the addition of overlapped waveform segments representing short speech elements. The system achieves improved output quality by using the moment of glottal closure to align the waveform fragments in the overlapped interval. Accurate estimates



of the moments of glottal closure are possible by using dyadic wavelet conversion and local peak search.—DLR

5,675,703

43.72.Ja APPARATUS FOR DECODING COMPRESSED AND CODED SOUND SIGNAL

Hirofumi Sato, assignor to Nippon Steel Corporation
7 October 1997 (Class 395/2.39); filed in Japan 12 April 1994

This is an improved method for decoding the audio channels of a motion picture experts group (MPEG) coded signal. Typically, the MPEG decoding operations are described as a sequence of separate steps, such as coefficient scaling, band expansion, and waveform reconstruction. In this apparatus the steps are integrated into a single waveform reconstruction operation, using bit shifts wherever possible to improve implementation efficiency.—DLR

5,677,985

43.72.Kb SPEECH DECODER CAPABLE OF REPRODUCING WELL BACKGROUND NOISE

Kazunori Ozawa, assignor to NEC Corporation
14 October 1997 (Class 395/2.29); filed in Japan 10 December 1993

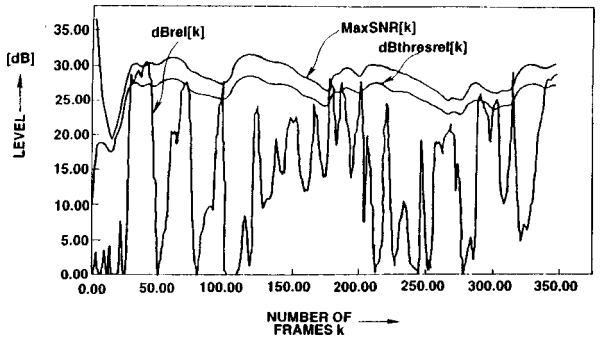
In speech communications systems, such as cellular telephones, the listener's overall perception can be improved by the introduction of a reasonable quality reproduction of the background noise during nonspeech intervals. This system does that with a very small bitrate increase by occasionally sending a code for the best choice from among a fixed set of random number generators which simulate various background noises.—DLR

5,668,927

43.72.Kb METHOD FOR REDUCING NOISE IN SPEECH SIGNALS BY ADAPTIVELY CONTROLLING A MAXIMUM LIKELIHOOD FILTER FOR CALCULATING SPEECH COMPONENTS

Joseph Chan and Masayuki Nishiguchi, assignors to Sony Corporation
16 September 1997 (Class 704/240); filed in Japan 13 May 1994

This patent describes a variety of noise reduction systems based on the concept of comparing a raw speech measure to a smoothed version of that measure. When the raw measure is smaller, the signal is judged to be background noise and is used to adapt a noise reduction filter. Measures discussed include rms energy, spectral band energies, and spectral probability estimates. A maximum likelihood noise filter adjusts band amplitudes based on band probabilities.—DLR



5,668,928

43.72.Ne SPEECH RECOGNITION SYSTEM AND METHOD WITH AUTOMATIC SYNTAX GENERATION

Gabriel F. Groner, assignor to Kor Team International, Incorporated
16 September 1997 (Class 704/243); filed 31 January 1995

This system is an automated syntax builder for use with a speech recognizer and voice-controlled application program. It accepts lists of possible words or phrases which could be used to activate each of the program's functions. The potential word sequences are then checked for redundancy and rated according to uniqueness of association with one of the functions. A grammar is then constructed suitable for use with a syntax-operated voice recognition system.—DLR

5,671,328

43.72.Ne METHOD AND APPARATUS FOR AUTOMATIC CREATION OF A VOICE RECOGNITION TEMPLATE ENTRY

Gregory P. Fitzpatrick *et al.*, assignors to International Business Machines Corporation
23 September 1997 (Class 395/2.55); filed 30 December 1992

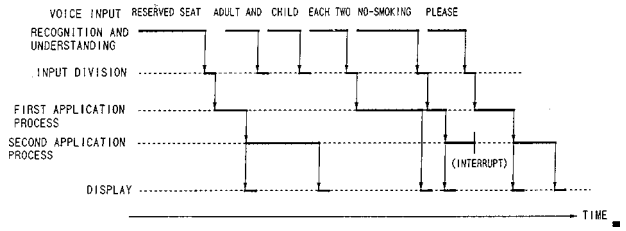
For use with a discrete utterance recognition system, this template builder constructs phrase templates by combining subtemplates representing individual words or short phrases. The template builder assigns a function command, such as a keystroke sequence, to each combined template and sets up training sessions to elicit the combined speech utterances.—DLR

5,671,329

43.72.Ne SPEECH DIALOGUE SYSTEM IN WHICH A RECOGNITION AND UNDERSTANDING PROCESS, APPLICATION PROCESS, AND VOICE INPUT RESPONSE ARE PERFORMED SIMULTANEOUSLY WITH VOICE INPUT

Kaichiro Hatazaki, assignor to NEC Corporation
23 September 1997 (Class 395/2.62); filed in Japan 9 March 1993

This automated reservations taking system uses a speech recognition system together with a decision stage for translating spoken input into system inputs, voice responses, or display controls. The recognizer is configured to provide partial outputs without waiting for the detected end of speech. The partial outputs are processed by a type of grammar driven parser, which then takes the appropriate system action.—DLR

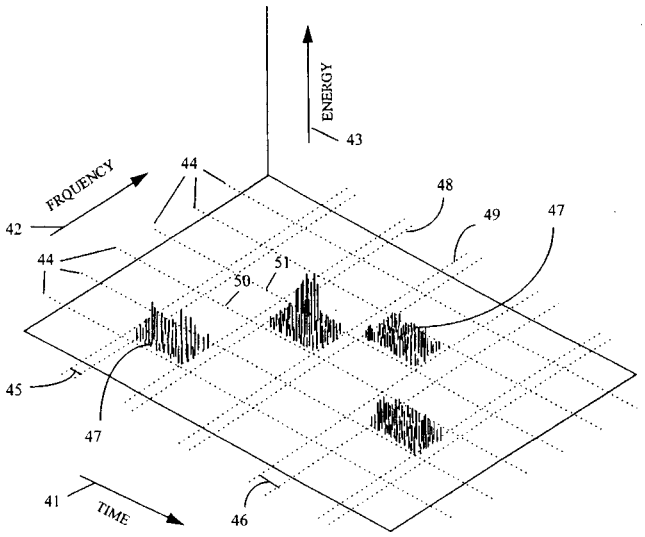


5,675,705

43.72.Ne SPECTROGRAM-FEATURE-BASED SPEECH SYLLABLE AND WORD RECOGNITION USING SYLLABIC LANGUAGE DICTIONARY

Tara Chand Singhal, Torrance, CA
7 October 1997 (Class 395/2.63); filed 27 September 1993

The field of large-vocabulary, continuous recognition has recently been dominated by the use of phoneme or segment level hidden Markov modeling. This recognizer uses a different approach based on the recognition of syllables from a set of roughly a thousand patterns. The syllables are



recognized by their known spectral peak structures, although formants are not explicitly extracted in the analysis. The patent text includes an interesting discussion of a hierarchy of elements of syllable structure.—DLR

5,675,706

43.72.Ne VOCABULARY INDEPENDENT DISCRIMINATIVE UTTERANCE VERIFICATION FOR NON-KEYWORD REJECTION IN SUBWORD BASED SPEECH RECOGNITION

Chin-Hui Lee and Rafid Antoon Sukkar, assignors to Lucent Technologies, Incorporated
7 October 1997 (Class 395/2.65); filed 31 March 1995

This extension to keyword spotting techniques uses a two-level verification strategy to improve accuracy. Keywords are recognized by subword portions, typically diphonelike in structure, adding a degree of vocabulary independence to the system. The verification methods include both HMM scoring and text string scoring.—DLR

5,675,704

43.72.Pf SPEAKER VERIFICATION WITH COHORT NORMALIZED SCORING

Biing-Hwang Juang *et al.*, assignors to Lucent Technologies, Incorporated
7 October 1997 (Class 395/2.55); filed 9 October 1992

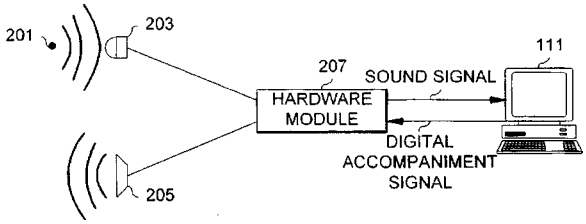
The patent describes a voice dialing system which allows both dialing and collect call charging by means of speaking either the party's name, a preassigned ''nickname,'' the actual dialed digits, or a speed dialing code. The system includes a ''cohort scoring'' method of speaker verification in which the input vectors are compared to both the purported speaker's reference and a reference for a small group of ''cohort'' speakers having similar voice characteristics. The cohort scores are used to normalize the speaker scores, providing for more robust speaker verification.—DLR

5,693,903

43.75.Rs APPARATUS AND METHOD FOR ANALYZING VOCAL AUDIO DATA TO PROVIDE ACCOMPANIMENT TO A VOCALIST

Allen J. Heidorn *et al.*, assignors to Coda Music Technology, Incorporated
2 December 1997 (Class 84/668); filed 4 April 1996

A computer **111** stores a digitized musical accompaniment for a specific published edition of a musical piece. It also stores parameters of a



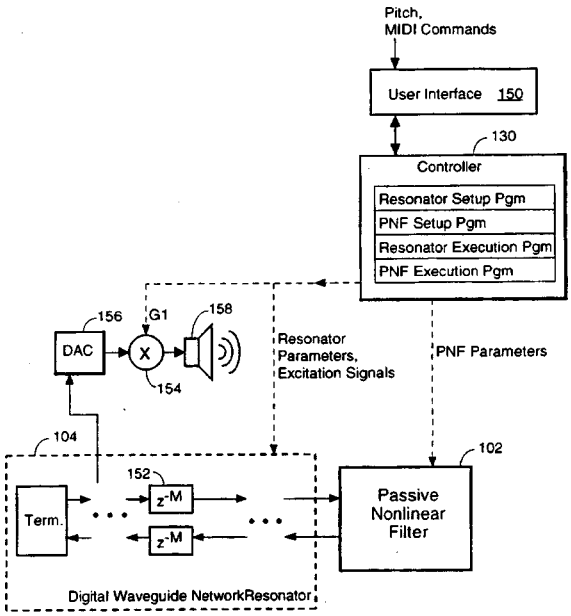
musical solo with which the accompaniment is to be synchronized. As a vocalist **201** sings the solo, microphone **203** transmits the sung tones to the computer for comparison to the stored solo tones, for triggering corresponding accompaniment signals to be mixed with the sung tones in hardware module **207** for reproduction by loudspeaker **205**.—DWM

5,703,313

43.75.Wx PASSIVE NONLINEAR FILTER FOR DIGITAL MUSICAL SOUND SYNTHESIZER AND METHOD

John R. Pierce and Scott A. Van Duyne, assignors to Leland Stanford Junior University
30 December 1997 (Class 84/622); filed 10 May 1994

This music synthesizer uses ''digital signal processing techniques which passively spread energy in a manner that mimics the frequency spreading in cymbals and Chinese gongs.'' It has a ''main resonator waveguide network (e.g., a loop or mesh) that is coupled to a digital passive



nonlinear filter. The passive nonlinear filter receives traveling wave signals propagating in the resonant network and passively modifies those received signals so as to generate modified traveling wave signals having a different frequency spectrum than the received traveling wave signals without chang-

ing the received traveling wave signals' energy content. The passive nonlinear filter then transmits the modified traveling wave signals back into the resonator.''—DWM

5,680,863

43.80.Vj FLEXIBLE ULTRASONIC TRANSDUCERS AND RELATED SYSTEMS

John A. Hossack *et al.*, assignors to Acuson Corporation
28 October 1997 (Class 128/662.03); filed 30 May 1996

Shape sensors such as strain gauges are used to produce signals that indicate the prevailing curvature of an ultrasonic transducer array. The curvature can be changed by a user-controlled actuator. The ultrasonic transducer can be positioned for different scan formats and the flexible part of the system may have a cross-sectional dimension of less than 10 mm to facilitate insertion into a tortuous vessel.—RCW

5,682,895

43.80.Vj THREE-DIMENSIONAL ULTRASOUND IMAGE-PROCESSING SYSTEM

Masaaki Ishiguro, assignor to Fuji Photo Optical Co.
4 November 1997 (Class 128/660.04); filed in Japan 12 October 1995

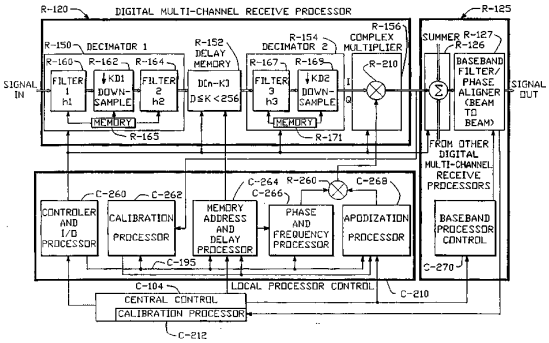
This system includes a cross-sectional ultrasonic image capture unit, a three-dimensional image generator to compile sequentially captured cross-sectional images, a processor to produce a three-dimensional image display on a screen, and another processor that produces a cut-away section of the original three-dimensional ultrasonic image.—RCW

5,685,308

43.80.Vj METHOD AND APPARATUS FOR RECEIVE BEAMFORMER SYSTEM

J. Nelson Wright *et al.*, assignors to Acuson Corporation
11 November 1997 (Class 128/660.07); filed 2 May 1994

This beamformer contains a number of programmable multichannel receivers, a local controller for each receiver processor, a unit that combines



signals from other receiver processors, and a central control unit. Each local controller receives focusing data from the central control unit and calculates dynamic focus delay, phase, apodization, and calibration values for each receiver signal. The unit that combines signals includes a baseband filter for post-beam-formation phase adjustments and signal shaping. The system can be used with different transducers and can implement various imaging modes.—RCW

The force between two parallel rigid plates due to the radiation pressure of broadband noise: An acoustic Casimir effect[a)]

Andrés Larraza, Christopher D. Holmes,[b),d)] and Robert T. Susbilla[c),d)]
Department of Physics—Code PH/La, Naval Postgraduate School, Monterey, California 93943

Bruce Denardo
Department of Physics and Astronomy, and National Center for Physical Acoustics, University of Mississippi, University, Mississippi 38677

(Received 21 July 1997; accepted for publication 26 November 1997)

Theoretical and experimental results are presented for the force law between two rigid, parallel plates due to the radiation pressure of broadband acoustic noise. The noise is in the band of 5–15 kHz and has an intensity of 133 dB ($re:\ 10^{-12}$ W/m^2). Excellent quantitative agreement is shown between theory and experiment. These results constitute an acoustic *analog* for the Casimir effect, which is the force experienced by two closely spaced uncharged conducting plates due to the quantum electromagnetic zero-point field. In contrast to this case, however, band-limited noise can cause the force to be *attractive* or *repulsive* as a function of separation between the plates. © *1998 Acoustical Society of America.* [S0001-4966(98)04003-X]

PACS numbers: 43.10.Ln, 43.25.Qp [MAB]

INTRODUCTION

In 1948, the Dutch physicist Hendrik Brugt Gerhard Casimir made the quantum electromagnetic prediction that two closely spaced uncharged parallel conducting plates would be mutually attractive.[1,2] Due to the presence of the plates, the mode structure of the quantum electromagnetic zero-point field (ZPF) changes relative to free space. The difference between the vacuum electromagnetic energy for infinite plate separation and a finite plate separation is the interaction energy between the plates, from which the force is calculated. If the plates are a distance d apart, the force per unit area is

$$f=\frac{\pi^2}{240}\frac{\hbar c}{d^4}, \tag{1}$$

where $\hbar$ is the reduced Planck's constant and c is the speed of light in vacuum. It can be easily shown by dimensional analysis that $\hbar c/d^4$ is the only possible combination in which $\hbar$, c, and d can be united to yield a force per unit area.

In spite of the large theoretical attention that this effect has received since Casimir's original article, it is only recently that a conclusive experimental confirmation has been made by Lamoreaux.[3] Previously, Sparnaay[4] attempted direct measurement of the Casimir force, achieving measurements of the correct shape, but with 100% uncertainty. Lamoreaux's results agree within 5% with the expression for the Casimir force.

Physically, the electromagnetic ZPF fluctuations induce local dipoles in both plates. As a result of the spatial correlations of the fluctuations, the interaction of these dipoles leads to a net attractive force. Based on this argument, there must also be a Casimir force between two parallel dielectric plates.[5]

The attraction between the two parallel plates can also be understood in terms of the radiation pressure exerted by the plane waves that comprise the homogeneous, isotropic ZPF spectrum. In the space between the conducting plates, the modes formed by reflections off the plates act to push the plates apart. The modes outside the cavity formed by the plates act to push the plates together. The total outward pressure and the total inward pressure are infinite, but it is only the difference that is physically meaningful. This difference leads to an attractive force. Because the energy per mode of the zero-point field has the same value $\frac{1}{2}\hbar\omega$ between and outside the plates, one may incorrectly be led to attribute the attractive character of the force as due to the fact that there are fewer modes between the plates. Surprisingly, as we will show in Secs. I and III, the force can be *repulsive* for band-limited noise.

The ZPF is homogeneous and isotropic. It is also Lorentz invariant (any two inertial observers do not see a difference). The ZPF can be said to be undetectable in any inertial frame because homogeneity, isotropy, and Lorentz invariance do not allow an observer to see any contrast between directions, times, and inertial frames. The Casimir effect is one of the several indirect manifestations of the ZPF. Other indirect manifestations include the van der Waals force between polarizable matter,[5] as well as spontaneous emission of radiation and the Lamb shift.[6]

A Lorentz transformation Doppler shifts the frequency and also alters the intensity of radiation. The only spectral shape that is the same for any inertial frame is proportional to ω^3.[7] This spectral shape implies an energy per mode proportional to the frequency, in agreement with the quantum zero-point energy $\frac{1}{2}\hbar\omega$ per mode of the electromagnetic vacuum.

---

a)''Selected research articles'' are ones chosen occasionally by the Editor-in-Chief that are judged (a) to have a subject of wide acoustical interest, and (b) to be written for understanding by broad acoustical readership.
b)LT, US Navy.
c)LCDR, US Navy.
d)Some of the experimental aspects of this article were submitted in partial fulfillment for the MS degree in Physics at the Naval Postgraduate School.

At finite temperatures, there must also be a force between the plates due to real photons. Thus the Casimir effect is the zero-temperature limit of the electromagnetic radiation pressure. However, thermal effects can be ignored when $\hbar\omega \gg k_B T$, where k_B is Boltzmann constant. The lowest frequency that contributes to the radiation pressure inside the plates is $\omega = \pi c/d$. Thus thermal effects can be ignored for separations $d < \pi\hbar c/k_B T$. At room temperature this corresponds to a distance of about 50 μm.

If the ZPF can be thought of as broadband noise of an infinite spectrum, it should be possible to use an acoustic broadband noise spectrum as an analog to some ZPF effects. An acoustic spectrum has several advantages. Because the speed of sound is six orders of magnitude less than the speed of light, the length and time scales are more manageable and measurable. Also, in an acoustic field, the shape of the spectrum as well as the field intensity can be controlled.

We refer to this new area of research as *Casimir acoustics*. By analogy to the Casimir effect, the measurement of the force between two plates in a homogeneous, isotropic, acoustic field is the acoustic Casimir effect.

An aspect of our analog is the extension of the Casimir effect due to zero-point phonons. Thus for a sufficiently large Debye temperature the force per unit area between two rigid plates due to zero-point phonons has the same form as Eq. (1), except that c should be replaced by the value of the speed of sound in the medium. As would be the case for liquid helium at low temperatures, there should also be a factor of $\frac{1}{2}$ because longitudinal phonons have no polarization. However, if the plates are also conducting, the Casimir force due to the electromagnetic zero-point field dominates the Casimir force due to zero-point phonons by at least six orders of magnitude. As shown in Sec. III, the band-limited acoustic noise due to commercial loudspeakers yields a force that is three orders of magnitude greater than the Casimir force due to the electromagnetic zero-point field, recently measured by Lamoreaux.[3]

I. THEORY

One of the key ideas in the derivation of the ZPF Casimir force is the fact that the energy per mode $\frac{1}{2}\hbar\omega$ is the same for modes both outside and between the plates, which can be understood with the adiabatic theorem.[8] To this purpose, imagine that the plates are initially far apart so that the spectral intensity (intensity per unit frequency interval) of the ZPF is that of free space: $I_\omega = \hbar\omega^3/2\pi^2 c^3$. If we now adiabatically move the walls toward each other, the modes comprising the ZPF will remain in their ground state; only their frequencies will be shifted in such a way that the ratio of the energy per mode E to the frequency ω remains constant, or $E/\omega = \hbar/2$. Thus the main effect of the boundaries is to redistribute ground state modes of which there is an infinite number. It should be noted that the energy per mode being a constant implies that the space between the plates cannot be considered as a resonant cavity unless the quality factor Q is unity.

In contrast to the ZPF Casimir effect, in the acoustic Casimir effect the adiabatic theorem does not apply both because of inherent losses in the system and because the spectrum can be arbitrary. Also, while external drivers can provide a steady state noise spectrum from which we can infer the energy per mode by dividing by the density of states $\omega^2/2\pi^2 c^3$, this energy may be different in the cavity formed by the plates as a result of Q amplification. However, for this open resonant cavity the quality factor is poor, so we may assume it to be equal to unity, which renders the energy per mode equal to its value in free space.

The shape of the acoustic spectrum is determined by the input to the drivers and their the frequency response. Good acoustic drivers have a flat response over a frequency band. With appropriate filters, one may shape the spectrum within the driver's frequency response and obtain, in principle, different force laws. We specialize the derivation of the force law below for a flat spectral shape over a finite frequency band.

In general, the radiation pressure of a wave incident at angle θ on a *rigid* plate is

$$P = \frac{2I}{c}\cos^2(\theta), \tag{2}$$

where I is the average intensity of the incident wave and c is the wave speed. The factor of 2 is due to perfect reflectivity assumed for the plate. Equation (2) follows from the time-averaged second-order acoustic pressure, which equals the time-averaged potential energy density minus the time-averaged kinetic energy density.[9] When the acoustic case is constrained to one dimension, mass conservation yields an explicit dependence of the radiation pressure on the elasticity of the medium characterized by γ, the ratio of specific heats, namely $P = (1+\gamma)I/c$.[10] However, for the three-dimensional open geometry in our case, the constraint due to mass conservation does not apply and the acoustic and electromagnetic expressions for the radiation pressure at a perfectly reflecting surface are identical.[11]

We consider broadband acoustic noise of constant spectral intensity I_ω in a band of frequencies from ω_1 to ω_2. The total intensity is then $I = (\omega_2 - \omega_1)I_\omega$. The spectral intensity in the wave vector space of traveling waves is

$$I_{\mathbf{k}} = \frac{cI_\omega}{4\pi k^2}, \tag{3}$$

where the wave vector is $\mathbf{k} = (k_x, k_y, k_z)$ whose magnitude is the wave number $k = \omega/c$. We choose the z axis to be normal to the plate, so that $k_z = k\cos(\theta)$. From Eq. (2), the total radiation pressure due to waves that strike the plate is then

$$\begin{aligned} P &= \frac{2}{c}\int dk_x\, dk_y\, dk_z\, I_{\mathbf{k}}\cos^2(\theta) \\ &= \frac{I_\omega}{2\pi}\int dk_x\, dk_y\, dk_z\, \frac{k_z^2}{k^4}, \end{aligned} \tag{4}$$

where the integration is over $\mathbf{k}$ values corresponding to waves that strike the plate. Converting Eq. (4) to polar coordinates, and performing the azimuthal integration, yields

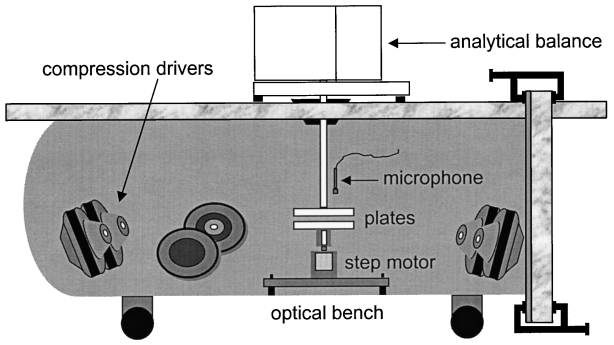


FIG. 1. Apparatus to measure the force between two rigid parallel plates due to the radiation pressure of broadband acoustic noise. The tank is made of 0.64-cm steel, and has length 1.5 m and diameter 0.5 m. One end is ellipsoidal while the other is flat with a steel faceplate to which is mounted a 5-cm-thick, 61.5-cm-square acrylic access cover secured by four C-clamps. A sliding 5-cm-thick acrylic bar allows the positioning of the balance at a desired location and also serves for spectral measurements of the noise along the tank. A microphone was positioned within 1 cm of the top plate to provide spectrum intensity measurements.

$$P = I_\omega \int_{\theta=0}^{\pi/2} d\theta \, \sin(\theta)\cos^2(\theta) \int_{k=k_1}^{k_2} dk = \frac{1}{3}(k_2 - k_1) I_\omega , \tag{5}$$

which can be expressed simply in terms of the total intensity as

$$P = \frac{I}{3c}. \tag{6}$$

Although Eq. (6) was derived for a constant spectral density I_ω, it can readily be shown to be valid for *any* isotropic spectral distribution. Equation (6) can be understood by a simplistic approach: If one-sixth of the intensity is considered to be propagating in each of six possible orthogonal directions, then the radiation pressure (2) becomes $2(I/6c)$, which is identical to Eq. (6).

Broadband noise outside two parallel rigid plates drives the discrete modes between the plates. For convenience, we continue to deal with the traveling wave (rather than standing wave) modes, which we label with wave vector components $k_x = n_x \pi / L_x$, $k_y = n_y \pi / L_y$, and $k_z = n_z \pi / L_z$, where n_x, n_y, and n_z are signed integers and L_x, L_y, and L_z are the dimensions between the plates. As in the previous analysis, the z axis is chosen to be normal to the plates. As a result of the quality factor of the modes being approximately unity, the intensity $I_{\text{in}}(\mathbf{k})$ of each mode between the plates is expected to be approximately the same as the outside broadband intensity in a bandwidth equal to the wave vector spacing of the inside modes: $I_{\text{in}}(\mathbf{k}) = I_{\mathbf{k}} \Delta k_x \Delta k_y \Delta k_z$, where $I_{\mathbf{k}}$ is given by Eq. (3) and where $\Delta k_i = \pi / L_i$. In the limit of large dimensions, this expression for the inside intensity yields the correct wave vector spectral intensity $I_{\mathbf{k}}$. From Eq. (2), the pressure on a plate due to the inside modes is

$$P_{\text{in}} = \frac{2}{c} \sum I_{\text{in}}(\mathbf{k}) \frac{k_z^2}{k^2}, \tag{7}$$

where the sum is over values of $\mathbf{k}$ corresponding to modes that strike the plate, which we can take to be those for which $n_z > 0$. We assume that the dimensions L_x and L_y of the plates are sufficiently large that the corresponding components of the wave vectors are essentially continuous. In comparison to Eq. (4), the total inside pressure (7) is therefore

$$P_{\text{in}} = \frac{I_\omega}{2\pi} \sum \Delta k_z \int dk_x \, dk_y \, \frac{k_z^2}{k^4}. \tag{8}$$

We now let $n = n_z$ and $d = L_z$, and define the fundamental transverse wave number (perpendicular to the plates) $k_0 = \pi / d$. As in Eq. (5), the range of k values in Eq. (8) is in a hemispherical shell of thickness $k_2 - k_1$. However, it is easier to calculate the radiation pressure (8) in this shell by performing an integration over a hemisphere of radius k_2 and subtracting the integration over a hemisphere of radius k_1. Employing polar coordinates for the parallel wave vector, and performing the azimuthal integration, yields

$$P_{\text{in}} = k_0^3 I_\omega \sum_{n=1}^{[k_2/k_0]} n^2 \int_0^{\sqrt{k_2^2 - k_0^2 n^2}} \frac{\kappa \, d\kappa}{(\kappa^2 + k_0^2 n^2)^2} - (k_2 \to k_1), \tag{9}$$

where $[k_2/k_0]$ is the greatest integer less than or equal to k_2/k_0. Transforming to the dimensionless variable $x = (\kappa/k_0)^2$ leads to

$$P_{\text{in}} = \frac{\pi I_\omega}{2d} \sum_{n=1}^{[k_2 d/\pi]} n^2 \int_0^{(k_2 d/\pi)^2 - n^2} \frac{dx}{(x+n^2)^2} - (k_2 \to k_1). \tag{10}$$

The difference $P_{\text{in}} - P_{\text{out}}$ between Eqs. (7) and (4) is the force f per unit area between the plates. Performing the integrals and sums in Eq. (10), and using Eq. (5), we obtain

$$f = \frac{\pi I_\omega}{2d} \left(N_2 - N_1 - \frac{N_2(N_2+1)(2N_2+1)}{6(k_2 d/\pi)^2} + \frac{N_1(N_1+1)(2N_1+1)}{6(k_1 d/\pi)^2} - \frac{2(k_2 - k_1)d}{3\pi} \right), \tag{11}$$

where $N_1 = [k_1 d/\pi]$ is the greatest integer equal to or less than $k_1 d/\pi$, and similarly for N_2. Equation (11) is continuous and piecewise differentiable, and can alternate between negative (attractive force) and positive (repulsive force) values. On the other hand, if the lower frequency in the band is zero, the force is always attractive. For the case when the spectrum extends over an infinite frequency interval $(0,\infty)$ the force (11) reduces to the value

$$f = -\frac{\pi I_\omega}{4d}, \tag{12}$$

corresponding to an attractive force. Equation (12) can be derived by noting that $k_1 = 0$ when the lower frequency is zero, so the second and fourth terms in Eq. (11) vanish, while in the limit of large values of k_2 the ratio $\pi N_2 / k_2 d$ equals unity.

Equations (1) and (12) are evidence of the direct relation that exists between the force law for two parallel plates and power law spectra. Thus while a flat spectrum yields an inverse distance force law (12), a spectrum proportional to the cube of the frequency yields force law (1) that is proportional to the inverse fourth power of the distance.

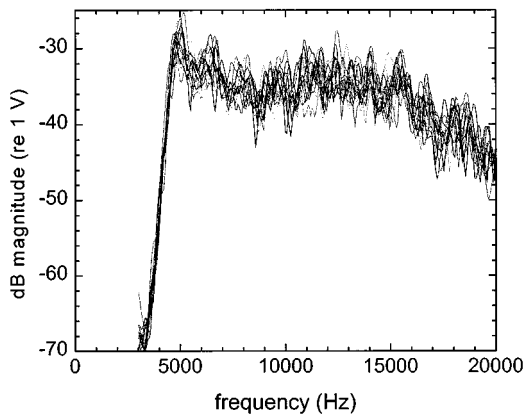


FIG. 2. Acoustic field intensity measurements for nine horizontal positions (left–center–right, 33–44–54 cm from the tank access) and three vertical positions (high–low–center), covering the internal location of the optical bench with step motor. The spectrum is essentially flat from 5 kHz to 15 kHz, and rolls off 20 dB from 15 to 20 kHz.

II. APPARATUS

In addition to the plates, the equipment for the acoustic Casimir effect experiment consists of three major components: an acoustic chamber, an optical bench with a step motor, and an analytical balance (Fig. 1). The acoustic chamber provides a nearly homogeneous, isotropic sound field. The internal optical bench with step motor provides precise plate separation, and the analytical balance provides the force measurement for the acoustic Casimir effect.

The acoustic chamber is a converted cylindrical air compressor tank mounted on wheels for mobility. The tank is made of 0.64-cm steel and has length 1.5 m and diameter 0.5 m, and is covered with 1.8-mm automotive sound deadening self-adhesive pads. The addition of the pads lowered the quality factor (Q) of flexural modes of the cylinder thereby minimizing external radiation. The drivers in the acoustic chamber are capable of providing an acoustic noise field of intensity as large as 135 dB (re: 20 μPa). One end of the cylinder is ellipsoidal while the other is flat with a steel faceplate that mounts a 5-cm-thick, 61.5-cm-square acrylic access cover plate secured by four C-clamps. There is a rubber gasket on the faceplate to improve the seal with the square acrylic plate. The ellipsoidal end is desirable because it improves isotropy of the noise.

Six J. B. Lansing (JBL) 2455J compression drivers provide the desired acoustic noise intensity within the acoustic chamber. The voltage source for the compression drivers is a General Radio Company 1390-B random noise generator, whose output is passed through three bandpass filters each followed by dual channel amplifiers. One filter is a 115-dB/octave bandpass filter [Stanford Research Systems (SRS) 650] set at a band of 5–15 kHz. The other two filters are 48-dB/octave bandpass filters (Krohn-Hite 3988) set at a band of 10–15 kHz. The separate bandpass filters allowed the flexibility of manipulating the spectral shape of the noise. The filter outputs are amplified by three dual channel Techron 5530 power amplifiers.

Acoustic field measurements were made within the tank using a sensing probe with a mounted Larson Davis $\frac{1}{4}$-in. microphone that has a flat response (± 1 dB to approximately 22 kHz). Measurements were taken at high, center, low, left and right positions along the length of the tank (14–94 cm), and showed that the six compression drivers were capable of providing a homogeneous isotropic field between 5 and 15 kHz.

Figure 2 is a plot of the acoustic field intensity measurements for nine horizontal positions (left–center–right, 33–44–54 cm from the tank access) and three vertical positions (high–low–center), covering the internal location of the optical bench with step motor. The spectrum is essentially flat from 5 kHz to 15 kHz, and rolls off 20 dB from 15 to 20 kHz.

The step motor was mounted on a small aluminum optical bench. The three-point leg adjustments of the bench compensated for the tank curvature allowing for leveling of the step motor and attached plate. The step motor is a Vexta PH266-02-A29, 2-phase, 1.8 degree/step, 12-V dc motor which provides accurate repeatable plate separation. The number of steps (1–255 steps) and direction (up or down) are varied with a microchip controller. Attached to a machined screw with 20 threads/in., the step motor mount yields a displacement ranging from 6.35 μm for one step (1.8 degrees) to 1.27 mm for 200 steps (360 degrees). The step motor displacement was verified through the range of the machined screw using a linear variable differential transformer.

Two 15.00-cm-diam plates were used for the acoustic Casimir force measurement. The bottom plate was 6.35-mm (1/4-in.)-thick aluminum, attached firmly to the top of the step motor mount. The top plate was 5.57-mm (7/32-in.) thick PVC which was vacuum aluminized and hung beneath the analytical balance by an aluminum bar screwed into the plate and attached to the balance by a hook (Fig. 1). Both plates were grounded to a common ground to eliminate electrostatic effects, thereby minimizing fluctuations in the force measurements.

For 20-cm by 20-cm plates separated less than 0.3 cm in a 120 dB (re: 20 μPa) noise field from 5- to 15-kHz frequency, the force acting on a plate due to acoustic radiation pressure is calculated from Eq. (6) to be approximately 0.1 mN, which corresponds to a weight of 10.2 mg. To meet such resolution weight measurements, we used a Mettler-Toledo AT261 Delta Range AT analytical balance with fully automatic self-calibration (FACT), ISO 9001 certification, and a resolution of 0.01 mg. The weight of a 15-cm-diam $\frac{1}{4}$ -in.-thick PVC plate is roughly 150 g, which is well within the 200-g weighing capability of the AT261.

III. MEASUREMENTS

In an experiment dealing with an acoustic analog to the Casimir effect, an important question is whether other nonzero time-averaged (dc) effects can play an important role. The only second order dc effects in acoustics are radiation pressure and streaming. Any other dc effect would be fourth order in the acoustic pressure, and at least 40 dB smaller in our case. Employing smoke in the apparatus described above, we detected no acoustic streaming when driving with broadband acoustic noise at the intensity levels used in the

experiment. Because the noise in our experiment can be thought of as collection of monochromatic waves over a band of frequencies with randomly varying phases, we would expect very little or no streaming when the characteristic time of phase variations is less that the diffusion time. Furthermore, because acoustic streaming is driven along the boundary from a pressure antinode to a pressure node, in the presence of broadband noise the pressure nodes and antinodes of the different noise components are densely distributed along the boundary, thus reducing or eliminating the streaming.

To measure the acoustic Casimir force, the optical bench with step motor was centered in the acoustic chamber 46 cm from the tank access. As indicated by a leveling bubble, the three height adjustment screws were used to level the step motor mount with the bottom plate attached. The top plate was hung from the Mettler-Toledo AT261 analytical balance and the bottom plate was raised beneath it. The plates were visually aligned to be parallel, and the initial plate spacing was determined with a spark plug gap gauge. The plate separation distance was verified to be uniform by measurements at different locations between the plates. A Larson Davis 2530 microphone was positioned within 1 cm of the top plate to provide spectrum intensity measurements. The acoustic chamber was sealed and the system allowed to stabilize. Once the oscillations of the top plate damped out and the balance reading locked in on the weight of this plate, which was 168.1653 g, the balance was tared to zero. All force measurements were made relative to this zero.

Force measurements were carried out sequentially as follows. The acoustic noise field was turned on for each step and the force measurement recorded once the balance readout first locked-in. Simultaneously, the internal microphone measured the sound field intensity which was averaged over 50 measurements with a HP 35670A dynamic signal analyzer. The sound field was then turned off and the reading of the balance was verified to return to zero. The plates were displaced an additional 1.27 mm (200 steps of the step motor), and the measurement procedure described above was repeated through the full 6.3-cm range of the step motor mount.

We first selected a uniform broadband noise spectrum between roughly 5 and 15 kHz. The lower limit was selected well above the lower modes of the tank in order to excite a noise distribution as homogeneous as possible. The upper limit was due to the JBL 2445J compression drivers rolling off 20 dB from 15 to 20 kHz. Figure 3(a) shows the measured noise spectrum which, except for 5-dB variations throughout, is nearly flat within the spectral range of 4.8–16 kHz. The total intensity is 133 dB (re: 10^{-12} W/m^2).

Using a spark plug gauge, we measured the initial plate separation to be 0.76 mm. As shown in Fig. 3(b), throughout a range of distances less than the smallest half-wavelength (no modes between the plates), the measured force is approximately independent of distance, and agrees with the expected value for an intensity level of 133 dB. When the half-wavelength of the highest frequency fits between the plates (10.63 mm for 16 kHz), the force begins to decrease nonmonotonically and becomes repulsive (i.e., negative) at

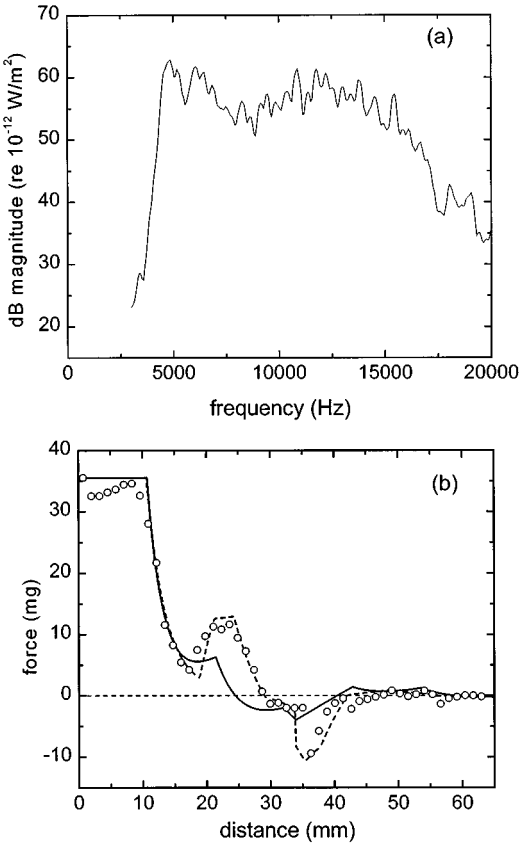


FIG. 3. (a) Experimental spectrum in a band of frequencies between 4.8 and 16 kHz. The spectrum is relatively flat and exhibits structure (dip) and an overall 5-dB roll off. The total intensity of the noise is 133 dB (re: 10^{-12} W/m^2). (b) Force between two parallel rigid plates as a function of the distance between them. The points are experimental data, the solid curve is from the theory (11) with no adjustable parameters, and the dashed curve is obtained from Eqs. (4) and (7) with a spectrum that approximates the experimental spectrum.

30–40 mm. Based on the experimental spectral range of 4.8–16 kHz for the theory, with a total intensity of 133 dB, the solid curve shows good quantitative agreement between experiment and theory [Eq. (11)] which has no adjustable parameters. The differences between experiment and theory could be a result of the lack of smoothness of the spectrum, the small roll off of 5 dB over the band, the lack of a sharp cutoff at the upper end of the band, and Q not necessarily being equal to unity. Based on Eqs. (4) and (7), calculations that incorporate power law variations from a uniform spectrum have been performed.[12] However, in contrast to the case of a flat spectrum [Eq. (11)], there is not a closed form expression for the force law. The dashed curve in Fig. 3(b) is the result of incorporating these modifications in the spectrum. The theory now fully accounts for the previous deviations, indicating that Q does not play a significant role, and that radiation pressure is the dominant mechanism.

The repulsive force can be understood as follows. When the distance between the plates is comparable to the half-wavelength associated with the lower edge of the frequency band, the corresponding modes inside the plates have wave vectors that are nearly perpendicular to the plates. However, the modes outside the plates corresponding to the same fre-

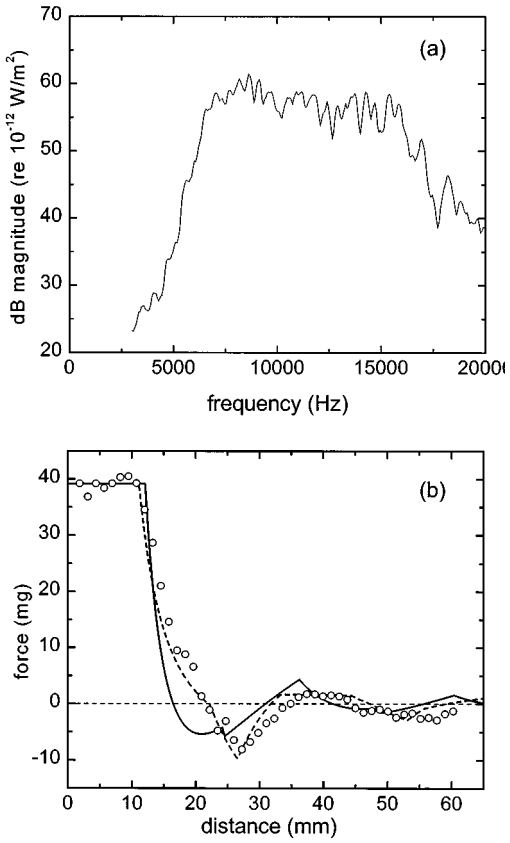


FIG. 4. (a) Experimental spectrum in a band of frequencies between 7.4 and 15.3 kHz. The spectrum shows less structure and a flatter spectral shape compared to Fig. 3(a). The total intensity of the noise is 133.5 dB (re: 10^{-12} W/m^2). (b) Force between two parallel rigid plates as a function of the distance between them. The points are experimental data, the solid curve is from the theory (11) with no adjustable parameters, and the dashed curve is obtained from Eqs. (4) and (7) with a spectrum that approximates the experimental spectrum.

quencies are spread over all possible angles of incidence. Thus for the same total intensity, the momentum transfer due to waves inside the plates is over a narrow cone while the momentum transfer due to waves outside the plates extends over all angles, leading to a repulsive force.

Next we selected an octave frequency band of roughly 7.5–15 kHz to investigate what effect a different spectrum would have on the acoustic Casimir force. Figure 4(a) shows the measured spectrum of the noise, which is nearly flat between 7.4 and 15.3 kHz. The total intensity is 133.5 dB.

With this spectrum, we set the initial plate separation to 0.96 mm. Figure 4(b) shows again that the force between the plates is independent of distance up to about 11.11 mm, corresponding to the half-wavelength of the highest frequency. Note that a repulsive force starts at a smaller distance (20 mm), covering a larger range (20–40 mm) than in the previous case, in agreement with theory. The solid curve corresponds to the theory (11) based on the experimental spectral range of 7.4–15.3 kHz, with a total intensity of 133.5 dB. The dashed curve is obtained from Eqs. (4) and (7) with a spectrum that approximates the experimental spectrum.

IV. CONCLUSIONS

In the presence of broadband acoustic noise, there is a force between two rigid parallel plates due to an imbalance in the radiation pressure between and outside the plates. The force can be attractive or repulsive as a function of the separation between the plates, and there is excellent quantitative agreement between theory and experiment. The theory reveals that the repulsive force is *not* due to possible Q-amplification of the modes between the plates, but rather to the noise being in band of frequencies with a nonzero lower cutoff. This successful experiment has opened up the field of Casimir acoustics as a viable means to investigate the probe by analogy some effects due to the zero point field.

Experimental evidence of attractive and repulsive forces within a finite acoustic bandwidth also suggest new means of acoustic levitation. The force between two objects can be manipulated by changing the distance between the objects and/or varying the spectrum. While the Casimir force is small compared to the force of the Earth's gravity, in a low gravity environment a method of material control through the manipulation of an acoustic noise spectrum or plate geometry may be possible.

The acoustic Casimir effect can also be a potential tool in noise transduction. Through direct measurement of the force between two plates, or any other cavity, it may be possible to determine background noise. The shape of the force over distance is an instantaneous time average over all frequencies and may provide an alternative to measurements of background noise.

ACKNOWLEDGMENTS

We are grateful to Metler-Toledo for the generous loan of their AT261 analytic balance for the duration of this experiment. This work was sponsored in part by the Naval Postgraduate School Merit Research Program, and by the Office of Naval Research.

[1] H. B. G. Casimir, ''On the attraction between two perfectly conducting plates,'' Proc. K. Ned. Akad. Wet. **51**, 793–796 (1948).
[2] E. Elizalde and A. Romeo, ''Essentials of the Casimir effect and its computation,'' Am. J. Phys. **59**, 711–719 (1991).
[3] S. K. Lamoreaux, ''Demonstration of the Casimir force in the 0.6 to 6 μm range,'' Phys. Rev. Lett. **78**, 5–8 (1997).
[4] M. J. Sparnaay, ''Measurements of attractive forces between flat plates,'' Physica (Utrecht) **24**, 751–764 (1958).
[5] E. M. Lifshitz, ''The theory of molecular forces between solids,'' Sov. Phys. JETP **2**, 73–83 (1956); J. Schwinger, L. DeRaad, and K. A. Milton, ''Casimir effect in dielectrics,'' Ann. Phys. (Paris) **115**, 1–23 (1978).
[6] J. J. Sakurai, *Advanced Quantum Mechanics* (Addison-Wesley, London, 1967), Secs. 2-4 and 2-8.
[7] T. Boyer, ''Derivation of the blackbody radiation without quantum assumptions,'' Phys. Rev. **182**, 1374–1383 (1969).
[8] F. S. Crawford, ''Elementary examples of adiabatic invariance,'' Am. J. Phys. **58**, 337–344 (1990).
[9] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon, London, 1959), Sec. 64.
[10] Lord Rayleigh, ''On the pressure of vibrations,'' Philos. Mag. **3**, 338–346 (19022); Lord Rayleigh, ''On the momentum and pressure of gaseous vibrations, and the connection with the virial theorem,'' Philos. Mag. **10**, 364–374 (1905).
[11] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon, London, 1959), Sec. 65, considered a plane wave incident at an arbitrary angle at the interface of two fluids. The radiation pressure at a rigid interface is obtained in the limit when the impedance of one of the fluids is infinite.
[12] A. Larraza and B. Denardo, ''An acoustic Casimir effect,'' Phys. Rev. Lett. (submitted).

The effect of compression ratio and release time on the categorical rating of sound quality[a)]

Arlene C. Neuman, Matthew H. Bakke,[b)] Carol Mackersie,[c)] Sharon Hellman, and Harry Levitt

Center for Research in Speech and Hearing Sciences, Graduate School, City University of New York, 33 West 42nd Street, New York, New York 10036

(Received 19 March 1996; accepted for publication 26 November 1997)

Two experiments were carried out to determine how manipulating the compression ratio and release time of a single-band wide dynamic range hearing aid affects sound quality. In experiment I, compression ratio was varied over the range from linear to 10:1 (low compression threshold, attack time=5 ms, release time=200 ms). In experiment II, compression ratios of 1.5, 2, and 3:1 were combined with release times of 60, 200, and 1000 ms (attack time=5 ms). Twenty listeners with sensorineural hearing loss rated the clarity, pleasantness, background noise, loudness, and the overall impression of speech-in-noise (Ventilation, Apartment, Cafeteria) processed through a compression hearing aid. Results revealed that increasing compression ratio caused decreases in ratings on all scales. Increasing release time caused ratings of pleasantness to increase, and ratings of background noise and loudness to decrease. At the 3:1 compression ratio, increasing the release time caused increases in ratings of clarity, pleasantness, and overall impression, and a decrease in background noise. Significant correlations were found between scales. Regression analysis revealed that the contributions of the scales of clarity, pleasantness, background noise, and loudness to the prediction of overall impression differed as a function of the competing noise condition. [S0001-4966(98)03903-4]

PACS numbers: 43.10.Ln, 43.66.Ts, 43.71.Ky, 43.38.Lc [JWH]

INTRODUCTION

In recent years, wide dynamic range compression has been incorporated into hearing aids with the purpose of achieving an increase in speech intelligibility and/or increased comfort for a wider range of input levels. At comfortable listening levels, a well-fit compression hearing aid and a well-fit linear hearing aid result in similar speech recognition performance (see meta-analysis of compression studies by Dillon, 1996). It is generally agreed that wide dynamic range compression amplification will result in better speech recognition performance than will linear amplification for low input levels, provided that the user does not adjust the gain of the linear hearing aid. The success of wide dynamic range compression in expanding the range of input signals that can be listened to comfortably without volume control adjustment is undisputed [see Dillon (1996) for an extensive review of the literature].

These two benefits of wide dynamic range compression can be achieved with either syllabic compression or with slow-acting compression. Release times <150 ms are considered to be syllabic, >150 ms are slow-acting (Johannson and Lindblad, 1971). Several researchers have argued the potential usefulness of slow-acting compression (Braida *et al.*, 1979; Walker and Dillon, 1982; Van Tasell, 1993). In the past, most compression hearing aids had fast attack and release times (syllabic compression). Hearing aids with longer release times were not commonly available. In the most recent generation of hearing aids, programmable and digital hearing aids allow control over many compression parameters. In some hearing aids, attack and release time may be controlled directly. When this option is available, release times may vary over a large range (e.g., 7–500 ms in one hearing aid currently in use). In some hearing aids, the option of using longer release times is now made more attractive by incorporating multiple attack and release times or adaptive attack and release times. This allows fast-acting gain control for high level transient sounds and slow-acting gain control for changes that are primarily due to long-term changes in input level. Multiple attack and release times alleviate the concern that the use of long release times will prevent the hearing of low level sounds that follow high level sounds.

Digitally programmable and digital wide dynamic range compression hearing aids may also allow control of compression ratio and compression threshold. Compression ratios most commonly can be varied over the range from 1:1 to 5:1 in either a single band or two bands. Compression thresholds may also be varied over a range (usually 45–65 dB SPL). Therefore wide dynamic range compression hearing aids may include either short or long time constants and a large range of compression ratios.

Since either slow-acting or syllabic compression can be used to increase user comfort and to expand the listener's dynamic range, knowledge of how compression parameters affect sound quality should be considered in the selection of compression parameters. However, the effect of compression amplification on speech quality has been largely overlooked.

[a)]''Selected research articles'' are ones chosen occasionally by the Editor-in-Chief that are judged (a) to have a subject of wide acoustical interest, and (b) to be written for understanding by broad acoustical readership.

[b)]Present address: Lexington Center, Inc., Research Division, 30th Avenue and 75th Street, Jackson Heights, NY 11370.

[c)]Present address: Department of Communication Disorders, San Diego State University, San Diego, CA 92182.

A recent study of Neuman and colleagues (1994) showed that sound quality preferences of hearing-impaired listeners were significantly affected by compression ratio, the competing noise type and level, and the dynamic range of the listener. Paired-comparison judgments of sound quality were obtained for a slow-acting (5-ms attack time, 200-ms release time) single-band compression amplifier. Lower compression ratios were judged to have significantly better sound quality than compression ratios greater than 3:1. Background noise level was found to interact with compression ratio. On average, subjects selected a linear hearing aid or the lowest compression ratio (1.5:1) for the noise having the highest level, and slightly more compression (up to 2:1) for two other noises presented at lower levels.

An early study by Nabelek (1984) found that hearing-impaired listeners were relatively insensitive to the manipulation of attack and release time (discrimination paradigm). Similarly, Bentler and Nelson (1997) failed to find significant differences in the preferences of hearing-impaired listeners for release time in a two-channel AGC hearing aid. In contrast, Neuman and colleagues (1995b) found that sound quality judgments were differentially affected by release time as a function of the background noise and compression ratio (release time was varied from 60 ms to 1000 ms, in combination with compression ratios of 1.5, 2, and 3:1). Longer release times were selected with the higher level background noises.

In the studies of the effect of compression on speech quality by Neuman and colleagues (1994, 1995b), the criterion of quality was not defined for the subject. Because sound quality is a multidimensional attribute, the instructions to the subject were purposely general in order to allow the listener to use his/her own definition of sound quality. Therefore, it is unknown what criterion subjects were using or whether common criteria were used among subjects.

The purpose of the present study was to determine: (1) how manipulating compression ratio and release time affects the perception of specific aspects of sound quality; i.e., clarity, pleasantness, background noise, loudness, and overall impression of sound; (2) the relation among these different aspects of sound quality; and (3) the relation between the global rating of overall impression and the more specific quality ratings. In experiment I, compression ratio was manipulated in a single-band compression hearing aid with a 200-ms release time. In experiment II, both release time and compression ratio were manipulated.

I. METHOD

A. Subjects

Twenty listeners with sensorineural hearing loss (mean age of 59.8 years) served as subjects. These subjects had served in two earlier experiments in which paired-comparison judgments of quality were obtained through the same experimental compression amplification system (Neuman *et al.*, 1994, 1995b). A complete description of the personal and audiometric characteristics of these subjects may be found in Neuman *et al.* (1994). Briefly, half of the subjects had a small dynamic range for speech ($\leqslant$30 dB between the speech reception threshold and loudness discomfort level) and half of the subjects had a large dynamic range ($>$30 dB). The small dynamic range group of subjects had moderately severe or severe degree of hearing loss with sloping configuration. The large dynamic range group had moderate sloping sensorineural hearing loss. All subjects were experienced hearing aid users. Fifteen of the subjects wore linear hearing aids. The remaining five subjects wore linear hearing aids with compression limiting.

B. Instrumentation and stimuli

A digital test system consisting of a personal computer and an Ariel DSP 16 board was used to simulate a hearing aid with input compression [model and formulas were those of Kates (1993)]. In the compression simulation, compression occurs when the control voltage exceeds the specified compression threshold (peak detection circuit). Post-compression filtering (64 tap finite impulse response filter) was used to implement the NAL-R frequency response (Byrne and Dillon, 1986). In calculating the target frequency response, the real-ear constants specified by Byrne and Dillon were included in the formula, as were real-ear-to-headphone corrections (Cox, 1986). These corrections were used in order to make the signal at the listener's eardrum as similar as possible to that which would have been obtained had a conventional hearing aid been used.

A two-channel digital audio tape recorder was the source of the speech and noise stimuli. The speech and noise levels were adjusted at the outputs of the tape recorder using two programmable attenuators. The speech and noise signals were then mixed and passed through an anti-aliasing filter (7.5-kHz low pass filter). Signals were digitized (20-kHz sampling rate, 16-bit quantization), compressed, filtered, and amplified by the simulated hearing aid. The processed signal was converted back to an analog signal, passed through an anti-imaging filter (7.5 kHz), attenuated, amplified, and sent to a TDH-50P earphone (mounted in a MX-41/AR cushion) for monaural presentation to the listener.

Both the static and dynamic characteristics of the simulated compression hearing aids were defined and measured using the procedures specified in ANSI S3.22-1987 (ANSI, 1987). Calibration of input levels was verified during each experimental session by checking the electrical input. Output levels were verified acoustically in a 6-cc coupler.

C. Experimental protocol

Continuous discourse (female talker) was presented at a level 20 dB above the compression threshold. This level was used to guarantee that for all conditions the speech would be fully compressed. Three different competing noises were used. The noises varied in their temporal characteristics as well as in level. The ventilation noise was used to represent a quiet environment with a constant background noise (ventilation noise), the apartment noise was used to represent an environment with fluctuating noise, and the cafeteria noise was used to represent a somewhat more difficult listening environment with both a constant level and fluctuations superimposed. Ventilation noise was 15 dB below the com-

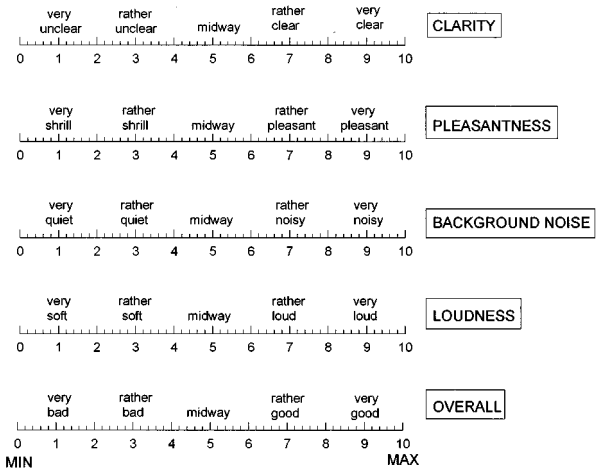


FIG. 1. A sample of the response form used for the rating procedure.

pression threshold, apartment noise was 5 dB below the compression threshold, and the cafeteria noise was 5 dB above the compression threshold. It is recognized that the signal-to-noise ratios used are better than would be found in most environments. The use of poorer signal-to-noise ratios would have degraded speech intelligibility. For the purposes of this experiment we wanted to maintain good speech intelligibility across conditions. Only then, can differences in sound quality be measured independently of changes in speech intelligibility (Preminger and Van Tasell, 1995).

D. Rating procedure

The method of categorical rating as developed and implemented by Gabrielsson and colleagues was used (Gabrielsson *et al.*, 1979b,c, 1988, 1990). Ratings are obtained on a scale ranging from 0 to 10. Five scales were used: clarity, pleasantness, background noise, loudness, and overall impression. The loudness scale was included as a check of the loudness matching procedure used in the experiment.

Subjects were instructed to listen to speech through the simulated hearing aid in a given background noise and to rate what they heard on each of the five rating scales that appeared on the response form. An example of the response form appears in Fig. 1. Subjects were given written instructions for the task (see the Appendix) and were given the opportunity to ask for clarification of the instructions.

Testing was preceded by a practice/training procedure designed to determine that subjects understood the task and would assign diverse ratings when the hearing aid parameters were changed. The training task included conditions that were expected to yield a range of scores on several rating scales. Conditions included a change in input level (linear hearing aid with the output at MCL, a hearing aid with the output 20 dB below MCL) and a large change in compression characteristics (linear hearing aid and a hearing aid with a 10:1 compression ratio). Speech presented in a background of cafeteria noise (15-dB signal-to-noise ratio) was the input signal for all practice conditions. The ratings of the subjects were examined to determine that the clarity, loudness, and background noise ratings decreased as the hearing aid output was lowered. Based on our previous study showing a decrease in the perceived quality of hearing aids with high compression ratios, we also expected that ratings on several of the scales would be poorer for the compression hearing aid than for the linear hearing aid. Subjects were allowed to ask questions during this practice/training phase. The collection of data was begun when the subject felt that he/she understood the task and when the experimenter had ascertained that the subject's ratings were consistent for a given listening condition.

TABLE I. Summary of the main effects from the analysis of variance done for each rating scale for experiment I. Significance levels are indicated as follows: $*p<0.05$, $**p<0.01$, $***p<0.001$.

		Clarity	Pleasant	Noise	Loud	Overall
Source	df	F	F	F	F	F
DR	1,1	0.04	0.02	0.64	5.49*	1.28
CR	5,90	17.00**	11.14**	46.33**	1.29	23.89**
Noise type	2,36	5.19**	3.95*	58.91**	0.01	17.74***

II. EXPERIMENT I

Compression ratio (CR) was varied (CR=1, 1.5, 2, 3, 5, and 10:1). All other parameters of the compression system were held constant (compression threshold was 20 dB below the rms level of the speech, attack time=5 ms, and release time=200 ms). The gain of each hearing aid was specified to place the processed speech at the midpoint of the listener's most comfortable listening range (Dirks and Kamm, 1976). Since loudness varied with compression ratio, loudness matching was achieved by equating the 90th percentile of the cumulative distribution of the compressed speech to the same point on the cumulative distribution of uncompressed speech (Levitt and Neuman, 1991; Bakke *et al.*, 1991).

Ratings were obtained for each of the six compression ratios (1, 1.5, 2, 3, 5, and 10:1) in each of three different competing noises (ventilation, apartment, and cafeteria). Data were collected in blocks of 18 trials (six hearing aids ×three replications for a particular background noise). A trial consisted of rating a particular hearing aid on the five scales. The order of presentation of the hearing aids for the three replications was randomized within a block. The order of testing in different background noise conditions was also randomly assigned to subjects.

III. RESULTS

A. Analysis of variance

The procedures for statistical analysis outlined by Gabrielsson (1979a) were followed. A separate analysis of variance was carried out for each rating scale to determine the effect of compression ratio (CR), noise type (N), dynamic range group (DR), and their interactions. A general mixed model analysis of variance was used with subjects nested within a group factor.

A summary of the five analyses appears in Table I. Compression ratio and noise type were found to be significant for all of the scales except for that of loudness. Dynamic range group was found to be a significant factor only for the loudness scale. None of the interactions was significant.

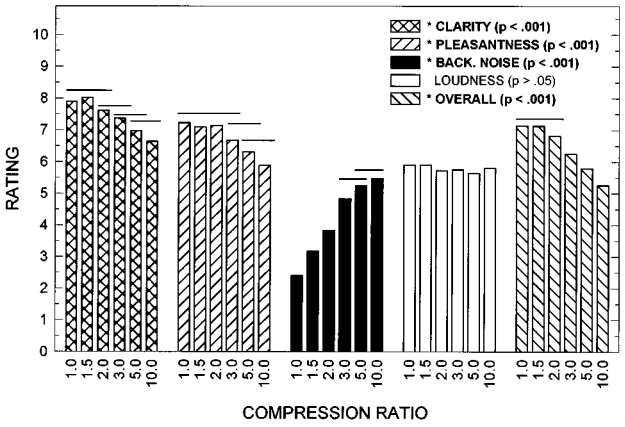


FIG. 2. Mean ratings on five quality scales as a function of compression ratio. Significance levels for each scale appear in parentheses.

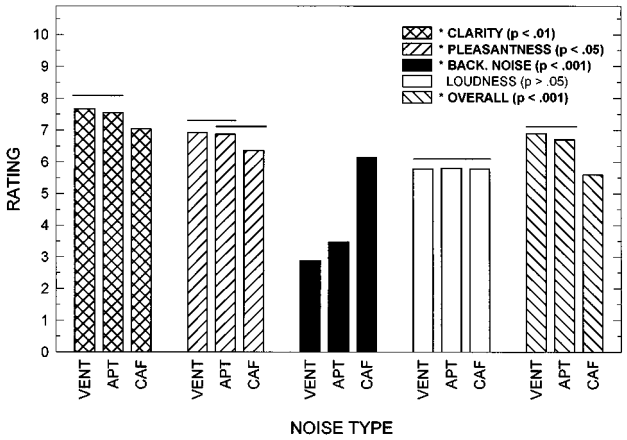


FIG. 3. Mean ratings on five quality scales as a function of noise type. Significance levels for each scale appear in parentheses. The horizontal lines above the graph indicate means that do not differ significantly ($p>0.05$).

The mean scores representing the main effect of compression ratio and the results of the tukey *post-hoc* analysis appear in Fig. 2. In general, increasing the compression ratio causes a decrease in the ratings of clarity, pleasantness, and overall impression. The rating of background noise increases with increasing compression ratio, i.e., the sound is perceived as being noisier and quality is degraded. For all four scales, ratings for compression ratios of 1, 1.5, and 2:1 were significantly better than for compression ratios of 5:1 or 10:1. The loudness ratings demonstrate that loudness has been matched successfully. Mean scores remain constant, approximately at the middle of the 11 point scale.

The main effect of noise type is illustrated in Fig. 3. Recall that the three different noises differ in level; the ventilation noise is lowest in level and the cafeteria noise is highest in level. The ratings of clarity, pleasantness, and overall impression decrease significantly in the cafeteria noise condition and there is a general trend for ratings to decrease with increasing noise level. The rating of background noise increases significantly as noise level increases. Loudness is the only scale whose ratings are unaffected by the noise type/level.

The factor of dynamic range group was found to be significant only for the loudness scale. The mean loudness rating for the small dynamic range group was 6.2 and for the large dynamic range group was 5.4. An examination of the means as a function of compression ratio reveals that the small dynamic range group consistently assigned higher loudness ratings than the large dynamic range group. This, despite the fact that for all subjects the output of the hearing aid was placed at the measured MCL (the midpoint of the MCL range). The finding is consistent with the results of a previous study on the same group of subjects that revealed that the preferred listening level of the small dynamic range group of subjects was slightly below the midpoint of the MCL range, while the preferred listening level of the large dynamic range group was slightly above the midpoint of the MCL range (Neuman *et al.*, 1995a).

B. Relationship among scales

A correlation analysis was done to determine the relationships among the different scales. The results of these analyses appear in Table II. For all three noise conditions, significant correlations were found between scales. Only the Loudness scale was not significantly correlated with the other scales. The pattern of correlations found for the apart-

TABLE II. Correlation between rating scales as a function of competing noise type (V=ventilation, A=apartment, C=cafeteria). Results for experiment I appear on the left and for experiment II on the right. Marked correlations are significant at $p<0.05$.

		Experiment I				Experiment II			
Noise type		Pleas.	Noise	Loud.	Over.	Pleas.	Noise	Loud.	Over.
V	Clarity	0.85*	−0.57*	0.11	0.80*	0.79*	−0.62*	0.14	0.79*
A		0.79*	−0.55*	0.22*	0.79	0.75*	−0.37*	0.03	0.76*
C		0.70*	−0.56*	0.12	0.67*	0.84*	−0.49*	−0.28*	0.54*
V	Pleas.		−0.59*	0.03	0.75*		−0.52*	0.10	0.66*
A			−0.52*	−0.06	0.66*		−0.44*	−0.21*	0.64*
C			−0.58*	−0.11	0.70*		−0.62*	−0.52*	0.56*
V	Noise			0.04	−0.66*			0.06	−0.48*
A				−0.02	0.41*			0.36*	−0.37*
C				0.32*	−0.56*			0.56*	−0.40*
V	Loud.				0.12				0.33*
A					0.41*				0.13
C					0.11				−0.13

TABLE III. Results of the ridge regression analysis as a function of noise type (V=ventilation, A=apartment, C=cafeteria). Overall impression is the dependent variable. Results for experiment I appear on the left and for experiment II appear on the right. Values followed by * are significant ($p<0.01$).

		Experiment I		Experiment II	
Noise	Scale	Beta	*t*	Beta	*t*
V	Clarity	0.40	4.89*	0.42	9.29*
V	Pleasantness	0.22	2.69*	0.32	7.83*
V	Noise	−0.28	−4.58*	−0.06	−1.77
V	Loudness	0.08	1.48	0.26	9.04*
A	Clarity	0.37	4.74*	0.41	11.23*
A	Pleasantness	0.26	3.47*	0.29	7.81*
A	Noise	−0.18	−3.08*	−0.14	−4.59*
A	Loudness	0.30	5.73*	0.16	5.51*
C	Clarity	0.22	2.68*	0.11	4.24*
C	Pleasantness	0.38	4.78*	−0.01	−0.51
C	Noise	−0.26	−3.38*	−0.01	0.31
C	Loudness	0.19	3.04*	0.78	28.78*

ment and cafeteria noise conditions were similar to those for the ventilation noise condition, but correlations were weaker.

When subjects rate a hearing aid for overall impression, presumably they are using several different criteria simultaneously. Significant correlations were found between the overall impression scale and clarity, pleasantness, and background noise for all three noise conditions. A separate multiple linear regression analysis was carried out for each noise condition to determine which scales were predictive of the overall impression scale. Because of the significant correlation between clarity and pleasantness, a ridge regression analysis was done to control for the problem of multicollinearity. In the analysis, a lambda of 0.10 was used (CSS Statistica, 1991). Results of the analysis appear in Table III.

All of the ridge regression analyses were significant [ventilation noise: $F(4,115)=61.48$, $p<0.0001$, $R^2=0.68$; apartment noise: $F(4,115)=63.74$, $p<0.0001$, $R^2=0.69$; cafeteria noise: $F(4,115)=39.26$, $p<0.0001$, $R^2=0.58$). The beta values indicate that the predictive value of the independent variables varies with the competing noise condition. Clarity is most heavily weighted in the ventilation and apartment noise conditions. In the cafeteria noise, pleasantness is more heavily weighted.

IV. EXPERIMENT II

The same subjects, instrumentation, and rating scales were used as for experiment I. Compression ratio (CR=1.5, 2, or 3:1) and release time (60, 200, or 1000 ms) were varied. All other parameters of the compression system were held constant. The 90th percentile of the cumulative distribution of the uncompressed and compressed speech signals were matched for the three compression ratios for the 200-ms release time. This allowed for small variations in level within a compression ratio among release times. All gain values placed the output within ±2 dB of the midpoint of the MCL range.

Data were collected in blocks of 27 trials (nine hearing aids×three replications for a particular background noise). A trial consisted of rating a particular hearing aid on the five scales. The order of presentation of the hearing aids for the three replications was randomized within a block. The order of testing in different background noise conditions was also randomly assigned to subjects.

V. RESULTS

A. Analysis of variance

A separate analysis of variance was carried out for each rating scale to determine the effect of compression ratio (CR), release time (RT), noise type (N), dynamic range group (DR), and their interactions. A general mixed model analysis of variance was used with subjects nested within a group factor.

A summary of these analyses appears in Table IV. The factors of release time, compression ratio, and noise type were all found to be statistically significant on several of the scales. The factor of dynamic range was not found to be statistically significant. A number of interactions (CR×RT, RT×N, and DR×N) were also statistically significant on various scales.

The main effect of release time is illustrated in Fig. 4. The factor of release time was found to be significant on the pleasantness, background noise, and loudness scales. Clarity and overall impression ratings were unaffected by changes in release time. The results of the tukey *post hoc* analysis revealed that ratings of pleasantness were significantly higher for the 200- and 1000-ms release time than for the 60-ms release time. Background noise and loudness ratings decreased significantly with each increase of release time.

The main effect of compression ratio is illustrated in Fig. 5. Significant effects were found for all scales. The tukey *post hoc* analysis revealed that ratings of clarity and overall impression decreased significantly with each increase of compression ratio. Ratings of pleasantness and loudness

TABLE IV. Summary of the main effects and significant interactions from the analyses for experiment II. Significance levels are indicated as follows: *$p<0.05$, **$p<0.01$, ***$p<0.001$.

Source		Clarity	Pleasant	Noise	Loud	Overall
	df	F	F	F	F	F
DR	1,18	0.00	0.50	0.31	4.92*	1.14
CR	2,36	17.09***	8.95***	46.03***	4.64*	25.70***
Noise type	2,36	9.92***	5.30**	39.91***	1.35	18.27***
Release	2,36	0.09	9.63***	39.02***	65.22***	0.55
DR×N	2,36	6.55**	4.62*	1.47	2.59	4.31*
CR×RT	4,72	3.17*	2.70*	6.12***	2.35	3.21*
RT×N	4,72	0.89	0.35	4.45***	0.86	0.34

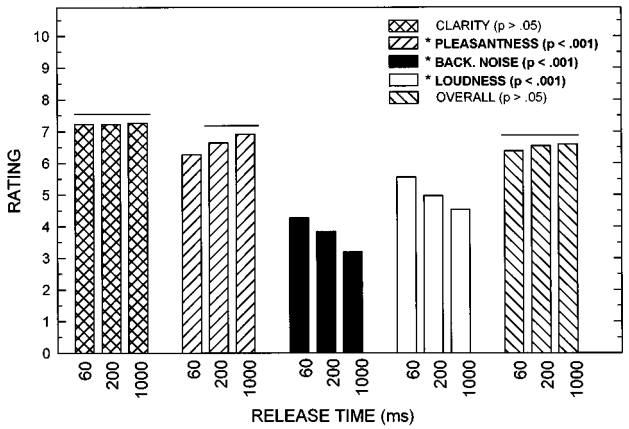


FIG. 4. Mean ratings on five quality scales as a function of release time. Significance levels for each scale appear in parentheses. The horizontal lines above the graphs indicate means that do not differ significantly ($p>0.05$).

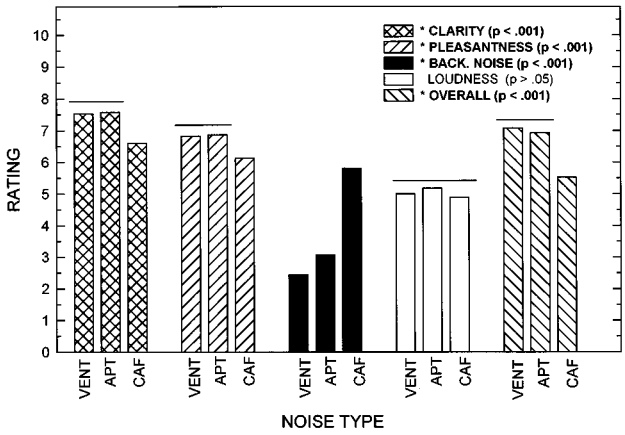


FIG. 6. Mean ratings on five quality scales as a function of noise type. Significance levels for each scale appear in parentheses. The horizontal lines above the graphs indicate means that do not differ significantly ($p>0.05$).

both decreased significantly at the 3:1 compression ratio. Background noise ratings increased significantly with each increase of compression ratio. Note that for the background noise scale, an increase in the rating indicates more noise and should be interpreted as a decrement in performance.

The main effect of noise type is illustrated in Fig. 6. *Post hoc* testing revealed that ratings on the scales of clarity, pleasantness, and overall impression were significantly lower with the cafeteria noise (highest noise level). Background noise ratings were significantly higher with cafeteria noise than with apartment noise, which in turn were significantly higher than with ventilation noise. Loudness did not differ significantly as a function of noise type.

The interactions of release time and compression ratio for the scales of clarity, pleasantness, background noise, and overall impression appear in Fig. 7. Mean scores within a compression ratio found to differ significantly on the tukey *post hoc* test ($p<0.05$) are indicated by filled symbols. This analysis reveals that the effect of release time is significant at the highest compression ratio utilized in this study. At the 3:1 compression ratio, the clarity rating for the 1000-ms release time is significantly higher than the rating for the 60-ms release time. Ratings change significantly with each change in release time on the pleasantness, background noise, and overall impression scales at the 3:1 compression ratio. On the background noise scale, the 1000-ms release time is rated significantly better (perceived as being less noisy) at all three compression ratios.

The interaction of release time and noise type was significant on the background noise scale. The interaction is illustrated in Fig. 8. Mean scores for a particular noise that differ significantly as a function of release time are indicated

FIG. 5. Mean ratings on five quality scales as a function of compression ratio. Significance levels for each scale appear in parentheses. The horizontal lines above the graphs indicate means that do not differ significantly ($p>0.05$).

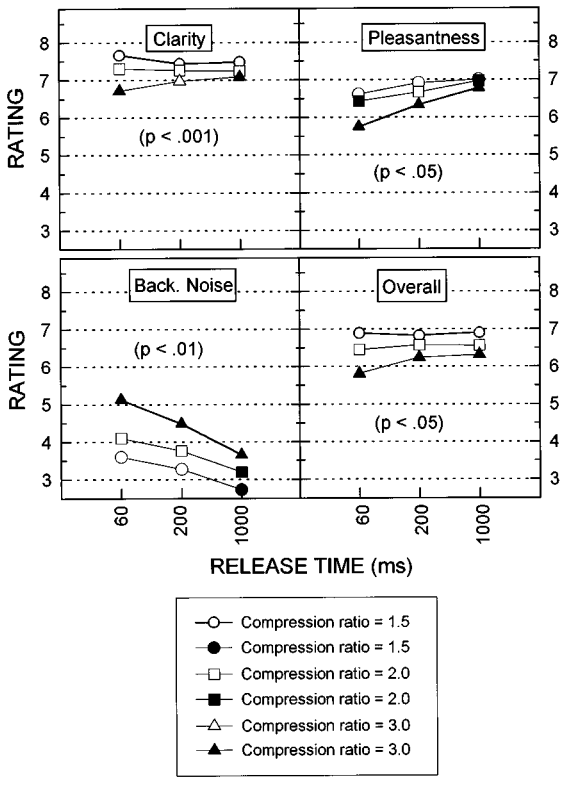


FIG. 7. Mean ratings of clarity, pleasantness, background noise, and overall impression as a function of compression ratio and release time. Significance levels for each scale appear in parentheses. Mean scores within a compression ratio found to differ significantly ($p<0.05$) are indicated by filled symbols.

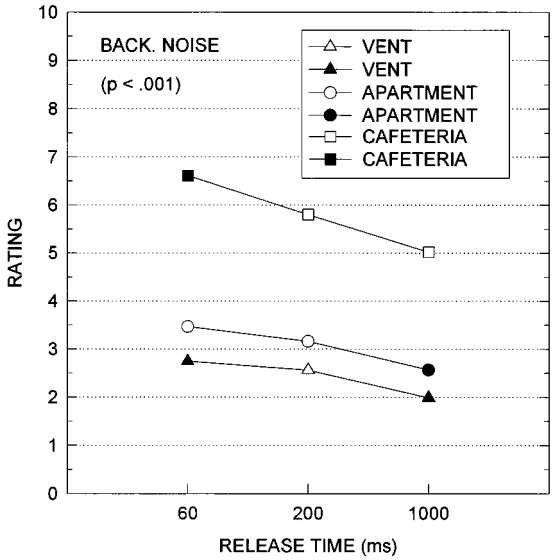


FIG. 8. Mean ratings of background noise as a function of noise type and release time. Significance levels appear in parentheses. Mean scores for a particular noise that differ significantly as a function of release time ($p < 0.05$) are indicated by filled symbols.

by filled symbols (tukey *post hoc* test, $p<0.05$). For all three noises, the ratings of background noise decreased with increasing release time, meaning that shorter release times are perceived as being noisier than longer release times. However, there are different patterns as a function of noise type.

The interaction of dynamic range group and noise type was significant on the clarity, pleasantness, and overall impression scales. The mean ratings of the small dynamic range group were significantly lower (poorer) for the cafeteria noise condition than for the other noises on all three scales. The mean ratings of the large dynamic range group did not differ significantly as a function of the noise.

B. Relationship among scales

A correlation analysis was done to determine the relationships between the different scales (see Table II). Many significant correlations are evident between the scales. The pattern of correlations is similar for the three noises.

A ridge regression analysis was done using a lambda of 0.10. In the analysis, overall impression was the dependent variable and clarity, pleasantness, background noise, and loudness were the independent variables. The results of these analyses appear in Table III. All of the ridge regression analyses were significant [ventilation noise: $F(4,535)=204.78$, $p<0.001$, $R^2=0.60$; apartment noise: $F(4,535)=160.31$, $p<0.0001$, $R^2=0.55$; cafeteria noise: $F(4,535)=218.05$, $p<0.0001$, $R^2=0.62$]. R^2values indicate that the significant scales accounted for approximately 60% of the variance. The beta values indicate that the predictive value of the independent variables varies with the competing noise condition. Whereas clarity has the most predictive weight for the ventilation and apartment noise conditions, loudness is most heavily weighted in the cafeteria noise condition.

VI. DISCUSSION AND CONCLUSIONS

The results of these two experiments reveal that manipulating either the compression ratio or the release time of a compression amplifier systematically affects the sound quality ratings of hearing-impaired listeners. The rating data supplements and helps to clarify the findings obtained in our two previous studies of the perceived sound quality of compression amplification (Neuman *et al.*, 1994, 1995b). Of the two compression parameters studied, compression ratio clearly has the larger effect on perceived sound quality. Manipulation of compression ratio significantly affects sound quality ratings whether the compression ratios cover the range from linear to compression limiting (experiment I) or only mild ($\leqslant$3:1) compression ratios (experiment II). As compression ratio is increased, ratings of clarity, pleasantness, and overall impression decrease. Ratings of background noise increase.

In contrast, the manipulation of release time has a more subtle effect on perceived sound quality. The main effect of release time was significant for the scales of pleasantness, background noise, and loudness. Clarity and overall impression are also affected under certain conditions (release time $\times$compression ratio interaction, release time$\times$noise type interaction). The effect of release time was most obvious at the highest compression ratio tested (3:1) and under certain noise conditions where longer release times result in higher ratings of clarity, pleasantness, and overall impression. Background noise ratings and loudness ratings decrease with longer release times.

The pattern of results obtained with the rating procedure in experiment I is similar to that obtained using paired-comparison judgments of sound quality (Neuman *et al.*, 1994). Both methods of measurement reveal a degradation in perceived speech quality with increasing compression ratio. However, the results of the rating procedure yield specific information about various aspects of sound quality—clarity, pleasantness, and overall impression decrease, and background noise increases with increasing compression ratio. It is noteworthy that in both the paired-comparison and in the present study, there were significant degradations in quality for compression ratios greater than 3:1.

The results in experiment II are in general agreement with the results of our previous experiment using paired-comparison judgments of sound quality (Neuman *et al.*, 1995b). In that experiment, we failed to find a significant main effect for release time on judgments of sound quality (the same compression ratios and release times were used). In this present experiment, we found that ratings of overall impression are not affected by release time. However, we did find that increasing release time causes positive changes in the ratings of pleasantness and noisiness. Increasing the release time also causes a decrease in the rating of loudness.

The finding of a decrease in loudness with increasing release time is consistent with physical measures of the signal. The variations in speech level due to differences in release time within a given compression ratio were not large. The largest variation was at the 3:1 compression ratio ($\pm$2 dB from the 200-ms condition). Nevertheless, these differences in level did result in differences in perceived loudness.

The ridge regression analysis revealed that the variation in loudness figured importantly as a predictor of overall impression in the cafeteria noise condition.

In experiment II, several significant interactions were found. A significant interaction between release time and noise type was found on the background noise scale. The mean ratings of noisiness decreased significantly with each increase in release time and was most dramatic for the cafeteria noise (the highest noise level). When these same subjects made paired-comparison judgments of sound quality for these same test conditions (Neuman *et al.*, 1995b), a significant interaction was found between release time and noise type. Subject preference for the 200- and 1000-ms release time was significantly greater than for the 60-ms release time in the cafeteria noise condition (and the same trend was seen in the apartment noise condition). We suggested that changes in the background noise due to the use of short release times was the underlying cause of the preferences. The results of the present experiment appear to confirm this hypothesis.

The compression ratio×release time interaction demonstrates that release time affects ratings of clarity, pleasantness, noisiness, and overall impression at the highest compression ratio (3:1). The compression ratios of 1.5, 2, and 3:1 used in this experiment were selected based on our findings that these compression ratios did not significantly degrade sound quality in a compression amplifier with a 200-ms release time [Neuman *et al.* (1994) and experiment I reported here]. A significant main effect of compression ratio was demonstrated in experiment II, with significant decreases in ratings of clarity, pleasantness, and overall impression being found as compression ratio increased. Background noise ratings increased significantly with increasing compression ratio, i.e., as compression ratio increased there was a perception of higher background noise levels. This finding, in combination with the compression ratio×release time interaction, indicates that the selection of compression ratio and release time should not be made in isolation. While the effect of release time is minimal at compression ratios $\leqslant$2:1, release time has a significant effect for the 3:1 compression ratio. At the 3:1 compression ratio, the shortest release time (60 ms) resulted in lower ratings of clarity, pleasantness, and overall impression. Similar compression characteristics are often used in single-band compression hearing aids.

In both experiments, testing was done at a single input level and with a single compression threshold. The results are indicative of what happens to sound quality when the speech is fully compressed and the noise level/type is changing. A more complete picture would be obtained with signals at multiple input levels or with several compression thresholds. The present results demonstrate that in a single-band compression hearing aid, compression ratios $<$3:1 will preserve good sound quality. Experiment I reveals that in a single-band compression hearing aid with a low compression threshold and a release time of 200 ms, compression ratios greater than 3:1 will degrade the overall impression of quality. The specific dimensions of clarity and pleasantness will be degraded. The perception of background noise will increase. Significant degradations of sound quality were observed at the 3:1 compression ratio in experiment II.

Experiment II reveals that release time does not significantly affect sound quality in a single-band compression hearing aid with compression ratios $\leqslant$2:1. However, when the compression ratio was 3:1, the 60-ms release time (syllabic) did result in significantly poorer ratings on all quality scales than did the 200-ms or the 1000-ms release times. In this experiment, we did not include release times longer than 1000 ms because we felt that using longer times would be detrimental to the hearing of soft sounds immediately following louder sounds in a hearing aid with a single set of time constants. Even longer release times can be safely implemented in hearing aids with multiple attack/release times or adaptive attack/release times. Overall, our results would suggest that superior sound quality will be obtained if multiple or adaptive release times are used in single-band compression hearing aids with compression ratios $>$2:1.

The rating procedure used in these two experiments allows a quantification of the relationship among different conditions. When paired-comparison judgments of sound quality were obtained in our previous experiments, we were able to rank the hearing aids with regard to sound quality but we had no indication of the magnitude of differences among hearing aids. The rating procedure gives an indication of how much sound quality changes with the manipulation of compression parameters or competing noise. Furthermore, the descriptors attached to the endpoints and to various intermediate points of the scales may be helpful in interpreting the importance of changes in rating. For example, in experiment I the largest changes in rating as a function of compression ratio occurred on the scales of background noise and overall impression. On the background noise scale, the mean rating for the linear hearing aid was 2.4, indicating the background noise was in the ''rather quiet'' range. For the 10:1 compression ratio, the mean rating of 5.5 is indicative that the background noise is approximately midway between ''very quiet'' and ''very noisy.'' The mean rating for the linear condition for overall impression was 7.1, indicating a judgment of the quality being ''rather good.'' The rating for the 10:1 compression condition was 5.3, or just above midway between ''very good'' and ''very bad.''

The rating procedure used in these experiments was quite sensitive to differences among hearing aids. In particular, the differences among the hearing aids in experiment II were quite subtle. This type of procedure may be useful in selecting the compression characteristics for individual listeners. Further research would be required to determine which rating scales would be most useful for this purpose. The rating scales used were a subset of those previously developed for studying the manipulation of frequency response in sound systems. We found these scales to be significantly correlated, although they all made significant contributions to the prediction of overall impression. Other scales not included in this study might have provided additional information about the effects of varying compression parameters. The development of a set of descriptors specifically relevant to the use of compression may be useful for future research and for clinical use.

ACKNOWLEDGMENTS

This research was supported by PHS Grant No. 5P50DC00178 from the National Institute on Deafness and Other Communication Disorders. We would like to thank Karen Bromberg for assistance in collection of data of experiment II. Portions of this paper were presented in a poster session at the First Biennial Conference on Hearing Aid Research and Development, September, 1995, National Institutes of Health, Bethesda, MD.

APPENDIX: INSTRUCTIONS GIVEN SUBJECTS FOR THE RATING TASK

You will be listening to speech along with background noise through a number of different hearing aids. You will be asked to judge five aspects of sound quality for each hearing aid using a rating scale. These qualities are

CLARITY
PLEASANTNESS
BACKGROUND NOISE
LOUDNESS
OVERALL IMPRESSION

Please listen carefully to each hearing aid and then rate it on each of the scales you see on the response form. There will be a separate scale for each of the five hearing aid qualities that you rate.

Each scale goes from 0 (minimum) to 10 (maximum). You may mark any place on the scale (including in between numbers).

Please be as accurate as you think necessary, and do your rating on each scale without looking at the other scales.

ANSI (**1987**). ''Specification of hearing aid characteristics,'' ANSI S3.22-1987 (American National Standards Institute, New York).

Bakke, M. H., Neuman, A. C., and Levitt, H. (**1991**). ''Loudness matching for compressed speech signals,'' J. Acoust. Soc. Am. **89**, 1974(A).

Bentler, R. A., and Nelson, J. A. (**1997**). ''Assessing release-time options in a two-channel AGC hearing aid,'' Am. J. Audiol. **6**, 43–51.

Braida, L. D., Durlach, N. I., Lippmann, R. P., Hicks, B. L., Rabinowitz, W. M., and Reed, C. M. (**1979**). *Hearing Aids: A Review of Past Research on Linear Amplification, Amplitude Compression, and Frequency Lowering*, ASHA Monograph No. 19, Rockville, MD: American Speech-Language-Hearing Assocation.

Byrne, D., and Dillon H. (**1986**). ''The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid,'' Ear Hear. **7**, 257–265.

Cox, R. M. (**1986**). ''NBS-9A coupler-to-eardrum transformation: TDH-39 and TDH-49 earphones,'' J. Acoust. Soc. Am. **79**, 120–123.

CSS (**1991**). *Statistica*, Volume 1 (Statsoft, Inc., Oklahoma), p. 315.

Dillon, H. (**1996**). ''Compression? Yes, but for low or high frequencies, for low or high intensities, and with what response times?'' Ear. Hear. **17**, 287–307.

Dirks, D. D., and Kamm, C. (**1976**). ''Psychometric functions for loudness discomfort and most comfortable loudness levels,'' J. Speech Hear. Res. **19**, 613–627.

Gabrielsson, A. (**1979a**). ''Statistical treatment of data from listening tests on sound-reproducing systems,'' Report No. TA92, Technical Audiology, KTH, Stockholm.

Gabrielsson, A., and Sjogren, H. (**1979b**). ''Perceived sound quality of sound reproducing systems,'' J. Acoust. Soc. Am. **65**, 1019–1033.

Gabrielsson, A., and Sjogren, H. (**1979c**). ''Perceived quality of hearing aids,'' Scand. Audiol. **8**, 159–169.

Gabrielsson, A., Hagerman, B., Bech-Kristensen, T., and Lundberg, G. (**1990**). ''Perceived sound quality reproductions with different frequency responses and sound levels,'' J. Acoust. Soc. Am. **88**, 1359–1366.

Gabrielsson, A., Schenkman, B. N., and Hagerman, B. (**1988**). ''The effects of different frequency responses on sound quality judgments of speech intelligibility,'' J. Speech Hear. Res. **31**, 166–177.

Johannson, B., and Lindblad, A. (**1971**). ''The use of compression and frequency transposition in hearing aids,'' Scand. Audiol. Suppl. **1**, 68–71.

Kates, J. M. (**1993**). ''Optimal estimation of hearing-aid compression parameters,'' J. Acoust. Soc. Am. **94**, 1–12.

Levitt, H., and Neuman, A. C. (**1991**). ''Evaluation of orthogonal polynomial compression,'' J. Acoust. Soc. Am. **90**, 241–252.

Nabelek, I. V. (**1984**). ''Discriminability of the quality of amplitude compressed speech,'' J. Speech Hear. Res. **27**, 571–577.

Neuman, A. C., Bakke, M. H., Hellman, S., and Levitt, H. (**1994**). ''Effect of compression ratio in a slow-acting compression hearing aid: Paired-comparison judgments of quality,'' J. Acoust. Soc. Am. **96**, 1471–1478.

Neuman, A. C., Bakke, M. H., Hellman, S., and Levitt, H. (**1995a**). ''Preferred listening levels for linear and slow-acting compression hearing aids,'' Ear. Hear. **16**, 407–416.

Neuman, A. C., Bakke, M. H., Mackersie, C., Hellman, S., and Levitt, H. (**1995b**). ''Effect of release time in compression hearing aids: Paired-comparison judgments of quality,'' J. Acoust. Soc. Am. **98**, 3182–3187.

Preminger, J., and Van Tasell, D. J. (**1995**). ''Quantifying the relation between speech quality and speech intelligibility,'' J. Speech Hear. Res. **38**, 714–725.

Van Tasell, D. J. (**1993**). ''Hearing loss, speech and hearing aids,'' J. Speech Hear. Res. **36**, 228–244.

Walker, G., and Dillon, H. (**1982**). *Compression in Hearing Aids: An Analysis, a Review and Some Recommendations*, NAL Report No. 90, Canberra, Australia: Australian Government Publishing Service.

Lamb waves as thickness vibrations superimposed on a membrane carrier wave

J. D. Achenbach[a)]
Center for Quality Engineering and Failure Prevention, Northwestern University, Evanston, Illinois 60208

(Received 2 September 1997; accepted for publication 5 February 1998)

Classical Lamb waves in an homogeneous, isotropic linearly elastic plate are reconsidered. The displacement components are obtained in terms of thickness motions superimposed on a membrane carrier wave which defines the propagation along the plate. The carrier wave can be any solution of the reduced wave equation for a membrane. The analysis of the thickness motions results in the usual Rayleigh–Lamb frequency equation. A number of special cases for the carrier wave are considered. © *1998 Acoustical Society of America.* [S0001-4966(98)04705-5]

PACS numbers: 43.20.Bi [ANN]

INTRODUCTION

The usual way to construct solutions to the elastodynamic equations of motion for homogeneous, isotropic, linearly elastic solids is to express the components of the displacement vector as the sum of the gradient of a scalar potential and the curl of a vector potential, where the potentials must be solutions of classical wave equations whose propagation velocities are the velocities of longitudinal and transverse waves, respectively.

This approach has generally also been used to investigate the propagation of waves in an elastic layer; see, for example, Refs. 1 and 2. For plane-strain conditions, time-harmonic waves in an elastic layer were first investigated by Lamb.[3] It is well known[1,2] that there are an infinite number of modes which may be split into antisymmetric and symmetric modes, corresponding to thickness shear and thickness stretch motions relative to the mid-plane of the layer.

As discussed in this paper, the particular nature of these layer motions suggests an alternative formulation in terms of a membrane motion of the mid-plane of the layer, which acts as a carrier wave of superimposed thickness motions. The corresponding formulation presented in this paper shows that the carrier wave satisfies a reduced membrane wave equation. The functions representing the superimposed thickness motions are solutions of ordinary differential equations. The formulation has the advantage that the carrier wave can be determined separately in a fairly general form. The analysis of the thickness motions results in the usual Rayleigh–Lamb frequency equations relating the circular frequency, ω, and a wave number-like quantity, k. The nature of the selected solution of the membrane equation affects the interpretation of k, but not the general form of the thickness motions.

Section I proposes general forms for the displacement components in terms of a single potential, and shows that these forms satisfy the displacement equations of motion. The connection with symmetric and antisymmetric modes of wave propagation is discussed in Sec. II. A number of examples for the solution of the membrane equation which governs the propagation of the Lamb wave along the plate, and its variation along wavefronts, are discussed in Sec. III.

a)Electronic mail: achenbach@nwu.edu

I. GOVERNING EQUATIONS

A layer of homogeneous, isotropic, linearly elastic material of thickness $2h$ is referred to a Cartesian coordinate system such that the x_1x_2 plane coincides with the mid-plane of the plate, as shown in Fig. 1. Greek indices exclusively refer to the x_1 and x_2 axes, while Latin indices refer to all three directions. We seek solutions for the displacement components in the forms

$$u_\alpha = \frac{1}{k} v(x_3) e^{i\omega t} \varphi_{,\alpha}(x_1, x_2), \tag{1}$$

$$u_3 = w(x_3) e^{i\omega t} \varphi(x_1, x_2). \tag{2}$$

Here k is a wave number-like quantity, i.e., its dimension is 1/length. The function $\varphi(x_1, x_2)$ is dimensionless. In what follows we will seek steady-state solutions, and we will omit $\exp(i\omega t)$.

Equations (1) and (2) yield for the divergence of the vector **u**

$$u_{k,k} = \frac{1}{k} v(x_3) \varphi_{,\beta\beta} + \frac{dw}{dx_3} \varphi. \tag{3}$$

Substitution of (1)–(3) into the equations of motion

$$(\lambda + \mu) u_{k,ki} + \mu u_{i,kk} = \rho \ddot{u}_i \tag{4}$$

yields for $i = \alpha = 1,2$:

$$(\lambda + \mu) \left[\frac{1}{k} v(x_3) \varphi_{,\beta\beta\alpha} + \frac{dw}{dx_3} \varphi_{,\alpha} \right] + \mu \left[\frac{1}{k} v(x_3) \varphi_{,\alpha\beta\beta} + \frac{1}{k} \frac{d^2 v}{dx_3^2} \varphi_{,\alpha} \right] = -\frac{\rho\omega^2}{k} v(x_3) \varphi_{,\alpha}, \tag{5}$$

and for $i = 3$:

$$(\lambda + \mu) \left[\frac{1}{k} \frac{dv}{dx_3} \varphi_{,\beta\beta} + \frac{d^2 w}{dx_3^2} \varphi \right] + \mu \left[w(x_3) \varphi_{,\beta\beta} + \frac{d^2 w}{dx_3^2} \varphi \right] = -\rho\omega^2 w(x_3) \varphi. \tag{6}$$

These equations can be rewritten as

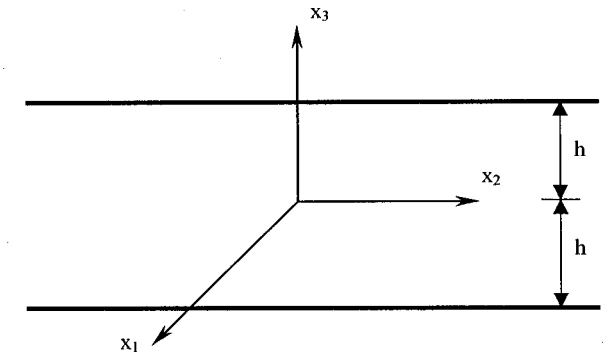


FIG. 1. Elastic plate.

$$\frac{\lambda+2\mu}{k}v(x_3)\varphi_{,\beta\beta\alpha}+\left[(\lambda+\mu)\frac{dw}{dx_3}+\frac{\mu}{k}\frac{d^2v}{dx_3^2}+\frac{\rho\omega^2}{k}v(x_3)\right]\varphi_{,\alpha}=0, \tag{7}$$

$$\left[\frac{\lambda+\mu}{k}\frac{dv}{dx_3}+\mu w(x_3)\right]\varphi_{,\beta\beta}+\left[(\lambda+2\mu)\frac{d^2w}{dx_3^2}+\rho\omega^2 w(x_3)\right]\varphi=0. \tag{8}$$

We now select equations for $v(x_3)$ and $w(x_3)$ such that both (7) and (8) are satisfied if $\varphi(x_1,x_2)$ is a solution of the reduced membrane wave equation

$$\varphi_{,\beta\beta}+k^2\varphi=0. \tag{9}$$

Clearly this can be achieved if

$$(\lambda+\mu)\frac{dw}{dx_3}+\frac{\mu}{k}\frac{d^2v}{dx_3^2}+\frac{\rho\omega^2}{k}v(x_3)=k^2\frac{\lambda+2\mu}{k}v(x_3), \tag{10}$$

$$(\lambda+2\mu)\frac{d^2w}{dx_3^2}+\rho\omega^2 w(x_3)=k^2\left[\frac{\lambda+\mu}{k}\frac{dv}{dx_3}+\mu w(x_3)\right]. \tag{11}$$

Equations (10) and (11) are a set of coupled ordinary differential equations for $v(x_3)$ and $w(x_3)$.

For the layer shown in Fig. 1 we seek solutions of (10) and (11) of the general forms

$$v(x_3)=Ve^{ipx_3} \tag{12a}$$

and

$$w(x_3)=We^{ipx_3}, \tag{12b}$$

where V and W are constants. Substitution of (12a) and (12b) into Eqs. (10) and (11) yields a set of homogeneous algebraic equations for the constants V and W. The condition that the determinant of the coefficients must vanish yields the following equation:

$$p^4+\left(2k^2-\frac{\omega^2}{c_L^2}-\frac{\omega^2}{c_T^2}\right)p^2+\left(\frac{\omega^2}{c_L^2}-k^2\right)\left(\frac{\omega^2}{c_T^2}-k^2\right)=0, \tag{13}$$

where c_L and c_T are the velocities of longitudinal and transverse waves, respectively,

$$c_L^2=\frac{\lambda+2\mu}{\rho} \tag{14a}$$

and

$$c_T^2=\frac{\mu}{\rho}. \tag{14b}$$

The two solutions of Eq. (13) are

$$p^2=\frac{\omega^2}{c_L^2}-k^2 \tag{15a}$$

and

$$q^2=\frac{\omega^2}{c_T^2}-k^2. \tag{15b}$$

The corresponding solutions of (10) and (11) may then be written as

$$v(x_3)=Ae^{ipx_3}+Be^{iqx_3}, \tag{16}$$

$$w(x_3)=\frac{ip}{k}Ae^{ipx_3}+\frac{k}{iq}Be^{iqx_3}. \tag{17}$$

In summary, it has been shown that Eqs. (1) and (2) are solutions of the elastodynamic equations (4), provided that $\varphi(x_1,x_2)$ is a solution of Eq. (9) and $v(x_3)$ and $w(x_3)$ are given by Eqs. (16) and (17).

In addition to a homogeneous layer, the formulation presented here also applies to homogeneous and layered full spaces, homogeneous and layered half-spaces and layered plates where the x_3 direction should be normal to the layering. It may also be applicable to certain simple cases of anisotropy. It will be of interest to investigate the applicability to cases of continuous inhomogeneity in the x_3 direction.

II. SYMMETRIC AND ANTISYMMETRIC MODES

Let us first consider displacements represented by the real parts of Eqs. (16) and (17), i.e.,

$$v(x_3)=A\cos(px_3)+B\cos(qx_3), \tag{18}$$

$$w(x_3)=-\frac{p}{k}A\sin(px_3)+\frac{k}{q}B\sin(qx_3). \tag{19}$$

These expressions represent motions that are symmetric relative to the mid-plane of the layer. They define the so called thickness-stretch modes of the layer. The displacement components given by Eqs. (1) and (2) with $v(x_3)$ and $w(x_3)$ defined by Eqs. (18) and (19) do not yet satisfy the conditions on the faces of the layer. These conditions are

$$\text{at } x_3=\pm h: \quad \tau_{3\beta}=0, \ \tau_{33}=0. \tag{20}$$

To obtain expressions for $\tau_{3\beta}$ and τ_{33} we use Hooke's law:

$$\tau_{ij}=\rho(c_L^2-2c_T^2)u_{k,k}\delta_{ij}+\rho c_T^2(u_{i,j}+u_{j,i}). \tag{21}$$

By using Eqs. (1) and (2) the relevant stress components may be expressed as

$$\tau_{3\beta}=\rho c_T^2\left[w(x_3)+\frac{1}{k}\frac{dv}{dx_3}\right]\varphi_{,\beta}, \tag{22}$$

$$\tau_{33}=\rho(c_L^2-2c_T^2)\left[\frac{1}{k}v(x_3)\varphi,_{\beta\beta}+\frac{dw}{dx_3}\varphi\right]+2\rho c_T^2\frac{dw}{dx_3}\varphi. \tag{23}$$

To satisfy the boundary conditions, Eq. (20), at $x_3=\pm h$, we must then have

$$w(h)+\frac{1}{k}\frac{dv}{dx_3}\bigg|_{x_3=h}=0, \tag{24}$$

$$-\rho k(c_L^2-2c_T^2)v(h)+\rho c_L^2\frac{dw}{dx_3}\bigg|_{x_3=h}=0, \tag{25}$$

where $\varphi,_{\beta\beta}$ has been eliminated by the use of Eq. (9). Substitution of $v(x_3)$ and $w(x_3)$ from Eqs. (18) and (19) yields a system of homogeneous algebraic equations for the constants A and B. The condition that the determinant of the coefficients must vanish yields after some manipulation

$$\frac{\tan(qh)}{\tan(ph)}=-\frac{4pqk^2}{(q^2-k^2)^2}. \tag{26}$$

Equation (26) is recognized as the well-known Rayleigh–Lamb frequency equation for symmetric layer modes, see, e.g., Ref. 2.

In a similar manner we can consider the imaginary parts of Eqs. (16) and (17), i.e.,

$$v(x_3)=A\ \sin(px_3)+B\ \sin(qx_3), \tag{27}$$

$$w(x_3)=\frac{p}{k}A\ \cos(px_3)-\frac{k}{q}B\ \cos(qx_3). \tag{28}$$

These expressions represent motions that are antisymmetric relative to the mid-plane of the layer. They are the so-called thickness-shear modes. Substitution of (27) and (28) into (24) and (25) yields the Rayleigh–Lamb frequency equation for anti-symmetric layer modes:

$$\frac{\tan(qh)}{\tan(ph)}=-\frac{(q^2-k^2)^2}{4pqk^2}. \tag{29}$$

The Rayleigh–Lamb frequency equations (26) and (29) relate the wave number, k, to the frequency, ω. To obtain detailed and precise numerical information the roots of (26) and (29) are obtained numerically by choosing a real-valued frequency ω and calculating the infinite number of corresponding values of k. Each value of k corresponds to a mode of wave propagation in the layer. The calculated values of k may be real, imaginary or complex. Real-valued wave numbers correspond to propagating wave modes. Imaginary and complex wave numbers correspond to standing waves with decaying amplitudes in the x_1x_2 plane. For small values of the frequency, as small as $\omega=0^+$, Eqs. (26) and (29) each have one real-valued solution for k. All the other solutions for k then are either imaginary or complex. The two real-valued k's correspond to the lowest symmetric and antisymmetric modes.

The interesting result of this section is that the thickness motions governed by Eqs. (26) and (29) which are well-known for the plane-strain case, apply to a general class of wave motions in the layer as long as $\varphi(x_1,x_2)$ is a solution of the reduced membrane wave equation (9).

III. EXAMPLES

The approach presented in this paper makes it possible to determine the carrier wave in a very simple manner as the solution of a reduced membrane wave equation. The corresponding thickness motion are independent of the choice of the carrier wave. In this section we present examples.

First we consider plane-strain waves in the layer. This case corresponds to a solution of the membrane equation of the form

$$\varphi(x_2)=e^{-ikx_2}. \tag{30}$$

Here k is the wave number, where $k=\omega/c$, c being the phase velocity. Equation (1) shows that the out-of-plane displacement, $u_1(x_2,x_3t)$, vanishes. The remainder of the analysis presented in Secs. I and II then reduces to the conventional Lamb-wave result as, for example, presented in Ref. 2.

As a generalization of Eq. (30) let us consider a wave standing in the x_1 direction but propagating in the x_2 direction, i.e.,

$$\varphi(x_1,x_2)=\sin(\xi x_1)e^{-i\eta x_2} \tag{31a}$$

$$=\frac{1}{2i}[e^{i(\xi x_1-\eta x_2)}-e^{i(-\xi x_1-\eta x_2)}]. \tag{31b}$$

Now we have

$$k^2=\xi^2+\eta^2. \tag{32}$$

For a specific frequency, the parameter k is still obtained from either (26) or (29), but now ξ and η must satisfy (32). Thus if a wave of a specific frequency and wave number η in the propagation direction is being considered, the variation with x_1 is fixed, and follows by obtaining ξ from (32).

Next let us consider axisymmetric motion. Equation (9) then becomes

$$\frac{\partial^2\varphi}{\partial r^2}+\frac{1}{r}\frac{\partial\varphi}{\partial r}+k^2\varphi=0. \tag{33}$$

The solution of (33) to form an outgoing wave compatible with $\exp(i\omega t)$ is

$$\varphi(r)=H_0^{(2)}(kr),\ \text{ where }\ r=(x_1^2+x_2^2)^{1/2}. \tag{34}$$

Equation (1) yields

$$u_1=v(x_3)e^{i\omega t}H_1^{(2)}(kr)\frac{x_1}{r}, \tag{35}$$

$$u_2=v(x_3)e^{i\omega t}H_1^{(2)}(kr)\frac{x_2}{r}. \tag{36}$$

Clearly the resulting displacement in the x_1x_2 plane is in the r direction

$$u_r=v(x_3)e^{i\omega t}H_1^{(2)}(kr). \tag{37}$$

Again the thickness vibrations follow from the analysis of Sec. II.

When the motion is not axially symmetric Eq. (9) becomes in polar coordinates

$$\frac{\partial^2\varphi}{\partial r^2}+\frac{1}{r}\frac{\partial\varphi}{\partial r}+\frac{1}{r^2}\frac{\partial^2\varphi}{\partial\theta^2}+k^2\varphi=0. \tag{38}$$

A solution of Eq. (38) is

$$\varphi(r,\theta)=H_n^{(2)}(kr)\sin(n\theta). \tag{39}$$

It may be checked that for this case we may define radial and circumferential displacements where

$$u_r=\frac{1}{k}v(x_3)e^{i\omega t}\frac{\partial\varphi}{\partial r}, \tag{40}$$

$$u_\theta=\frac{1}{k}v(x_3)e^{i\omega t}\frac{1}{r}\frac{\partial\varphi}{\partial\theta}, \tag{41}$$

while u_3 is given by

$$u_3=w(x_3)e^{i\omega t}\varphi(r,\theta). \tag{42}$$

The approach presented here also lends itself to investigate by an asymptotic approach the propagation of more general waves. As an example, we consider a wave which propagates out radially over a sector spanned by an angle $2\hat{\theta}$, such that the distribution along the wavefronts is represented by a Gaussian. For large value of kr we seek a solution of Eq. (38) of the form

$$\varphi(r,\theta)=A\frac{e^{-i(kr-\pi/4)}}{(kr)^{1/2}}e^{-\theta^2/2\hat{\theta}^2}\left[1+\frac{b_0+b_2\theta^2}{kr}+\frac{c_0+c_2\theta^2}{(kr)^2}+\cdots\right]. \tag{43}$$

This expression will satisfy Eq. (38) with greater accuracy, the more terms are included in the brackets. Substitution of Eq. (43) into (38) yields a sequence of terms of orders $r^{-(n+1)/2}$, where $n=0,1\ldots$. Here we will keep terms up to order $r^{-5/2}$. It can be verified that the terms of orders $r^{-1/2}$ and $r^{-3/2}$ vanish identically. The condition that the terms of order $r^{-5/2}$ should be equal to zero leads to the equation

$$\left(2ib_0+\frac{1}{4}-\frac{1}{\hat{\theta}^2}\right)+\theta^2\left(2ib_2+\frac{1}{\hat{\theta}^4}\right)=0. \tag{44}$$

The requirement that the two terms should vanish separately yields expressions for b_0 and b_2. It follows that $\varphi(r,\theta)$ may be written as

$$\varphi(r,\theta)=A\frac{e^{-i(kr-\pi/4)}}{(kr)^{1/2}}e^{-\theta^2/2\hat{\theta}^2}\left[1+\frac{i}{kr}\left(\frac{1}{8}-\frac{1}{2\hat{\theta}^2}+\frac{\theta^2}{2\hat{\theta}^4}\right)+O(kr)^{-2}\right]. \tag{45}$$

The actual displacements for this case then follow from Eqs. (40)–(42), with $\varphi(r,\theta)$ defined by Eq. (45).

It is of interest to note that for $\hat{\theta}\to\infty$, i.e., for a uniform distribution along the wavefront, Eq. (45) reduces to the first two terms of the asymptotic expansion of $H_0^{(2)}(kr)$, in accordance with Eq. (34).

IV. CONCLUSIONS

This paper has shown that time-harmonic wave motion in an elastic plate can be analyzed on the basis of a carrier wave and superimposed thickness motions. The carrier wave, which can be any solution of a reduced membrane wave equation, describes the propagation along the plate. The superimposed thickness motions have the same general form for any choice of the carrier wave. The relation between the frequency, ω, and a wave number-like quantity, k, is governed by the classical Rayleigh–Lamb frequency equations, but the interpretation of k depends on the choice of the carrier wave. The approach is attractive because it simplifies the consideration of a broad class of Lamb waves.

ACKNOWLEDGMENTS

This work was carried out in the course of research sponsored by the Office of Naval Research under Contract No. N00014-89-J-1362.

[1] R. D. Mindlin, ''Waves and vibrations in isotropic elastic plates,'' in *Structural Mechanics*, edited by J. N. Goodier and N. J. Hoff (Pergamon, New York, 1960), pp. 199–232.
[2] J. D. Achenbach, *Wave Propagation in Elastic Solids* (North Holland/American Elsevier, New York, 1973).
[3] H. Lamb, ''Waves in an elastic plate,'' Proc. R. Soc. London, Ser. A **93**, 114–128 (1917).

Aperture synthesis of time-limited X waves and analysis of their propagation characteristics

Argyrios A. Chatzipetros
Paging Group, Motorola, Boyton Beach, Florida 33426

Amr M. Shaarawi
Department of Engineering Physics and Mathematics, Faculty of Engineering Cairo University, Giza, Egypt

Ioannis M. Besieris and Mohamed A. Abdel-Rahman
The Bradley Department of Electrical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061

(Received 20 June 1997; revised 8 January 1998; accepted 16 January 1998)

The feasibility of exciting a localized X-wave pulse from a finite aperture is addressed. Also, the possibility of using a finite-time excitation of a dynamic aperture to generate a finite-energy approximation to an X-wave pulse is explored. The analysis is carried out by using a Gaussian time window to time limit the infinite X-wave initial excitation. Huygens' construction is used to calculate the amplitude of the radiated wave field away from the finite-time source. The decay rate of the peak of the X wave is compared to that of a quasi-monochromatic signal. It is shown that the finite-time X-wave propagates to much farther distances without significant decay. Furthermore, the decay pattern of the radiated X-wave pulse is derived for a source consisting of an array of concentric annular sections. The decay behavior of the radiated pulse is similar to that of an X-wave launched from a finite-time aperture. This confirms the fact that time windowing the infinite energy X-wave excitation is a viable scheme for constructing finite apertures. A discussion of the diffraction limit of the X-wave pulse is also provided. © *1998 Acoustical Society of America.* [S0001-4966(98)00605-5]

PACS numbers: 43.20.Bi, 43.20.Px [JEG]

INTRODUCTION

In this paper, we address the possibility of exciting localized X-wave pulses from finite apertures. A large class of localized wave (LW) solutions to the scalar wave and Maxwell equations have been studied extensively.[1–12] A number of these pulsed wave solutions are finite-energy variants of the focus wave mode (FWM).[13,14] They are usually derived either as superpositions of infinite-energy FWM pulses,[8–11,14] or by time limiting various FWM-like excitations to construct finite-energy aperture fields.[2–7] Alternatively, such LW pulses can be synthesized as superpositions over Bessel beams with appropriately chosen spectra.[15] Our approach is based mainly on the proven possibility of exciting Bessel beams from infinite apertures.[1–4] It has been shown that a Bessel beam defined initially on an infinite plane (e.g., $z=0$) propagates under the effect of Huygens' operator in the positive z direction.[1–4] Any possible contributions from acausal (negative z-directed) components sum up to zero.

In contradistinction to other solutions excited from infinite apertures, such as plane waves and Bessel beams, the excitation of the FWM pulse does not require infinite power. This is the case because as the generating aperture becomes infinitely large, the intensity of the field exciting it decreases to zero as $(ct)^{-2}$, while the area of the aperture increases as $(ct)^2$. These two effects balance each other and the power of the excitation wave field remains constant. The energy of the generated pulse becomes infinite because the aperture is illuminated for an infinitely long time. Time limiting the excitation wave field results in the generation of finite-energy LW pulses.

Other pulsed solutions exhibiting distinct similarities to FWM-like pulses are the X-wave pulses introduced by Lu and Greenleaf.[16–23] Similarly to the FWM-like pulses, the X-waves have extended ranges of localization and acquire large focus depths in the near field. These properties make these pulsed wave fields ideal for applications involving detection of objects buried at unknown depths, as well as in high-resolution imaging. The use of X-wave pulses in ultrasonic medical imaging has been investigated by Lu and Greenleaf.[16–23] They have shown that X-waves can be generated using ultrasonic transducers consisting of annular arrays of concentric piezoelectric rings excited by independent waveforms applied at different times (Refs. 16, 18; cf. also Ref. 24). Furthermore, Lu and Greenleaf have studied experimentally the improvement in pulse-echo imaging using X-waves.[21] Applying X-waves to tissue-equivalent phantoms and standard test objects, they have shown that resolution of pulse-echo images obtained using X-waves are superior to that obtained with focused Gaussian beams.[21,22] In addition to their large focus depths, X-waves and FWM-like pulses are both wideband wave fields. This property renders them suitable for applications involving parameter characterization using backscattered signals from objects of varying sizes. Along this direction, Donnelly and Power[25] have shown that an acoustical LW pulse backscattered from compressible spheres immersed in water can accurately provide information about the materials as well as the sizes of the spheres.

Similar results can be obtained using X-wave acoustical pulses. As a consequence, these pulsed wave fields can have applications in parameter characterization and nondestructive testing of materials.

The excitation wave field of an X-wave pulse has the form of an annulus whose radius varies with time. The speed of expansion v_{ap} of the aperture field is related to the speed of the pulse center v_p and can be greater than the wave speed c of the medium. The temporal and spatial frequency content of an X-wave pulse will be calculated in order to improve our understanding of the decay pattern of the radiated wavefield. Similarly to the FWM-like solutions, we shall explore the possibility of using a finite-time excitation of an X-wave by imposing a Gaussian time window on the infinite time initial X-wave field. The amplitude of the radiated pulse will be calculated in the near-to-far field range and its decay pattern will be compared to that of a quasi-monochromatic pulse characterized by a specific carrier frequency. It will be shown that the finite-time version of the X-wave propagates without significant decay much further than quasi-monochromatic signals.

I. THE X-WAVE APERTURE

The X-wave solution to the 3-D scalar wave equation may be expressed as

$$\Psi_{XW}(\rho,z,t)=\frac{z_0}{[(z_0\rho)^2+(a_1z_0+i(z-ct\ \coth\beta))^2]^{1/2}}, \tag{1}$$

where a_1 and β are free real parameters and $z_0=1/\sinh\beta$. Note that $c\ \coth\beta$ is the propagation speed of the pulse. To ensure that the field will propagate in the forward direction, we restrict β to be positive. The solution given in Eq. (1) is identical to that derived by Lu and Greenleaf[16,17] if the following substitutions are made: $a_1z_0\rightarrow a_0/\cos\eta$, $z_0\rho\rightarrow\rho\tan\eta$ and $c\ \coth\beta\rightarrow c/\cos\eta$. In the notation used by Lu and Greenleaf,[16] α_0 is a real parameter and η is the axicon angle.

The infinite energy solution given in Eq. (1) was used in Ref. 16 to excite an aperture of a finite size situated at $z=0$ and then to compute the radiated field. In this paper, we construct a finite aperture in a different manner. We use the fact that the X-wave excitation at $z=0$ acquires the form of an annulus whose radial position is time dependent. By time limiting the excitation wave field, we restrict the radius of the time-dependent annulus to be smaller than a finite maximum value. The advantage of such an approach is that the spatio-temporal spectral structure of the radiated pulse becomes very transparent. One can then have a better understanding of the depletion of the spectral content of such a pulse, as well as its behavior as it decays away from its source.[2,4–7] Subsequently, this information can be used to design other localized variants of X waves. First, we shall present some of the aspects of the X-wave excitation given in Eq. (1) without time-limiting it. In Sec. II, the finite-time X-wave will be studied in detail.

Consider the X-wave excitation of a planar aperture positioned at $z=0$, viz.,

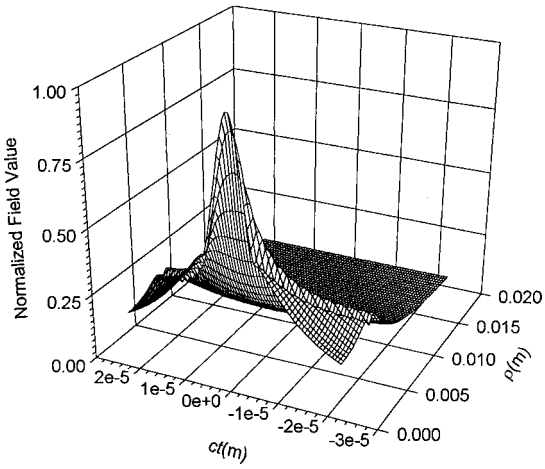


FIG. 1. Surface plot of the real part of the X-wave pulse on the aperture plane ($z=0$) for the parameters $a_1=0.004\ \mathrm{m}^{-1}$, $\beta=7.4$, and $c=1500$ m/s.

$$\Psi_{XW}(\rho,z=0,t)=\frac{1}{[\rho^2+(a_1-iy_0t)^2]^{1/2}}, \tag{2}$$

where $y_0=c\ \cosh\beta$. The surface plot of the real part of the aperture field given in (2) is shown in Fig. 1 for $a_1=0.004$ m and $\beta=7.4$. The wave velocity of the medium is chosen to be $c=1500$ m/s. The intensity of the infinite energy excitation of an X-wave aperture is given as

$$I(\rho,0,t)=\Psi(\rho,0,t)\Psi^*(\rho,0,t)=\frac{1}{[(\rho^2-(y_0t)^2)^2+a_1^4+2a_1^2\rho^2+2a_1^2(y_0t)^2]^{1/2}}. \tag{3}$$

After rearrangement, the denominator in (3) becomes

$$D=[\rho^4+(y_0t)^4-2\rho^2(y_0t)^2+a_1^4+2a_1^2\rho^2+2a_1^2(y_0t)^2]^{1/2}, \tag{4}$$

and for $\rho\gg a_1$

$$D\cong[\rho^4+(y_0t)^4-2\rho^2(y_0t)^2]^{1/2}=[\rho^2-(y_0t)^2]. \tag{5}$$

Expressions (3) and (5) show how the intensity of the illumination wave field varies with time at the various radial positions ρ on the aperture plane. At the centroid of the excitation field ($\rho=0$ and $t=0$) the intensity I is finite and is equal to $1/a_1^2$. For large values of ρ and t, the highest intensity occurs at $\rho=y_0t$. Thus, the illumination of the X-wave aperture acquires the form of an annulus whose radius changes linearly with time. The intensity at $\rho=y_0t$ varies as $1/a_1\sqrt{a_1^2+4\rho^2}$, as it can be deduced from Eq. (3). Since the area of an annulus is proportional to ρ, it follows that the power of the initial X-wave on the aperture (=intensity $\times$area) remains constant as $ct\rightarrow\infty$. The energy of the excitation wave field is infinite because the aperture is illuminated for an infinite time duration.

The time dependence of the radius of the X-wave aperture is found from the real part of Eq. (2). The amplitude of the real part is maximum on the aperture at $\rho=0$ and $t=0$. The highly focused central portion of the aperture has a ra-

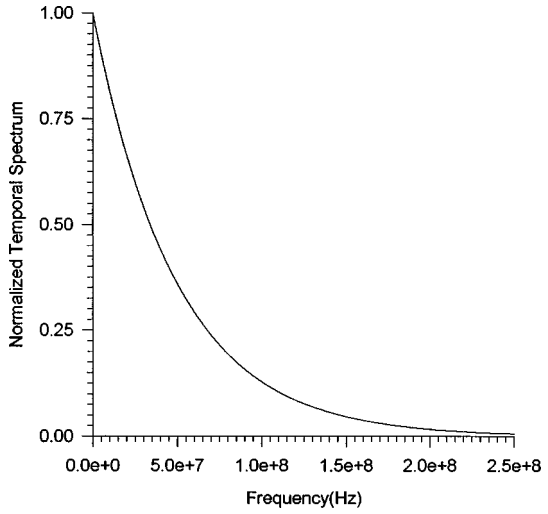


FIG. 2. Temporal frequency content, $\Phi_t(\rho,\omega)$, of the X-wave pulse at $\rho=0$ for the parameters $a_1=0.004\ \mathrm{m}^{-1}$, $\beta=7.4$, and $c=1500$ m/s.

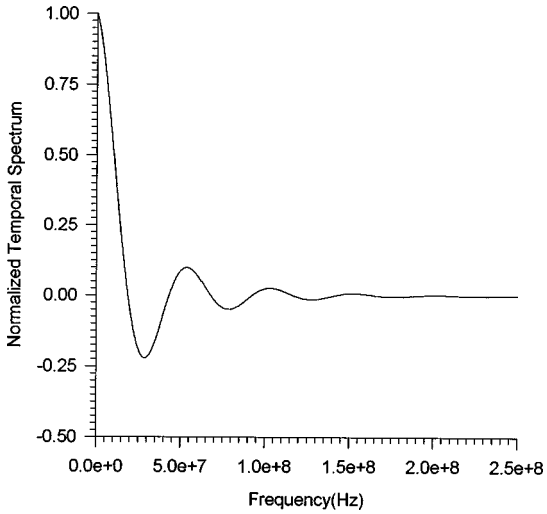


FIG. 3. Temporal frequency content, $\Phi_t(\rho,\omega)$, of the X-wave pulse at $\rho=2.5$ cm for the parameters $a_1=0.004\ \mathrm{m}^{-1}$, $\beta=7.4$, and $c=1500$ m/s.

dius $R_{\min}=R(t=0)=a_1$. As y_0t becomes larger than a_1, the annulus forms at $R(t)=y_0t$ with a width of order $a_1 \ll R(t)$. As a consequence, a good estimate of the radius of the excitation wave field as a function of time is given by

$$R(t)=y_0t+R(0). \tag{6}$$

For $y_0t \gg a_1$, the speed of the expansion of the aperture is given by $v_{\mathrm{ap}}=y_0=c\ \cosh\beta$. The speed of the centroid of the radiated pulse and that of the expansion of the aperture are always greater than c. However, the faster the centroid of the pulse travels, the slower the expansion of the aperture becomes.

II. FREQUENCY SPECTRA OF THE X-WAVE

In order to gain a better understanding of the spectral content of the X-wave pulse, we need to define the temporal and spatial frequency contents of the excitation wave field given in Eq. (2). Both spectra have to be studied because the X-wave illumination is a pulsed wave field exciting different parts of the generating aperture at different times. The spectrum of the initial excitation given in Eq. (2) is deduced from the Fourier inversion

$$\Phi(\chi,\omega)=\int_{-\infty}^{+\infty}dt\int_0^{+\infty}d\rho\ \rho J_0(\chi\rho) \times e^{-i\omega t}\frac{1}{[\rho^2+(a_1-iy_0t)^2]^{1/2}}. \tag{7}$$

The integration over ρ is carried out using (6.554.1) of Ref. 26; as a consequence, we obtain

$$\Phi(\chi,\omega)=\frac{1}{\chi}\int_{-\infty}^{+\infty}dt\ e^{-i\omega t}e^{-\chi(a_1-iy_0t)}. \tag{8}$$

The integration over t yields the following Fourier spectrum:

$$\Phi(\chi,\omega)=\frac{2\pi}{\chi}\,e^{-\chi a_1}\delta(\omega-\chi y_0). \tag{9}$$

One should note that the argument of the Dirac δ function couples the temporal frequency ω to the spatial wave number χ. This is a characteristic attribute of most LW pulses. A similar δ function appears in the Fourier spectrum of the infinite-time FWM excitation of a flat aperture studied in Refs. 1 and 2. For the FWM spectrum we have $\omega \propto \chi^2$, while the X-wave exhibits an $\omega \propto \chi$ coupling. The appearance of a Dirac δ function in the spectrum signifies that the initial wave field illuminates the aperture for an infinite time.

To identify the temporal frequency content of the excitation wave field at various positions on the aperture we define the temporal spectrum as follows:

$$\Phi_t(\rho,\omega)=\int_0^{\infty}d\chi\ \chi J_0(\chi\rho)\Phi(\chi,\omega) =2\pi\int_0^{\infty}d\chi\ \chi J_0(\chi\rho)\,\frac{e^{-\chi a_1}}{\chi}\,\delta(\omega-\chi y_0). \tag{10}$$

The above integration gives

$$\Phi_t(\rho,\omega)=\frac{2\pi}{y_0}\,J_0(\omega\rho/y_0)e^{-\omega a_1/y_0}\mathbf{H_s}(\omega), \tag{11}$$

where $\mathbf{H_s}(\omega)$ is the Heaviside unit-step function. The highest contribution to the temporal frequency content comes from values close to $\omega=0$ rad/s. The maximum frequency is determined from the exponential and is given explicitly as

$$\omega_{\max}=(4y_0/a_1)=(4c\ \cosh\beta/a_1). \tag{12}$$

For the parameters $a_1=0.004$ m, $\beta=7.4$, and $c=1500$ m/s, the maximum frequency becomes $\omega_{\max}=1.227\times10^9$ rad/s or $f_{\max}=195.3$ MHz. This implies that the minimum wavelength in this case is $\lambda_{\min}=1.54$ m. The temporal spectrum given in Eq. (11) is plotted in Fig. 2 for $\rho=0$. The same spectrum is shown in Fig. 3 at $\rho=0.025$ m. The oscillations appearing in the latter case indicate that an integration over ω should yield a smaller amplitude. This is consistent with the field given in Eq. (2), where the amplitude is maximum at $\rho=0$ and decreases as ρ becomes larger.

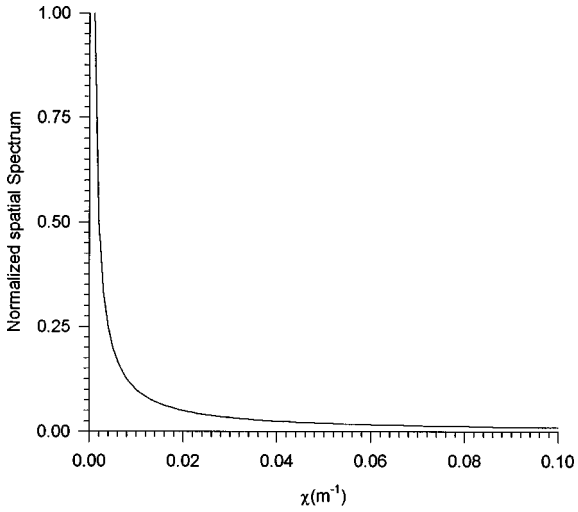


FIG. 4. Spatial frequency content, $\Phi_s(\chi,t)$, of the X-wave pulse at $t=0$ for the parameters $a_1=0.004$ m^{-1}, $\beta=7.4$, and $c=1500$ m/s.

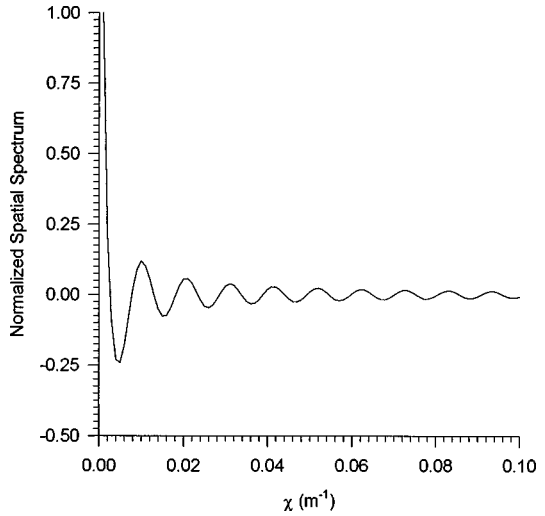


FIG. 5. Spatial frequency content, $\Phi_s(\chi,t)$, of the X-wave pulse at $t=5\times10^{-4}$ s for the parameters $a_1=0.004$ m^{-1}, $\beta=7.4$, and $c=1500$ m/s.

The spatial spectrum, on the other hand, describes the spatial distribution of the source elements on the aperture plane for various times. We define the spatial spectrum as

$$\Phi_s(\chi,t)=\frac{1}{2\pi}\int_0^\infty d\omega\; e^{i\omega t}\,\frac{2\pi}{\chi}\,e^{-\chi a_1}\delta(\omega-\chi y_0) \qquad (13)$$

or

$$\Phi_s(\chi,t)=\frac{1}{\chi}\,e^{-\chi a_1}e^{i\chi y_0 t}. \qquad (14)$$

The spectrum given in Eq. (14) is shown in Figs. 4 and 5 for $t=0$ and $t=5\times10^{-4}$ s, respectively. The oscillations appearing in the latter case indicate that the amplitude of the X wave decreases after the highly focused central portion of the pulse has been launched from the aperture.

III. FINITE TIME EXCITATION

We have shown in Sec. I that the power required to excite an X-wave aperture is always finite and that the excitation energy is infinite simply because the source is excited for an infinite duration of time. As a consequence, the excitation wave field, having the shape of an annulus, is allowed to expand to infinite dimensions. The illumination given in Eq. (2) thus requires an infinite aperture to be realized. Because of the time dependence of the radius of the significant annular contributions to the excitation wave field, we can construct a finite-energy source using one of the following two methods. First we can apply the X-wave excitation given in Eq. (2) to an aperture having a finite radius R and calculate the amplitude of the pulse radiated into the $z>0$ half-space. Lu and Greenleaf[16] used this method to construct a finite X-wave aperture. One should note that at some instant in time we will have $R(t)>R$; subsequently, the excitation wave field inside the finite radius of the aperture becomes very small and can be neglected for all practical purposes. One can then presume that the illumination wave field is turned off at that time instant. Alternatively, we can start by time limiting the infinite excitation given in Eq. (2). This allows the radius of the annular excitation to expand to a finite radius before the initial wave field is turned off. The same approach has been used extensively in defining finite FWM excitations.[2,3,5,6] The latter approach has the advantage of providing a deeper insight into the spectral structure of the radiated LW pulses. Within such a framework, we have been able to demonstrate that the spectral depletion of LW pulses is essentially different from that of the usual quasi-monochromatic pulses.[6] This spectral approach has also helped us in deriving a diffraction-limit formula capable of characterizing the decay behavior of ultra-wideband LW pulses.[7] Because of these advantages, we seek to use this approach to time limit the X-wave excitation (2), thus constructing a finite-energy X-wave aperture. The decay patterns of the radiated fields calculated using both approaches will be compared and the differences between the two methods will be discussed.

A convenient time window to use is the Gaussian $\exp(-t^2/4T^2)$. In this case, the Fourier spectrum of the initial excitation can be calculated as follows:

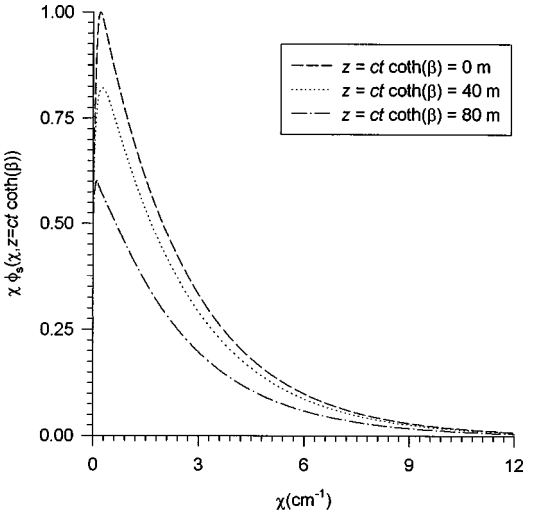


FIG. 6. Normalized spatial spectrum of the finite-time X wave at different distances $z=ct$ coth $\beta=0$, 40, and 80 m for $R_{\max}=0.15$ m.

$$\Phi(\chi,\omega)=\int_{-\infty}^{+\infty}dt\int_{0}^{+\infty}d\rho\ \rho J_0(\chi\rho)$$

$$\times e^{-i\omega t}\frac{e^{-t^2/4T^2}}{[\rho^2+(a_1-iy_0t)^2]^{1/2}}. \tag{15}$$

The integration over ρ is carried out using (6.554.1) of Ref. 26, yielding

$$\Phi(\chi,\omega)=\frac{e^{-\chi a_1}}{\chi}\int_{-\infty}^{+\infty}dt\ e^{-it(\omega-\chi y_0)}e^{-t^2/4T^2} \tag{16}$$

or

$$\Phi(\chi,\omega)=\frac{2\pi}{\chi}\,e^{-\chi a_1}\hat{\delta}(\omega-\chi y_0), \tag{17a}$$

where $\hat{\delta}(\omega-\chi y_0)$ is a Gaussian function given explicitly by

$$\hat{\delta}(\omega-\chi y_0)=\frac{T}{\sqrt{\pi}}\,e^{-T^2(\omega-\chi y_0)^2}. \tag{17b}$$

Note that the parameter T characterizes the cutoff time of the illumination of the aperture. As a consequence, it determines the maximum radius of the aperture, i.e., the maximum radius of the annular illumination before the field is cutoff. Even though for a Gaussian window we cannot define a precise cutoff time, it is convenient to use such a function because its Fourier transform is a smooth Gaussian. A rectangular time window provides a definite cutoff time, but its Fourier transform is highly oscillatory and more spread out than the Gaussian ω-window given in Eq. (17b). As a consequence, the Gaussian time window is easier to handle numerically, while its e^{-2} point provides a reasonable estimate of the cutoff time. The use of other time windows and their effect on the decay patterns of FWM-like pulses have been investigated.[5,7] We have shown that the decay of the centroid of LW pulses can be altered significantly in the near-to-far-field range by changing the time sequence of the initial excitation. We believe that similar results can be obtained for finite-time X-wave apertures. For large T values, the Gaussian ω window in (17b) reduces to a narrow distribution centered around $\omega\sim\chi y_0=\chi c\cosh\beta$. The spectral width of such a window is chosen to be smaller than $\omega_{\max}$. Therefore, the temporal spectrum can be viewed as a juxtaposition of these windows centered around all the significant χ components determined by the spatial spectrum given in Eq. (14).

Using Huygen's formula,[27] the field radiated into the $z>0$ half-space can be expressed as[1–3]

$$\Psi(\rho,z,t)=\int_{0}^{+\infty}d\chi\ \chi J_0(\chi\rho)$$

$$\times\int_{0}^{+\infty}d\omega\ \Phi(\chi,\omega)e^{i\omega t}e^{-i\sqrt{(\omega/c)^2-\chi^2}z}. \tag{18}$$

Substituting the spectrum given in Eq. (17a), the amplitude of the radiated field may be expressed as a superposition over the spatial spectral content of the X-wave pulse; specifically,

$$u(\rho,z,t)=\Re e\{\Psi(\rho,z,t)\}$$

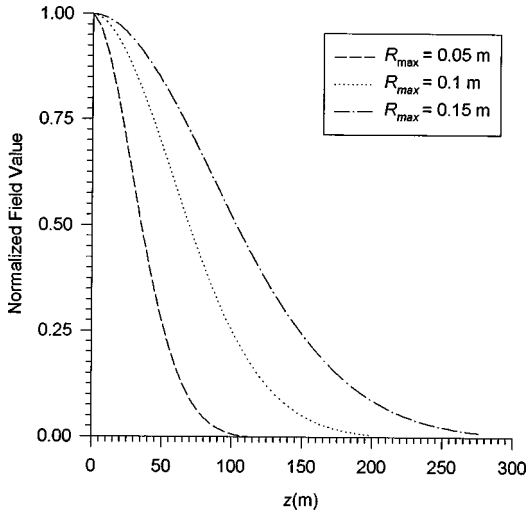


FIG. 7. Decay pattern of the centroid of the finite time X wave for different values of the maximum aperture radius.

$$=\int_{0}^{+\infty}d\chi\ \chi J_0(\chi\rho)\hat{\Phi}_s(\chi,z,t). \tag{19a}$$

Here, $\hat{\Phi}_s(\chi,z,t)$ is the spatial spectral content at a distance z and time t. It is given explicitly by

$$\hat{\Phi}_s(\chi,z,t)=\int_{0}^{+\infty}d\omega\ \frac{2\pi}{\chi}\,e^{-\chi a_1}\hat{\delta}(\omega-\chi y_0)$$

$$\times\cos(\sqrt{(\omega/c)^2-\chi^2}z-\omega t). \tag{19b}$$

For the field generated from an infinite aperture ($T\to\infty$), the Gaussian in Eq. (19) becomes a Dirac δ function forcing ω to equal χy_0. In this limit, the argument of the sinusoidal term becomes equal to $\chi\sinh\beta(z-ct\coth\beta)$, leading to the diffraction-free X wave moving with speed $v_p\equiv c\coth\beta$. For the finite-time excitation, the fine spectral balance that yields the exact $(z-ct\coth\beta)$ functional dependence is upset. Instead, as the pulse travels away from the source plane, the sinusoidal term progressively introduces oscillations into the Gaussian windows $\hat{\delta}(\omega-\chi y_0)$ centered around all the significant spectral χ components.[3,5–7] Since these oscillations increase with distance, the integration over ω yields smaller spectral contributions at the different χ frequencies. As a consequence, the integration over χ in Eq. (19a) gives a field amplitude that decreases with distance. The depletion of the spatial spectral components of the X wave is depicted in Fig. 6, where we have plotted the spatial spectrum given in Eq. (19b) for $cT=91.69$ μm and at $z=ct\coth\beta=0$, 40, and 80 m. This figure shows that the spatial spectral components acquire smaller amplitudes as the pulse travels away from the aperture. Consequently, the integration over χ in Eq. (19) yields the expected decay pattern of the centroid of the radiated pulse. The decay of the peak ($\rho=0$ and $z=ct\coth\beta$) of the X wave, normalized with respect to the peak amplitude at the aperture, is shown in Fig. 7. The curves are plotted for the cutoff times $cT=30.56$, 61.13, and 91.69 μm. Such cutoff times correspond to time-dependent apertures acquiring maximum radii $R_{\max}=R(2cT)=5$, 10, and 15 cm.

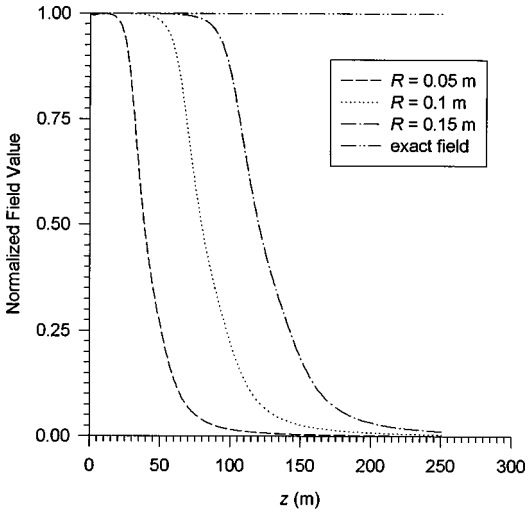


FIG. 8. Decay pattern of the centroid of the X-wave pulse generated by an array of concentric annular sections for different radii.

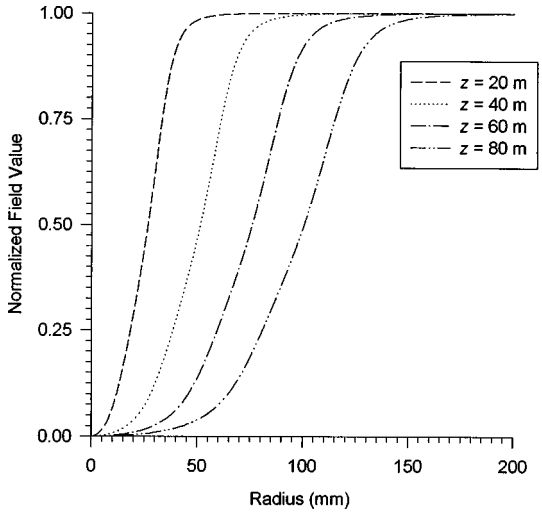


FIG. 9. Contribution of the different annular sections to the field amplitudes at various distances.

To verify the validity of the estimated maximum radii, we deduce the decay pattern of the peak of the X-wave using a finite-radius aperture. For such a construction the source is assumed to consist of a large number of concentric annular array elements. The amplitude of the initial excitation is specified on each distinct element and the radiated field along the axis of propagation ($\rho=0$) is evaluated using the following discrete version of Huygens' formula:

$$u_H(0,z,t)=-\sum_{n=1}^{N}\frac{A_n}{4\pi R_n}\left\{[\partial_{z'}u]-[\partial_{ct'}u]\frac{z-z_0}{R_n}-[u]\frac{z-z_0}{R_n^2}\right\}. \tag{20}$$

Here, A_n is the area of the annular section labeled n, N is the total number of annular sections, $\rho_n'=n\Delta\rho'$ specifies the radial position of the annular section labeled n, and z is the distance at which the field of the radiated pulse is reconstructed. The distance from the source point to the observation point is denoted by $R_n=\sqrt{(n\Delta\rho')^2+(z-z_0)^2}$ for $\rho=0$. The square brackets in Eq. (20) indicate that the enclosed quantities are evaluated at the retarded time $t'=t-(R_n/c)$. The main purpose of the numerical reconstruction procedure embodied in Eq. (20) is to focus on the possibility of using space–time sources on the aperture plane $z=0$ in order to generate LW pulses in the region $z\geqslant 0$. A second objective is to gain more physical insight into the propagation characteristics of the reconstructed pulses.

For annular elements of width $\Delta\rho'=1$ mm, we can construct apertures having radii $R=5$, 10, and 15 cm by using $N=50$, 100, and 150 sections, respectively. The decay patterns of the LW pulses radiated by these three apertures are plotted in Fig. 8 for the values of a_1 and β used earlier. A comparison between Figs. 7 and 8 shows that the ranges of the two decay patterns are similar. For the finite-time dynamic apertures, the centroids of the radiated X waves start decaying at close ranges and their rolloff is smoother than that of X waves generated from discrete circular arrays. The centroids of the X-wave pulses generated by discrete circular arrays exhibit a very slow decay in the near field. The decay in the far field, however, is more abrupt than that corresponding to the finite-time pulses. The smoother decay is attributed to the continuous turnoff of the initial finite-time excitation due to the application of the Gaussian time window. The excitation wave field, therefore, acquires smaller amplitudes at larger radial positions as the illumination of the aperture approaches its turn off point. We would like to emphasize, however, that the finite-time aperture performance is very close to that of a physically bounded source even though the former has no physical edge containing the illumination wave field. This is a crucial point allowing us to use finite-time dynamic sources in our analysis instead of physical apertures. The former lend themselves nicely to the advocated spectral depletion approach. In Fig. 9 we display a set of curves showing how the amplitude of the peak of the propagating X wave, at a specific distance, consists of contributions from the illumination of a finite domain of the aperture. This information can be used to determine the size of the aperture needed to radiate a decay-free pulse up to a specific distance z (cf. Ref. 20 for an experimental procedure along this path).

Finally, we provide in Figs. 10 and 11 a comparison between the shapes of the longitudinal envelopes of the propagating X waves at $z=100$. In Fig. 10 we display the time history of the X-wave pulse launched from a finite-time aperture and compare it to the initial pulse excitation at the aperture. On the other hand, in Fig. 11 we compare an X-wave pulse generated by a discrete circular array to the exact wave solution. The oscillations appearing in the front part of the radiated X wave are due to the discretization of the excitation wave field. Otherwise, the time histories of the pulses generated using the two suggested frameworks are almost identical to the exact wave solution.

IV. DIFFRACTION LENGTH ANALYSIS

For a quasi-monochromatic CW signal, the diffraction length is determined in terms of the Rayleigh distance $\pi R^2/\lambda_c$, where R is the radius of the source and λ_c is the

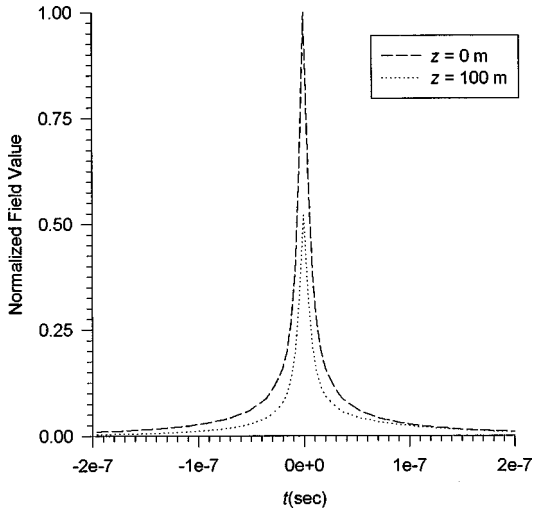


FIG. 10. The time history of the finite-time X-wave pulse at $z=0$, 100 m due to an aperture having radius $R_{\max}=0.15$ m.

wavelength of the carrier wave field. In contradistinction, any attempt to define the diffraction length of a LW pulse generated by a dynamic aperture antenna is more involved. The difficulty can be attributed to the ultra-wideband temporal bandwidth of the LW pulse, as well as the fact that the aperture is characterized by a time-varying radius. One of the earlier attempts to define a diffraction length for LW pulses generated by dynamic aperture antennas was undertaken by Hafizi and Sprangle.[28] Their diffraction length is given as $Z_{\mathrm{HS}}=(2\pi w_0 R_{\max})/\lambda_{\min}$, where w_0 is the beam waist, $R_{\max}$ is the maximum dimension of the aperture or antenna, and $\lambda_{\min}$ is the minimum wavelength. Specific applications of this formula to FWM-like LWs can be found in Ref. 7.

The Hafizi–Sprangle diffraction length gives a good *broad sense* estimate of the behavior of LW wave pulses as they propagate away from their generating aperture planes. However, it does not include special *local features* arising from different types of dynamic aperture excitations. In this section, the decay pattern of an X-wave pulse will be characterized using a diffraction length criterion that has been proposed recently.[7] Such a definition of the diffraction limit is based on the spatial depletion of LW pulses generated by ''finite-time'' dynamic apertures. One can argue that a pulse enters its far-field region when most of the initially significant spatial spectral components are depleted. These spectral components obviously characterize the spatial distribution of the source elements on the aperture. Hence, the loss of information through the depletion of these components signals the beginning of the far-field region. For ultra-wideband pulses one is faced with the ambiguity of determining what are the initially ''significant'' spatial spectral components. Should one use a 3-dB point as the highest ''significant'' frequency, or would a $(1/e^2)$ or $(1/e^4)$ point be a more suitable choice? This ambiguity reflects the fact that there is no sharp limit between the near- and far-field regions of ultra-wideband pulses. A definition of a diffraction limit can only be used to compare the relative performance of different LW pulses. The crucial issue here is to be consistent in our use of definitions of the radii and the bandwidths of the excitation wave fields. Since any of the above definitions can be used to specify the highest ''significant'' spatial spectral components, we shall use a generic cutoff frequency denoted by χ_d. In this case, we may refer to the following expression for the diffraction length of LW pulses [cf. Eq. (12) in Ref. 7]:

$$Z_d=\pi\left(\frac{4}{\eta(\omega/c)}\right)\left(\frac{\omega_0(\chi_d)}{\chi_d c}\right)^2 = \mu(\chi_d)R_{\max}R_{\min}(\omega_{\max}/c). \tag{21}$$

Here, $\eta(\omega/c)$ is the effective width of the spectral ω-window given in Eq. (17b) and $\omega_0(\chi_d)$ is the ω value of the center of $\hat{\delta}(\omega-\chi y_0)$ at $\chi=\chi_d$. The quantity $\mu(\chi_d)$ is a numerical factor of $O(1)$ that depends on the specific choice of χ_d, in general. The diffraction length in Eq. (21) is expressed in terms of spectral quantities depicting the excitation wave field,[7] or as a function of the dimensions of the utilized dynamic aperture and the bandwidth of its source elements.[2,3] The two expressions may be deduced from each other if one recalls that $R_{\max}=R(2T)$, $R_{\min}=a_1$, and $(\omega_{\max}/c)=(4\cosh\beta)/a_1$. Furthermore, one can choose the ''significant'' spatial components as those exhibiting amplitudes larger than $(1/e^4)$ of the maximum of the spatial spectrum. As a consequence, explicit expressions for the diffraction length given in Eq. (21) can be derived. For the X-wave pulse we define $\chi_d=4/a_1$, $\eta(\omega/c)=4/cT$, and $\omega_0(\chi_d)=c\chi_d\cosh\beta$. In this case, Eq. (21) reduces to

$$Z_d=\pi cT(\cosh\beta)^2, \tag{22}$$

which is independent of the choice of χ_d. This means that the source of the X-wave radiated field has the unique property of yielding a diffraction length that does not depend on the particular criterion chosen to define χ_d. In contrast, the diffraction length of the time-limited FWM pulse is proportional to χ_d^2.[7] Hence, X waves can serve to test the relative performance of other LW pulses of comparable source dimensions and temporal spectral bandwidths. For the parameter values used in Fig. 7, Eq. (22) gives $Z_d=64.25$, 128.5,

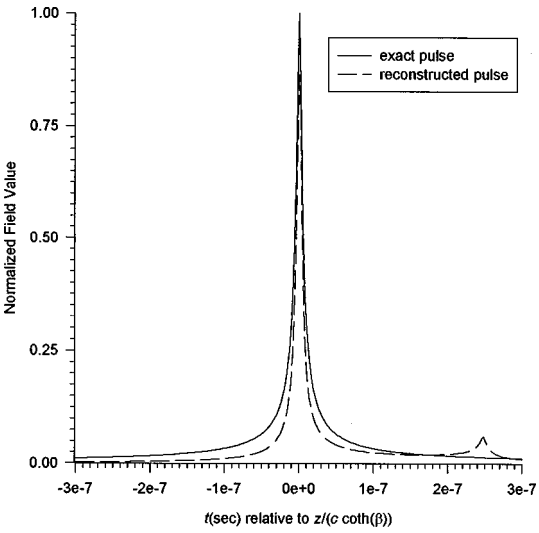


FIG. 11. The time history of the X-wave pulse reconstructed at $z=100$ m from a circular array and compared to the exact solution at the same distance.

and 192.75 m. Since the maximum dynamic radius is given by $R_{\max}=2cT\cosh\beta$, the diffraction limit given in Eq. (22) may be rewritten as

$$Z_d=(\pi/2)R_{\max}\cosh\beta=(\pi/4)D\cosh\beta, \tag{23}$$

where D is the diameter of the source. Apart from a small difference in the numerical constant multiplying the $D\cosh\beta$ dependence of the diffraction limit, our result is the same as the one derived by Lu and Greenleaf.[16–18] The difference in the numerical constant is due to the different criteria used to define the quantities involved in the two approaches.

The expression given in Eq. (21) illustrates the kind of performance enhancement that LW pulses exhibit when compared to quasi-monochromatic Gaussian wave fields. Consider, for example, the case of a quasi-monochromatic beam of waist $w_0=2R_{\min}$ equal to that of the highly focused central portion of the X wave. For a static aperture having $D\gg w_0$, the associated diffraction length is defined as $Z_w=\pi w_0^2/4\lambda_{\min}$. Using $(\omega_{\max}/c)=2\pi/\lambda_{\min}$, the enhancement in the diffraction range of the X wave over an equivalent quasi-monochromatic beam, of the same waist, is given by the ratio $(Z_d/Z_w)\propto(R_{\max}/R_{\min})$. For the parameters used in this paper, the enhancement in the depth of propagation of the X wave over a comparable quasi-monochromatic signal is of order $O(10)\rightarrow O(40)$. This demonstrates that LW pulses become advantageous when narrow pulsed wave fields are required to have extended focused ranges. A dynamic aperture illuminated by a wave field having a time-dependent radius provides an efficient scheme to generate narrow LW pulses from much larger sources. Finally, one should note that the diffraction limit given in Eq. (22) is independent of the parameter a_1. This indicates that the diffraction length can be kept constant for pulses having different temporal bandwidths. As such, an X-wave pulse can travel to the same diffraction-free range even if we use an excitation wave field having a smaller bandwidth. This is true; however, the radius of the highly focused central portion of the radiated field is equal to $R_{\min}=a_1$ and $\omega_{\max}\propto 1/a_1$. Hence, a smaller bandwidth corresponds to a larger focus radius $R_{\min}$ for which the enhancement ratio $(R_{\max}/R_{\min})$ acquires smaller values. Consequently, the use of smaller bandwidths reduces the enhancement achieved over quasi-monochromatic signals.

V. CONCLUDING REMARKS

In this paper, we have demonstrated that an X wave exhibits all the features characterizing other LW pulses. It has been shown that it is characterized by the spatio-temporal spectral coupling that distinguishes LW pulses from other pulsed wave fields.[6,7] A direct consequence of such a coupling is the unique depletion of the spectral components and, hence, the exhibited slow decay pattern of such LW pulses in the near-field range. The main difference between an X wave and other FWM-like wave fields is the functional dependence of the coupling between the temporal and spatial spectral components. Whereas the FWM displays an $\omega\propto\chi^2$ coupling, the X wave has an $\omega\propto\chi$ coupling. Once the similarity between an X wave and other LW pulses is recognized, we can bring together the apparently disparate results of the two research areas into a single framework.

It has also been shown that an X wave may be generated from finite-time dynamic apertures analogous to those proposed as sources for launching FWM-like pulses. One should note that such sources acquire finite sizes by turning on/off their illumination wave fields at finite times. They exhibit most of the essential features of physical apertures and at the same time are much easier to handle mathematically. The illumination of the dynamic source of an X wave acquires the shape of an annulus whose size varies with time. This should be contrasted with the excitation of the FWM aperture that has the shape of a disc having a time-dependent radius. A finite-time dynamic aperture does not have a physical edge bounding its illumination wave field. This might shed some doubt on the validity of the results deduced using such a construction. To examine the origin of any possible discrepancies, we have studied the decay properties of an X wave generated from a finite radius source consisting of discrete concentric annular sections. The resulting decay patterns of the X wave have been shown to compare nicely with those derived using a finite-time dynamic source.

Similarly to other LW pulses generated from dynamic apertures, the spectral depletion of an X wave differs from conventional quasi-monochromatic pulsed wave fields. The decay in the spectral components contributing to the amplitude of the centroid of the propagating pulse results from the increasing oscillations introduced into the spectral ω-windows with the distance z. One can, thus, utilize the knowledge gained from earlier investigations[2–7] to derive the diffraction limit characterizing an X wave. Our spectral depletion approach has established that an X wave has the attractive property that its diffraction length does not depend on the particular choice of χ_d. As a consequence, the decay behavior of any other LW pulses can be compared to that of an X wave.[29] Both fields should have equal aperture sizes, with illumination wave fields exhibiting comparable bandwidths. Furthermore, the two radiated fields should attain equal focused waists. Finally, it has been shown that the enhancement in the focused depth of an X-wave dynamic source, in comparison to a static aperture used to generate quasi-monochromatic signals of a comparable waist is proportional to the ratio $(R_{\max}/R_{\min})$.

[1] A. M. Shaarawi, R. W. Ziolkowski, and I. M. Besieris, ''On the evanescent fields and the causality of the focus wave modes,'' J. Math. Phys. **36**, 5565–5587 (1995).

[2] A. M. Shaarawi, I. M. Besieris, R. W. Ziolkowski, and S. M. Sedky, ''The generation of approximate focus wave mode pulses from wide-band dynamic Gaussian apertures,'' J. Opt. Soc. Am. A **12**, 1954–1964 (1995).

[3] A. M. Shaarawi, S. M. Sedky, R. W. Ziolkowski, and I. M. Besieris, ''The spatial distribution of the illumination of dynamic apertures and its effect on the decay rate of the radiated localized pulses,'' J. Phys. A **29**, 5157–5179 (1996).

[4] R. W. Ziolkowski, I. M. Besieris, and A. M. Shaarawi, ''Aperture realizations of the exact solutions to homogeneous-wave equations,'' J. Opt. Soc. Am. A **10**, 75–87 (1993).

[5] A. M. Shaarawi, S. M. Sedky, R. W. Ziolkowski, and F. M. Taiel, ''Effect of the switching pattern of the illumination of dynamic apertures on the ranges of the generated waves,'' J. Opt. Soc. Am. A **13**, 1712–1718 (1996).

[6] A. M. Shaarawi, S. M. Sedky, F. M. Taiel, R. W. Ziolkowski, and I. M.

Besieris, ''Spectral analysis of time-limited pulsed Gaussian wavefields,'' J. Opt. Soc. Am. A **13**, 1817–1835 (1996).
[7] S. M. Sedky, A. M. Shaarawi, F. M. Taiel, and I. M. Besieris, ''On the diffraction length of localized waves generated by dynamic apertures,'' J. Opt. Soc. Am. A **13**, 1719–1727 (1996).
[8] R. W. Ziolkowski, ''Localized transmission of electromagnetic energy,'' Phys. Rev. A **39**, 2005–2033 (1989).
[9] R. W. Ziolkowski, ''Localized wave physics and engineering,'' Phys. Rev. A **44**, 3960–3987 (1991).
[10] R. W. Ziolkowski, I. M. Besieris, and A. M. Shaarawi, ''Localized wave representations of acoustic and electromagnetic radiation,'' Proc. IEEE **79**, 1371–1378 (1991).
[11] A. M. Vengsarkar, I. M. Besieris, A. M. Shaarawi, and R. W. Ziolkowski, ''Closed-form localized wave solutions in optical fiber waveguides,'' J. Opt. Soc. Am. A **9**, 937–949 (1992).
[12] A. M. Shaarawi, I. M. Besieris, and R. W. Ziolkowski, ''Localized energy pulse trains launched from an open, semi-infinite, circular waveguide,'' J. Appl. Phys. **65**, 805–813 (1989).
[13] J. N. Brittingham, ''Focus wave modes in homogeneous Maxwell equations: transverse electric mode,'' J. Appl. Phys. **54**, 1179–1189 (1983).
[14] R. W. Ziolkowski, ''Exact solutions of the wave equation with complex source locations,'' J. Math. Phys. **26**, 861–863 (1985).
[15] J. Durnin, ''Exact solutions for nondiffracting beams. I. scalar theory,'' J. Opt. Soc. Am. A **4**, 651–654 (1987).
[16] J. Y. Lu and J. F. Greenleaf, ''Nondiffracting X waves—exact solutions to free space scalar wave equation and their finite aperture realization,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 19–31 (1992).
[17] J. Y. Lu, H. Zou, and J. F. Greenleaf, ''A new approach to obtain limited diffraction beams,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42**, 850–853 (1995).
[18] J. Y. Lu and J. F. Greenleaf, ''Experimental verification of nondiffracting X waves,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 441–446 (1992).
[19] J. Y. Lu and J. F. Greenleaf, ''Sidelobe reduction for limited diffraction pulse-echo system,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **40**, 735–746 (1993).
[20] J. Y. Lu and J. F. Greenleaf, ''Ultrasonic nondiffracting transducer for medical imaging,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **37**, 438–447 (1990).
[21] J. Y. Lu, M. Fatemi, and J. F. Greenleaf, ''Pulsed-echo imaging with X wave,'' Acoust. Imaging **22**, 191–196 (1996).
[22] J. Y. Lu and J. F. Greenleaf, ''Diffraction-limited beams and their applications for ultrasonic imaging and tissue characterization,'' Proc. SPIE **1733**, 92–119 (1992).
[23] J. Y. Lu, ''Designing limited diffraction beams,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **44**, 181–193 (1997).
[24] F. Hsu, J. Margetan, and D. O Thomson, ''Bessel beam ultrasonic transducer: Fabrication method and experimental results,'' Appl. Phys. Lett. **55**, 2066–2068 (1989).
[25] D. Power, R. Donnelly, and R. MacIsaac, ''Spherical scattering of superpositions of localized waves,'' Phys. Rev. E **48**, 1410–1417 (1993).
[26] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products* (Academic, New York, 1965).
[27] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw–Hill, New York, 1953), Sec. 11.3.
[28] B. Hafizi and P. Sprangle, ''Diffraction effects in directed radiation beams,'' J. Opt. Soc. Am. A **8**, 705–717 (1991).
[29] A. M. Shaarawi, ''Comparison of two localized wave fields generated from dynamic apertures,'' J. Opt. Soc. Am. A **14**, 1804–1816 (1997).

Wave propagation in a wavy fiber–epoxy composite material: Theory and experiment

Kwang Yul Kim,[a] Wei Zou, and Wolfgang Sachse
Department of Theoretical and Applied Mechanics, Thurston Hall, Cornell University, Ithaca, New York 14853-1503

(Received 21 July 1997; accepted for publication 5 February 1998)

In this paper analytic formulas are developed for the ray path and travel time of a ray propagating in a wavy fiber–epoxy composite material and calculate them for rays initiating at various points with wave normals of differing directions. The arrival times observed by using various combinations of pointlike sources and pointlike detectors are found in good agreement with those predicted by the theory of geometrical acoustics. © *1998 Acoustical Society of America.* [S0001-4966(98)04805-X]

PACS numbers: 43.20.Dk [ANN]

INTRODUCTION

Fiber-waviness is introduced in composite materials during manufacturing processes because of uneven curing and shrinkage of resin, and it has detrimental effects on mechanical properties, such as stiffness reduction and degradation of compressive strength.[1,2] Layer waviness also occurs in thick crossply or multidirectional laminates due to the lamination residual stresses built up during curing process. Therefore, it is important to detect and characterize fiber-waviness during and after fabrication of a fiber-reinforced composite, and nondestructive ultrasonic evaluation of the wavy composite remains a challenging problem.

There have been the investigations on the unidirectional wavy composite by means of ultrasonic waves. Wooh and Daniel[3] used a ray tracing method in an attempt to understand the waveforms detected by the piezoelectric transducers of finite size attached on one side of the flat surfaces. The waveforms were launched by the piezoelectric transducers of finite size on the same or opposite side of the flat surfaces. Using the finite difference method, McIntyre *et al.*[4] theoretically investigated the harmonic wave propagation with wave normal directed in the mean fiber-direction, and found the variation of amplitude in tune with the energy conservation principle implied by the transport equation.

In this paper we apply the results of geometrical acoustics to investigate ray paths and travel times along the various ray paths of ultrasonic waves propagating in a wavy graphite fiber–epoxy composite. Even though experimental verification of the theoretical ray paths is difficult to achieve because of the opaque composite specimen, it will be demonstrated that the observed travel times of the quasilongitudinal (QL) and shear horizontally (SH) polarized pure transverse (PT) modes propagating along the different ray paths are found in good agreement with those predicted by the theory of geometrical acoustics.

I. THEORY

A schematic diagram of a wavy composite specimen is drawn in Fig. 1, where the reinforcing fibers are running in a wavy pattern in the xy plane with their mean direction pointing along the x axis. We specify the waviness of the fiber-reinforced epoxy composite by the following sine wave function

$$y_f = A_f \sin(2\pi x/L_f), \tag{1}$$

where y_f is the fiber height in the y direction, A_f is the amplitude of fiber-waviness, and L_f is the periodical length of waviness. A local fiber-direction is specified by the angle β_f the x axis makes with the fiber. Here, $\tan\beta_f$ is given by

$$\tan\beta_f = \frac{dy_f}{dx} = \frac{2\pi A_f}{L_f}\cos\frac{2\pi x}{L_f}. \tag{2}$$

The imbedded wavy fibers notwithstanding, the fabricated wavy composite specimens are observed to have the uniform density ρ throughout the specimens, and we assume that the density of the wavy composite is constant. The wavy composite is also uniform along the z axis normal to the xy plane on which the waviness of the fibers is confined. The composite is also uniform along the y direction. As a result, both anisotropy and heterogeneity of the wavy composite is a function of one variable x only. Therefore, the wave normals, which are initially directed in the xy plane, are confined on the xy plane and refraction of the wave normals is governed by Snell's law during the propagation of rays in the xy plane. A composite material which has the fibers running straight in the x direction ideally has transversely isotropic symmetry about the x axis. However, because of a less-than-ideal fabrication procedure, it is better characterized as possessing weak orthorhombic symmetry with proximity to transverse isotropy. Let the x, y, and z directions be represented by indices 1, 2, and 3, respectively. Then, the composite specimen with straight fibers aligned in the x direction is characterized by nine elastic constants: C_{11}, C_{22}, C_{33}, C_{12}, C_{13}, C_{23}, C_{44}, C_{55}, and C_{66}.

Consider a small element of the wavy composite at a typical local point (x,y_0,z_0). The small local element is considered to have orthorhombic symmetry whose x and y symmetry axes are rotated about the third symmetry axis z_0z by the angle β_f from the x axis. Let's specify a wave normal **n** of the wave propagating in the xy plane by an angle θ mea-

[a]Electronic mail: kykim@msc.cornell.edu

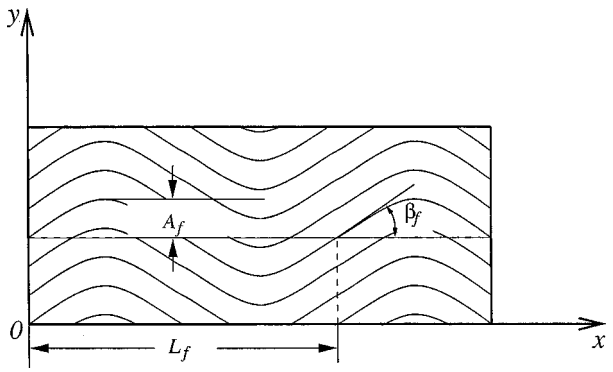


FIG. 1. Schematics of a wavy fiber composite specimen.

sured from the x axis. Then, the angle α which the wave normal **n** makes from the local fiber direction is

$$\alpha(x,\theta)=\theta-\beta_f=\theta-\tan^{-1}\left(\frac{2\pi A_f}{L_f}\cos\frac{2\pi x}{L_f}\right). \tag{3}$$

We invoke Snell's law and the formulas for the arrival time τ and ray path of a wave initiating from the point (x_0,y_0,z_0) and arriving at the point (x_1,y_1,z_0). Referring to the results established in the geometrical acoustics of inhomogeneous anisotropic media,[5,6] they can be written as

$$h=\frac{\sin\theta_0}{v_0}=\frac{\sin\theta}{v}=\text{constant}, \tag{4}$$

$$\tau=\int_{x_0}^{x_1}\frac{dx}{v\cos\theta-v'\sin\theta}, \tag{5}$$

$$y_1=y_0+\int_{x_0}^{x_1}\frac{v\sin\theta+v'\cos\theta}{v\cos\theta-v'\sin\theta}\,dx, \tag{6}$$

where $v(x,\theta)$ is the phase velocity and $v'=\partial v/\partial\theta$.

For the waves propagating in the xy plane, the eikonal and phase velocity equations are factored into three modes whose vibration directions are mutually perpendicular to each other: shear horizontally (SH) polarized pure transverse (PT) mode vibrating in the z direction, quasilongitudinal (QL) mode and quasitransverse (QT) mode. The QL and QT modes are both polarized in the xy plane. The SH polarized PT mode is uncoupled from the QL and QT modes. We first deal with the PT mode.

A. Pure transverse mode

The phase velocity v of the PT mode is given by[7]

$$\rho v^2=C_{44}\sin^2(\theta-\beta_f)+C_{55}\cos^2(\theta-\beta_f), \tag{7}$$

which we differentiate with respect to θ to obtain

$$\rho v v'=(C_{44}-C_{55})\sin(\theta-\beta_f)\cos(\theta-\beta_f). \tag{8}$$

Using Snell's law Eq. (4), Eq. (7) is expressed as

$$\begin{aligned}\rho\sin^2\theta&=h^2[C_{44}\sin^2(\theta-\beta_f)+C_{44}\cos^2(\theta-\beta_f)]\\&=h^2[(C_{44}-C_{55})\sin^2(\theta-\beta_f)+C_{55}],\end{aligned} \tag{9}$$

where $h=\sin\theta/v$ is the constant in Snell's law. Consider a case in which rays are initiating from a broadband source in virtually every direction inside the specimen. We choose an arbitrary ray whose wave normal at the source is directed at an angle θ_0 from the x direction. The fiber direction β_f at the source is obtained from Eq. (2). Then, we calculate the initial velocity v_0 corresponding to the wave normal θ_0 via Eq. (7) to determine the Snell's law constant h from Eq. (4). For a given x, the fiber angle β_f can be calculated via Eq. (2) and from Eq. (9), one obtains the solutions for θ, $\sin\theta$, and $\cos\theta$. Then the values of v and v' are obtained from Eqs. (7) and (8). Thus, one can calculate the values of $\sin\theta$, $\cos\theta$, and v and v' at many different values of x to determine the arrival time τ and ray path through Eqs. (5) and (6). One can repeat this procedure for rays with various initial wave normal directions.

For the special case of the initial wave normal **n** pointing in the x direction, $h=\sin\theta/v=0$, i.e., $\theta=0$ all along the ray path, and the arrival time and ray path integrals are simplified to

$$\tau=\rho\int_{x_0}^{x}\frac{dx}{[(C_{44}-C_{55})\sin^2\beta_f+C_{55}]^{1/2}}, \tag{10}$$

$$\begin{aligned}y-y_0&=\int_{x_0}^{x}\frac{(C_{55}-C_{44})\sin\beta_f\cos\beta_f}{C_{44}\sin^2\beta_f+C_{55}\cos^2\beta_f}dx\\&=\frac{C_{55}-C_{44}}{C_{44}}\frac{L_f^2}{4\pi^2A_f\,p}\\&\quad\times\tanh^{-1}\frac{p[\sin(2\pi x/L_f)-\sin(2\pi x_0/L_f)]}{p^2-\sin(2\pi x_0/L_f)\sin(2\pi x/L_f)},\end{aligned} \tag{11}$$

where

$$p=\sqrt{1+\frac{L_f^2C_{55}}{4\pi^2A_f^2C_{44}}}. \tag{12}$$

Equation (11) indicates that the ray with initial wave normal **n** directed along the x-axis follows a path with the same periodic length that the wavy fiber has.

B. Quasilongitudinal and quasitransverse modes

For simplicity of notation we introduce the following identities

$$C_{11\pm}=C_{11}\pm C_{66},\quad C_{22\pm}=C_{22}\pm C_{66},\quad C_{12\pm}=C_{12}\pm C_{66}. \tag{13}$$

The phase velocities of the QL and QT modes are given by[7]

$$2\rho\,v^2=C_{11+}\cos^2\alpha+C_{22+}\sin^2\alpha\pm\sqrt{D}, \tag{14}$$

where α is defined in Eq. (3), the upper and lower signs in $\pm$ in front of $\sqrt{D}$ correspond to the QL and QT modes, respectively, and

$$D=(C_{11-}\cos^2\alpha-C_{22-}\sin^2\alpha)^2+4C_{12+}^2\sin^2\alpha\cos^2\alpha. \tag{15}$$

Differentiating Eq. (14) with respect to θ, one obtains

$$\frac{v'}{v}=\frac{1}{4\rho v^2}\Bigg\{(C_{22+}-C_{11+})\sin 2\alpha\pm\frac{1}{\sqrt{D}}\times[(C_{22-}\sin^2\alpha-C_{11-}\cos^2\alpha)(C_{11-}+C_{22-})\times\sin 2\alpha+C_{12+}^2\sin 4\alpha]^{1/2}\Bigg\}. \tag{16}$$

Substitution of Snell's law Eq. (4) into Eq. (14) yields

$$2\rho\ \sin^2\ \theta=h^2[C_{11+}\cos^2(\theta-\beta_f)+C_{22+}\sin^2(\theta-\beta_f)\pm\sqrt{D}]. \tag{17}$$

Determination of the arrival time and ray path of a wave with initial wave normal θ_0 at the source can be carried out in a similar way to that described in the case of the PT mode. In the special case of $h=\sin\theta=0$, the evaluation of the integrals for the arrival time and ray path is also simplified as in the PT mode.

II. EXPERIMENTAL METHOD

Two kinds of specimens are fabricated using graphite-fiber prepregs: one is a unidirectional, straight graphite fiber–epoxy composite and the other is a wavy graphite fiber–epoxy specimen with sine wave period 40 mm and wavy amplitude 2 mm. Both specimens have the same density 1524 kg/m^3. The straight fiber direction in the unidirectional composite and the mean fiber direction in the wavy composite are taken as the x direction. The wavy fiber plane is denoted as the xy plane.

In order to fabricate the wavy composite, the top surface of an aluminum plate, which is 20 cm wide, 60 cm long, and 12.7 mm thick, is first machined in the desired sine wave pattern (2 mm amplitude and 40 mm period) on a numerically controlled machine. Then, the machined aluminum plate is split into two identical pieces, which are next aligned against each other with a gap, where many layers of graphite–fiber epoxy prepregs are sandwiched and cured.

The unidirectional composite, which is weakly orthorhombic, is prepared to determine its nine elastic constants from measurements of the travel times of ultrasonic waves propagating through the specimen. A cube of the unidirectional composite which has been cut perpendicular to the three symmetry axes is mechanically polished. The PT waves are launched from one side of the specimen and they are detected on the opposite side, while the PL (pure longitudinal) waves are generated and detected by the same transducer. Denoting the propagation and polarization directions of the phase velocity v of an ultrasonic wave by subscript indices i and j, respectively, the pure index elastic constants or the diagonal elements of C_{ij} matrix are expressed as

$$\rho[v_{ij}^2]=\begin{bmatrix}C_{11} & C_{66} & C_{55}\\ C_{66} & C_{22} & C_{44}\\ C_{55} & C_{44} & C_{33}\end{bmatrix}. \tag{18}$$

Thus, all six pure index elastic constants can be determined via Eq. (18) from the phase velocity measurements along the three symmetry directions. Given the values of the pure-index elastic constants, the mixed index elastic constant C_{ij} $(i\neq j)$ can be obtained from the wave speed measurements of the QL mode propagating in an oblique direction in the symmetry plane. For this purpose two opposite faces of three additional specimens cut from the large unidirectional composite are polished with one in 45° to the x axis and normal to the xy plane, another in 45° to the y axis and normal to the yz plane, and the other in 45° to the z axis and normal to the zx plane. From the measurement of the QL phase velocity propagating in 45° direction to the x axis in the xy plane, one uses Eq. (14) by setting $\beta_f=0$ and $\theta=45°$ to obtain the mixed index elastic constant C_{12}. One proceeds similarly for determination of C_{13} and C_{23}, using the other two obliquely cut specimens. The nine elastic constants of the orthorhombic unidirectional composite can also be determined from the group velocity measurements and this is described in detail elsewhere.[8,9]

One of the fabricated wavy composites has two opposite sides machined flat, which are normal to the wavy fiber plane and parallel to the x axis. It has next been polished with thickness 15.93 mm between the flat surfaces. Then another sample is prepared with two flat opposite faces normal to the x axis with thickness 30 mm between them. A schematic of a typical sample is shown in Fig. 1. First, a 28 μm thick, piezoelectric polyvinylidene fluoride (PVDF) film is laid on the top surface of the flat wavy composite and a glass capillary of size 0.10 mm inside diameter and 0.17 mm outside diameter is mounted on the PVDF film. Pointlike sources of the QL and QT modes are generated by breaking the glass capillary with a razor blade at various x positions on the top surface. The generated elastic waves propagate in virtually every direction inside the specimen and are detected at various x positions on the bottom surface by a pointlike piezoelectric detector of 0.75 mm diameter. The output of the PVDF film indicates a time of source generation and serves as a trigger for the output of the piezoelectric detector. From the arrival-time difference between the two signals, the travel time of the QL mode is determined. Identification of the QL ray arrival in a detected signal is easy, because the arrival is found at the point from which the signal first jumps from noise level. The QT mode in the detected signal is very difficult to identify and we did not attempt to measure its arrival time.

For measurements of the arrival times of the PT mode propagating on the xy wavy fiber plane, a line-type piezoelectric shear transducer, whose active area is 0.75 mm wide and 5 mm long, is mounted on the bottom face of the specimen at different x positions with its polarization aligned along the z direction. The polarization direction of the shear transducer is along the 5 mm lengthwise direction. It generates not only QL modes but also shear horizontally (SH) polarized pure transverse (PT) mode propagating along virtually every direction on the xy plane. The propagated PT modes are detected by a small piezoelectric shear transducer with active area of 0.75 mm diameter, which is attached on the various x positions of the top surface with its polarization aligned along the z direction. Identification of the PT mode arrival in the detected signals that travel through composite materials is not so easy, because its arrival trails that of the QL mode. How to identify the arrival time of the SH polar-

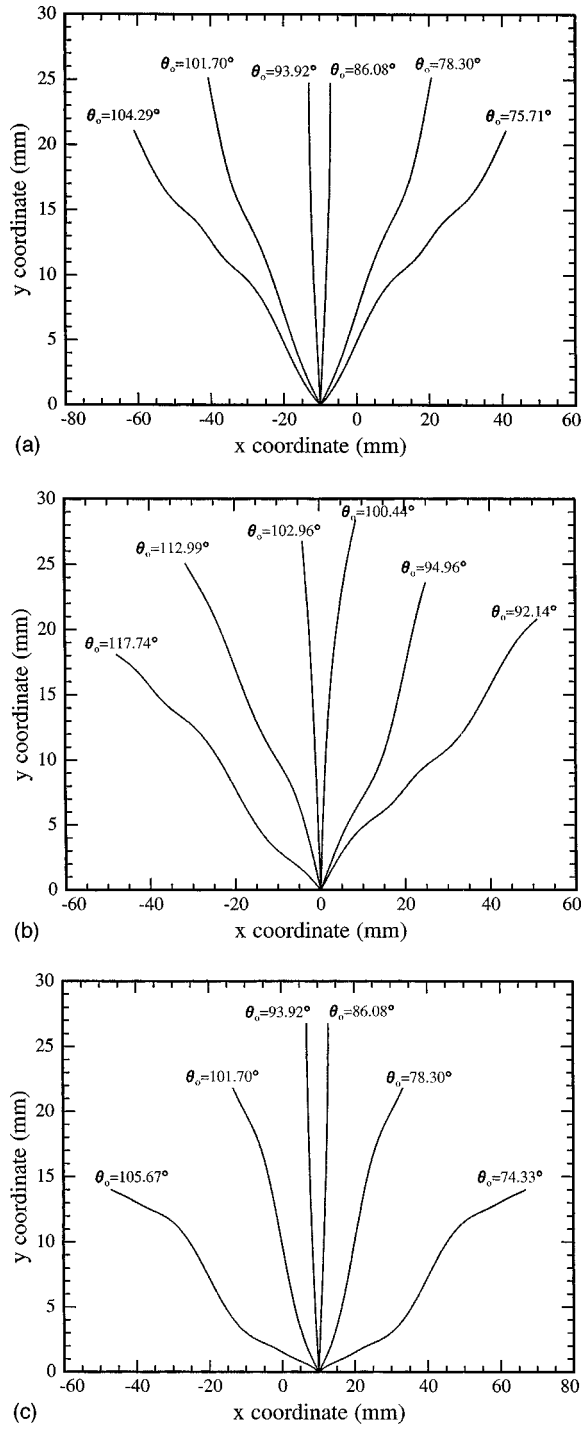


FIG. 2. Ray paths of the QL mode with sources at (a) $x=-10$ mm, (b) $x=0$ mm, and (c) $x=10$ mm.

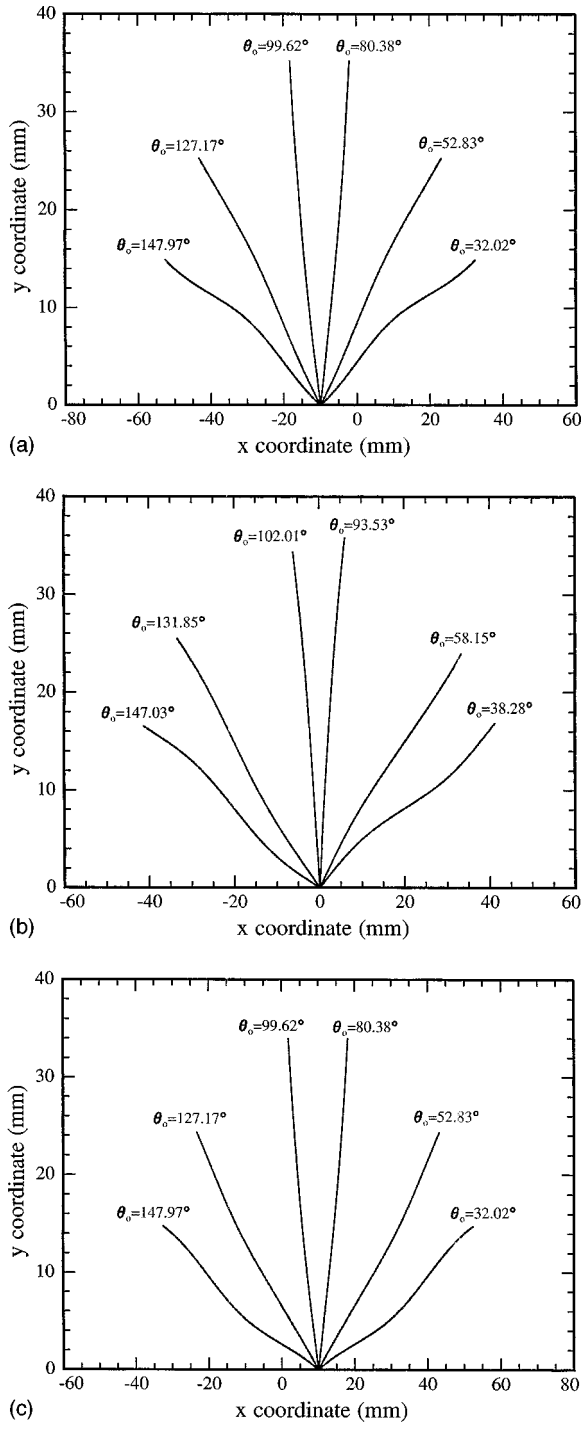


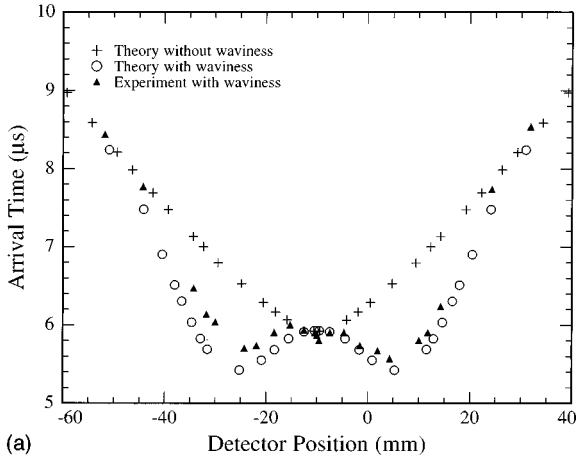
FIG. 3. Ray paths of the PT mode with sources at (a) $x=-10$ mm, (b) $x=0$ mm, and (c) $x=10$ mm.

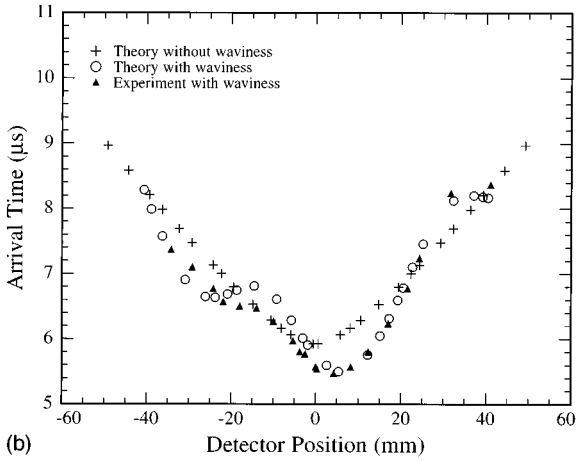
ized PT mode in the detected signal is described in Ref. 9.

It is noted that both the capillary fracture source and piezoelectric detector used for the QL mode have a frequency bandwidth extending well above 10 MHz, while both piezoelectric source and detector used for the PT mode have a broadband frequency spectrum which is centered around 5 MHz. These correspond to a wavelength of the QL mode less than 1 mm and that of the PT mode less than 0.2 mm. These wavelengths are much less than the wavy period of the composite, 40 mm and one expects that a high-frequency geometrical acoustics approximation would hold well in describing the propagation of wave fronts.

III. RESULTS

The obtained nine elastic constants of the unidirectional composite in units of GPa are: $C_{11}=130.3\pm1.5$, $C_{22}=11.00\pm0.28$, $C_{33}=12.46\pm0.44$, $C_{44}=2.95\pm0.08$, $C_{55}=5.47\pm0.22$, $C_{66}=4.96\pm0.08$, $C_{12}=6.01\pm0.36$, $C_{13}=1.04\pm0.09$,





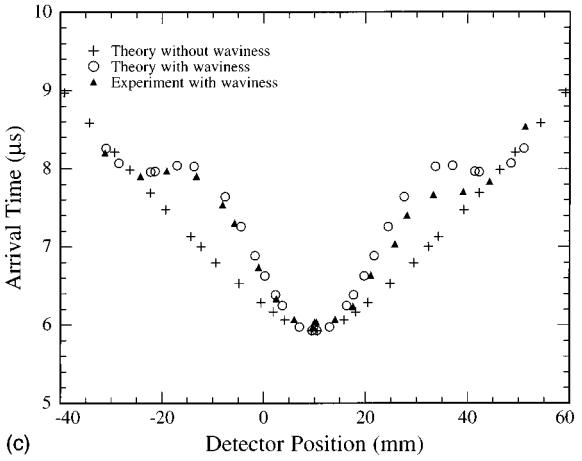


FIG. 4. Comparison of predicted arrival times of the QL ray with measured values. (a) Source at $x=-10$ mm, (b) source at $x=0$ mm, and (c) source at $x=10$ mm.

and $C_{23}=6.27\pm0.37$. We recall that the density and thickness of the wavy composite specimen are 1524 kg/m^3 and 15.93 mm, respectively.

Using these data, the ray paths of the QL mode with sources at $x=-10$, 0, and 10 mm on the bottom face are plotted in Fig. 2(a), (b), and (c), respectively. Those of the PT mode with sources at the same x positions are displayed in Fig. 3(a), (b), and (c), respectively. Although the capillary fracture QL sources are located on the top surface of the wavy composite, the ray paths are drawn as if they were

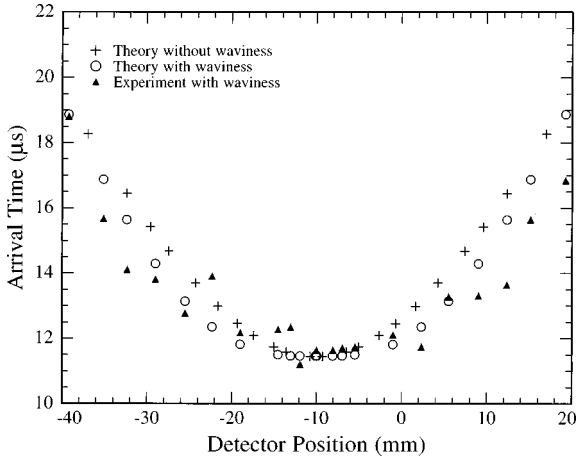


FIG. 5. Comparison between predicted arrival times of the PT ray with a source at $x=-10$ mm and measured values.

located on the bottom surface, which is given the coordinate of $y=0$. As expected, the ray paths of both QL and PT modes with the sources at $x=-10$ and 10 mm exhibit a symmetry pattern about the y axis passing through the source, while the symmetry is broken for the ray paths with a source at $x=0$ mm. θ_0 in these figures indicates the direction of the wave normal associated with a particular ray at the source point. Note that the rays are not only curved but also multiply curved for those substantially away from the vertical direction. QT ray paths are not plotted, because they can be refracted into QL modes anywhere on their path and the arrival of the QT rays are almost impossible to identify in the detected signals.

Comparison of theory with experiment on the ray arrival times of the QL mode is displayed in Fig. 4(a), (b), and (c) with sources at $x=-10$, 0, and 10 mm, respectively. Theoretical values of arrival times in a straight unidirectional fiber composite of the same thickness with similar source-to-detector configurations are also drawn with + (plus) symbols to elucidate the effect of fiber-waviness. It is seen that predicted values generally agree well with experimentally observed values. This signifies that the geometrical acoustics developed for a heterogeneous anisotropic medium provides a good approximation to the travel times of ultrasonic rays and it can be applied to the study of wave propagation in the wavy composite. It also indicates that geometrical acoustics, which is a high-frequency approximation, remains quite valid in the ultrasonic frequency range for description of wave front behavior in a wavy composite material. The deviations of predicted values from measured travel times at some detector positions are likely to be a result of rays passing through a local region where fiber arrangement deviates from the regular wavy pattern before the rays arrive at the detector. As expected, the QL ray arrival times with sources at $x=-10$ and 10 mm exhibit symmetry about their source positions. It is, however, interesting to see that the symmetry patterns are quite different in two cases.

Predicted arrival times of the PT mode are compared with measured values in Fig. 5. Theoretical arrival times in a corresponding straight fiber composite without waviness are also drawn with + (plus) symbols for comparison. The mea-

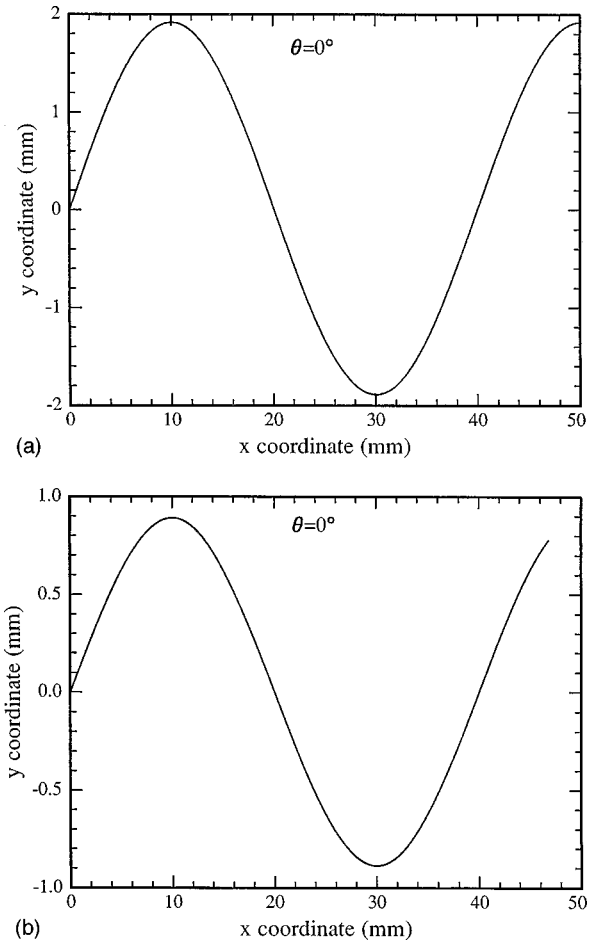


FIG. 6. The path of a ray initiating from origin with wave normal pointing in the x direction. (a) QL mode and (b) PT mode.

sured arrival times exhibit symmetry about the y axis that passes a source located at $x=-10$ mm, but they are more scattered than those of the observed QL mode. This is partly due to the difficulty of correctly identifying the PT mode arrival and as a result, the measured arrival times are considered approximate values with several % error.

Finally, the ray paths of the QL and PT modes initiating at origin ($x=0$ mm, $y=0$ mm) with initial wave normal pointing in the x direction are plotted in Fig. 6(a) and (b), respectively. In this case the direction of wave normal is constant, always in the x direction, all along the ray path, as Snell's law Eq. (4) implies. As can be seen from Fig. 6(a) and (b), both ray paths have the same periodic length as the wavy fiber has, but they have different amplitudes. In the case of the PT mode, this result is already predicted by Eq. (11). It is interesting to notice that the ray path of the QL mode with wave normal in the x direction closely follows a wavy fiber, while that of the PT mode deviates somewhat from the wavy fiber. One finds in Fig. 6(a) that the y coordinate of the QL ray at $x=30$ mm is -1.89 mm. We have prepared a specimen with two parallel sides terminating at $x=0$ mm and $x=30$ mm. The measured arrival time of the QL mode at the point ($x=30$ mm, $y=-1.89$ mm) with a source at origin is 3.33 μs, which agrees remarkably well with the predicted value 3.32 μs.

IV. CONCLUSIONS

We have established closed-form analytic solutions for the ray path and travel time of the QL, QT, and PT rays propagating in a wavy fiber composite. Predicted values of travel time of QL and PT modes are in good agreement with experimentally observed values. Geometrical acoustics developed for a heterogeneous anisotropic material can be useful for investigation of a wavy composite material. We have also shown that a high-frequency geometrical acoustics approximation holds well in the ultrasonic frequency range in describing the propagation of wave fronts in a wavy fiber composite. The method used in this paper may well be adapted to investigating wave propagation in various inhomogeneous anisotropic structures, such as devices with gradient of material properties (e.g., diffused concentration), depth-dependent media, and structures of inhomogeneous stress field, etc.

ACKNOWLEDGMENTS

This work was initiated with the support of the Office of Naval Research and completed with the support of the Materials Science Center at Cornell University, which is funded by the National Science Foundation.

[1] A. L. Highsmith, J. J. Davis, and K. L. E. Helms, ''The influence of fiber waviness on the compressive behavior of unidirectional continuous fiber composites,'' in *Composite Materials: Testing and Design*, Vol. 10, ASTM STP 1120, edited by G. C. Grimes (ASTM, Philadelphia, 1992), pp. 20–36.

[2] H. M. Hsiao, I. M. Daniel, and S. C. Wooh, ''Effect of fiber waviness on the compressive behavior of thick composites,'' in *Failure Mechanics in Advanced Polymeric Composites*, AMD-Vol. 196 (ASTM, Philadelphia, 1994), pp. 141–159.

[3] S. C. Wooh and I. M. Daniel, ''Wave propagation in composite materials with fiber waviness,'' Ultrasonics **33**, 3–10 (1995).

[4] J. S. McIntyre, C. W. Bert, and R. A. Kline, ''Wave propagation in a composite with wavy reinforcing fibers,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 14, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1995), pp. 1311–1318.

[5] N. J. Vlaar, ''Ray theory for an anisotropic inhomogeneous medium,'' Bull. Seismol. Soc. Am. **58**, 2053–2072 (1968).

[6] V. Cerveny and I. Psencik, ''Rays and travel time curves in inhomogeneous anisotropic media,'' Z. Geophys. **38**, 565–577 (1972).

[7] M. J. P. Musgrave, *Crystal Acoustics* (Holden-Day, San Francisco, 1970).

[8] K. Y. Kim, ''Analytic relations between the elastic constants and the group velocity in an arbitrary direction of symmetry planes of media with orthorhombic or higher symmetry,'' Phys. Rev. B **49**, 3713–3724 (1994).

[9] K. Y. Kim, R. Sribar, and W. Sachse, ''Analytical and optimization procedures for determination of all elastic constants of anisotropic solids from group velocity data measured in symmetry planes,'' J. Appl. Phys. **77**, 5589–5600 (1995).

Sound transmission through a periodic cascade with application to drill pipes

Niels J. C. Lous, Sjoerd W. Rienstra, and Ivo J. B. F. Adan
Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

(Received 21 March 1997; revised 8 January 1998; accepted 16 January 1998)

Acoustical data transmission through the wall of drill pipes is considered. Drill pipes are known to behave like bandpass filters; the position of the pass bands can be determined analytically. This work extends the frequency domain drill pipe models presented by Barnes and Kirkwood [J. Acoust. Soc. Am. **51**, 1606–1608 (1972)], and more recently by Drumheller [J. Acoust. Soc. Am. **85**, 1048–1064 (1989)]. The approach discussed in this paper has the advantage that it yields explicit expressions for the fine structure of the drill pipe's frequency response in the pass bands. It furthermore allows the effect of energy dissipation and pipe segment length variations to be included in the model. The emphasis of the paper, however, lies on the time domain modeling of the drill pipe. The propagation of sound energy pulses through its wall, and the effect of multiple reflections and/or transmissions during this propagation, are described using a Markov chain. Explicit expressions result for the expected duration of an energy pulse's trip from one end of the drill pipe to the other, depending upon the number of drill pipe segments and the transmission coefficient at the tool-joints connecting them. The results are applicable to any situation where sound or energy transmission through a cascade of acoustic components occurs. © *1998 Acoustical Society of America.* [S0001-4966(98)00705-X]

PACS numbers: 43.20.Mv, 43.20.Tb, 43.40.Ph [JEG]

INTRODUCTION

When performing soil drillings, one needs to know on the platform what happens at the drill bit. Adverse soil conditions and a too high temperature, among others, may invoke the necessity to stop drilling in order to avoid material damage. Relevant information may be gathered by equipment installed in the drill bit, but conveying the measured data to the platform is a problem as the severe circumstances in the bore hole do not allow wires to be installed.

The only communication method between drill bit and platform applied in practice is modulation of the mud stream inside the drill pipe. This stream is pumped downward through the pipe to cool or to drive the bit. At the lower pipe end, the drill pipe can be closed or opened by means of a flap. Closing the flap causes the pressure in the mud stream to increase, which can be measured at the upper drill pipe end. One information bit thus can be conveyed by respectively closing and opening the flap during one time unit. This method allows only a few bits to be transmitted each second.

In the present study, an alternative is considered where acoustical data are transmitted through the drill pipe wall (see also Barnes and Kirkwood, 1972; Drumheller, 1989). The drill pipe, consisting of steel, hollow, cylindrical segments, each about 10 m long, is excited at the lower end by a piezoelectrical transducer or by a hammer. The resulting sound signals travel through the wall of these segments, but are partly reflected at the tool-joints connecting them. Depending upon the reflection and transmission properties of the tool-joints, where the wall is thicker and includes a screw thread, sound signals of given frequencies do or do not reach the upper drill pipe end, with or without distortion and delay.

We first briefly discuss our drill pipe model in the frequency domain, which is similar to the one in Drumheller (1989), but allows a further-going analytical description of the data transmission process. We list the main modeling assumptions, give a short description of the behavior of tool-joints and pipe segments, and come to equations representing the transmission properties of the whole drill pipe. We then apply our model to consider energy dissipation and pipe segment length variations.

In the time domain, we introduce a few more modeling assumptions, and describe the configuration in terms of a Markov chain. This yields expressions for the time energy (or sound) pulses need to travel through the pipe.

We refer to Fig. 1 for names of drill pipe parts.

Both the frequency domain analysis and the time domain expressions are applicable to any situation where sound or energy transmission through a cascade of acoustic components occurs. In speech production models, for instance, the vocal tract shape is usually approximated by a series of small cylindrical tubes, not unlike a drill pipe. One of the main contributions of this paper is that we derive explicit expressions for the transmission behavior of a cascade of (regular) acoustic components. The emphasis, however, lies on the Markov chain analogue to the drill pipe, which extends the frequency domain results (Drumheller, 1989; Barnes and Kirkwood, 1972) to the time domain.

I. MODELING ASSUMPTIONS AND THEIR MOTIVATION

In this section, we introduce and motivate a number of modeling assumptions to obtain a mathematical description of the drill pipe.

We assume the whole configuration to be ''ideal'' in the sense that all pipe segments and tool-joints are exactly cylin-

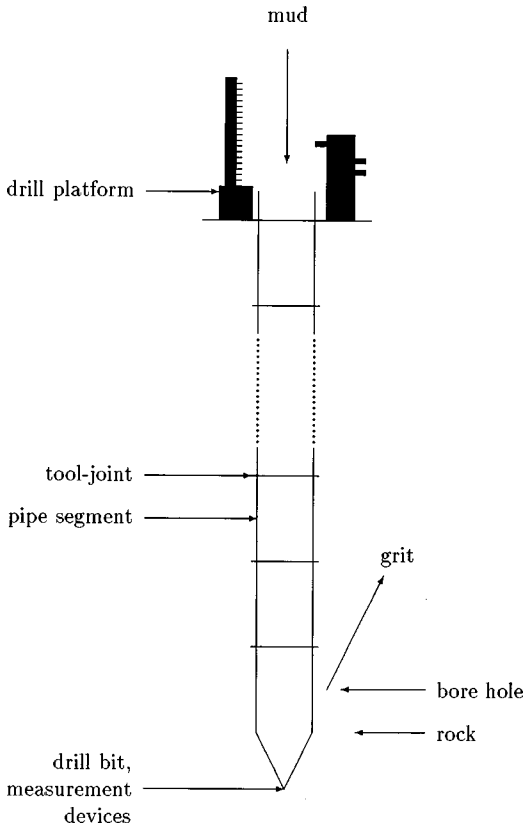


FIG. 1. Drill pipe.

drical, with a homogeneous wall of constant thickness, and that the screw thread of one tool-joint fits precisely into that of the next.

When the drill pipe is excited at its lower end, longitudinal, transversal, and spiraling sound waves result. The latter two cause wall displacements perpendicular to the pipe axis, and thus interact with the environment. As a consequence, they lose part of their energy by radiation. Furthermore, for the low frequencies considered (see below), any spiraling modes are cut off and will eventually vanish. Longitudinal waves, corresponding to wall displacements parallel to the pipe axis, are therefore likely to prevail in long drill pipes. In the sequel, we restrict ourselves to longitudinal sound waves. The cylindrical symmetry of the configuration implies that a one-dimensional mathematical description suffices.

We choose the co-ordinate system such that the z axis corresponds to the drill pipe axis (see Fig. 2). In drill pipe

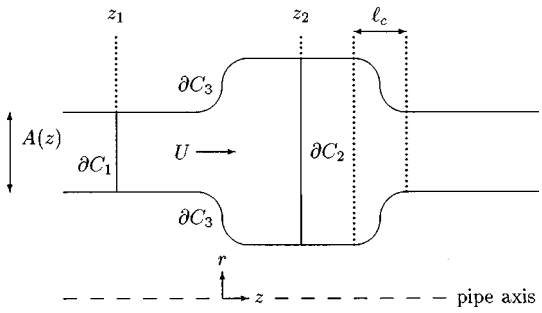


FIG. 2. Drill pipe wall of variable surface.

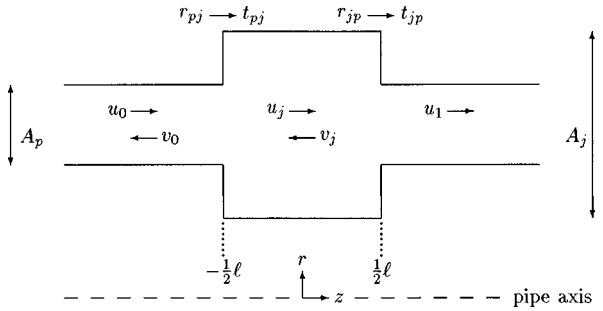


FIG. 3. Tool-joint modeled for large wave lengths.

parts where the configuration wall surface is constant, a combination of Newton's and Hooke's laws (Timoshenko *et al.*, 1974) yields an ordinary one-dimensional wave equation for the longitudinal displacement $U(z,t)$,

$$\frac{\partial^2 U}{\partial t^2} - c^2 \frac{\partial^2 U}{\partial z^2} = 0. \tag{1}$$

Its harmonic solutions are combinations of sound waves traveling to the left and to the right,

$$U(z,t) = \{u e^{i\kappa z} + v e^{-i\kappa z}\} e^{-i\omega t}, \tag{2}$$

in which $\kappa = \omega/c$ is the ratio of angular frequency and speed of sound in steel.

To simplify the model, we neglect the (cutoff) higher-order modes, and assume the plane waves (2) to be stepwise discontinuously connected at the joints. This assumption is reasonable for frequencies at which the wavelength λ is sufficiently larger than the drill pipe wall thicknesses h_p in the pipe segments, and h_j in the tool-joints. The length ℓ_c of the curved tool-joint part should typically not be much larger than h_p or h_j. Furthermore, to avoid tunneling effects, both the pipe segment length L and the tool-joint length ℓ should exceed the geometrical parameters h_p, h_j, and ℓ_c at the curved tool-joint parts. For the considered drill pipe, $h_p = 1$ cm, $h_j = 3.5$ cm, $\ell_c = 5$ cm, $L = 9.1$ m, and $\ell = 0.5$ m. The model is therefore restricted to frequencies satisfying $f \ll c/\ell_c \approx 100$ kHz for $c = 5100$ m/s in steel. Results are presented for frequencies up to 10 kHz.

We thus have reduced the real configuration to a series of straight segments and tool-joints in which only longitudinal, harmonic sound waves propagate. The drill pipe is driven at the lower end by a transducer; the upper end is assumed to reflect all sound waves completely, which follows from the observation that the normal force at that end equals zero, and Hooke's law.

We proceed by capturing tool-joints, transducer, and pipe segments in mathematical expressions. Their combination will yield the mathematical model.

II. FREQUENCY DOMAIN MODELING

In the sequel, we consider a drill pipe configuration consisting of $N+1$ pipe segments, numbered $n = 0,...,N$, and N tool-joints, numbered $n = 1,...,N$ (see Fig. 4). The outer ends may be counted as ''tool-joints'' 0 and $N+1$. The length of the nth pipe segments is L_n.

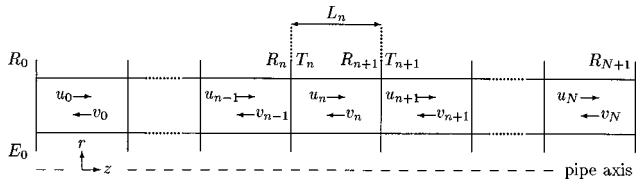


FIG. 4. Drill pipe modeled using reflection and transmission coefficients.

A. A tool-joint, transducer, and pipe segment model

We first describe a tool-joint in terms of reflection and transmission coefficients $R=v_0/u_0$ and $T=u_1/u_0$ (see Fig. 3) (Dowling and Ffowcs Williams, 1983, pp. 66–67).

We choose the origin of the coordinate system to correspond with the middle of a tool-joint with length ℓ. Continuity of displacement and normal force implies that at $z=-\frac{1}{2}\ell$, $U_0=U_j$, and $A_p\partial U_0/\partial z=A_j\partial U_j/\partial z$; A_p and A_j denote the wall surface of pipe segment and tool-joint, respectively. Similar equations result at $z=\frac{1}{2}\ell$. Substitution of (2)—with the correct wave amplitudes—into these expressions yields

$$R=2i\,\frac{r_{pj}\,\sin(\kappa\ell)}{1-r_{jp}^2\vartheta^4} \quad \text{and} \quad T=\frac{t_{pj}t_{jp}}{1-r_{jp}^2\vartheta^4}, \tag{3}$$

where $\vartheta=e^{i\kappa\ell/2}$. The coefficients $r_{jp}=(A_j-A_p)/(A_j+A_p)$, $t_{jp}=1-r_{jp}$, $r_{pj}=-r_{jp}$, and $t_{pj}=1-r_{pj}$ describe sound reflection at each of the wall area discontinuities[1] (Dowling and Ffowcs Williams, 1983, p. 64; Fletcher and Rossing, 1991, pp. 147–148).

Denoting the displacement amplitude in the pipe segment to which the transducer is connected by U_0, and assuming U_0 to satisfy (2), a mathematical model of a transducer (van de Ven, 1975) yields

$$u_0=\phi_0E_0+\phi_0^2R_0v_0, \tag{4}$$

where R_0 and E_0 are two transducer-dependent constants, and $\phi_0=e^{i\kappa L_0/2}$ is a phase change. The factor R_0 is similar to a reflection coefficient, but the transducer also keeps feeding energy into the drill pipe, which is described by the factor E_0.

If we substitute $z+\Delta z$ for z in Eq. (2), it appears that for sound waves that travel a distance Δz, only the phase changes by a factor $e^{i\kappa\Delta z}$. As a consequence, the pipe segments can be modeled as phase shifts.[2] We introduce the factor $\phi_n=e^{i\kappa L_n/2}$, describing the phase change of a sound wave traveling through half of the pipe segment n, i.e., over a distance $\frac{1}{2}L_n$.

B. A segmented drill pipe model

We denote the amplitude of the sound wave traveling to the right in the middle of pipe segment n by u_n; the wave moving to the left at that location has amplitude v_n. The reflection and transmission coefficients at tool-joint n are denoted R_n and T_n. We choose $z=0$ in the middle of the pipe segment connected to the transducer. We find

$$u_n=\phi_{n-1}T_n\phi_nu_{n-1}+\phi_nR_n\phi_nv_n, \quad n=1,...,N,$$

$$v_n=\phi_{n+1}T_{n+1}\phi_nv_{n+1}+\phi_nR_{n+1}\phi_nu_n, \quad n=0,...,N-1,$$

or, after rewriting in terms of matrices and applying induction,

$$\begin{pmatrix}u_N\\v_N\end{pmatrix}=\left(\prod_{n=1}^{N}(\Phi_nM_n\Phi_{n-1})\right)\begin{pmatrix}u_0\\v_0\end{pmatrix}=:\mathcal{T}\begin{pmatrix}u_0\\v_0\end{pmatrix}, \tag{5}$$

with

$$\Phi_n=\begin{pmatrix}\phi_n & 0\\ 0 & \phi_n^{-1}\end{pmatrix} \quad \text{for } n=0,...,N, \tag{6}$$

and

$$M_n=\begin{pmatrix}(T_n^2-R_n^2)/T_n & R_n/T_n\\ -R_n/T_n & 1/T_n\end{pmatrix} \quad \text{for } n=1,...,N. \tag{7}$$

The phase-shift matrices Φ_n and Φ_{n-1} describe the propagation of sound waves through (half) pipe segments; M_n models sound reflection and transmission at the joint between segments $n-1$ and n. $\mathcal{T}$ describes the behavior of sound waves traveling from segment 0 to segment N, and $\mathcal{T}^{-1}$ how sound waves propagate in the opposite direction. Symmetry thus implies that $\det(\mathcal{T})=\det(\mathcal{T}^{-1})$, and therefore $\det(\mathcal{T})=1$.

Assuming all pipe segments to have the same length L, and all tool-joints to have identical properties ℓ, ϕ, R, T, and M, we have $\mathcal{T}=(\Phi M\Phi)^N=:\mathcal{M}^N$. The powers of $\mathcal{M}$ can be determined explicitly, allowing explicit analytical results for the case $N\to\infty$.

The Cayley–Hamilton theorem (Tropper, 1969, pp. 72–73) states that each square matrix satisfies its characteristic equation, so that $\mathcal{M}^2-2\tau\mathcal{M}+\mathcal{I}=0$, where

$$\tau=\tfrac{1}{2}\,\mathrm{tr}(\mathcal{M})=\tfrac{1}{2}(\mathcal{M}_{11}+\mathcal{M}_{22}),$$

and we used the fact that the determinant of $\mathcal{M}$ is given by $\det\mathcal{M})=\mathcal{M}_{11}\mathcal{M}_{22}-\mathcal{M}_{12}\mathcal{M}_{21}=\det(\Phi)\det(M)\det(\Phi)=1\cdot1\cdot1=1$. The characteristic equation implies that $\mathcal{M}^n$ must be a linear combination of $\mathcal{M}$ and the unity matrix $\mathcal{I}$, i.e., $\mathcal{M}^n=a_n\mathcal{M}+b_n\mathcal{I}$. Similarly, the eigenvalues $\lambda_{1,2}$ of $\mathcal{M}$ satisfy its characteristic equation, and therefore $\lambda_{1,2}^n=a_n\lambda_{1,2}+b_n$. Some straightforward computations give

$$\lambda_{1,2}=\tau\pm\sqrt{\tau^2-1}, \quad \text{while } \lambda_1\lambda_2=1.$$

Combination of the latter two expressions yields two equations for a_n and b_n, from which both coefficients can be determined (assuming $\tau^2\neq1$, so that $\lambda_1\neq\lambda_2$). This yields the following expressions for the coefficients of $\mathcal{M}^N$:

$$\mathcal{M}_{11}^N=\frac{1}{\lambda_1-\lambda_2}\{(\mathcal{M}_{11}-\lambda_2)\lambda_1^N-(\mathcal{M}_{11}-\lambda_1)\lambda_2^N\}, \tag{8}$$

$$\mathcal{M}_{12}^N=\frac{1}{\lambda_1-\lambda_2}\{\mathcal{M}_{12}\lambda_1^N-\mathcal{M}_{12}\lambda_2^N\}=-\mathcal{M}_{21}^N, \tag{9}$$

$$\mathcal{M}_{22}^N=\frac{1}{\lambda_1-\lambda_2}\{(\mathcal{M}_{22}-\lambda_2)\lambda_1^N-(\mathcal{M}_{22}-\lambda_1)\lambda_2^N\}. \tag{10}$$

If $\tau^2=1$, similar expressions for the components of $\mathcal{M}^N$ can be found by taking limits for τ tending to ±1 in the latter equations.

At the platform end of the drill pipe, we have

$$v_N = \phi_N^2 R_{N+1} u_N . \tag{11}$$

Since all pipe segments are assumed to have the same length, $\phi_N = \phi$. Furthermore, in all numerical calculations performed in the sequel, we set $R_{N+1} = 1$, which corresponds to total reflection at the platform end of the configuration.

We thus have derived the four linear equations (4), (5), and (11) for the unknown amplitudes u_0, v_0, u_N, and v_N. Some calculations yield the total amplitude $U_N = u_N + v_N$ of the sound wave in the middle of pipe segment N:

$$U_N = \frac{\phi E_0 (1+\phi^2 R_{N+1}) \mathcal{M}_{11}^N}{1-(\phi^2 R_0 \mathcal{M}_{11}^N + \mathcal{M}_{12}^N)(\phi^2 R_{N+1} \mathcal{M}_{11}^N + \mathcal{M}_{12}^N)} . \tag{12}$$

A similar expression[3] can be derived for segment $n=0$; note that the coefficients $\mathcal{M}_{11}^N$ and $\mathcal{M}_{12}^N$ are known explicitly.

To study the behavior of (12) for $N \to \infty$, we distinguish two cases.

If τ has nonzero imaginary part, or if τ is real with $|\tau| > 1$, one of the eigenvalues λ_1, λ_2 of $\mathcal{M}$ has norm exceeding one. The Nth powers of both eigenvalues occur in (8)–(10), and so $\mathcal{M}_{11}^N$ tends to infinity if N does. Division of the denominator and the numerator of (12) by $(\mathcal{M}_{11}^N)^2$ thus implies $\lim_{N\to\infty} U_N = 0$.

If τ is real, and $|\tau| \leqslant 1$, both λ_1 and λ_2 (and their powers) have modulus 1, and are located on the boundary of the complex unit circle. In this case, (12) shows that U_N does not tend to zero for $N \to \infty$. Combining Eq. (3) for the tool-joint reflection and transmission coefficients with (5), and using the definition of τ, the following resonance condition results:[4] only sound waves with frequencies for which

$$\mathrm{Im}(\tau) = 0$$

and

$$|\tau| = \left| \cos \kappa \ell \cos \kappa (L-\ell) - \frac{1}{2}\left(\frac{A_j}{A_p} + \frac{A_p}{A_j}\right) \times \sin \kappa \ell \sin \kappa (L-\ell) \right| \leqslant 1 \tag{13}$$

are able to travel through long drill pipes, all segments and joints of which are identical.

This resonance condition (13), which is also reported in Barnes and Kirkwood (1972) and Drumheller (1989), explains the bandpass filter behavior of the drill pipe visualized in Fig. 5. However, it does not by itself give any information about the fine structure of the frequency response of the pipe's pass bands, which have only been considered numerically by Drumheller (1989) and Barnes and Kirkwood (1972). This fine structure is more apparent in Fig. 6, in which only one pass band is displayed.

Equation (12) gives an analytical expression for this spectrum structure, and allows energy dissipation and pipe segment length variations to be included in the model.

C. Energy dissipation

We model energy dissipation in the drill pipe wall by introducing a linear, dissipative term in the ordinary wave equation:

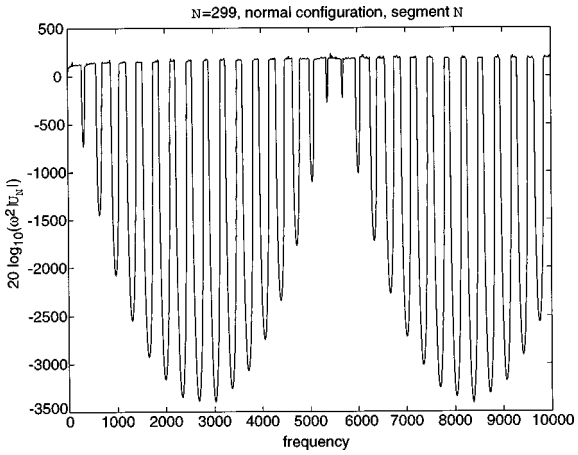


FIG. 5. $20 \log_{10} (\omega^2 |U_N|)$ in the middle of the last segment of a 300-segment drill pipe ($N=299$).

$$\frac{\partial^2 U}{\partial t^2} - c^2 \frac{\partial^2 U}{\partial z^2} = -\beta \frac{\partial U}{\partial t} .$$

Restriction to harmonic solutions as in (2) then gives[5] $(-\omega^2 + c^2 \kappa_e^2 - i\beta\omega) = 0$, or $\kappa_e^2 = \kappa^2 (1 + i\beta/\omega)$. For small dissipative terms β, we may write $\kappa_e = \pm(\kappa + i\alpha)$, with

$$\alpha = \frac{\beta/c}{1+\sqrt{1+i\beta/\omega}} \approx \frac{\beta}{2c} .$$

Equation (2) can be rewritten as

$$U(z,t) = (u e^{-\alpha z} e^{i\kappa z} + v e^{\alpha z} e^{-i\kappa z}) e^{-i\omega t},$$

i.e., the sound wave amplitude of sound waves traveling to the left or to the right decreases exponentially with the distance traveled.

Figure 7 shows the consequences of energy dissipation for sound transmission through the drill pipe: the resonance peaks vanish. We have chosen an amplitude attenuation of 10 dB/km, corresponding with $\alpha = 10^{-3}/\log_{10} e$. The configuration consists of 20 pipe segments, connected by $N=19$ tool-joints. Comparing with Fig. 6, we see that in a configu-

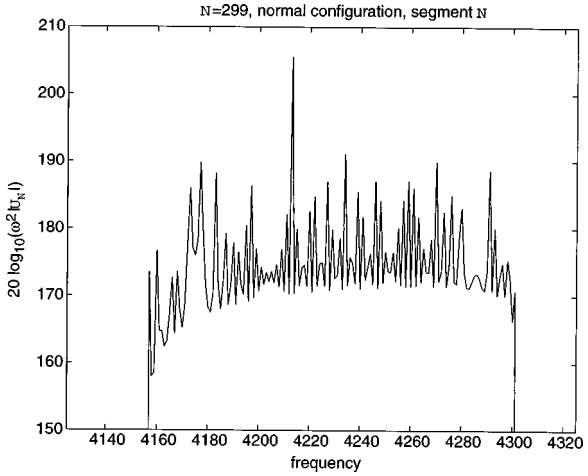


FIG. 6. $20 \log_{10} (\omega^2 |U_N|)$ in the middle of the last segment of a 300-segment drill pipe ($N=299$).

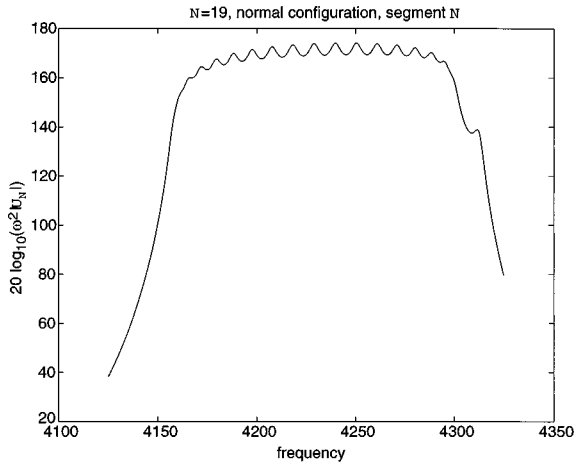


FIG. 7. $20 \log_{10} (\omega^2 |U_N|)$ in the middle of the last segment of a 20-segment drill pipe ($N=19$), with energy dissipation (10 dB/km).

ration with only 19 tool-joints the diffusion has removed the sharp peaks almost completely. This effect becomes even more pronounced for larger values of N.

An alternative way to model the effect of energy losses in the drill pipe is to assume them to occur at the tool-joints only. Instead of introducing a complex wave number κ_e, it then suffices to slightly change the tool-joint reflection and transmission coefficients R and T, such that the determinant Δ of $\mathcal{M}$ becomes smaller than one. Expression (12) remains valid if the second term in its numerator is divided by Δ^N. Furthermore, the eigenvalues λ_1 and λ_2 have to be changed to $\lambda_{1,2}=\tau \pm \sqrt{\tau^2-\Delta}$ and the conditions $\tau^2=1$ and $\tau^2 \neq 1$ have to be modified to $\tau^2=\Delta$ and $\tau^2 \neq \Delta$ (similarly for inequalities).

D. Pipe segment length variations

To roughly estimate the effect of small pipe segment length variations on the amplitude U_N in the rightmost drill pipe segment, the pipe segment lengths L_n occurring in the phase-shift matrices Φ_n [see (5)] have been disturbed with a normally distributed error l_n with mean zero and standard deviation σ_p.

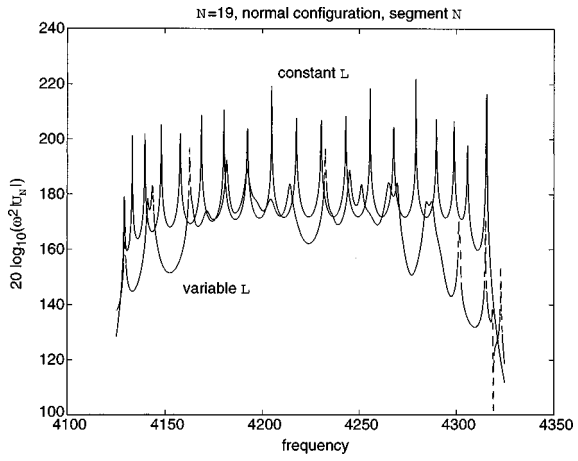


FIG. 8. $20 \log_{10} (\omega^2 |U_N|)$ in the middle of the last segment of a 20-segment drill pipe ($N=19$), with variable pipe segment length ($\sigma_p=0.05$).

The results for a drill pipe with 20 pipe segments and $N=19$ tool-joints are shown in Fig. 8; we have chosen $\sigma_p=0.05$. The peaks within the pass bands are less sharp compared to those in Fig. 6 (notice that the vertical scales differ in both figures). The stop bands appear to be less pronounced. In fact, the conditions for resonance are never fulfilled exactly if L varies, and neither are the conditions that make different parts of the drill pipe move in exact counterphase.

Small pipe segment length variations appear to affect mainly the drill pipe's sound transmission properties at higher frequencies. The location of the resonance peaks is slightly affected as well.

III. TIME DOMAIN MODELING

A. Additional simplifications and their motivation

As we want to study sound transmission through the drill pipe in the time domain, we must switch from a stationary situation with harmonic sound waves to a dynamical one, where sound signals are used of which the time duration is limited. We assume them to be so short that they do not interact, i.e., their duration is short compared to the time sound needs to propagate through a pipe segment.[6] In the sequel, we refer to such short time-duration signals as ''pulses'' $\hat{u}(t)$.

In order to make sure that the reflection and transmission of a pulse remains pulse shaped, we have to make restrictions on the reflection coefficient $R(\omega)$ at the tool-joints. Assume that an incident signal $u(\omega)$ yields a reflection $v(\omega)=R(\omega)u(\omega)$. After inverse Fourier transformation, this gives the time-domain expression

$$\hat{v}(t)=\int_{-\infty}^{\infty} R(\omega)u(\omega)e^{-i\omega t}\,d\omega$$

$$=\frac{1}{2\pi}\int_{-\infty}^{t} \hat{R}(t-t')\hat{u}(t')dt',$$

in which we used causality of $\hat{R}(t)$, the inverse Fourier transform of $R(\omega)$ (Rienstra, 1988). If $\hat{u}(t)=\delta(t)$, this reduces to $\hat{v}(t)=(1/2\pi)\hat{R}(t)$, which is in general not pulselike because $R(\omega)$ is in general not a constant or of the form $\sim e^{i\omega s}$. In fact, substitution of Eq. (3) for $R(\omega)$ yields

$$\hat{R}(t)=-2\pi r_{jp}\delta\left(t+\frac{l}{c}\right)+2\pi(1-r_{jp}^2)$$

$$\times\sum_{n=1}^{\infty} r_{jp}^{2n-1}\delta\left(t-(2n-1)\frac{l}{c}\right),$$

i.e., the reflection of a pulse is a pulse train. This calls for further modeling assumptions.

$\hat{R}(t)$ would be pulselike if $(1-r_{jp}^2)r_{jp} \ll r_{jp}$, or $r_{jp} \approx 1$, an uninteresting trivial case. Alternatively, $\hat{R}(t)$ can be considered pulselike if its energy is concentrated in the first N_0 pulses of the pulse train, such that $(1-r_{jp}^2)r_{jp}^{2N_0-2}$ is small enough, and the effective duration $N_0 \ell/c$ of $\hat{R}(t)$ and the corresponding length of the reflected signal $N_0 \ell$ are short enough.

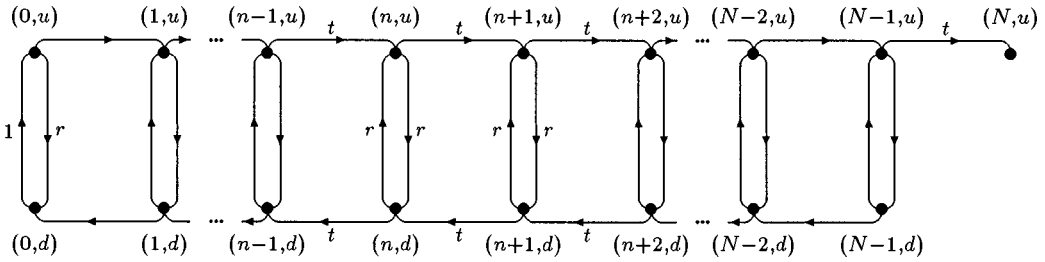


FIG. 9. Drill pipe modeled as a Markov chain.

Implementing these considerations, we assume $R(\omega)$ to be constant in the sequel, and neglect the tool-joint length.

The bandwidth of the sound signals used for data transmission will not exceed the width of the drill pipe's pass bands. According to Eq. (3), the values of R and T generally vary only slightly in such limited spectral intervals. Within the signal's frequency band, it therefore seems realistic indeed to approximate R by a constant.

Furthermore, since $u(\omega)=0$ outside this band, the reflected signal $R(\omega)u(\omega)$ does not change if we assume R to have the same constant value for all ω. By choosing a real value for the reflection coefficient, we neglect phase jumps at reflection.

A similar argument holds for the transmission coefficient T and, at the transducer end, for R_0. For the sake of demonstration we set $R_0=1$.

Clearly, with respect to the original equation (1), the results presented in the sequel will have to be interpreted as approximations.

B. A Markov model for the drill pipe

Consider an energy pulse of amplitude 1 traveling toward a tool-joint of which the constant reflection and transmission coefficients with respect to energy[7] are denoted by r and t. At the tool-joint, part of the pulse reflects and part of it is transmitted. Conservation of energy implies

$$1=r+t.$$

This observation leads to a statistical analogue for (sound) energy pulse propagation through the drill pipe.

We model energy pulses as marbles, rolling from left to right or vice versa within the pipe segments. At each tool-joint, they change direction with probability r. With probability t, they keep rolling in the same direction.

The drill pipe can now be modeled as a Markov chain, introducing two states for each pipe segment.[8] A ''marble'' is said to be in state (n,u) if it rolls upward in segment n. If it rolls downward, it is in state (n,d). The Markov chain is shown in Fig. 9.

Consider one ''marble'' which starts to roll upward in the first pipe segment at clock unit $k=0$. Let $\pi_k(n,j)$ be the probability that at clock unit k it moves in direction $j\in\{u,d\}$ in pipe segment n. Hence, $\pi_0(0,u)=1$ and for $k=1,2,...$ the probabilities $\pi_k(n,j)$ satisfy (see Fig. 9)

$$\pi_k(0,u)=\pi_{k-1}(0,d), \tag{14}$$

$$\pi_k(n,u)=t\,\pi_{k-1}(n-1,u)+r\,\pi_{k-1}(n,d), \quad n=1,...,N, \tag{15}$$

$$\pi_k(n,d)=t\,\pi_{k-1}(n+1,d)+r\,\pi_{k-1}(n,u), \quad n=0,...,N-1, \tag{16}$$

where $\pi_k(N,d)=0$ by definition. From these equations we can easily compute for consecutive values of k the probability distribution $\{\pi_k(n,j)\}$. Figure 10 displays this distribution for some values of k in a configuration with $N=15$ tool-joints (and 16 pipe segments), with $r=0.1$ and $t=0.9$. The value of $\pi_k(N,u)$ for consecutive values of k is shown in Fig. 11; it shows how much energy reaches the upper drill pipe end at each clock unit. We have chosen the same reflection and transmission coefficients, but $N=63$.

In terms of the physical problem, two effects influence the energy pulse propagation. The first will be called **direct transmission**: part of the pulse travels through the drill pipe without reflecting anywhere. This explains the rightmost peak in Fig. 10, for $k=4$ and $k=12$, and the small peak in Fig. 11 at time unit $k=63$. The second effect is similar to **diffusion**. Energy that reflects at a number of tool-joints appears to be distributed more or less regularly throughout the configuration, and diffuses to the upper end of the configuration. This is visible in Fig. 10, for $k=20$ and $k=28$, and explains the ''hump'' in Fig. 11 following the first peak at $k=63$.

The effect of decreasing t is similar to that of increasing N, and vice versa. For $t=1$, the ''hump'' disappears: all energy travels through the drill pipe at once. For small values of t, the diffusion effect is large compared to the direct trans-

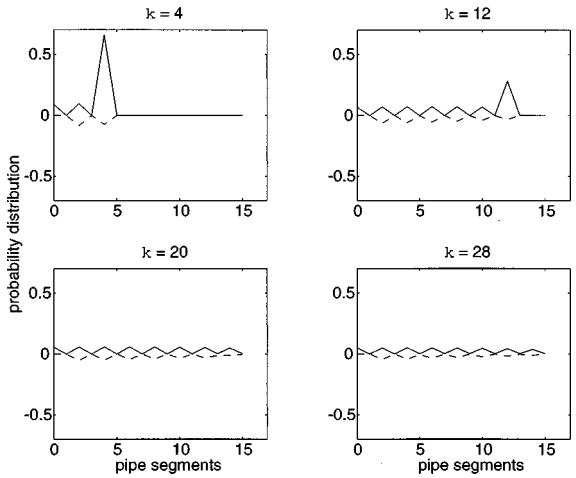


FIG. 10. Probability distribution of the energy pulse (''marble'') position after k time units; $N=15$, $r=0.1$, $t=0.9$ (downward probabilities with minus sign).

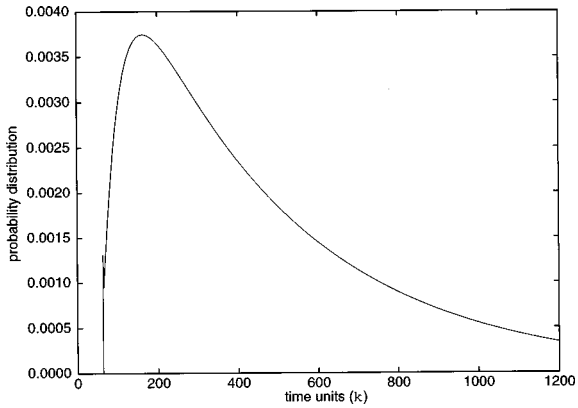


FIG. 11. Probability distribution of the energy pulse (''marble'') arrival after k time units; $N=63$, $r=0.1$, $t=0.9$.

mission, and only the ''hump'' reaches the upper drill pipe end.

In the following paragraphs, we show that the ''hump'' is a mixture of N geometric functions. This observation will allow exact determination of the expected traveling time of the energy pulse through an $N+1$-segment drill pipe. It will also yield an expression for the variance of this traveling time, i.e., an estimate for the width of the ''hump.'' We will finally show that two geometric functions suffice to closely approximate the shape of the ''hump.''

C. The ''hump'' as a mixture of geometric functions

We introduce the (probability) generating functions

$$F(n,j;z)=\sum_{k=0}^{\infty}\pi_k(n,j)z^k; \tag{17}$$

$F(N,u;z)$ is the generating function of the ''hump.'' We now show how to determine $F(N,u;z)$ explicitly. This will lead to a mathematical description of the ''hump'' in terms of geometric functions.

Equations (14)–(16) combined with the initial condition $\pi_0(0,u)=1$ yield

$$F(0,u;z)=zF(0,d;z)+1, \tag{18}$$

$$F(n,u;z)=tzF(n-1,u;z)+rzF(n,d;z), \quad n=1,...,N, \tag{19}$$

$$F(n,d;z)=tzF(n+1,d;z)+rzF(n,u;z), \quad n=0,...,N-1. \tag{20}$$

We rewrite expressions (19) and (20) to

$$\begin{pmatrix}F(n,u;z)\\F(n,d;z)\end{pmatrix}=\mathcal{W}(z)\begin{pmatrix}F(n-1,u;z)\\F(n-1,d;z)\end{pmatrix}, \quad n=1,...,N, \tag{21}$$

where

$$\mathcal{W}(z)=\begin{pmatrix}(t^2-r^2)z/t & r/t\\ -r/t & 1/tz\end{pmatrix}. \tag{22}$$

Applying induction to (21), we obtain

$$\begin{pmatrix}F(N,u;z)\\F(N,d;z)\end{pmatrix}=\mathcal{W}^N(z)\begin{pmatrix}F(1,u;z)\\F(1,d;z)\end{pmatrix}.$$

We combine this result with (18), and substitute $F(N,d;z)=0$. This leads to

$$F(N,u;z)=\frac{\det(\mathcal{W}^N(z))}{\mathcal{W}_{22}^N(z)+z\mathcal{W}_{21}^N(z)}=\frac{\det(\mathcal{W}(z))^N}{\mathcal{W}_{22}^N(z)+z\mathcal{W}_{21}^N(z)},$$

in which $\mathcal{W}_{ij}^N(z)$ denotes the ij-component of $\mathcal{W}^N(z)$ for $i,j=1,2$. Since $\det(\mathcal{W}(z))=1$ we eventually find

$$F(N,u;z)=\frac{1}{\mathcal{W}_{22}^N(z)+z\mathcal{W}_{21}^N(z)}=:\frac{1}{f(z)}. \tag{23}$$

As the components of $\mathcal{W}(z)$ are either constants, or multiples of z or $1/z$, it is readily verified that the function $f(z)$ may be written as

$$f(z)=\sum_{k=-N}^{N}f_kz^k,$$

hence $z^Nf(z)$ is a polynomial of degree $2N$.

With Eqs. (17) and (23) we have two different expressions for $F(N,u;z)$, one in terms of the probabilities $\pi_k(N,u)$, the other as the reciprocal of $f(z)$. We now express the probabilities $\pi_k(N,u)$ in terms of the zeros of $z^Nf(z)$. These expressions give rise to the mathematical description of the ''hump'' in Fig. 11 as a mixture of N geometric functions.

To compare (17) and (23), we first exploit some properties of $F(N,u;z)$. Note that it takes at least N clock units for a ''marble'' to reach pipe segment N and each segment passed in the downward direction must be passed again in the upward direction to ever reach the end of the drill pipe. This implies that $\pi_k(N,u)=0$ for $k<N$ and $\pi_{N+1+2k}(N,u)=0$ for $k\geqslant 0$. Hence

$$F(N,u;z)=z^N\sum_{k=0}^{\infty}\pi_{N+2k}(N,u)z^{2k},$$

from which we can conclude that $F(N,u;z)/z^N$ is even, and convergent for $|z|\leqslant 1$; it has no poles within the closed unit disk. Thus also $z^Nf(z)$ is even. So each zero of $z^Nf(z)$ has a companion zero with opposite sign. Furthermore, all zeros of $z^Nf(z)$ must be situated outside the closed unit disk. Denote the zeros of $z^Nf(z)$ by $z_1,-z_1,...,z_N,-z_N$, which we assume to be distinct. We may then write

$$z^Nf(z)=f_{-N}(1-(z/z_1)^2)\cdots(1-(z/z_N)^2).$$

Partial fraction decomposition of $F(N,u;z)$ now yields

$$F(N,u;z)=\frac{z^N}{z^Nf(z)}=\frac{z^N}{f_{-N}}\sum_{i=1}^{N}\frac{a_i}{1-(z/z_i)^2}=z^N\sum_{k=0}^{\infty}z^{2k}\sum_{i=1}^{N}\frac{a_i}{f_{-N}}\left(\frac{1}{z_i}\right)^{2k}, \tag{24}$$

with $a_i=1/\Pi_{j=1,j\neq i}^{N}(1-z_i/z_j)$. From (17) and (24) we finally get

$$\pi_{N+2k}(N,u)=\sum_{i=1}^{N}b_i(1-p_i)p_i^k, \quad k=0,1,2,... . \tag{25}$$

To ease the interpretation of (25), we introduced

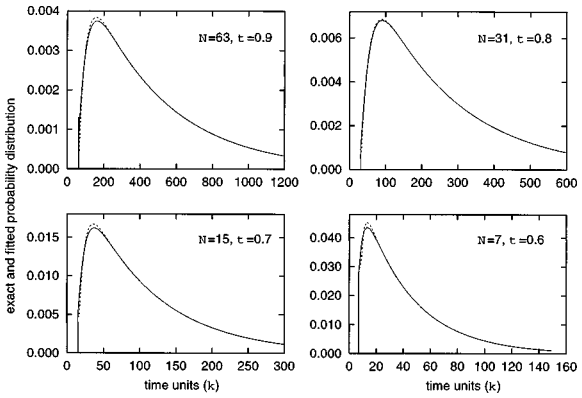


FIG. 12. Expected duration of the energy pulse (''marble's'') trip through the drill pipe, and its standard deviation as a function of N for various values of t.

$$p_i = 1/z_i^2, \quad b_i = \frac{a_i}{f_{-N}(1-p_i)}, \quad i = 1,...,N.$$

It follows that the probabilities $\pi_k(N,u)$ are mixtures of N geometric distributions for $k=N,\ N+2,...$; the ''hump'' in Fig. 11 is a mixture of N geometric functions.

D. Expectation and variance of an energy pulse's traveling time

Expression (25) is not useful for practical purposes, since it requires the determination of all zeros of the polynomial $z^N f(z)$. An approximation that appears to be remarkably accurate is obtained by constructing a suitable distribution that fits the mean μ_t and variance σ_t^2 of the time it takes the ''marble'' to reach the right end of the drill pipe. The distribution that will be used is a discrete version of the continuous K_2-distribution (Tijms, 1986). Such an approximation, of course, only makes sense if μ_t and σ_t^2 can be easily determined. This is indeed the case. Below we will derive explicit expressions for μ_t and σ_t^2, by utilizing again the generating function $F(N,u;z)$.

The mean and variance of the traveling time can be found by differentiating the generating function $F(N,u;z)$, yielding[9]

$$\mu_t = \sum_{k=1}^{\infty} k\pi_k(N,u) = F'(N,u;1) = -\frac{f'(1)}{f^2(1)} = -f'(1),$$

where the prime denotes the derivative with respect to z. For the variance of the traveling time, a similar derivation yields

$$\sigma_t^2 = f'(1)^2 - f''(1) - f'(1).$$

To determine $f'(1)$ and $f''(1)$ note that the format of $\mathcal{W}(z)$ in (22) is extremely similar to that of $\mathcal{M}$ in the frequency domain analysis. The components of $\mathcal{W}^N(z)$ are given by (8)–(10) after proper substitution of the components of $\mathcal{W}(z)$ for the ones of $\mathcal{M}$. So $f(z)$, and thus also $f'(1)$ and $f''(1)$, can be determined explicitly! After some algebra we find

$$\mu_t = \frac{1}{t} N(1+rN),$$

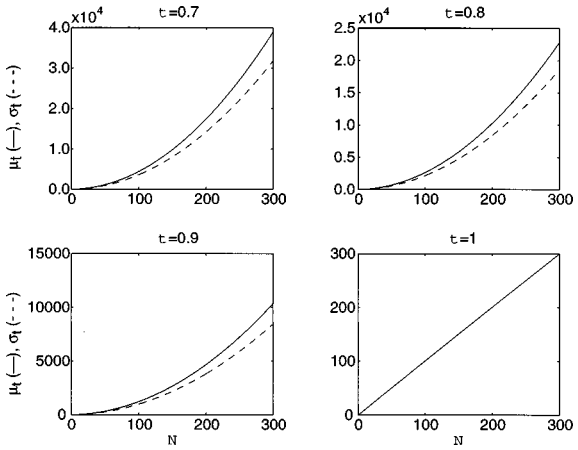


FIG. 13. Fitting of the probability distribution of the energy pulse (''marble'') arrival after k time units for various values of N and t; exact: —, fitted: ---.

$$\sigma_t^2 = \frac{2r}{3t^2} N(N+1)(1+(1+t)N+rN^2).$$

As could be expected for a diffusionlike process, the time energy pulses need to travel through the drill pipe increases quadratically with N, and so does its standard deviation. Figure 12 displays the results for different transmission coefficients t.

E. Approximation of the ''hump'' using two geometric functions only

We saw that the ''hump,'' following the ''direct-transmission-peak'' at $k=N$, is a mixture of N geometric functions. We now show that, by appropriate parameter choice, only two geometric functions suffice to approximate its shape.

In analogy to Eq. (25), we write

$$\pi_{N+2k}(N,u) = c_1(1-q_1)q_1^k + c_2(1-q_2)q_2^k, \quad k=1,2,... \tag{26}$$

thus replacing the N known geometric distributions by two distributions, the parameters of which have to be fitted to the mean and variance of the traveling time of a pulse. Note that the ''direct-transmission-peak'' is given by

$$\pi_N(N,u) = t^N. \tag{27}$$

The probabilities q_1 and q_2 and the coefficients c_1 and c_2 in (26) are chosen to satisfy

$$q_1 = \frac{a-(1+1/\hat{\mu})+\sqrt{2a-1}/\hat{\mu}}{a-1/\hat{\mu}^2-(1+1/\hat{\mu})^2},$$

$$q_2 = \frac{\hat{\mu}-q_1/(1-q_1)}{1+\hat{\mu}-q_1/(1-q_1)},$$

$$q_1 c_1 = (1-t^N) - q_2 c_2 = (1-t^N)\frac{q_1/(1-q_1)}{q_1/(1-q_1)-q_2/(1-q_2)},$$

where[10]

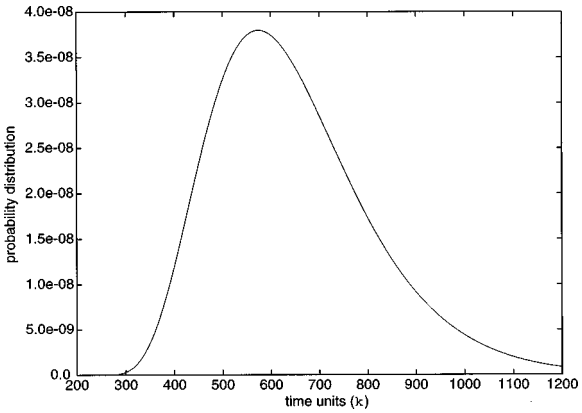


FIG. 14. Probability distribution of the energy pulse (''marble'') arrival after k time units, with energy losses at tool-joints; $N=255$, $r=0.1$, $t=0.9$, $d=0.01$.

$$a=\frac{\hat{\sigma}^2}{\hat{\mu}^2}-\frac{1}{\hat{\mu}},\quad \hat{\mu}=\frac{1}{2}\left(\frac{\mu_t-Nt^N}{1-t^N}-(N+2)\right),$$

$$\hat{\sigma}^2=\frac{1}{4}\left(\frac{\sigma_t^2}{1-t^N}-\frac{t^N}{(1-t^N)^2}(\mu_t-N)^2\right).$$

The resulting distribution, a discrete version of the continuous K_2-distribution (see Tijms, 1986, pp. 399–400), fits the mean and variance of the traveling time of a pulse, provided $\frac{1}{2}<a<1$. The fitting result is shown in Fig. 13. Apparently, only two geometrical distributions suffice to provide an excellent description of the ''hump.''

The mentioned condition on a holds when N is sufficiently large.[11] In case it does not hold, we refer to the recipe described by Adan *et al.* (1995) to find an appropriate discrete distribution to fit the ''hump.''

F. Energy dissipation

Energy dissipation in the drill pipe wall can be easily included in the model. Note that $\pi_k(N,u)$ indicates how much energy reaches the upper drill pipe end after k clock units in case there is no energy dissipation. If we assume a fraction d of the energy to dissipate in each clock cycle, the amount of energy reaching the drill pipe end after k clock units is $(1-d)^k\pi_k(N,u)$. This observation is used to obtain the results shown in Fig. 14. Only part of the signal energy reaches the upper drill pipe end. The ''hump'' therefore becomes less high, but sharper.

IV. CONCLUSIONS

In addition to the results found by Barnes and Kirkwood (1972) and Drumheller (1989), the frequency domain drill pipe model yields explicit expressions for the fine structure of the frequency response within the pass bands, of which the position along the frequency axis is found from a resonance condition as in the work of Barnes and Kirkwood (1972) and Drumheller (1989). After including energy dissipation in the model, the resonance peaks in the pass-bands become less pronounced. Small variations in pipe segment length have the same effect, but also slightly affect the position of these peaks on the frequency axis. The effect of length variations increases with frequency.

In the time domain, the drill pipe can be modeled as a Markov chain, from which expressions result for the expected time a sound signal needs to travel from one side to the other, and its standard deviation. Both increase quadratically with the number of pipe segments. Sound (pulse) propagation through drill pipe and tool-joints is comparable to a diffusion process. Energy pulses input at one end of the drill pipe are received at the other as energy ''humps'' that become flatter for decreasing transmission coefficient or for increasing drill pipe length. The ''humps'' appear to be mixtures of geometric functions. For practical purposes, their shape can be approximated well by two geometric functions only, i.e., by a discrete version of the K_2 distribution. Energy dissipation results in less flat, but smaller, output ''humps.''

ACKNOWLEDGMENTS

The research leading to this paper was done for, and partly funded by, Hollandse Signaalapparaten (HSA), Hengelo, The Netherlands, under supervision of Emiel Stolp, HSA. The derivation of the wave equation in the drill pipe wall, and the transducer model, are due to Fons van de Ven, TUE.

[1]For $|r_{jp}^2\phi^4|<1$, T can be written as $t_{pj}t_{jp}\{1+r_{jp}^2\phi^4+r_{jp}^4\phi^8+\cdots\}$. All sound waves that travel through the tool-joint are attenuated at both discontinuities, and therefore by $t_{pj}t_{jp}$. Waves that make a detour inside the tool-joint, and travel hence and forth once, have a resulting amplitude which is a factor r_{jp}^2 smaller, while their phase has changed by a factor ϕ^2. The remaining terms in the power expansion of T have a similar interpretation.

[2]The factors ϕ introduced to describe the tool-joints have a similar interpretation; we refer to footnote 1.

[3]The expression for U_N is similar to the one for T in (3). The numerators of both equations represent the effect on a sound wave of traveling through the whole configuration at once. The denominators represent the influence of sound waves that travel up and down the configurations a number of times before leaving them at the rightmost end.

[4]We assume $A_j/A_p+A_p/A_j$ to be smaller than 2. If it equals its upper bound 2, which is the case if $A_j=A_p$, then $|\tau|=|\cos\kappa L|\leqslant 1$ for all frequencies.

[5]Index e for energy.

[6]Given the pass-band width of some 250 Hz, and the pipe segment length of about 10 m, this assumption is not completely realistic. The Uncertainty Principle (Papoulis, 1984, p. 273) states that the time duration of sound signals with a bandwidth of 250 Hz is lower bounded by $\frac{1}{500}=0.002$ s; this is just equal to the time L/c sound needs to travel through a pipe segment.

[7]These are different from the amplitude reflection and transmission coefficients!

[8]We assumed complete reflection at the transducer end, and complete absorption at the upper end of the drill pipe.

[9]Note that $f(1)=\mathcal{W}_{22}^N(1)+\mathcal{W}_{21}^N(1)=1$. In terms of energy pulses, $f(1)=1$ corresponds to conservation of energy.

[10]$\hat{\mu}$ and $\hat{\sigma}^2$ are the mean and variance of the random variable $\frac{1}{2}(X-(N+2))$, where X is the traveling time a pulse needs to reach the upper end of the drill pipe, *given* that the traveling time is greater than or equal to $N+2$.

[11]Using the formulas for μ_t and σ_t^2 it is readily verified that $a\to\frac{2}{3}$ as $N\to\infty$. In most cases the condition $\frac{1}{2}<a<1$ appears to hold for small values of N already.

Adan, I. J. B. F., van Eenige, M. J. A., and Resing, J. A. C. (**1995**). ''Fitting discrete distributions on the first two moments,'' Prob. Eng. Inf. Sci. **9**, 623–632.

Barnes, T. G., and Kirkwood, B. R. (**1972**). ''Passbands for acoustic transmission in an idealized drill string,'' J. Acoust. Soc. Am. **51**, 1606–1608.

Dowling, A. P., and Ffowcs Williams, J. E. (**1983**) *Sound and Sources of Sound* (Wiley, New York).
Drumheller, D. S. (**1989**). ''Acoustical properties of drill strings,'' J. Acoust. Soc. Am. **85**, 1048–1064.
Fletcher, N. H., and Rossing, T. D. (**1991**). *The Physics of Musical Instruments* (Springer-Verlag, New York).
Papoulis, A. (**1984**). *Signal Analysis* (McGraw–Hill, London).
Rienstra, S. W. (**1988**). ''1-D Reflection at an Impedance Wall,'' J. Sound Vib. **125**, 43–51.
Tijms, H. C. (**1986**). *Stochastic Modeling and Analysis: A Computational Approach* (Wiley, New York).
Timoshenko, S., Young, D. H., and Weaver, W. (**1974**). *Vibration Problems in Engineering* (Wiley, New York).
Tropper, A. M. (**1969**). *Linear Algebra* (Nelson, London).
van de Ven, A. A. F. (**1975**). ''Interaction of electromagnetic and elastic fields in solids,'' Ph.D. thesis, Eindhoven University of Technology.

A finite element algorithm for the study of nonlinear standing waves

L. Elvira-Segura and E. Riera-Franco de Sarabia
Instituto de Acústica, C.S.I.C. Serrano 144, Madrid 28006, Spain

(Received 9 December 1996; accepted for publication 7 February 1998)

A finite element method (FEM) is presented for the study of nonlinear acoustic standing waves inside a chamber. The method is developed using the Galerkin–Bubnov weighted residual formulation and applied to solve the Lagrangian second-order wave equation, including thermal and viscous dissipation terms. A one-dimensional problem is studied in the frequency domain. Numerical data are compared with analytical results obtained by using a perturbation scheme for the cases of resonance and antiresonance in a rigid-walled tube with one-sided rigid or pressure release boundary condition. The FEM algorithm is shown to be well suited for the study of high-frequency standing wave fields in which the effect of absorption cannot be described by simple analytical expressions. © *1998 Acoustical Society of America.* [S0001-4966(98)05105-4]

PACS numbers: 43.25.Gf [MAB]

LIST OF SYMBOLS

a	Lagrangian spatial coordinate
c_0	velocity of sound
c_P, c_V	specific heats at constant pressure and constant volume
C	absorption matrix
dg_i, dh	nodal coefficients for the Fourier expansion terms of displacements
dg, **dh**	vectors of nodal coefficients for the Fourier expansion terms of displacements
e	Euler number (2.71828...)
fc_i, fs_i	nonlinear terms
FCc, **FCs**	boundary condition vectors
FNc, **FNs**	nonlinear term vectors
g_i, h_i	Fourier expansion coefficients
k_1	fundamental frequency wave number
K	stiffness matrix
l	tube length
M	mass matrix
p	acoustic pressure
p_ν, p_t	pressure terms due to viscous and thermal effects
P, P_0	absolute pressure, atmospheric pressure
P_t	tube perimeter
s, s_0	specific entropy, specific equilibrium entropy
S_t	tube section
t	time
α_1	$\omega\sigma_1/2c_0$
β	$(\gamma+1)/2$
γ	c_P/c_V
η	shear viscosity
κ	thermal conductivity
ξ	particle displacement
Ξ_0	excitation displacement amplitude
ρ, ρ_0	density, equilibrium density
σ	absorption coefficient
σ_i	absorption coefficient for the ith harmonic
φ^i	interpolating polynomials
ω	angular frequency
ω_i	$i\omega$

INTRODUCTION

A numerical algorithm based on the finite element method is proposed for the study of standing wave fields. This procedure avoids some of the inherent difficulties in calculating a solution for the nonlinear acoustic wave equation. The method is specially suited to solving confined wave problems.

The normal modes theory is the main analytical tool employed for the study of a standing wave field when the wavelength and the size of the chamber are of the same order. However, if the wavelength is small in comparison with the chamber dimensions, the ray theory or statistical approaches are generally used. Nevertheless, there are intermediate cases in which none of these methods is applicable, and analytical solutions are difficult to find. If the shape of the boundaries is not regular or the geometry of the chamber is not simple, analytical solutions are also difficult to find, although modal or ray theories could be applied. In these situations, numerical methods are becoming the best way to proceed, as it is impossible to find an explicit solution for the wave equation. Finite element methods (FEM) and boundary element methods (BEM) are probably the most important of the numerical methods, for solving partial differential equations with boundary conditions. Nowadays, commercial programs based on these two methods are being developed and used for acoustic applications.

However, when the acoustic wave intensity increases, nonlinear effects become important and a change in the shape of the excitation wave is noticed. This change is related to the increase of higher harmonic amplitudes. So, linear equations are no longer valid for describing the acoustic field. An accurate general solution for the nonlinear wave equation has not yet been found and different approaches are used depending on the particular problem to be solved. The

numerical procedures are also made more complicated by nonlinearities.

Different numerical algorithms have been proposed to solve the nonlinear wave equation using Fourier series. Van Buren[1] used iterative wave superposition to compute the standing wave field inside a tube. Tjøtta and Aanonsen,[2,3] developed a numerical algorithm based on a finite difference scheme to solve the nonlinear wave equation for progressive waves. Tsuchiya and Kagawa[4] developed a finite element method to solve the nonlinear Kuznetsov equation. This method has proved to be an excellent tool for the computation of nonlinear focusing fields, improving the solutions provided by analytical methods.

In this paper, a finite element method is proposed to solve the second-order nonlinear equation in Lagrangian coordinates with thermal and viscous dissipation effects taken into account.[5] The Lagrangian coordinates are suitable for furnishing the boundary conditions, and it is precisely for this reason that these coordinates are commonly chosen for nonlinear acoustics problems in the description of standing waves. A one-dimensional problem is solved and the results are compared with the analytical results. These analytical solutions are obtained by means of a perturbation scheme for the cases of resonance and antiresonance in a rigid-walled tube with rigid or pressure release boundary conditions at one end. A piston vibrating sinusoidally is assumed at the other end.

I. THEORY

The nonlinear equation for the acoustic field inside a tube can be derived from a set of five equations with five unknowns:

The continuity equation

$$\rho_0=\rho\left(1+\frac{\partial\xi}{\partial a}\right). \tag{1}$$

The force equation

$$\rho_0\ddot{\xi}=-p'-p'_\nu. \tag{2}$$

The state equation for an ideal gas

$$P=P_0\left(\frac{\rho}{\rho_0}\right)^\gamma e^{(s-s_0/c_V)}. \tag{3}$$

The viscous dissipation equation on the walls of the tube

$$p_\nu=\frac{P_t}{S_t}\sqrt{\frac{\eta}{2\omega\rho_0}}\left(\frac{1}{\omega}\frac{\partial p}{\partial t}-p\right). \tag{4}$$

The thermal dissipation equation on the walls of the tube

$$\begin{aligned}p_t&=c_0^2\frac{\rho_0}{\gamma c_V}(s-s_0)\\&=\frac{P_t}{S_t}\sqrt{\frac{\kappa}{2\omega\rho_0 c_P}}(\gamma-1)\left(\frac{1}{\omega}\frac{\partial p}{\partial t}-p\right).\end{aligned} \tag{5}$$

It is assumed in Eqs. (4) and (5) that dissipative processes are weak. These equations, derived by Weston,[6] are based on Kirchhoff–Rayleigh's theory for wall absorption in a tube.

For second-order pressure changes, the following expansion can be substituted for Eq. (3):

$$\begin{aligned}p=P-P_0\approx c_0^2(\rho-\rho_0)&\\+c_0^2\frac{\gamma-1}{2\rho_0}(\rho-\rho_0)^2+c_0^2\frac{\rho_0}{\gamma c_V}(s-s_0).&\end{aligned} \tag{6}$$

The combination of Eqs. (1)–(6) yields to the wave equation in terms of displacement ξ,

$$(1-\sigma)\xi''-\frac{1}{c_0^2}\ddot{\xi}+\frac{\sigma}{\omega}\dot{\xi}''-2\beta\xi'\xi''=0, \tag{7}$$

where σ is the dissipation coefficient including both viscous and thermal absorption effects. This coefficient has the form

$$\sigma(\omega)=\frac{P_t}{S_t}\left(\sqrt{\frac{\eta}{2\omega\rho_0}}+(\gamma-1)\sqrt{\frac{\kappa}{2\omega\rho_0 c_P}}\right). \tag{8}$$

Equation (7) is similar to the wave equation obtained by Coppens and Sanders[7] for velocity. This nonlinear equation is simplified if there is a harmonic solution. In this case, the displacement can be expressed in terms of a Fourier expansion and Eq. (7) becomes a set of equations,

$$(1-\sigma_k)g_k''+\frac{\omega_k^2}{c_0^2}g_k+\sigma_k h_k''=fc_k, \tag{9a}$$

$$(1-\sigma_k)h_k''+\frac{\omega_k^2}{c_0^2}h_k-\sigma_k g_k''=fs_k,\quad k=0,1,2,..., \tag{9b}$$

where g_k and h_k are the Fourier coefficients of the expansion

$$\xi(a,t)=g_0(a)+\sum_{i=1}^{\infty}g_i(a)\cos(i\omega t)+h_i(a)\sin(i\omega t), \tag{10}$$

and the nonlinear terms fc_k and fs_k take the form

$$fc_0=\beta\left(2g_0'g_0''+\sum_{j=1}^{\infty}g_j'g_j''+h_j'h_j''\right), \tag{11a}$$

$$\begin{aligned}fc_k=\beta\Bigg(&\frac{1}{2}\sum_{j=0}^{k}(g_j'g''_{k-j}+g_j''g'_{k-j}-h_j'h''_{k-j}-h_j''h'_{k-j})\\&+\sum_{j=k}^{\infty}(g_j'g''_{j-k}+g_j''g'_{j-k}+h_j'h''_{j-k}+h_j''h'_{j-k})\Bigg),\end{aligned} \tag{11b}$$

$$\begin{aligned}fs_k=\beta\Bigg(&\frac{1}{2}\sum_{j=0}^{k}(h_j'g''_{k-j}+h_j''g'_{k-j}+g_j'h''_{k-j}+g_j''h'_{k-j})\\&+\sum_{j=k}^{\infty}(h_j'g''_{j-k}+h_j''g'_{j-k}-g_j'h''_{j-k}-g_j''h'_{j-k})\Bigg),\\&k=1,2,...\,.\end{aligned} \tag{11c}$$

If monochromatic excitation is assumed, higher harmonics tend to vanish. In fact, harmonics of third and larger order may be disregarded in a second-order approach. In this case, the infinite set of Eqs. (9) is reduced to a set of five equations

$$(1-\sigma_1)g_1''+\frac{\omega^2}{c_0^2}\,g_1+\sigma_1 h_1''=0, \tag{12a}$$

$$(1-\sigma_1)h_1''+\frac{\omega^2}{c_0^2}\,h_1-\sigma_1 g_1''=0, \tag{12b}$$

$$(1-\sigma_2)g_2''+\frac{(2\omega)^2}{c_0^2}\,g_2+\sigma_2 h_2''=\beta(g_1'g_1''-h_1'h_1''), \tag{12c}$$

$$(1-\sigma_2)h_2''+\frac{(2\omega)^2}{c_0^2}\,h_2-\sigma_2 g_2''=\beta(h_1'g_1''+g_1'h_1''), \tag{12d}$$

$$g_0''=\beta(g_1'g_1''+h_1'h_1''). \tag{12e}$$

In the following, a solution for this set of equations is proposed using a finite element algorithm.

II. FINITE ELEMENT EQUATIONS

To overcome the difficulty in finding a general solution for these differential equations, finite element methods propose a discretization of the continuum problem. In addition, it is possible to approach the true continuum solution as closely as desired as the number of discrete variables increases.

Two different procedures are available to integrate the discrete approximation proposed: the methods of weighted residuals and the determination of variational functionals.[8] The former was chosen in the present work as it is quite general and easy to apply. There are various types of weighted residual methods. Of these, the Bubnov–Galerkin method is widespread as it frequently leads to symmetric matrices, which increases the speed of the numerical algorithm computation. This is therefore the method used in our work.

Once our problem is discretized into a finite number of elements; the unknown functions g_k and h_k in each element of the domain are approximated by the expansions:

$$g_k=dg_{ki}\;\varphi^i(a), \tag{13a}$$

$$h_k=dh_{ki}\;\varphi^i(a), \tag{13b}$$

where the Einstein summation criterion is used. Functions φ^i are the interpolating polynomials. In these expansions, the number of terms is the same as the number of nodes in each element.

Using Eqs. (13) and integrating, the differential equations (9) are replaced by a set of integral equations

$$\int (1-\sigma_k)dg_{ki}\;\varphi''^i\varphi^j\,da+\int \frac{\omega^2}{c_0^2}\,k^2\,dg_{ki}\;\varphi^i\varphi^j\,da + \int \sigma_k\,dh_{ki}\;\varphi''^i\varphi^j\,da-\int fc_k\varphi^j\,da=0, \tag{14a}$$

$$\int (1-\sigma_k)dh_{ki}\;\varphi''^i\varphi^j\,da+\int \frac{\omega^2}{c_0^2}\,k^2\,dh_{ki}\;\varphi^i\varphi^j\,da - \int \sigma_k\,dg_{ki}\;\varphi''^i\varphi^j\,da-\int fs_k\varphi^j\,da=0, \tag{14b}$$

where constant terms can be taken out of the integrals.

Using a matrix formulation, Eqs. (14) can be rewritten as

$$(\mathbf{M_k}+\mathbf{K_k})\mathbf{dg_k}+\mathbf{C_k}\mathbf{dh_k}=\mathbf{FCc_k}+\mathbf{FNc_k}, \tag{15a}$$

$$(\mathbf{M_k}+\mathbf{K_k})\mathbf{dh_k}-\mathbf{C_k}\mathbf{dg_k}=\mathbf{FCs_k}+\mathbf{FNs_k}, \tag{15b}$$

where

$$\mathbf{K_k}=(1-\sigma_k)\int [\varphi'^i][\varphi'^j]da, \tag{16a}$$

$$\mathbf{M_k}=-\frac{\omega^2}{c_0^2}\,k^2\int [\varphi^i][\varphi^j]da, \tag{16b}$$

$$\mathbf{C_k}=\sigma_k\int [\varphi'^i][\varphi'^j]da, \tag{16c}$$

$$\mathbf{FCc_k}=(1-\sigma_k)[dg_{ki}\;\varphi'^i\varphi^j|_{a_{ini}}^{a_{fin}}]+\sigma_k[dh_{ki}\;\varphi'^i\varphi^j|_{a_{ini}}^{a_{fin}}], \tag{16d}$$

$$\mathbf{FCs_k}=(1-\sigma_k)[dh_{ki}\;\varphi'^i\varphi^j|_{a_{ini}}^{a_{fin}}]-\sigma_k[dg_{ki}\;\varphi'^i\varphi^j|_{a_{ini}}^{a_{fin}}], \tag{16e}$$

$$\mathbf{FNc_k}=-\int [fc_k][\varphi^j]da, \tag{16f}$$

$$\mathbf{FNs_k}=-\int [fs_k][\varphi^j]da. \tag{16g}$$

In these relations the subscript k refers to the harmonic, **M** is the mass matrix, **K** is the stiffness matrix, **C** is the absorption matrix, **FCc** and **FCs** include information about the boundary conditions and, finally, **FNc** and **FNs** are the nonlinear terms.

To find the solution for the system, Eqs. (15) must be first solved for $k=1$. In this case, the nonlinear terms are zero. The results obtained for g_1 and h_1 are then used to compute **FNc** and **FNs**, making it possible to find the solutions for $k=0$ and $k=2$.

III. ANALYTICAL SOLUTIONS

Analytical results may be used to validate the numerical algorithm developed. The case of a closed rigid-walled tube excited at one end is studied. At the other end two different boundary conditions (a rigid end and a pressure-release end) are investigated; the case of. The differential equations relating to the fundamental components are linear [see Eqs. (12)]. This implies that an exact solution for these terms can be found. The boundary conditions applied for the rigid termination are

$$\xi(0,t)=\Xi_0\,\cos(\omega t), \tag{17a}$$

$$\xi(l,t)=0, \tag{17b}$$

and the solutions for the components relating to the fundamental frequency are:

$$g_1=\frac{\Xi_0}{\sinh^2(\alpha_1 l)\cos^2(k_1 l)-\cosh^2(\alpha_1 l)\sin^2(k_1 l)} \times[\sinh(\alpha_1(l-a))\sinh(\alpha_1 l)\cos(k_1(l-a))\cos(k_1 l) -\cosh(\alpha_1(l-a))\cosh(\alpha_1 l)\sin(k_1(l-a))\sin(k_1 l)], \tag{18a}$$

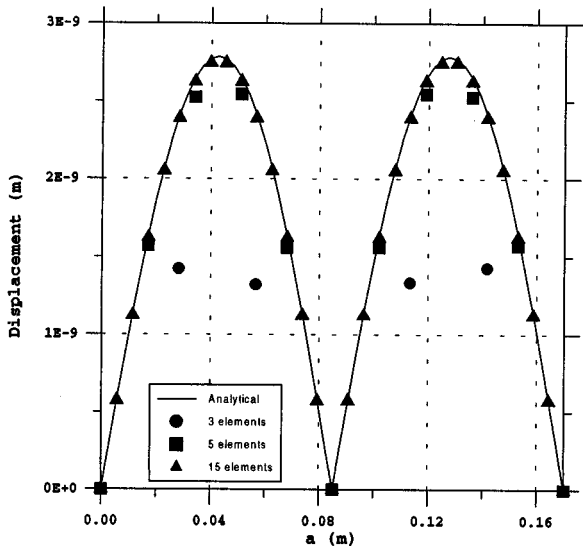


FIG. 1. Tube with rigid boundary condition at the end. Displacement for $k=0$. Resonance.

$$h_1=\frac{\Xi_0}{\sinh^2(\alpha_1 l)\cos^2(k_1 l)-\cosh^2(\alpha_1 l)\sin^2(k_1 l)}\times[\sinh(\alpha_1(l-a))\cosh(\alpha_1 l)\cos(k_1(l-a))\sin(k_1 l)-\cosh(\alpha_1(l-a))\sinh(\alpha_1 l)\sin(k_1(l-a))\cos(k_1 l)], \tag{18b}$$

where

$$k_1=\frac{\omega}{c_0}\left(1+\frac{\sigma_1}{2}\right), \tag{19a}$$

$$\alpha_1=\frac{\omega\sigma_1}{2c_0}. \tag{19b}$$

The nonlinear terms are calculated for resonance and antiresonance, giving the following expressions for the rigid-end case:

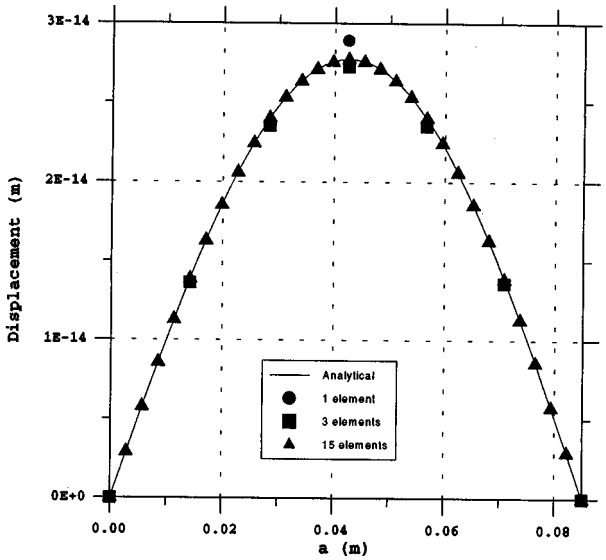


FIG. 2. Tube with rigid boundary condition at the end. Displacement for $k=0$. Antiresonance.

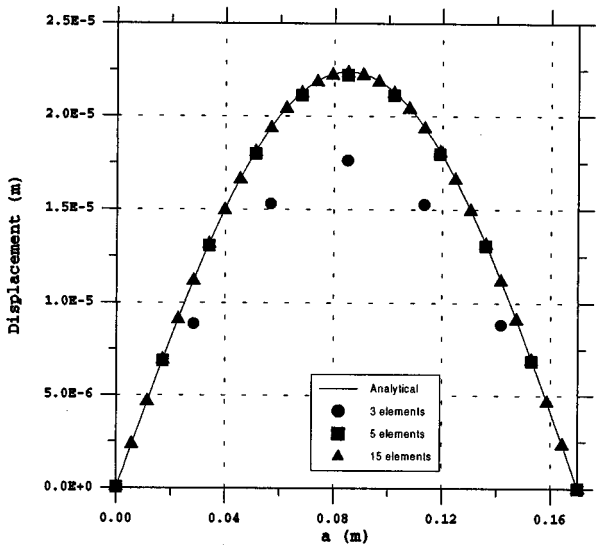


FIG. 3. Tube with rigid boundary condition at the end. Displacement for $k=1$. Resonance.

$$\text{Resonance}\quad k_1 l=n\pi, \tag{20a}$$

$$\text{Antiresonance}\quad k_1 l=(2n-1)(\pi/2)\quad n=1,2,\ldots. \tag{20b}$$

In both cases, Eqs. (18) are reduced to a simpler form. Retaining only terms of first and second order, the solutions found are:

Resonance

$$g_1=\frac{\Xi_0(-1)^n(l-a)}{l}\cos(k_1(l-a)), \tag{21a}$$

$$h_1=-\frac{\Xi_0(-1)^n}{\alpha_1 l}\sin(k_1(l-a)), \tag{21b}$$

$$g_2=\frac{\beta}{16}\left(\frac{\Xi_0}{\alpha_1 l}\right)^2\frac{k_1}{\sigma_1}\sin(2k_1(l-a)), \tag{21c}$$

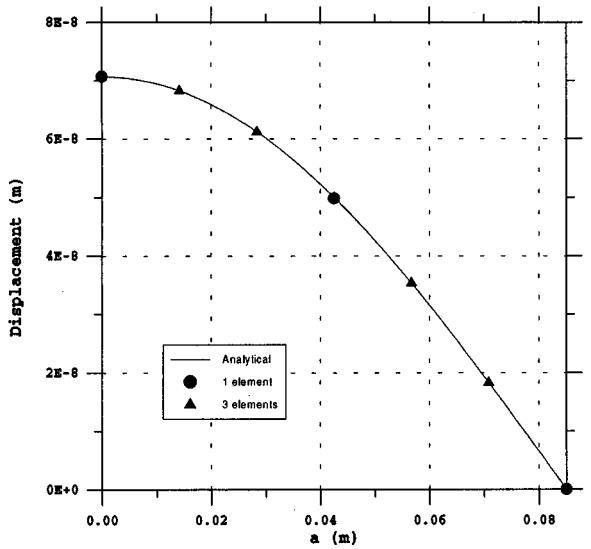


FIG. 4. Tube with rigid boundary condition at the end. Displacement for $k=1$. Antiresonance.

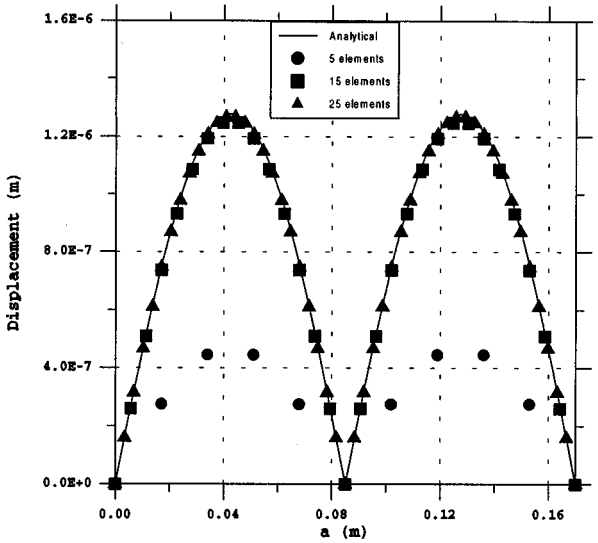


FIG. 5. Tube with rigid boundary condition at the end. Displacement for $k=2$. Resonance.

$$h_2=\frac{\beta}{16}\left(\frac{\Xi_0}{\alpha_1 l}\right)^2\frac{k_1}{\sigma_1}(\sqrt{2}+1)\sin(2k_1(l-a)), \tag{21d}$$

$$g_0=-\frac{\beta}{8}\left(\frac{\Xi_0}{\alpha_1 l}\right)^2 k_1\ \sin(2k_1(l-a)). \tag{21e}$$

These solutions are consistent with the results obtained by Coppens and Sanders,[7]

Antiresonance

$$g_1=-\Xi_0(-1)^n\ \sin(k_1(l-a)), \tag{22a}$$

$$h_1=-\Xi_0(-1)^n\alpha_1(l-a)\cos(k_1(l-a)), \tag{22b}$$

$$g_2=-\frac{\beta}{16}\Xi_0^2\frac{k_1}{\sigma_1}\sin(2k_1(l-a)), \tag{22c}$$

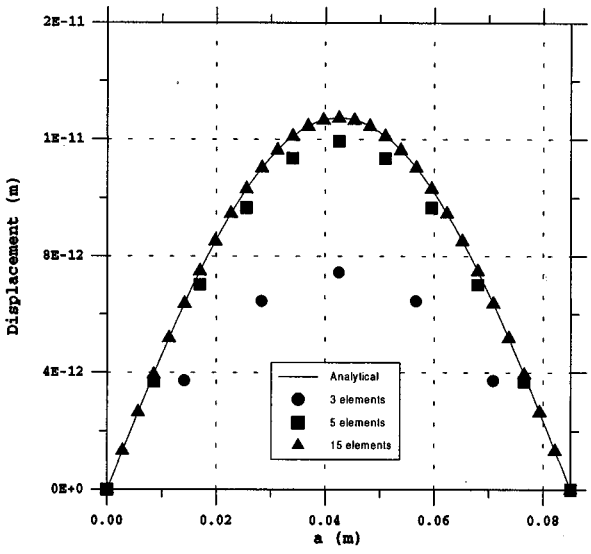


FIG. 6. Tube with rigid boundary condition at the end. Displacement for $k=2$. Antiresonance.

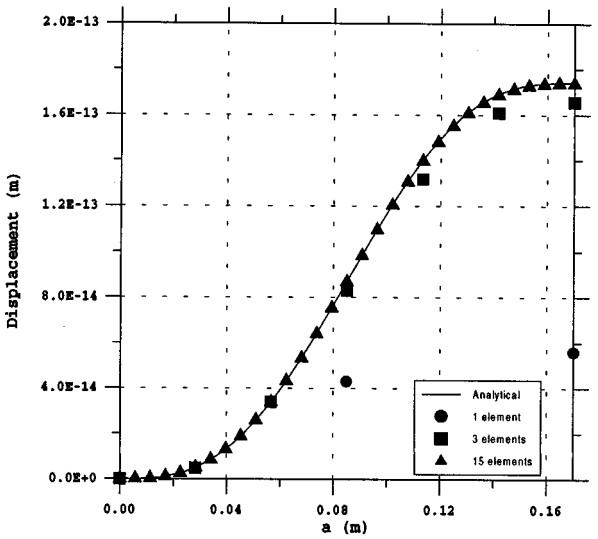


FIG. 7. Tube with pressure release boundary condition at the end. Displacement for $k=0$. Antiresonance.

$$h_2=-\frac{\beta}{16}\Xi_0^2\frac{k_1}{\sigma_1}(\sqrt{2}+1)\sin(2k_1(l-a)), \tag{22d}$$

$$g_0=-\frac{\beta}{8}\Xi_0^2 k_1\ \sin(2k_1(l-a)). \tag{22e}$$

The pressure-release boundary condition was also studied for the end of the tube. The boundary conditions applied in this case were:

$$\xi(0,t)=\Xi_0\ \cos(\omega t), \tag{23a}$$

$$\left.\frac{\partial\xi(a,t)}{\partial a}\right|_{a=l}=0. \tag{23b}$$

However, the conditions for resonance and antiresonance are now reversed:

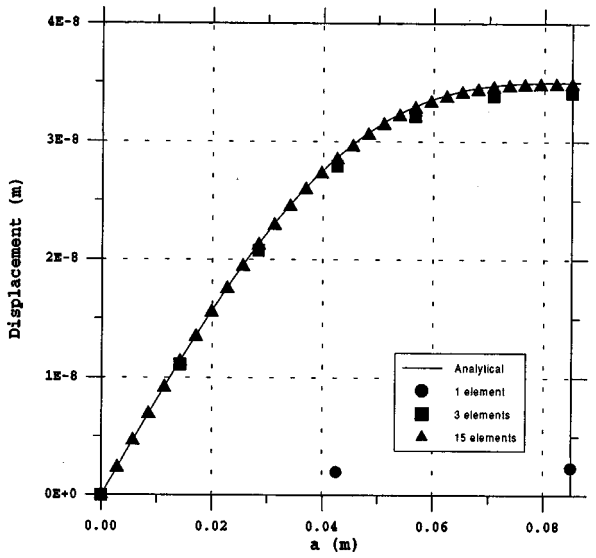


FIG. 8. Tube with pressure release boundary condition at the end. Displacement for $k=0$. Resonance.

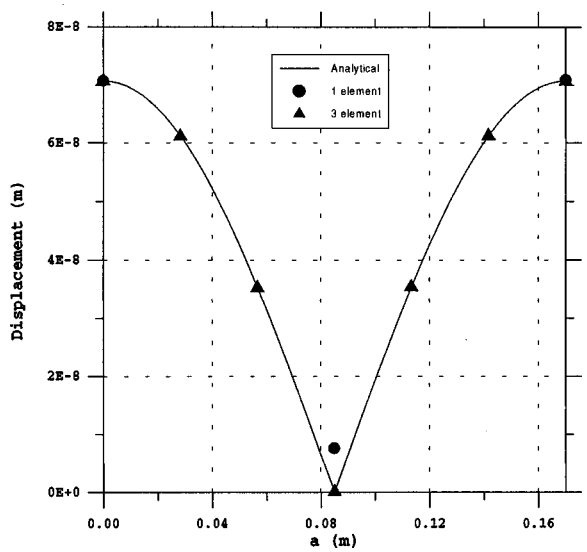


FIG. 9. Tube with pressure release boundary condition at the end. Displacement for $k=1$. Antiresonance.

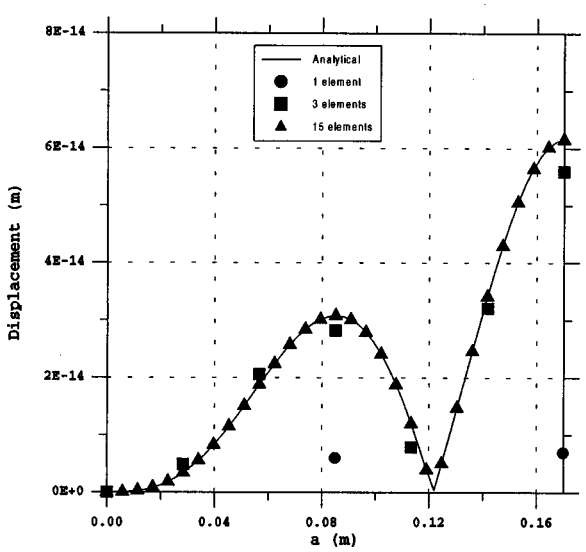


FIG. 11. Tube with pressure release boundary condition at the end. Displacement for $k=2$. Antiresonance.

$$\text{Resonance} \quad k_1 l=(2n-1)(\pi/2), \tag{24a}$$

$$\text{Antiresonance} \quad k_1 l=n\pi \quad n=1,2,\ldots . \tag{24b}$$

The analytical results are:

Resonance

$$g_1=-\frac{\Xi_0(-1)^n(l-a)}{l}\sin(k_1(l-a)), \tag{25a}$$

$$h_1=-\frac{\Xi_0(-1)^n}{\alpha_1 l}\cos(k_1(l-a)), \tag{25b}$$

$$g_2=\frac{\beta}{16}\left(\frac{\Xi_0}{\alpha_1 l}\right)^2 k_1[\sin(2k_1(l-a))+2k_1 a \times\cos(2k_1(l-a))], \tag{25c}$$

$$h_2=0, \tag{25d}$$

$$g_0=\frac{\beta}{8}\left(\frac{\Xi_0}{\alpha_1 l}\right)^2 k_1[2k_1 a+\sin(2k_1(l-a))]. \tag{25e}$$

Antiresonance

$$g_1=\Xi_0(-1)^n\cos(k_1(l-a)), \tag{26a}$$

$$h_1=-\Xi_0(-1)^n\alpha_1(l-a)\sin(k_1(l-a)), \tag{26b}$$

$$g_2=-\frac{\beta}{16}\Xi_0^2 k_1[\sin(2k_1(l-a))+2k_1 a \times\cos(2k_1(l-a))], \tag{26c}$$

$$h_2=0, \tag{26d}$$

$$g_0=\frac{\beta}{8}\Xi_0^2 k_1[2k_1 a+\sin(2k_1(l-a))]. \tag{26e}$$

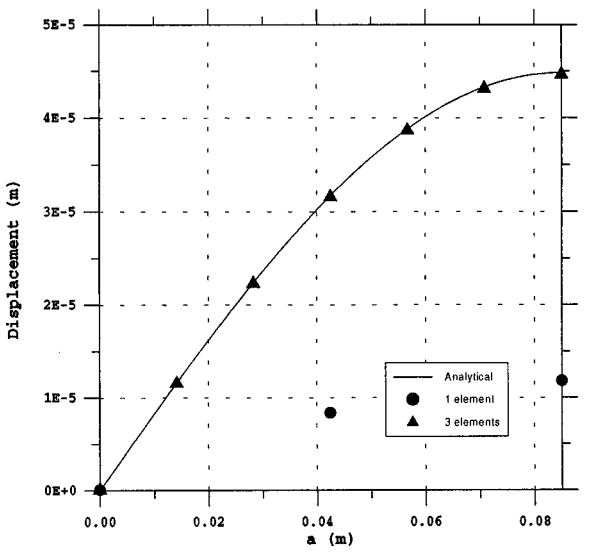


FIG. 10. Tube with pressure release boundary condition at the end. Displacement for $k=1$. Resonance.

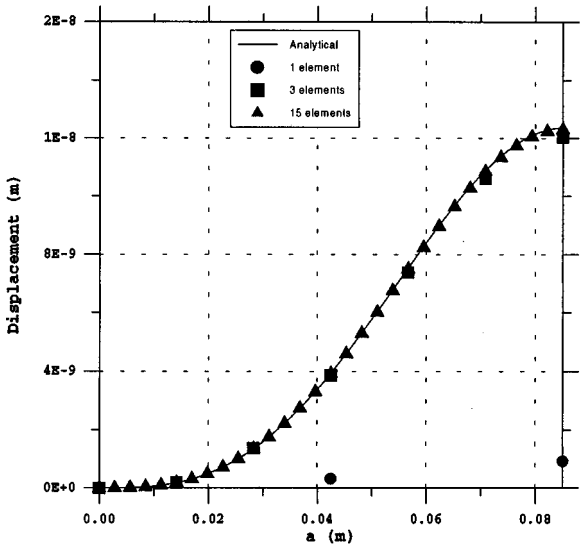


FIG. 12. Tube with pressure release boundary condition at the end. Displacement for $k=2$. Resonance.

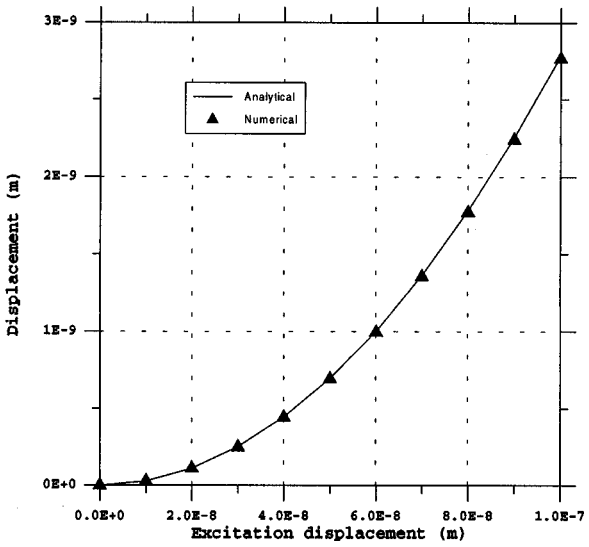


FIG. 13. Displacement of the $k=0$ component versus excitation level.

Further interesting acoustical magnitudes can be obtained from these results. Velocity can be found, for example, by derivating displacement with respect to time. Other magnitudes such as pressure or density can be calculated using Eqs. (1)–(5).

IV. NUMERICAL RESULTS

The same four cases were studied numerically, using the new finite element algorithm. Different meshes or discretizations were tested. It should be noted that the derivatives which relate to the nonlinear terms within the integrals are of second order. Accordingly, interpolating polynomials φ^i of second or higher order must be used.

A. Displacement distribution

The displacement distribution inside the chamber was studied for the cases mentioned above. The results presented

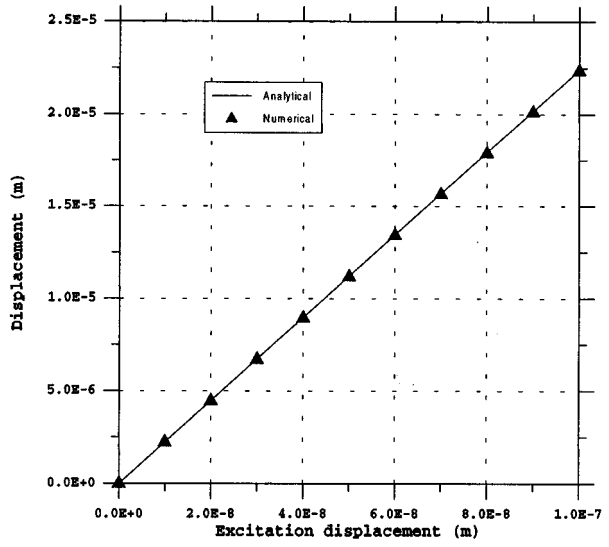


FIG. 14. Displacement of the $k=1$ component versus excitation level.

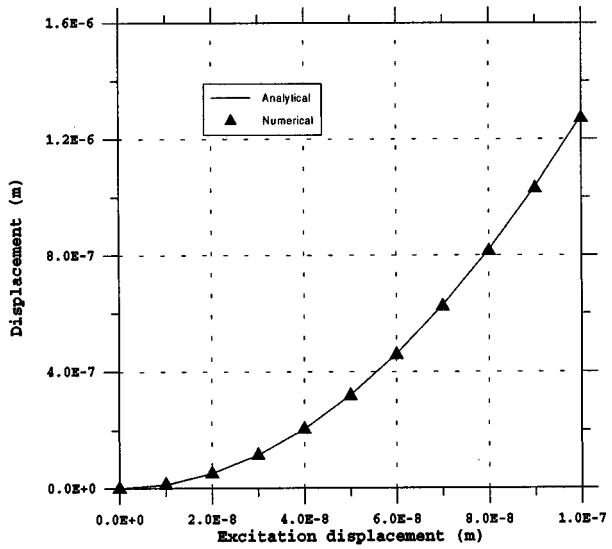


FIG. 15. Displacement of the $k=2$ component versus excitation level.

were obtained for a 1-kHz sinusoidal excitation and 0.1-μm displacement amplitude. The displacement distributions for $k=0,1,2$ are displayed separately. These results are plotted in terms of the rms displacement of the ith harmonic

$$\langle \xi_0 \rangle_{\mathrm{rms}} = g_0 , \tag{27a}$$

$$\langle \xi_i \rangle_{\mathrm{rms}} = \frac{\sqrt{g_i^2 + h_i^2}}{\sqrt{2}} \quad i=1,2. \tag{27b}$$

Figures 1–6 relate to the tube with a rigid end and Figs. 7–12 plot the results obtained assuming a pressure-release end. The length of the tube is adjusted to achieve either resonance or antiresonance

It can be seen from Figs. 1–12 that a good fit between analytical (continuous line) and numerical data (discrete points) is obtained. Establishing the number of elements needed to obtain accurate solutions is usually done by running the numeric algorithm with different meshes. An element length of $\lambda/4$ is often chosen for wave problems. This criterion gives for a nonlinear, second-order approach a minimum element length of $\lambda/8$. Figures 1–12 show that this criterion adequately describes the qualitative behaviour of the acoustic field but the absolute values obtained are not

TABLE I. Tube radii, absorption and tube length considered as a function of the driving frequency.

Frequency (Hz)	Tube radii (m)	Absorption coefficient σ	Tube length (m) first resonance
1000	0.049	0.00209	0.171
2000	0.024	0.00301	0.0856
3000	0.016	0.00369	0.0571
4000	0.012	0.00426	0.0428
5000	0.0099	0.00463	0.0342
6000	0.0083	0.00504	0.0285
7000	0.0071	0.00543	0.0244
8000	0.0062	0.00584	0.0214
9000	0.0055	0.00618	0.0190
10 000	0.0050	0.00646	0.0171

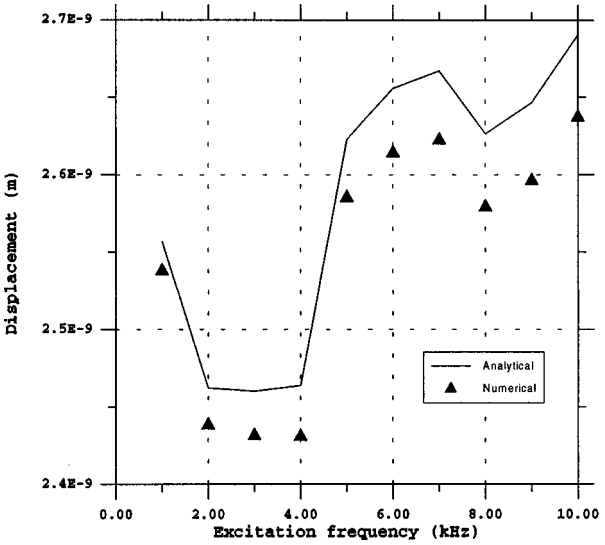


FIG. 16. Maximum displacement as a function of the frequency for the $k=0$ component.

accurate, which is due to the important changes caused by discretization near resonance conditions. This explanation is consistent with the results obtained which, for the antiresonance cases, converge faster than those for the cases of resonance. As was expected, the fundamental frequency displacement converges faster than the displacements relating to the rest of the frequency components. In any case, 30 elements per wavelength were enough to obtain accurate numerical data for the problems studied.

B. The effect of displacement amplitude of the excitation

Figures 13–15 show the relation between the amplitude of the excitation and the level obtained for the different harmonics. These results correspond to the maximum rms displacement inside the chamber when the rigid-end tube is excited at the resonance frequency (1 kHz).

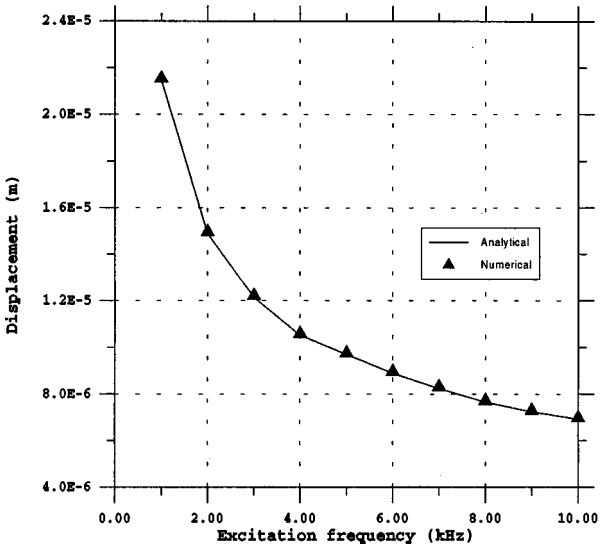


FIG. 17. Maximum displacement as a function of the frequency for the $k=1$ component.

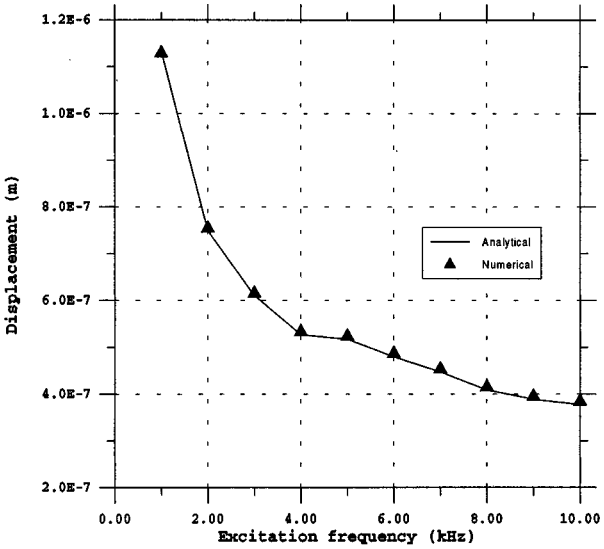


FIG. 18. Maximum displacement as a function of the frequency for the $k=2$ component.

It can be seen from these figures (13–15) that the relationship between the excitation level and the fundamental frequency displacement, follows a straight line. The other two components show a parabolic behavior related to the square exponent appearing in the corresponding Fourier coefficients [see Eqs. (21)].

C. The effect of absorption at high frequency

At higher frequencies, the absorption in the tube increases. This is a consequence of the frequency dependence of this coefficient. Besides this, the tube has to be narrower to avoid propagation of transverse modes. For small tube radii, the absorption increases even further (see Table I). Figures 16–18 show the pressure levels obtained at different frequencies for $k=0,1,2$. The resonance is kept by choosing an adequate tube length.

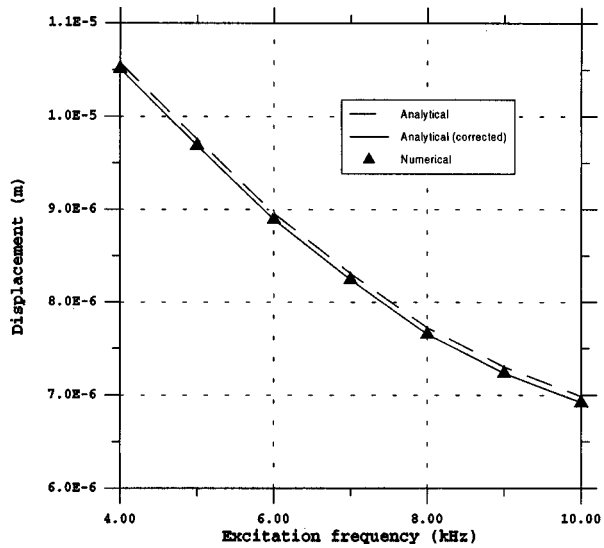


FIG. 19. Maximum displacement as a function of the frequency for the $k=1$ component.

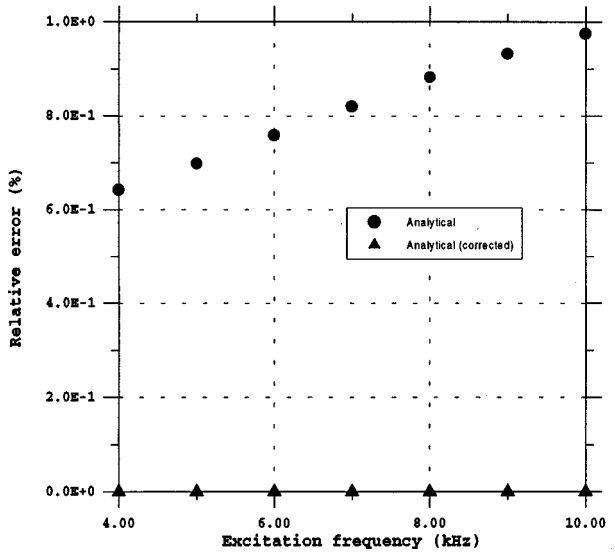


FIG. 20. Error in the numeric results when they are compared to the analytic solution or the analytic corrected one.

Discrepancies appear between the numerical and analytical results when the frequency (and the absorption) increases. Although the difference is obvious for $k=0$, a careful study shows that there are also differences between analytical and numerical data for the other two components. These discrepancies derive from the approximations made to obtain the analytical solutions [Eqs. (18)–(21)]. This can be easily shown for the fundamental, because in this case the exact solution can be calculated analytically [Eqs. (17)]. For the rest of the harmonics it is extremely complicated.

Figures 19 and 20 show these results. It can be seen that the difference between the corrected analytical solution and the approximated analytical solution increases with frequency. However, numerical data fit the former exactly, and there is no appreciable error.

V. CONCLUSIONS

A finite element algorithm was developed to describe nonlinear one-dimensional standing waves. The method incorporates perturbation techniques and it leads to the expression of displacement as a Fourier expansion. This numerical approach was checked against analytical solutions obtained for several special situations (resonance and antiresonance, with rigid and pressure-release boundary conditions), and proved to be accurate. The method can also be applied to problems where the tube is out of resonance or out of antiresonance and different boundary conditions are established at the end.

This numerical procedure is well suited for the study of high frequency standing fields, for which the effect of absorption leads to inaccurate analytic predictions (this inaccuracy was due to the approximations made to obtain these solutions).

Future work would give an expansion of the finite element method to simulate 2-D and 3-D nonlinear fields, where there are even fewer analytical results than in the 1-D case. It is precisely for these 2-D and 3-D problems where the finite element method is clearly advantageous over other numerical approaches, giving accurate solutions for a large number of different cases, no matter how irregular the geometry of the problem. For this reason, this method will probably play an important role in the study and design of finite-amplitude wave systems in the near future.

ACKNOWLEDGMENTS

The work presented in this paper was supported by the research Grant No. PN91-02621490 from the Spanish Ministry of Education and Science; by two research Projects: CICYT-TAP93-0230 and CICYT-AMB96-1211-C02-01.

[1] A. L. Van Buren, ''Mathematical model for nonlinear standing waves in a tube,'' J. Sound Vib. **42**, 273–280 (1975).
[2] S. I. Aanonsen, ''Numerical computation of the nearfield of a finite amplitude sound beam,'' Int. Report No 73, University of Bergen, Norway (1983).
[3] S. I. Aanonsen, T. Barkve, J. N. Tjøtta, and S. Tjøtta, ''Distortion and harmonic generation in the nearfield of a finite amplitude sound beam,'' J. Acoust. Soc. Am. **75**, 749–768 (1984).
[4] T. Tsuchiya and Y. Kagawa, ''Finite Element Analysis of focusing field in nonlinear acoustic waves,'' Proc. 11th Symposium on Ultrasonic Electronics, pp. 51–53 (1990).
[5] L. Elvira, ''Desarrollo y validación de un método de elementos finitos para el estudio de ondas acústicas estacionarias no lineales,'' Ph.D. dissertation, Universidad Complutense de Madrid, 1996.
[6] D. E. Weston, ''The theory of the propagation of plane sound waves in tubes,'' Proc. Phys. Soc. London, Sect. B **66**, 695–709 (1953).
[7] A. B. Coppens and J. V. Sanders, ''Finite-amplitude standing waves in rigid-walled tubes,'' J. Acoust. Soc. Am. **43**, 516–529 (1968).
[8] O. C. Zienkiewicz, *The Finite Element Method* (McGraw-Hill Book, London 1977), 3rd ed., Chap. 3, pp. 42–65.

The acoustic radiation force in lossless fluids in Eulerian and Lagrangian coordinates

K. Beissner
Physikalisch-Technische Bundesanstalt, 38116 Braunschweig, Germany

(Received 21 November 1996; accepted for publication 7 February 1998)

A general theory of the (Langevin) acoustic radiation force in three-dimensional fields in lossless fluids in Eulerian and Lagrangian coordinates is given. It is based on Brillouin's approach of an acoustic radiation stress identified in Eulerian coordinates with the negative time average of the momentum flux density and in Lagrangian coordinates with the time average of the Piola–Kirchhoff–Boussinesq stress. It is shown in comparison to other statements from the literature that these stresses are not in general identical but that in the linear-acoustics approximation, the difference between the results in the two representations vanishes if the radiation force acting on an object entirely surrounded by the sound-propagating fluid is considered. It is just this situation which is of experimental relevance.
[S0001-4966(98)05205-9]

PACS numbers: 43.25.Qp [MAB]

INTRODUCTION

The acoustic radiation force is the time-average force exerted by an acoustic field on an object (''target'') or on a boundary surface. In scientific practice it occurs as an unintentional effect under various circumstances, but its importance lies in its intentional and quantitative use for which at least three main areas can be identified, namely: (a) The radiation force on a large target as measured by a radiation force balance[1,2] and serving as a measure of the ultrasonic power; (b) the radiation force on a small target as measured by a sphere radiometer[3] and serving as a measure of the ultrasonic intensity; (c) the radiation force on small liquid and solid samples[4] for purposes of separation techniques and for levitation and positioning purposes, with the aim of a containerless processing.[5] In all these cases, it is important to theoretically and quantitatively understand the acoustic radiation force. This applies particularly to application (a) which is regarded as a primary measurement method and for which theoretical knowledge is necessary to obtain the ultrasonic power from the radiation force measured. This means that the acoustic radiation force theory is one of the cornerstones of the exposimetry of current medical ultrasonic equipment.

Lord Rayleigh,[6] Paul Langevin (as reported by P. Biquard[7]), Léon Brillouin,[8] and Gustav Hertz (in cooperation with H. Mende[9]) in particular influenced the early development of the acoustic radiation force theory. Hertz and Mende[9] introduced the formal distinction between the Rayleigh and the Langevin radiation force (cf. also Ref. 10). The former is defined as the force acting on the wall of a closed vessel containing the sound-propagating medium. The latter is defined as the force exerted on a target in an open vessel in which the general hydrostatic pressure (which is determined by the atmospheric pressure in this case) acts on both sides of the target and does, therefore, not enter into the radiation force result. The present paper deals with the time-average force exerted by an arbitrary, steady-state sound field on an arbitrary target entirely surrounded by the sound-propagating fluid. This situation falls under the Langevin case and covers all three application areas mentioned above. It should be noted that the field is here generally understood in the presence of the target. If this is to be traced back to the field in the absence of the target, a suitable calculation or conclusion has to be added, unless this is already contained in the results as, for example, in the so-called case 3 application below where $\hat{p}$ is the pressure amplitude of the incident wave. The present paper does not deal with particular field/target configurations but only with the fundamentals and with rather simple examples involving certain basic field types. The general results obtained can be applied to any specific problem of interest.

While radiation force methods based on the Langevin situation are in current use and the radiation force balance method (a) has been successfully tested in international intercomparisons,[11] the Rayleigh radiation force discussion so far appears to be in a more difficult state. A number of (controversial) plane-wave calculations based on particular field solutions under various boundary conditions have been made (for a review and a subsequent discussion, cf. Refs. 12 and 13), but the problem has not yet been solved by experiment and will not be dealt with here.

The fundamentals of the theory of the Langevin radiation force given in this paper are taken from Brillouin and, in a certain sense, from Langevin. According to the Brillouin approach,[8] the radiation force is due to an acoustic radiation stress (tensor) which is identified in Eulerian coordinates with the negative time average of the momentum flux density (with some terminological differences in the literature) and in Lagrangian coordinates with the time average of a quantity which has been referred to by Brillouin[8,14] as the Boussinesq stress and elsewhere[14–16] as the (first) Piola–Kirchhoff stress and which will here be simply regarded as the stress in Lagrangian coordinates and referred to as the Piola–Kirchhoff–Boussinesq stress. The latter quantity has been used so far in the radiation stress literature only for calculations concerning solids.[8,14] Most radiation force theories for

fluids have started from the Eulerian version of the Brillouin approach (see Sec. II below). The question of a Lagrangian radiation force theory for fluids according to the Brillouin approach has until now remained open[17] and will be answered here. The answer is not straightforward but makes it necessary to more closely consider the stress in Lagrangian coordinates and a number of formal relations important for the theory. In addition and as a reference, the Eulerian version will be given for comparison purposes.

The Eulerian (i.e., spatial) and the Lagrangian (i.e., material) description of fluid (and solid) motions both have specific advantages and disadvantages. The Lagrangian one is generally regarded as more cumbersome but may be advantageous if the boundary surface between two media moves in the Eulerian space but is at rest in Lagrangian coordinates. The Lagrangian version of Brillouin's radiation force theory for fluids is worked out here first, for having a complete theory in both representations and second, for comparing it with another Lagrangian radiation force theory which can sometimes be found in the literature and will be briefly quoted below and discussed in the Appendix.

The important comparison of the Eulerian with the Lagrangian results of the Brillouin approach will be dealt with in this paper but has also been discussed in the literature. Brillouin[8] found agreement between his plane-wave results for solids. This is in line with the plane-wave agreement found here (although the calculation methods are completely different) but it cannot be extended to arbitrary fields which Brillouin may have had in mind. Cantrell[14] postulated a general identity between the two versions of radiation stresses for arbitrary, three-dimensional fields.[18] The present paper, on the other hand, produces the result that the two versions of radiation stresses are in general different but that the difference becomes irrelevant if the radiation force acting on an object entirely surrounded by the sound-propagating fluid is considered.

The expressions obtained directly from the Brillouin approach are not sufficient for practical radiation force calculations as they contain in both representations (see below) the time-average acoustic pressure at least in the second-order approximation. Brillouin and a number of more recent authors in addition used certain plane-wave restrictions and boundary conditions[19] to arrive at practical results. It is better to use modifications of the field equations themselves in order to obtain results which are valid for any three-dimensional field type. This way of proceeding has been shown by the author for the Eulerian version[2,19,20] and it will here again be applied to the Lagrangian version, however along different lines, of course. The modified field equations to be used for this purpose go back to Langevin[7,21] and have on occasion[22] been referred to as ''Langevin's first relation'' for the Lagrangian version and as ''Langevin's second relation'' for the Eulerian version. So the general radiation force theory given below is a combination of the two representations of the Brillouin approach with Langevin's relations, but the benefits of such a combination were obviously not realized by the two authors themselves.[22]

Langevin[7] himself used his first relation in a different way. He considered the radiation force to be due to an acoustic radiation ''pressure'' and identified it with the Lagrangian time average of the acoustic pressure which according to his first relation is equal to the time average of the total acoustic energy density plus a constant which is not relevant to the Langevin situation. This approach has been repeated in the literature[12,23,24] and is briefly discussed in the Appendix.

This paper deals with a number of matrix quantities. The question of their behavior in a transformation of coordinates will not be discussed in detail but answered in summary. The paper does not consider curvilinear coordinates and covariant and contravariant vector and tensor components but only rectangular Cartesian coordinate systems. The tensor behavior in a transformation from one Cartesian system to another has been described, for example, in Refs. 25–27. Matrix quantities following these rules have been referred to as Cartesian tensors.[25,26] The following can be derived from the references mentioned: (1) The Cauchy stress[15,28] is known to be a tensor and thus also a Cartesian tensor, of course. (2) The determinant of a Cartesian tensor is an invariant. (3) The unit matrix is a Cartesian tensor. (4) The product of a scalar with a Cartesian tensor is a Cartesian tensor. (5) The product of two matrices each of which is a Cartesian tensor is also a Cartesian tensor. (6) The transpose of a Cartesian tensor is a Cartesian tensor. (7) The inverse of a Cartesian tensor is a Cartesian tensor. (8) The dyadic product of two vectors is a Cartesian tensor. (9) The gradient of a vector (cf. below) is a Cartesian tensor.

It follows from these rules that all the matrices appearing below, namely **B**, **B**′, **F**, **G**, **H**, **K**, **M**, **S**, **T**, and **U**, are Cartesian tensors.[29]

I. DEFINITIONS AND TERMINOLOGY

The physical quantities are understood here with reference to a (rectangular) Cartesian coordinate system whose axes are numbered from 1 to 3. Stresses are represented by matrices. For example, the Cauchy stress which is the usual mechanical stress in Eulerian coordinates is given by

$$\mathbf{T}=\begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix}. \tag{1}$$

The indices are understood according to the following example:[27,28,30,but cf. 31] T_{12} is the force component in direction No. 2 per unit area of the surface perpendicular to direction No. 1. There are two implications of this convention. First, the infinitesimal force vector **df** exerted across an infinitesimal surface element which is characterized by its normal, infinitesimal vector **da** is given by

$$\mathbf{df}=\mathbf{da}\cdot\mathbf{T}, \tag{2}$$

i.e., the surface element vector is postmultiplied by the stress matrix. Second, this convention implies the special form of the vector gradient, the matrix divergence and the matrix curl given below. A different definition of the stress indices would imply different forms of these operators. The foregoing discussion might appear superfluous as the Cauchy stress is known to be symmetric, i.e., $\mathbf{T}=\mathbf{T}^T$ where the superscript T characterizes the transpose, but with respect to the other

stresses to be discussed, allowance is made for unsymmetric stresses, at least formally.

Products of physical quantities are represented here as follows. If at least one of the two factors is a scalar quantity, no multiplication symbol is used. If both factors are higher-rank quantities, the following symbols are used. The dot $\cdot$ characterizes the scalar product of two vectors or the product of a vector with a matrix[32] or the product of two matrices. The symbol $\otimes$ characterizes the dyadic product of two vectors. The symbol $\times$ characterizes the vector product of two vectors or the vector product of a vector with a matrix.

Let q be a general symbol for a scalar quantity, $\mathbf{q}$ for a vector quantity, and $\mathbf{Q}$ for a matrix, all being regarded as functions of the position vector $\mathbf{x}=(x_1,x_2,x_3)$. The well-known, conventional differential operators grad, div, and curl can be extended to a higher order by defining the gradient of a vector, grad $\mathbf{q}=\boldsymbol{\nabla}\otimes\mathbf{q}$, which is a matrix according to[25,33]

$$(\text{grad } \mathbf{q})_{ij}=\frac{\partial q_j}{\partial x_i}, \tag{3}$$

the divergence of a matrix, div $\mathbf{Q}=\boldsymbol{\nabla}\cdot\mathbf{Q}$, which is a vector according to[27,33]

$$(\text{div } \mathbf{Q})_j=\sum_i \frac{\partial Q_{ij}}{\partial x_i}, \tag{4}$$

and the curl of a matrix, curl $\mathbf{Q}=\boldsymbol{\nabla}\times\mathbf{Q}$, which is a matrix according to[33,34]

$$\text{curl } \mathbf{Q}=\begin{pmatrix} \frac{\partial Q_{31}}{\partial x_2}-\frac{\partial Q_{21}}{\partial x_3} & \frac{\partial Q_{32}}{\partial x_2}-\frac{\partial Q_{22}}{\partial x_3} & \frac{\partial Q_{33}}{\partial x_2}-\frac{\partial Q_{23}}{\partial x_3} \\ \frac{\partial Q_{11}}{\partial x_3}-\frac{\partial Q_{31}}{\partial x_1} & \frac{\partial Q_{12}}{\partial x_3}-\frac{\partial Q_{32}}{\partial x_1} & \frac{\partial Q_{13}}{\partial x_3}-\frac{\partial Q_{33}}{\partial x_1} \\ \frac{\partial Q_{21}}{\partial x_1}-\frac{\partial Q_{11}}{\partial x_2} & \frac{\partial Q_{22}}{\partial x_1}-\frac{\partial Q_{12}}{\partial x_2} & \frac{\partial Q_{23}}{\partial x_1}-\frac{\partial Q_{13}}{\partial x_2} \end{pmatrix}, \tag{5}$$

where $\boldsymbol{\nabla}$ is the nabla operator and where the vector product of a vector and a matrix is given by $(\mathbf{q}\times\mathbf{Q})_{ij}=\Sigma_m\Sigma_n\epsilon_{imn}q_mQ_{nj}$, with the Levi-Civita symbol ϵ representing the well-known third-order permutation tensor.[26,28]

The symbols grad, div, and curl will be used for both the conventional and the higher-order versions, so it is the order of the argument quantity which determines the true meaning of a differential operator symbol. Although it will not be shown in detail here, the following is true for the Cartesian transformation behavior of the quantities of Eqs. (3)–(5): The gradient of a vector is a Cartesian tensor. The divergence of a Cartesian tensor is a vector. The curl of a Cartesian tensor is a Cartesian tensor, but only in the case of a right-handed Cartesian system. In the case of a transition to a left-handed Cartesian system, the curl changes to its negative, which is similar to the behavior of the conventional vector curl. For this reason, the treatment is restricted to right-handed Cartesian systems.[35]

The following formulas, which are needed for this paper, can be derived from the above definitions by direct calculation:

$$(\text{grad } \mathbf{q})\cdot\mathbf{q}=\tfrac{1}{2}\text{ grad}(\mathbf{q}\cdot\mathbf{q}), \tag{6}$$

$$\tfrac{1}{2}\text{ grad grad } q^2=\text{grad}(q\text{ grad } q)$$
$$=\text{grad } q\otimes\text{grad } q+q\text{ grad grad } q, \tag{7}$$

$$\text{div}(q\mathbf{Q})=q\text{ div } \mathbf{Q}+(\text{grad } q)\cdot\mathbf{Q}, \tag{8}$$

$$\text{div}(q\mathbf{U})=\text{grad } q, \tag{9}$$

where $\mathbf{U}$ is the unit matrix with

$$U_{ij}=\delta_{ij}, \tag{10}$$

with δ_{ij} being the Kronecker delta,

$$\text{div}(\mathbf{q}\otimes\mathbf{q})=(\mathbf{q}\cdot\text{grad})\mathbf{q}+\mathbf{q}\text{ div } \mathbf{q}, \tag{11}$$

$$\text{curl grad } \mathbf{q}=0, \tag{12}$$

$$\text{div curl } \mathbf{Q}=0, \tag{13}$$

$$\text{curl curl}(q\mathbf{U})=\text{grad grad } q-\mathbf{U}\text{ div grad } q, \tag{14}$$

$$\text{adj}(\text{grad } \mathbf{q})=\tfrac{1}{2}\text{curl}(\mathbf{q}\times(\text{grad } \mathbf{q})^T)^T, \tag{15}$$

where the adjoint, adj, is the matrix of the cofactors[33] according to

$$\text{adj}(\mathbf{Q})=(\mathbf{Q}^T)^{-1}\text{ det } \mathbf{Q} \tag{16}$$

and, since grad $\mathbf{x}=\mathbf{U}$, Eq. (15) has a special application, namely

$$\text{adj}(\mathbf{U})=\mathbf{U}=\tfrac{1}{2}\text{curl}(\mathbf{x}\times\mathbf{U})^T, \tag{17}$$

and two algebraic identities can be derived from the usual vector and matrix calculus,

$$\mathbf{q}^{(a)}\cdot(\mathbf{q}^{(b)}\otimes\mathbf{q}^{(c)})=(\mathbf{q}^{(a)}\cdot\mathbf{q}^{(b)})\mathbf{q}^{(c)}, \tag{18}$$

$$(\mathbf{q}^{(a)}\cdot\mathbf{Q})\times(\mathbf{q}^{(b)}\cdot\mathbf{Q})=(\mathbf{q}^{(a)}\times\mathbf{q}^{(b)})\cdot\text{adj}(\mathbf{Q}). \tag{19}$$

The well-known Gauss and Stokes integral theorems can also be extended to a higher order.[33,34,36] Their extended meaning is: (Gauss:) The vectorial surface integral of a matrix over a closed surface is equal to the volume integral of the divergence of the matrix over the volume bounded by the surface. (Stokes:) The vectorial surface integral of the curl of a matrix is equal to the vectorial line integral of the matrix along the boundary line of the surface. The following helpful rule (mentioned in Ref. 37 in the nonextended meaning) can be derived from the Stokes integral theorem: The vectorial surface integral of the curl of a matrix over a closed surface is zero. To see this, consider a closed surface as, for example, a sphere, and a curve on it such as, for example, the equator. The curve divides the surface into two parts. The total surface integral is the sum of the surface integrals over the two parts. The Stokes integral theorem can be applied to both parts, thus the total surface integral of the curl of a matrix is given by a sum of two line integrals of the matrix over the same curve but in opposite directions. This is zero.[38]

The radiation force on an object (target) is the vectorial surface integral of the radiation stress over a closed surface enclosing the target. If the radiation stress is a zero-divergence quantity throughout the sound-propagating fluid, this has two important implications: First, the result is independent of the specific choice of the surface of integration, as

long as the entire target is contained in the interior of that surface. This means that the surface of integration can be freely chosen so as to minimize the calculation difficulties, even far away from the target or from the transducer (far-field approximation[39,40]). Second, if there is only fluid and no target inside the surface of integration, the result is zero, i.e., there is no net radiation force on the sound-propagating fluid itself.[41]

If, in addition, the radiation stress (or a certain part of it) can be expressed as the curl of some other quantity, the radiation force (or the corresponding part of it) is zero, even with a target inside the surface of integration. This means that any part of the radiation stress representing a curl can be regarded as ineffective and negligible for the Langevin radiation force. A radiation force which is zero can, therefore, be due to the following conditions: (a) The divergence of the radiation stress is zero and (b) the radiation stress can be formulated as a curl, but with an important difference: in case (a) the divergence must be zero everywhere inside the surface of integration (which means: no target), whereas the Stokes integral theorem does not refer to the volume bounded by the surface of integration, and so, in case (b), the condition that the radiation stress is the curl of something (and is divergenceless) need not be fulfilled everywhere inside the closed surface.[42]

Consideration is given to an isentropic motion of a lossless, nondispersive, compressible, sound-propagating fluid (gas or liquid) which is described by the primary field quantities pressure p, (mass) density ρ, and particle velocity $\mathbf{v}$ as functions of space and time. As usual in acoustics, these quantities are considered with reference to an initial state without sound field, characterized by constants p_0, ρ_0, and $\mathbf{v}_0$ with $\mathbf{v}_0=0$. The gravitational field is neglected. $p-p_0$ is the acoustic pressure. In addition, the specific enthalpy h is considered according to

$$dh=\frac{1}{\rho}\,dp, \quad h-h_0=\int_{p_0}^{p}\frac{dp}{\rho(p)}. \tag{20}$$

The equation of state is assumed in the usual Taylor-series form[43]

$$p-p_0=c_0^2(\rho-\rho_0)\pm\cdots, \tag{21}$$

where c_0 is the low-amplitude speed of sound and terms with higher powers $(\rho-\rho_0)^n$ and containing the nonlinearity parameters such as the Fox–Wallace nonlinearity parameter[44] can be neglected here.

In the following, certain secondary field quantities such as the energy densities and particularly the radiation stresses will be calculated as functions of the primary field quantities. There will be two types of formulas: (a) exact formulas which are strictly valid and for which the primary field quantities are understood to be exact solutions of the underlying, nonlinear differential equations given in the subsequent sections; (b) formulas for secondary field quantities valid up to and including the second order in the primary field quantities, i.e., containing products and powers up to the order of two. These type (b) equations will be characterized by (l.a.) standing for ''linear acoustics.'' For example, the exact expression for the kinetic energy density is

$$E_{\mathrm{kin}}=\tfrac{1}{2}\rho\mathbf{v}\cdot\mathbf{v}, \tag{22}$$

but its linear-acoustics approximation is

$$E_{\mathrm{kin}}=\tfrac{1}{2}\rho_0\mathbf{v}\cdot\mathbf{v}. \quad \text{(l.a.)} \tag{23}$$

The potential acoustic energy density in the linear-acoustics approximation is

$$E_{\mathrm{pot}}=\frac{(p-p_0)^2}{2\rho_0 c_0^2}. \quad \text{(l.a.)} \tag{24}$$

For its exact expression, cf. Ref. 43. The total acoustic energy density is $E=E_{\mathrm{pot}}+E_{\mathrm{kin}}$.

Another useful Taylor series now involving powers of $(p-p_0)$ should be given, namely

$$\rho_0(h-h_0)=p-p_0-E_{\mathrm{pot}}\pm\cdots, \tag{25}$$

which follows from Eqs. (20), (21), and (24) (cf. also Ref. 43).

The fluid motion can be described in Eulerian or Lagrangian coordinates and the terminology chosen here is as follows. If the definition of a physical quantity is the same in both representations (which can be regarded as the rule; for the exceptions, cf. below), the same symbol is used for this quantity in both descriptions. The difference lies, however, in the operations with respect to the coordinates. So $\partial/\partial t$ is used for the time derivative in the Eulerian representation, i.e., at constant Eulerian coordinates. The symbol D/Dt stands for the time derivative in the Lagrangian representation, i.e., at constant Lagrangian coordinates. The overbar is used for the temporal average in the Eulerian representation, i.e., at constant Eulerian coordinates. The pointed brackets $\langle\ \rangle$ are used for the temporal average in the Lagrangian representation, i.e., at constant Lagrangian coordinates. The operators grad, div, and curl are understood with reference to Eulerian coordinates. When understood as referring to Lagrangian coordinates, they are written as Grad, Div, and Curl. Everything written above about vectors and matrices is of course also valid for the Lagrangian representation, but with the formal difference that the Eulerian coordinates x_i should be replaced by the Lagrangian coordinates X_i. The infinitesimal vectorial surface element in the Eulerian space is referred to as **da**; the infinitesimal vectorial surface element in the Lagrangian space is referred to as **dA**.

For all those physical quantities which express a certain quantity per unit area or per unit volume, it must be distinguished whether the area or the volume is understood in the Eulerian or in the Lagrangian space. One group of such quantities is given by the energy densities. Strictly speaking, formulas (22)–(24) above refer to the energy per unit volume of Eulerian space. The Eulerian energy densities must be multiplied by ρ_0/ρ to obtain the corresponding expressions giving the energy per unit volume of Lagrangian space. So it can be seen, for example, that Eq. (23) gives the exact expression for the kinetic energy per unit Lagrangian volume. Strictly speaking, different names and different symbols should be used for the Eulerian and for the Lagrangian energy densities, but in this paper the energy densities will be used only in the linear-acoustics approximation. So the dif-

ferences between the Eulerian and the Lagrangian forms are negligible in this case and formulas (23) and (24) will be used throughout.

A closer examination, however, is advised when dealing with the stress. The Cauchy stress **T** describes the contact force per unit Eulerian area. The stress expressing the contact force per unit Lagrangian area is a different quantity and is here referred to as the Piola–Kirchhoff–Boussinesq stress **K**. Its relation with **T** will be studied in Sec. IV.

In both representations, a transition from a time-dependent quantity to its temporal average will often be necessary in the following. The paper is not restricted to sinusoidal waveforms but applies to arbitrary waveforms provided two conditions are fulfilled: (1) The time-average value exists. (2) The temporal average of a time derivative is zero (cf., for example, Ref. 7). The latter condition is referred to in the text by assuming that the quantity in question is ''in a steady state.'' This can, for example, be fulfilled by a periodic quantity. The fields from ultrasonic diagnostic instruments are periodic inasmuch as the pulses are repeated at a constant rate. ''Periodic'' does not necessarily mean that the quantity averages to zero, but its time derivative averages to zero.

II. THE RADIATION STRESS IN EULERIAN COORDINATES

In a lossless fluid, the Cauchy stress is determined by the scalar pressure according to

$$\mathbf{T}=-p\mathbf{U}. \tag{26}$$

The fundamental Eulerian field equations[28,45] are the equation of motion

$$\frac{\partial \mathbf{v}}{\partial t}+(\mathbf{v}\cdot\mathrm{grad})\mathbf{v}=\frac{1}{\rho}\,\mathrm{div}\,\mathbf{T}=-\frac{1}{\rho}\,\mathrm{div}(p\mathbf{U})=-\frac{1}{\rho}\,\mathrm{grad}\,p \tag{27}$$

[cf. Eq. (9)] and the equation of continuity (mass conservation law)

$$\frac{\partial \rho}{\partial t}=-\mathrm{div}(\rho\mathbf{v}), \tag{28}$$

which can be combined using Eq. (11) to yield the momentum conservation law

$$\frac{\partial(\rho\mathbf{v})}{\partial t}=-\mathrm{div}\,\mathbf{M}, \tag{29}$$

with **M** being the momentum flux density,[45]

$$\mathbf{M}=-\mathbf{T}+\rho\mathbf{v}\otimes\mathbf{v}=p\mathbf{U}+\rho\mathbf{v}\otimes\mathbf{v}. \tag{30}$$

If the momentum density $\rho\mathbf{v}$ is in a steady state, it follows from Eq. (29) that

$$\mathrm{div}\,\overline{\mathbf{M}}=0. \tag{31}$$

Following the Brillouin approach, we understand the radiation stress in Eulerian coordinates as the negative temporal average of the momentum flux density and thus arrive at the Brillouin stress **S** with

$$\mathbf{S}=-\overline{\mathbf{M}}+p_0\mathbf{U}=-\overline{(p-p_0)}\mathbf{U}-\overline{\rho\mathbf{v}\otimes\mathbf{v}}, \tag{32}$$

where the addition of $p_0\mathbf{U}$ does not change the result of the closed-surface integral, which is a trivial statement but can also be formally inferred from the fact that according to Eq. (17), the unit matrix **U** can be given as a curl. Owing to Eq. (31),

$$\mathrm{div}\,\mathbf{S}=0. \tag{33}$$

The radiation force **f** is given by the vectorial surface integral

$$\mathbf{f}=\oiint(\mathbf{da}\cdot\mathbf{S}) \tag{34}$$

over a closed surface which is at rest in the Eulerian space; the surface element vector is understood to point outward.[46]

S is at least of the second order and when passing to the linear-acoustics approximation, a second-order expression for $\overline{p-p_0}$ is needed. This is briefly derived here as follows. The equation of motion (27) can be reformulated with the help of a well-known formula from conventional vector analysis[33] and of Eq. (20) according to

$$\frac{\partial \mathbf{v}}{\partial t}+\frac{1}{2}\,\mathrm{grad}(\mathbf{v}\cdot\mathbf{v})-\mathbf{v}\times\mathrm{curl}\,\mathbf{v}=-\mathrm{grad}\,h=-\mathrm{grad}(h-h_0). \tag{35}$$

If this is time averaged and multiplied by ρ_0 and if it is assumed that **v** is in a steady state and that the fluid motion is irrotational, the following is obtained with the help of Eqs. (23) and (25):

$$\mathrm{grad}(\rho_0\overline{(h-h_0)}+\overline{E_{\mathrm{kin}}})=\mathrm{grad}(\overline{p-p_0}-\overline{E_{\mathrm{pot}}}+\overline{E_{\mathrm{kin}}})=0. \quad \text{(l.a.)} \tag{36}$$

This leads to the following second-order result for $\overline{p-p_0}$ (the exact version can be found in Ref. 43):

$$\overline{p-p_0}=\overline{E_{\mathrm{pot}}}-\overline{E_{\mathrm{kin}}}+\mathrm{const}, \quad \text{(l.a.)} \tag{37}$$

which is ''Langevin's second relation.'' [21,22] It has repeatedly been given in the literature, cf. the citations in Ref. 20. The free constant const is important when this result is compared with other results from the literature,[2,19] but it makes no contribution to the closed-surface integral and can be neglected for the Langevin radiation force. The final, linear-acoustics result for the Brillouin stress which is the acoustic radiation stress in the Eulerian representation then is[47]

$$\mathbf{S}=(\overline{E_{\mathrm{kin}}}-\overline{E_{\mathrm{pot}}})\mathbf{U}-\rho_0\overline{\mathbf{v}\otimes\mathbf{v}}. \quad \text{(l.a.)} \tag{38}$$

This result, given here with reference to Brillouin and Langevin, can also be found in the literature under different circumstances. It is referred to as the ''wave-stress tensor'' in Ref. 48 and also appears, either explicitly or implicitly, in Refs. 39, 49, and 50 which deal with small targets and the application areas (b) and (c) from the Introduction (cf. also Ref. 2). The result of Eq. (38) has been used in Refs. 20 and 51 to calculate the radiation force in nonplane fields for application area (a) from the Introduction.

III. EXAMPLES OF THE RADIATION STRESS IN EULERIAN COORDINATES

Equation (38) is applied to three basic field types as follows.

A. Case 1: Axial field

It is assumed that in the field point considered, the particle velocity vector **v** lies only in one direction (including its negative), namely in direction No. 1, so that $\nu_2=0$ and $\nu_3=0$ all the time, which means, $\overline{\nu_2^2}=0$ and $\overline{\nu_3^2}=0$. In this case

$$\overline{E_{\mathrm{kin}}}=\tfrac{1}{2}\rho_0\overline{\nu_1^2} \qquad \text{(l.a.)} \quad (39)$$

and therefore

$$\mathbf{S}=\begin{pmatrix} -\overline{E} & 0 & 0 \\ 0 & \overline{E_{\mathrm{kin}}}-\overline{E_{\mathrm{pot}}} & 0 \\ 0 & 0 & \overline{E_{\mathrm{kin}}}-\overline{E_{\mathrm{pot}}} \end{pmatrix}. \qquad \text{(l.a.)} \quad (40)$$

The negative of the normal stress S_{11} in the axial direction, i.e., in direction No. 1, is given here by the temporal average of the (total) acoustic energy density E, a classic result. Its limitations, however, should be clearly noted as follows.

First, this is not a scalar, omnidirectional radiation ''pressure'' but it applies only to normal incidence in the axial direction. Second, this applies only to restricted field types, namely to plane waves of the progressive, standing or mixed type, for example. It is also valid on the symmetry axis of a nonplane field of the rotationally symmetrical type.[20] It is not valid under oblique incidence and for general, three-dimensional fields (particularly in off-axis positions) where often phase differences between the particle velocity components occur so that the above conditions concerning the $\overline{\nu_i^2}$ cannot be fulfilled, not even locally or with a rotation of the coordinate axes.

In the following examples a continuous, sinusoidal waveform is assumed but only to simplify things. Other periodic waveforms are also possible, of course.

B. Case 2: Plane-progressive wave in an arbitrary direction

It is assumed that

$$p-p_0=\hat{p}\,\cos(\omega t-\mathbf{k}\cdot\mathbf{x}) \qquad \text{(l.a.)} \quad (41)$$

and

$$\mathbf{v}=\frac{\hat{p}}{\rho_0\omega}\,\mathbf{k}\,\cos(\omega t-\mathbf{k}\cdot\mathbf{x}), \qquad \text{(l.a.)} \quad (42)$$

where $\hat{p}$ is the pressure amplitude, ω the circular frequency, and **k** the wave propagation vector, with $|\mathbf{k}|=k=\omega/c_0$ being the circular wave-number. The final result is

$$\mathbf{S}=-\frac{\hat{p}^2}{2\rho_0 c_0^2}\,\frac{\mathbf{k}\otimes\mathbf{k}}{k^2}. \qquad \text{(l.a.)} \quad (43)$$

If the wave is incident on a plane-perfect absorber of arbitrary orientation as expressed by its surface element vector **da**, the infinitesimal radiation force following from Eq. (43) is, due to Eq. (18),

$$\mathbf{df}=\mathbf{da}\cdot\mathbf{S}=-\frac{\hat{p}^2}{2\rho_0 c_0^2}\,\frac{\mathbf{da}\cdot\mathbf{k}}{k}\,\frac{\mathbf{k}}{k}, \qquad \text{(l.a.)} \quad (44)$$

where the surface element vector is understood as pointing outwards from the target; at the target front surface, the scalar product $\mathbf{da}\cdot\mathbf{k}$ thus has a negative value. The force vector **df** of Eq. (44) is in the direction of **k**, not of **da**. This means that irrespective of the absorber orientation, the radiation force on a perfect absorber lies in the direction of the incident wave.[52] This reflects the tensorial character of the underlying theory, whereas any scalar radiation ''pressure'' theory would lead to the result that the radiation force is normal to the time-averaged target surface.

The dependence of the radiation force on the absorber orientation has been measured by Herrey,[53] confirming the tensorial theory and Eq. (44). Experimental evidence for this is also reported in Ref. 54.

The plane-wave result obtained here is in line with the formula used in radiation-force practice in the case of a perfect absorber.[1,2]

C. Case 3: Superposition of two oblique plane waves of equal amplitude

It is assumed that

$$p-p_0=\hat{p}\{\cos(\omega t-\mathbf{k}^{(a)}\cdot\mathbf{x})+\cos(\omega t-\mathbf{k}^{(b)}\cdot\mathbf{x}-\alpha)\} \qquad \text{(l.a.)} \quad (45)$$

and that

$$\mathbf{v}=\frac{\hat{p}}{\rho_0\omega}\,\{\mathbf{k}^{(a)}\,\cos(\omega t-\mathbf{k}^{(a)}\cdot\mathbf{x})+\mathbf{k}^{(b)}\,\cos(\omega t-\mathbf{k}^{(b)}\cdot\mathbf{x}-\alpha)\}, \qquad \text{(l.a.)} \quad (46)$$

where, in order to simulate a reflection at a plane $x_1=\text{const}$,

$$\mathbf{k}^{(a)}=(k\,\cos\,\gamma,k\,\sin\,\gamma,0)$$

and (l.a.) (47)

$$\mathbf{k}^{(b)}=(-k\,\cos\,\gamma,k\,\sin\,\gamma,0)$$

is chosen, with γ being the angle of incidence and α a phase constant to allow for hard and soft reflection. In this case the final result is

$$\mathbf{S}=-\frac{\hat{p}^2}{\rho_0 c_0^2}\begin{pmatrix} \cos^2\,\gamma & 0 & 0 \\ 0 & \sin^2\,\gamma+\cos\,\beta & 0 \\ 0 & 0 & \cos^2\,\gamma\,\cos\,\beta \end{pmatrix}, \qquad \text{(l.a.)} \quad (48)$$

where β is an abbreviation for $2kx_1\,\cos\,\gamma-\alpha$. It follows from the result that in contrast to an absorber, no tangential radiation forces are exerted on a perfect reflector (cf. Ref. 53). Although it will not be shown in detail, the result obtained here for the normal stress S_{11} on the reflecting surface is in line with the common perfect-reflector formula[1,2] which is well established in radiation-force practice.

IV. THE STRESS IN LAGRANGIAN COORDINATES

Consider the transition of a medium from a reference (''undeformed'') state to the actual (''deformed'') state. The position of a particle in the reference state is given by the (Lagrangian) position vector $\mathbf{X}=(X_1,X_2,X_3)$; the position

of that particle in the actual state is given by the (Eulerian) position vector $\mathbf{x}=(x_1,x_2,x_3)$. The transition from **X** to **x** is described by

$$\mathbf{x}=\mathbf{x}(\mathbf{X})=\mathbf{X}+\mathbf{s}(\mathbf{X}), \tag{49}$$

s being the displacement vector and Eq. (49) forming the basis for two matrices, namely the displacement gradient[16]

$$\mathbf{H}=\text{Grad }\mathbf{s} \text{ with } H_{ij}=\frac{\partial s_j}{\partial X_i} \tag{50}$$

and the deformation gradient[16,55]

$$\mathbf{F}=\text{Grad }\mathbf{x}=\mathbf{U}+\mathbf{H} \text{ with } F_{ij}=\frac{\partial x_j}{\partial X_i}=\delta_{ij}+H_{ij}. \tag{51}$$

Consider a material surface in the reference state of the medium, a point **X** on it and two infinitesimal vectors $\mathbf{dX}^{(a)}$ and $\mathbf{dX}^{(b)}$ tangent to the surface and defining the vectorial surface element $\mathbf{dA}=\mathbf{dX}^{(a)}\times\mathbf{dX}^{(b)}$. Consider the actual, deformed surface, the point **x** on it and the two corresponding vectors $\mathbf{dx}^{(a)}$ and $\mathbf{dx}^{(b)}$ now forming the actual surface element $\mathbf{da}=\mathbf{dx}^{(a)}\times\mathbf{dx}^{(b)}$. The transition $\mathbf{dX}\Rightarrow\mathbf{dx}$ of an infinitesimal vector is generally given by[33,56]

$$\mathbf{dx}=(\mathbf{dX}\cdot\text{Grad})\mathbf{x}=\mathbf{dX}\cdot\mathbf{F}. \tag{52}$$

The transition law for the surface element vector now follows from Eqs. (19) and (52) to be

$$\begin{aligned}\mathbf{da}&=\mathbf{dx}^{(a)}\times\mathbf{dx}^{(b)}=(\mathbf{dX}^{(a)}\cdot\mathbf{F})\times(\mathbf{dX}^{(b)}\cdot\mathbf{F})\\&=(\mathbf{dX}^{(a)}\times\mathbf{dX}^{(b)})\cdot\mathbf{G}=\mathbf{dA}\cdot\mathbf{G},\end{aligned} \tag{53}$$

with the transition matrix **G** being the adjoint of the deformation gradient, namely

$$\mathbf{G}=\text{adj}(\text{Grad }\mathbf{x})=\text{adj}(\mathbf{F})=(\mathbf{F}^T)^{-1}\text{ det }\mathbf{F} \tag{54}$$

or, in detail,

$$\begin{aligned}\mathbf{G}&=\begin{pmatrix}F_{22}F_{33}-F_{23}F_{32} & F_{23}F_{31}-F_{21}F_{33} & F_{21}F_{32}-F_{22}F_{31}\\ F_{13}F_{32}-F_{12}F_{33} & F_{11}F_{33}-F_{13}F_{31} & F_{12}F_{31}-F_{11}F_{32}\\ F_{12}F_{23}-F_{13}F_{22} & F_{13}F_{21}-F_{11}F_{23} & F_{11}F_{22}-F_{12}F_{21}\end{pmatrix}\\&=(1+\text{trace }\mathbf{H})\mathbf{U}-\mathbf{H}^T+\text{adj}(\mathbf{H}),\end{aligned} \tag{55}$$

where Eq. (51) has been used and where trace $\mathbf{H}=\text{Div }\mathbf{s}$.

If a stress **T** exists at the actual point **x**, the actual contact force **df** exerted across the actual surface element **da** is $\mathbf{df}=\mathbf{da}\cdot\mathbf{T}$. If, on the other hand, the actual force is to be described by a stress acting across the reference surface element **dA**, this leads to the Piola–Kirchhoff–Boussinesq stress **K** according to $\mathbf{df}=\mathbf{dA}\cdot\mathbf{K}$. A comparison with Eq. (53) immediately yields the result[57] $\mathbf{K}=\mathbf{G}\cdot\mathbf{T}$ or, more specifically,

$$\mathbf{K}(\mathbf{X})=\mathbf{G}(\mathbf{X})\cdot\mathbf{T}(\mathbf{x}(\mathbf{X})), \tag{56}$$

which, for a lossless fluid and due to Eq. (26), means that

$$\mathbf{K}=-p\mathbf{G}. \tag{57}$$

It can be seen that the stress in the Lagrangian representation in a lossless fluid does not in general reduce to a scalar quantity. So strictly speaking, the scalar pressure p remains a Eulerian quantity (=force per unit Eulerian area) even if used in the Lagrangian representation, i.e., if its value is considered as a function of Lagrangian coordinates.

The transition matrix **G** has the following properties: It is symmetric if the deformation gradient is symmetric. Due to Eqs. (15) and (54), it can be expressed as a Curl (i.e., its surface integral over a closed surface in the Lagrangian space vanishes); as a result and owing to Eq. (13),

$$\text{Div }\mathbf{G}=0 \tag{58}$$

which can also be directly obtained by applying Eq. (4) to Eq. (55).

V. THE RADIATION STRESS IN LAGRANGIAN COORDINATES

In the Lagrangian representation, the displacement vector **s** is considered to be the fourth primary field quantity in addition to pressure, density, and particle velocity. The basic relation reads

$$\mathbf{v}=\frac{D\mathbf{s}}{Dt}=\frac{D\mathbf{x}}{Dt}. \tag{59}$$

The fundamental field equations[26,28,37] are the equation of motion

$$\mathbf{F}\cdot\frac{D\mathbf{v}}{Dt}=-\frac{1}{\rho}\text{ Grad }p=-\text{Grad }h \tag{60}$$

and the equation of continuity

$$\rho\text{ det }\mathbf{F}=\rho_0. \tag{61}$$

When the equation of motion is premultiplied by $\mathbf{F}^{-1}$ and the equation of continuity and Eq. (54) are taken into account, the following relationship is obtained:

$$\begin{aligned}\frac{D\mathbf{v}}{Dt}&=-\frac{1}{\rho}\mathbf{F}^{-1}\cdot\text{Grad }p=-\frac{1}{\rho_0}(\text{Grad }p)\cdot(\mathbf{F}^T)^{-1}\text{ det }\mathbf{F}\\&=-\frac{1}{\rho_0}(\text{Grad }p)\cdot\mathbf{G}.\end{aligned} \tag{62}$$

Owing to Eqs. (8), (57), and (58), this equation can be reformulated:

$$\frac{D\mathbf{v}}{Dt}=-\frac{1}{\rho_0}\text{Div}(p\mathbf{G})=\frac{1}{\rho_0}\text{ Div }\mathbf{K}. \tag{63}$$

This equation is completely equivalent to its analogue in Eulerian coordinates, Eq. (27). The left-hand side of Eq. (27) contains the material time derivative of **v** in Eulerian coordinates, the left-hand side of Eq. (63) contains the (material) time derivative of **v** in Lagrangian coordinates. ρ in Eq. (27) is the mass per unit volume of Eulerian space. ρ_0 in Eq. (63) is the mass per unit volume of Lagrangian space. **T** in Eq. (27) is the Eulerian stress, and **K** in Eq. (63) is the Lagrangian stress.

If **v** is in a steady state (in the Lagrangian representation), it follows from Eq. (63) that

$$\text{Div}\langle\mathbf{K}\rangle=0. \tag{64}$$

Following the Brillouin approach, we understand the radiation stress in Lagrangian coordinates as the temporal average of the Piola–Kirchhoff–Boussinesq stress **K** and thus obtain the PKB stress **B** with[58]

$$\mathbf{B}=\langle\mathbf{K}\rangle+p_0\langle\mathbf{G}\rangle=-\langle(p-p_0)\mathbf{G}\rangle, \tag{65}$$

where, as **G** can be given as a Curl, the addition of $p_0\langle\mathbf{G}\rangle$ does not change the result of the closed-surface integral. Owing to Eqs. (58) and (64),

$$\text{Div } \mathbf{B}=0 \tag{66}$$

[Eq. (58) is valid at any instant and thus also on the temporal average]. The radiation force **f** is given by the vectorial surface integral

$$\mathbf{f}=\oiint (\mathbf{dA}\cdot\mathbf{B}) \tag{67}$$

over a closed surface which is at rest in the Lagrangian space; the surface element vector is again understood to point outward.

Equation (65) is exact. In order to get its second-order approximation to be used in linear acoustics, the first-order approximation of **G** is needed. Using Eq. (55) up to the first order, the following can be obtained for **B** up to the second order:

$$\mathbf{B}=-\langle p-p_0\rangle\mathbf{U}-\langle(p-p_0)\text{Div }\mathbf{s}\rangle\mathbf{U}+\langle(p-p_0)\mathbf{H}^T\rangle. \tag{68}$$

(l.a.)

For the first term of the right-hand side, the second-order approximation of $\langle p-p_0\rangle$ is needed. This can be derived as follows. As

$$\frac{D\mathbf{F}}{Dt}=\frac{D\mathbf{H}}{Dt}=\frac{D}{Dt}\text{ Grad }\mathbf{s}=\text{Grad }\mathbf{v} \tag{69}$$

and, due to Eq. (6) (and cf. Ref. 24),

$$\frac{D\mathbf{F}}{Dt}\cdot\mathbf{v}=(\text{Grad }\mathbf{v})\cdot\mathbf{v}=\frac{1}{2}\text{ Grad}(\mathbf{v}\cdot\mathbf{v}), \tag{70}$$

the equation of motion (60) can be reformulated as

$$\begin{aligned}\mathbf{F}\cdot\frac{D\mathbf{v}}{Dt}&=\frac{D}{Dt}(\mathbf{F}\cdot\mathbf{v})-\frac{D\mathbf{F}}{Dt}\cdot\mathbf{v}\\&=\frac{D}{Dt}(\mathbf{F}\cdot\mathbf{v})-\frac{1}{2}\text{ Grad}(\mathbf{v}\cdot\mathbf{v})\\&=-\text{Grad }h=-\text{Grad}(h-h_0).\end{aligned} \tag{71}$$

If this is time averaged and multiplied by ρ_0 and if it is assumed that $\mathbf{F}\cdot\mathbf{v}$ is in a steady state (in the Lagrangian representation), the following is obtained with the help of Eqs. (23) and (25):

$$\begin{aligned}\text{Grad}(\rho_0\langle h-h_0\rangle-\langle E_{\text{kin}}\rangle)&=\text{Grad}(\langle p-p_0\rangle-\langle E_{\text{pot}}\rangle\\&\quad-\langle E_{\text{kin}}\rangle)=0.\end{aligned} \tag{72}$$

(l.a.)

This leads to the following second-order result for $\langle p-p_0\rangle$:

$$\langle p-p_0\rangle=\langle E_{\text{pot}}\rangle+\langle E_{\text{kin}}\rangle+\text{Const}=\langle E\rangle+\text{Const}, \tag{73}$$

(l.a.)

which is ''Langevin's first relation.''[7,22,cf. also 12,24] The free constant Const does not make a contribution to the closed-surface integral and can thus be neglected for the Langevin radiation force. The linear-acoustics result for the PKB stress now reads

$$\mathbf{B}=-\langle E\rangle\mathbf{U}-\langle(p-p_0)\text{Div }\mathbf{s}\rangle\mathbf{U}+\langle(p-p_0)\mathbf{H}^T\rangle. \tag{74}$$

(l.a.)

This can be further simplified. Due to Eqs. (21), (50), (51), and (61), the following holds approximately (first-order approximation):

$$\text{Div }\mathbf{s}=\text{trace }\mathbf{H}=\det\mathbf{F}-1=\frac{\rho_0}{\rho}-1=-\frac{\rho-\rho_0}{\rho_0}=-\frac{p-p_0}{\rho_0 c_0^2} \tag{75}$$

(l.a.)

and therefore

$$(p-p_0)\text{Div }\mathbf{s}=-2E_{\text{pot}}. \tag{76}$$

(l.a.)

The final, linear-acoustics result for the PKB stress which is the acoustic radiation stress in the Lagrangian representation then is

$$\mathbf{B}=(\langle E_{\text{pot}}\rangle-\langle E_{\text{kin}}\rangle)\mathbf{U}+\langle(p-p_0)\mathbf{H}^T\rangle. \tag{77}$$

(l.a.)

Another remark should be made. In a lossless fluid, acoustic streaming usually does not occur, but it is interesting to consider the potential influence of such a streaming on the ''steady-state'' assumptions in the two representations. If the streaming itself is in a steady state, i.e., if it is temporally constant at constant Eulerian coordinates, the total field (acoustic field plus streaming) is in a steady state in the Eulerian sense, but not in the Lagrangian sense. In this case the fluid particles move through the field, which means through field points exhibiting different amplitudes of the acoustic quantities, and even leave the beam into a space where the acoustic quantities are practically zero, and so the Lagrangian time average of any field quantity is meaningless.[59] It might appear that this is an advantage of the Eulerian representation but it had to be assumed there that the field is irrotational and this is usually not the case in an acoustic streaming, strictly speaking.[60] Conditions of no streaming thus are obviously important for both of Langevin's relations.

VI. EXAMPLES OF THE RADIATION STRESS IN LAGRANGIAN COORDINATES

Equation (77) is applied to the three basic field types of Sec. III. Equivalent formulas where only $\mathbf{x}=(x_1,x_2,x_3)$ is to be replaced by $\mathbf{X}=(X_1,X_2,X_3)$ are not explicitly repeated. In addition, the displacement vector **s** has to be considered here as it appears in Eq. (77) in the displacement gradient **H**.

A. Case 1: Axial field

It is assumed that ν_1 and s_1 are the only nonzero components. It is also assumed that $H_{11}=\partial s_1/\partial X_1$ is the only

nonzero element of the displacement gradient, i.e., that all other elements vanish. This can be seen as follows. In the case of a (progressive or standing or mixed) plane wave in the X_1 direction, it is obvious. In the case of points on the symmetry axis of a nonplane rotationally symmetrical field, the derivatives with respect to X_2 and X_3 are zero for symmetry reasons. As s_2 and s_3 are zero on the axis, their derivatives with respect to X_1 vanish.

Consequently, $H_{11}=\text{trace}\ \mathbf{H}=\text{Div}\ \mathbf{s}$ here, and it follows from Eq. (76) that $(p-p_0)H_{11}=-2E_{\text{pot}}$. This leads to the final result

$$\mathbf{B}=\begin{pmatrix} -\langle E\rangle & 0 & 0 \\ 0 & \langle E_{\text{pot}}\rangle-\langle E_{\text{kin}}\rangle & 0 \\ 0 & 0 & \langle E_{\text{pot}}\rangle-\langle E_{\text{kin}}\rangle \end{pmatrix}. \qquad \text{(l.a.)} \quad (78)$$

B. Case 2: Plane-progressive wave in an arbitrary direction

In addition to the Lagrangian equivalents of Eqs. (41) and (42),

$$\mathbf{s}=\frac{\hat{p}}{\rho_0\omega^2}\,\mathbf{k}\,\sin(\omega t-\mathbf{k}\cdot\mathbf{X}) \qquad \text{(l.a.)} \quad (79)$$

is assumed. The final result is

$$\mathbf{B}=-\frac{\hat{p}^2}{2\rho_0 c_0^2}\,\frac{\mathbf{k}\otimes\mathbf{k}}{k^2}. \qquad \text{(l.a.)} \quad (80)$$

This is completely equivalent to Eq. (43) and so the orientation dependence of the radiation force here is completely equivalent to that of Eq. (44) which need not be repeated.

C. Case 3: Superposition of two oblique plane waves of equal amplitude

In addition to the Lagrangian equivalents of Eqs. (45) and (46),

$$\mathbf{s}=\frac{\hat{p}}{\rho_0\omega^2}\,\{\mathbf{k}^{(a)}\,\sin(\omega t-\mathbf{k}^{(a)}\cdot\mathbf{X})+\mathbf{k}^{(b)}\,\sin(\omega t-\mathbf{k}^{(b)}\cdot\mathbf{X}-\alpha)\} \qquad \text{(l.a.)} \quad (81)$$

is assumed. The final result is

$$\mathbf{B}=-\frac{\hat{p}^2}{\rho_0 c_0^2}\begin{pmatrix} \cos^2\gamma & 0 & 0 \\ 0 & \sin^2\gamma+(\sin^2\gamma-\cos^2\gamma)\cos\beta & 0 \\ 0 & 0 & -\cos^2\gamma\cos\beta \end{pmatrix}. \qquad \text{(l.a.)} \quad (82)$$

VII. INTERCOMPARISON AND UNIFIED RESULT

In the linear-acoustics approximation, the solutions for the primary field quantities in the Eulerian and Lagrangian representations are identical, cf. the linear wave equations (84) and (89) given below. The results obtained in the two representations can therefore be directly compared with each other. In case 2 which deals with the plane-progressive wave, the radiation stresses of Eqs. (43) and (80) have turned out to be identical. In cases 1 [Eqs. (40) and (78)] and 3 [Eqs. (48) and (82)], the normal stress numbered 11 which is relevant to the radiation force under the conditions assumed is also identical in the two representations, but the other normal stresses numbered 22 and 33 are not. So it appears that the Brillouin and the PKB stress are not in general identical but that fortunately, they differ only in ''unimportant'' details.[61] This can also be seen in other field examples which are not shown here for the sake of brevity. The reasons for this behavior will finally be clarified as follows.

The fluid motion is assumed to be irrotational and described by a velocity potential. Although this will only be considered in the linear-acoustics approximation, the formal distinction between the Eulerian and Lagrangian versions is maintained.

The first-order Eulerian velocity potential is ϕ with

$$p-p_0=\rho_0\,\frac{\partial\phi}{\partial t},\quad \mathbf{v}=-\,\text{grad}\ \phi \qquad \text{(l.a.)} \quad (83)$$

and the well-known linear wave equation

$$\frac{\partial^2\phi}{\partial t^2}=c_0^2\ \text{div grad}\ \phi, \qquad \text{(l.a.)} \quad (84)$$

which follows from Eqs. (27) and (28) by linearization. The following identity holds under the steady-state condition:[62]

$$\overline{\frac{\partial}{\partial t}\left(\phi\,\frac{\partial\phi}{\partial t}\right)}=\overline{\frac{\partial\phi}{\partial t}\,\frac{\partial\phi}{\partial t}}+\overline{\phi\,\frac{\partial^2\phi}{\partial t^2}}=0, \qquad \text{(l.a.)} \quad (85)$$

which, where some formulas from conventional vector analysis[33] and Eq. (84) are used, leads to

$$\begin{aligned}\overline{E_{\text{kin}}}-\overline{E_{\text{pot}}}&=\frac{1}{2}\,\rho_0\left\{\overline{\text{grad}\ \phi\cdot\text{grad}\ \phi}-\frac{1}{c_0^2}\,\overline{\frac{\partial\phi}{\partial t}\,\frac{\partial\phi}{\partial t}}\right\}\\ &=\tfrac{1}{2}\,\rho_0\{\overline{\text{grad}\ \phi\cdot\text{grad}\ \phi}+\overline{\phi\ \text{div grad}\ \phi}\}\\ &=\tfrac{1}{2}\,\rho_0\ \text{div}(\overline{\phi\ \text{grad}\ \phi})\\ &=\tfrac{1}{4}\,\rho_0\ \text{div grad}\ \overline{\phi^2}\end{aligned} \qquad \text{(l.a.)} \quad (86)$$

and, when Eqs. (7), (14), and (38) are made use of, to

$$\begin{aligned}\mathbf{S}&=\rho_0\{\tfrac{1}{4}\,\mathbf{U}\ \text{div grad}\ \overline{\phi^2}-\overline{\text{grad}\ \phi\otimes\text{grad}\ \phi}\}\\ &=\rho_0\{\overline{\phi\ \text{grad grad}\ \phi}-\tfrac{1}{4}\,\mathbf{U}\ \text{div grad}\ \overline{\phi^2}\\ &\quad-\tfrac{1}{2}\ \text{curl curl}(\overline{\phi^2}\mathbf{U})\}.\end{aligned} \qquad \text{(l.a.)} \quad (87)$$

The first-order velocity potential formally considered in Lagrangian coordinates is ψ with

$$p-p_0=\rho_0\frac{D\psi}{Dt},\quad \mathbf{v}=-\text{Grad }\psi \qquad \text{(l.a.)} \quad (88)$$

and the linear wave equation

$$\frac{D^2\psi}{Dt^2}=c_0^2\ \text{Div Grad}\ \psi \qquad \text{(l.a.)} \quad (89)$$

which follows from Eqs. (60) and (61) by linearization. The following identity holds under the steady-state condition:

$$\left\langle \frac{D}{Dt}(\psi\mathbf{H}^T)\right\rangle=\left\langle\frac{D\psi}{Dt}\mathbf{H}^T\right\rangle+\left\langle\psi\frac{D\mathbf{H}^T}{Dt}\right\rangle=0, \qquad \text{(l.a.)} \quad (90)$$

leading to

$$\begin{aligned}\langle(p-p_0)\mathbf{H}^T\rangle&=\rho_0\left\langle\frac{D\psi}{Dt}\mathbf{H}^T\right\rangle\\&=-\rho_0\left\langle\psi\frac{D\mathbf{H}^T}{Dt}\right\rangle\\&=-\rho_0\langle\psi(\text{Grad }\mathbf{v})^T\rangle. \qquad \text{(l.a.)}\end{aligned} \quad (91)$$

Due to Eq. (88) (right), the matrix Grad $\mathbf{v}$ is symmetric, and

$$\begin{aligned}\langle(p-p_0)\mathbf{H}^T\rangle&=-\rho_0\langle\psi\ \text{Grad}\ \mathbf{v}\rangle\\&=\rho_0\langle\psi\ \text{Grad Grad}\ \psi\rangle. \qquad \text{(l.a.)}\end{aligned} \quad (92)$$

On the other hand, exactly like Eqs. (85) and (86),

$$\langle E_{\text{pot}}\rangle-\langle E_{\text{kin}}\rangle=-\tfrac{1}{4}\rho_0\ \text{Div Grad}\langle\psi^2\rangle \qquad \text{(l.a.)} \quad (93)$$

and, making use of Eqs. (7), (14), and (77),

$$\begin{aligned}\mathbf{B}&=\rho_0\{\langle\psi\ \text{Grad Grad}\ \psi\rangle-\tfrac{1}{4}\mathbf{U}\ \text{Div Grad}\langle\psi^2\rangle\}\\&=\rho_0\{\tfrac{1}{4}\mathbf{U}\ \text{Div Grad}\langle\psi^2\rangle-\langle\text{Grad }\psi\otimes\text{Grad }\psi\rangle\\&\quad+\tfrac{1}{2}\text{Curl Curl}(\langle\psi^2\rangle\mathbf{U})\}. \qquad \text{(l.a.)}\end{aligned} \quad (94)$$

Equations (87) and (94) give the final results for the radiation stress in the Eulerian and Lagrangian representations in the linear-acoustics approximation. It can be seen that the difference between the two expressions is a curl, i.e., it is ''unimportant'' inasmuch as it is irrelevant to the radiation force on a target entirely surrounded by the sound-propagating fluid. The two results can be unified by adding or subtracting, respectively, a curl as follows, taking into account Eqs. (7) and (14):

$$\begin{aligned}&\mathbf{S}+\tfrac{1}{4}\rho_0\ \text{curl curl}(\overline{\phi^2}\mathbf{U})\\&\quad=\tfrac{1}{2}\rho_0\{\overline{\phi\ \text{grad grad}\ \phi}-\overline{\text{grad }\phi\otimes\text{grad }\phi}\},\\&\qquad\qquad\text{(l.a.)}\end{aligned} \quad (95)$$

$$\begin{aligned}&\mathbf{B}-\tfrac{1}{4}\rho_0\ \text{Curl Curl}(\langle\psi^2\rangle\mathbf{U})\\&\quad=\tfrac{1}{2}\rho_0\{\langle\psi\,\text{Grad Grad }\psi\rangle-\langle\text{Grad }\psi\otimes\text{Grad }\psi\rangle\}.\\&\qquad\qquad\text{(l.a.)}\end{aligned} \quad (96)$$

Equations (95) and (96) give the unified result for the (Langevin) acoustic radiation stress in a lossless fluid in the linear-acoustics approximation.[63]

APPENDIX: THE LANGEVIN APPROACH

According to the Langevin approach[7,12,24] and contrary to Eq. (65), the acoustic radiation stress in Lagrangian coordinates is assumed to be

$$\mathbf{B}'=-\langle p-p_0\rangle\mathbf{U} \quad (A1)$$

which represents the first term of Eq. (68) and which according to Eq. (73) is

$$\mathbf{B}'=-\langle E\rangle\mathbf{U} \qquad \text{(l.a.)} \quad (A2)$$

in the linear-acoustics approximation. Due to Eq. (9), its divergence is

$$\text{Div }\mathbf{B}'=-\text{Grad}\langle E\rangle, \qquad \text{(l.a.)} \quad (A3)$$

which indeed vanishes in a plane wave but is not in general zero in other fields, cf., for example, the fields from modern ultrasonic diagnostic instruments which are highly focussed and in which, as a result, the energy density is very different at different points in the field.

According to this approach, the radiation force is given by the integral[64]

$$\begin{aligned}\mathbf{f}=\oiint(\mathbf{dA}\cdot\mathbf{B}')&=-\oiint\langle p-p_0\rangle\mathbf{dA}\\&=-\oiint\langle E\rangle\mathbf{dA} \qquad \text{(l.a.)}\end{aligned} \quad (A4)$$

over the target surface.

Equation (A2) is applied to the three basic field types of Secs. III and VI as follows.

Case 1: Axial field

This is no special case of the Langevin approach, i.e., the general result of Eq. (A2) is retained. A comparison with Eqs. (40) and (78) shows that there is agreement in the ''important'' stress component which is the normal stress numbered 11. It is probably for this reason that the Langevin approach has been regarded as meaningful.

Case 2: Plane-progressive wave in an arbitrary direction

The result in this case is

$$\mathbf{B}'=-\frac{\hat{p}^2}{2\rho_0c_0^2}\mathbf{U}, \qquad \text{(l.a.)} \quad (A5)$$

to be compared with Eqs. (43) and (80). If the wave is incident on an absorber surface of orientation **dA** the infinitesimal radiation force is

$$\mathbf{df}=\mathbf{dA}\cdot\mathbf{B}'=-\frac{\hat{p}^2}{2\rho_0c_0^2}\mathbf{dA}, \qquad \text{(l.a.)} \quad (A6)$$

which means that the radiation force given here is normal to the absorber surface and not in the direction of the incident wave as in the corresponding results of Secs. III and VI.

Case 3: Superposition of two oblique plane waves of equal amplitude

The result in this case is

$$\mathbf{B}' = -\frac{\hat{p}^2}{\rho_0 c_0^2}(1+\sin^2\gamma\cos\beta)\mathbf{U}. \qquad \text{(l.a.)} \quad \text{(A7)}$$

It can be seen that with the exception of the special case $\cos\beta=-1$, there is no agreement with Eqs. (48) and (82), not even for the ''important'' stress component which is the normal stress numbered 11.

The Langevin approach is obviously in contradiction to important experimental facts (cf. also Ref. 53) and does not in general agree with the Brillouin approach worked out in this paper and shown in Sec. II to be in line with other radiation force theories of the literature.

Finally, the following is worth mentioning. In Ref. 20 the radiation stress of Eq. (38) has been integrated over a field cross section of infinite extent in a piston field including diffraction effects, in order to obtain a value for the radiation force on an absorber of infinite size as a theoretical reference result for practical radiation force measurements. If, however, the radiation stress according to the Langevin approach had been used for this purpose, no result would have been obtained as the improper integral would not exist, or in other words, any result might have been obtained if only the radius of the integration area had been chosen large enough.

[1] IEC 1161: ''Ultrasonic power measurement in liquids in the frequency range 0.5 MHz to 25 MHz'' (International Electrotechnical Commission, Geneva, 1992).

[2] K. Beissner, in *Ultrasonic Exposimetry*, edited by M. C. Ziskin and P. A. Lewin (CRC, Boca Raton, 1993), pp. 127–142.

[3] F. Dunn, A. J. Averbuch, and W. D. O'Brien, Jr., Acustica **38**, 58–61 (1977); R. C. Chivers and L. W. Anson, J. Acoust. Soc. Am. **72**, 1695–1705 (1982).

[4] R. E. Apfel, in *Fortschritte der Akustik–DAGA'90* (DPG-GmbH, Bad Honnef, 1990), pp. 19–36.

[5] P. Collas, M. Barmatz, and C. Shipley, J. Acoust. Soc. Am. **86**, 777–787 (1989); E. G. Lierke, Acust. Acta Acust. **82**, 220–237 (1996). The latter reference is in German, but with a comprehensive list of international references.

[6] Lord Rayleigh, Philos. Mag. **3**, 338–346 (1902); Philos. Mag. **10**, 364–374 (1905).

[7] P. Biquard, Rev. Acoust. **1**, 93–109 (1932) (in French).

[8] L. Brillouin, Ann. Phys. (Paris) **4**, 528–586 (1925) (in French); *Tensors in Mechanics and Elasticity* (Academic, New York, 1964).

[9] G. Hertz and H. Mende, Z. Phys. **114**, 354–367 (1939) (in German).

[10] J. A. Rooney and W. L. Nyborg, Am. J. Phys. **40**, 1825–1830 (1972).

[11] C. E. Tschiegg, M. Greenspan, and D. G. Eitzen, J. Res. Natl. Bur. Stand. **88**, 91–103 (1983) (the author was one of the participants in this intercomparison); K. Beissner, W. A. Oosterbaan, R. T. Hekkenberg, and A. Shaw, Acust. Acta Acust. **82**, 450–458, 671 (1996).

[12] B.-T. Chu and R. E. Apfel, J. Acoust. Soc. Am. **72**, 1673–1687 (1982). As far as the Langevin radiation force is concerned, their claim to deal with an exact, three-dimensional theory is expressed in particular in the ''Notes added in proof'' on p. 1686.

[13] W. L. Nyborg and J. A. Rooney, J. Acoust. Soc. Am. **75**, 263–264 (1984); B.-T. Chu and R. E. Apfel, J. Acoust. Soc. Am. **75**, 1003–1004 (1984).

[14] J. H. Cantrell, Jr., Phys. Rev. B **30**, 3214–3220 (1984).

[15] C. Truesdell and R. A. Toupin, in *Handbuch der Physik (Encyclopedia of Physics)*, edited by S. Flügge (Springer-Verlag, Berlin, 1960), Vol. III/1, pp. 226–793.

[16] C. Truesdell and W. Noll, in *Handbuch der Physik (Encyclopedia of Physics)*, edited by S. Flügge (Springer-Verlag, Berlin, 1965), Vol. III/3, pp. 1–579.

[17] Cf., however, the discussion of the influence of radiation stresses (not on targets but) on flows as, for example, by D. G. Andrews and M. E. McIntyre, J. Fluid Mech. **89**, 609–646 (1978).

[18] He stated that the closed surface integrals of the two versions of radiation stresses are both zero (in the no-target case) and then concluded that the radiation stresses themselves are equal to each other, but this conclusion is not necessarily true (zero-divergence quantities are not necessarily equal to one another; cf., for example, the various zero-divergence quantities appearing in this paper, including the unit matrix and, in addition, the zero matrix, too).

[19] K. Beissner, Acustica **57**, 1–4 (1985).

[20] K. Beissner, J. Sound Vib. **93**, 537–548 (1984); cf. also J. Acoust. Soc. Am. **76**, 1505–1510 (1984).

[21] P. Biquard, Rev. Acoust. **1**, 315–335 (1932) (in French).

[22] K. Beissner, J. Acoust. Soc. Am. **79**, 1610–1612 (1986).

[23] E. B. Miller and D. G. Eitzen, IEEE Trans. Sonics Ultrason. **26**, 28–37 (1979).

[24] Z.-Y. Jiang and J. F. Greenleaf, J. Acoust. Soc. Am. **100**, 741–747 (1996).

[25] H. Jeffreys, *Cartesian Tensors* (Cambridge U.P., Cambridge, 1931).

[26] R. R. Long, *Mechanics of Solids and Fluids* (Prentice-Hall, Englewood Cliffs, NJ, 1961).

[27] S. F. Borg, *Matrix-Tensor Methods in Continuum Mechanics* (World Scientific, Singapore, 1990).

[28] J. Serrin, in *Handbuch der Physik (Encyclopedia of Physics)*, edited by S. Flügge and C. Truesdell (Springer-Verlag, Berlin, 1959), Vol. VIII/1, pp. 125–263.

[29] This does not automatically imply that all these quantities are general tensors in curvilinear coordinates, i.e., in the broader sense of tensor analysis.

[30] D. Mintzer, P. Tamarkin, and R. B. Lindsay, in *American Institute of Physics Handbook*, edited by D. E. Gray (McGraw-Hill, New York, 1972), 3rd ed., pp. 2-2–2-18.

[31] In Ref. 15 and obviously also in Ref. 16, the stress indices are understood in the reverse sense.

[32] The dot product of a vector $\mathbf{q}$ and a matrix $\mathbf{Q}$ is a vector according to $(\mathbf{q}\cdot\mathbf{Q})_j=\Sigma_i q_i Q_{ij}$ or $(\mathbf{Q}\cdot\mathbf{q})_i=\Sigma_j Q_{ij} q_j$.

[33] I. N. Bronshtein and K. A. Semendyayev, *Handbook of Mathematics* (Harri Deutsch, Frankfurt/Van Nostrand Reinhold, New York, 1985), 3rd ed.

[34] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics, Part I* (McGraw-Hill, New York, 1953).

[35] It should, however, be noted that the curl here often appears in combination with another curl or with a vector product, cf. Eqs. (14), (15), (17), and others. So this restriction might not be necessary.

[36] J. L. Ericksen, in *Handbuch der Physik (Encyclopedia of Physics)*, edited by S. Flügge (Springer-Verlag, Berlin, 1960), Vol III/1, pp. 794–858.

[37] H. Lamb, *Hydrodynamics* (Dover, New York, First American Edition 1945).

[38] Conditions for the applicability of the Stokes integral theorem are that the matrix and its curl exist, and are continuous, in a region containing the surface under consideration and its border line. The theorem is applied here twice, namely separately to both hemispheres (or generally, to both parts of the closed surface), and not to the whole sphere and its entire interior. So the rule mentioned is valid even with a target in the interior of the sphere (or generally, of the closed surface).

[39] P. J. Westervelt, J. Acoust. Soc. Am. **29**, 26–29 (1957).

[40] K. Beissner, Acustica **62**, 255–263 (1987).

[41] *Nota bene:* as long as the radiation stress is a zero-divergence quantity which, in general, is not the case in lossy fluids, but this need not be discussed here as lossy fluids are not dealt with in this paper.

[42] The divergence of the radiation stress is assumed to be nonzero in the interior of the target or at its surface. The target usually represents a discontinuity with respect to the sound-propagating medium and its acoustic properties. The case that the divergence of the radiation stress is zero even throughout the target and its surface is trivial and is equivalent to the no-target case.

[43] K. Beissner and S. N. Makarov, J. Acoust. Soc. Am. **97**, 898–905 (1995). K. Beissner, J. Acoust. Soc. Am. **99**, 1244–1247 (1996). K. Beissner, in *Proc. Ultrasonics World Congress 1995*, edited by J. Herbertz (GEFAU, Duisburg, 1995), pp. 41–49.

[44] F. E. Fox and W. A. Wallace, J. Acoust. Soc. Am. **26**, 994–1006 (1954). As far as Langevin's relations, Eqs. (37) and (73), are concerned, the Fox–Wallace nonlinearity parameter appears in their third-order approximation only or, on the other hand, it may appear in the constant which depends on the boundary conditions. When the Langevin case (where the

constant is zero or can be neglected) is considered up to the second order, the nonlinearity parameter is irrelevant.

[45] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon, Oxford, 1963), 2nd imp.

[46] If it were understood as pointing inwards, a minus sign would appear as, for example, in Ref. 20.

[47] In the literature it is often stated that the potential acoustic energy density is equal to the kinetic one (their difference being zero) or in a more indirect way, that the total acoustic energy density is twice the kinetic one and so on. This is an indication that only plane waves and not general fields are dealt with in these cases.

[48] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968). The wave-stress tensor appears as the wave momentum flux density in their Eq. (6.2.17), but then they speak of a radiation pressure (a vector in their notation) exerted by the wave, and associate it with the wave momentum density. Already Brillouin criticized Rayleigh for attributing the radiation forces to the momentum density and not to the momentum flux density. [In the present paper, the momentum refers to the fluid and not to the wave; cf. also M. E. McIntyre, J. Fluid Mech. **106**, 331–347 (1981)].

[49] K. Yosioka and Y. Kawasima, Acustica **5**, 167–173 (1955); T. Hasegawa and K. Yosioka, J. Acoust. Soc. Am. **46**, 1139–1143 (1969).

[50] L. P. Gor'kov, Sov. Phys. Dokl. **6**, 773–775 (1962).

[51] P. B. Nagy, J. Acoust. Soc. Am. **79**, 1794–1797 (1986).

[52] In a lossless fluid the instantaneous contact force is of course normal to the instantaneous surface under consideration, but the orientation of a material surface is in general time dependent, and the time-average force is, therefore, not necessarily normal to the time-average surface, cf. Refs. 2 and 22 and J. Awatani, J. Acoust. Soc. Am. **27**, 278–281 (1955). Note in particular that the acoustic pressure is an alternating quantity. Consider the average of a positive force vector and a negative one, both being of the same order of magnitude but with a slightly different orientation. The superposition of these vectors leads to a small resultant vector whose direction depends on the details of the situation but can even be tangent to the surface (the radiation force is generally small in comparison with the force amplitudes of the instantaneous field).

[53] E. M. J. Herrey, J. Acoust. Soc. Am. **27**, 891–896 (1955).

[54] K. Beissner, in *Proceedings of the Ultrasonics World Congress 1997*, edited by K. Takagi, distributed by S. Ueha (Tokyo Institute of Technology, Yokohama, 1997), pp. 198–199.

[55] Grad $\mathbf{x}$ to be distinguished from grad $\mathbf{x}=\mathbf{U}$ and from Grad $\mathbf{X}=\mathbf{U}$.

[56] Equivalents of Eqs. (52) and (53/54) in a more generalized tensor notation can be found as Eqs. (20.10) and (20.8) in Ref. 15. Equivalents of Eqs. (56) and (63) in a more generalized tensor notation can be found as Eqs. (210.4) and (210.8) in Ref. 15, but here a different nomenclature and the different definition of the stress indices should be taken into account.

[57] Strictly speaking, this is one solution but not the only one. Any matrix $\mathbf{K}'$ for which $\mathbf{dA}\cdot\mathbf{K}'=0$ is valid can be added to $\mathbf{K}$. However, we are interested here in the resulting force values, and as these other matrices give no force contribution they can be ignored here.

[58] The symbols $\mathbf{F}$ and $\mathbf{H}$ have been adopted from Ref. 16. The symbols $\mathbf{K}$ and $\mathbf{B}$ stand for ''Kirchhoff'' and ''Boussinesq.'' The matrix symbols $\mathbf{H}$ and $\mathbf{B}$ here have nothing to do with the electromagnetic field vectors.

[59] In addition, the following may occur in an acoustic streaming: Fluid particles which were close together in the reference state may get more and more distant from one another with time. This leads to the Lagrangian spatial derivatives assuming increasingly higher values and therefore, $\mathbf{F}$ and $\mathbf{H}$ are not in a steady state.

[60] It may be, however, that in the temporal average of Eq. (35), the deviation of $\mathbf{v}\times\operatorname{curl}\mathbf{v}$ from zero in an acoustic streaming is only a higher-order effect, but this need not be discussed here.

[61] In the older literature the lateral stresses have sometimes been considered important for the ''boundary conditions'' mentioned in the Introduction. Note, however, that all results obtained here refer to the ''relaxed'' condition as set out in the discussions in Ref. 19 and that in this context the lateral stresses are unimportant for the radiation force on the target in the cases considered.

[62] It might appear that with the velocity potential being in a steady state, its time derivative, the acoustic pressure, should generally by zero on the time average. It should, however, be noted that Eq. (83) (left) is only valid in the first-order approximation and that in fact Eq. (37) [as well as Eq. (73)] does not contain first-order contributions to the time-average acoustic pressure.

[63] Equations (83) and (84) involve only temporal and spatial derivatives of the velocity potential, i.e., ϕ is indeterminate to the extent of an additive constant. In the first term in the curly brackets on the right-hand side of Eq. (95), however, ϕ itself appears. So it might appear that an additive constant in ϕ could have an influence on the radiation force, but this is not the case. If $\phi(\mathbf{x},t)$ is replaced by $\phi_0+\phi(\mathbf{x},t)$, with ϕ_0 being constant in space and time, it is easy to show with the aid of Eqs. (14) and (84) that all terms with ϕ_0 are unimportant for the radiation force as they either have the form of a curl or vanish in the time average, due to the steady-state condition. The similar applies to ψ and Eq. (96).

[64] Cf. Eqs. (A23) and (A26) of Ref. 23 (where, however, the minus signs do not appear as the surface element vector is understood there to point inwards).

Measurement and localization of interface wave reflections from a buried target

Eric Smith, Preston S. Wilson,[a] Fred W. Bacon, Jason F. Manning, John A. Behrens, and Thomas G. Muir

Applied Research Laboratories, The University of Texas at Austin, P.O. Box 8029, Austin, Texas 78713-8029

(Received 10 October 1996; revised 1 August 1997; accepted 31 December 1997)

It is demonstrated that seismic interface waves on the surface of a natural beach can be used to identify the position of a buried object. For this experiment, the waves were created with a sediment-coupling transducer and received on a three-element horizontal line array of triaxial geophones. The source and its coupling to the medium provided a high degree of signal repeatability, which was useful in improving signal-to-noise ratio. Reception of all three directions of particle velocity made it possible to augment conventional beamforming techniques with polarization filters to enhance interface-wave components. Reverberation in the beach was found to be large, though, and coherent background subtraction was required to isolate the component of the sound field reflected by the target. Propagation loss measurements provided comparisons of reflected signal power with predictions made previously, and the two were found to agree closely. [S0001-4966(98)03304-9]

PACS numbers: 43.30.Gv, 43.30.Vh, 43.20.Gp [JHM]

INTRODUCTION

One approach to the problem of detecting and locating buried objects is echo location, using direct excitation and reception of seismic waves. Because it is basically an acoustic method, it offers an advantage over magnetic detection that it can be applied to nonmagnetic objects. At the same time, unlike applications based on fluid or bulk-medium sonars, it circumvents the complexity associated with transferring sound energy and information across a fluid/solid interface. Furthermore, seismic excitations can have shear as well as compressional components, so they comprise many independent degrees of freedom that can carry a rich spectrum of information.

Interface waves, such as the free-surface Rayleigh wave, are of particular interest in the context of seismic echo location, because they contain energy, and are sensitive to medium characteristics, near the surface of the solid in which they are excited. These are useful features for detection of shallowly buried objects in unconsolidated sediments, where restriction of the search field to the upper sediment region is appropriate, and attenuation in the sediment makes energy containment near the surface desirable.

It has been shown that interface waves propagating along the sea floor can be detected and processed with seismic arrays to clearly identify the direction of the source.[1] This demonstration incorporated beamforming techniques which correct for the dispersion that is a common feature of such waves. The work reported here extends many of the results of the previous investigation to include the interface-wave reflections from a passive target insonified with an active sonar.

In many applications, the buried objects of interest may be small compared to the wavelength of the excitations used. In this limit, it is possible to predict the scattering strength from such targets in terms of their material properties and those of the medium, using efficient algebraic expressions. A perturbative method for generating such predictions is presented in a companion article.[2]

Finite-element numerical calculations for similar systems have been made and show good agreement with the perturbative methods,[2,3] and hybrid boundary-element/global-matrix[4] and finite-difference methods[5,6] have been applied to the related problem of facet reverberation. Experimental investigations, similar in sensitivity and scale to that reported here, have previously been presented,[7–9] but prior to now experimental data have not been available in the literature to test the perturbative predictions in detail. The work reported here also differs from that of Refs. 7–9 in the nature of the source used, and in the employment of polarization filters to process received signals.

An experiment was performed in May 1995, on a beach of the Gulf of Mexico, to address these issues. The data presented here show that it is possible to extract the interface-wave reflections induced by a buried target from noise backgrounds, with sufficient precision to quantitatively test the predictions of the perturbative scattering theory.

For this experiment, an active seismo-acoustic interface-wave sonar was designed and implemented. Its source was an electromechanical transducer that coupled directly to the beach surface, and its receiver was a three-element line array of triaxial geophone sensors. Special signal processing methods were employed, to make use of the large amount of directional and phase information present in the three-dimensional particle velocity field provided by this seismic sensor array. The interface-wave reflections from the target were found to possess sufficient coherence that beamforming

[a]Present affiliation: Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215.

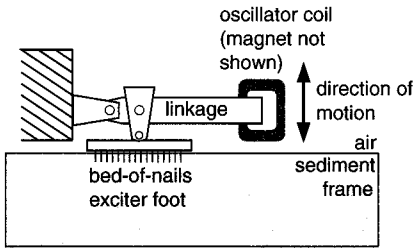


FIG. 1. Schematic of the transducer and coupling to the sediment.

methods could be used to produce a sonar output in plan position indicator (PPI) format, clearly identifying the target location. However, the seismic response of the beach led to a large reverberant component in the signal, and it was necessary to perform coherent background subtraction of a measured reverberation reference signal from measurements with a target present, to extract the target-induced reflection.

A second part of the experiment consisted of seismo-acoustic characterization of the medium with broadband sources. These were used to obtain information about dispersion, geometric spreading, and attenuation of seismic signals. The dispersion and attenuation characteristics were used to determine the preferred range of operating frequencies for the sonar, and the spreading and attenuation measurements were used to obtain transmission loss parameters for this site. With these, it was possible to compute interface-wave target strengths from the reflected signal power measured at the array. The resulting target strength values will be shown to be in good agreement with the results of the first-order perturbative calculation.

I. SOURCE AND SIGNAL CHARACTERISTICS

Two inertial impact sources and a specially designed electromechanical transducer were used as seismo-acoustic sources in this experiment. The impact sources will be described separately below. The transducer consisted of an electro-mechanical oscillator, linked to the sediment frame at the beach surface through an inverted ''bed of nails'' foot. Electrical inputs to the transducer were generated on a personal computer and amplified with a 500-W audio power amplifier. The oscillator and linkage assembly was mounted on one end of a cantilever, with the other end supported by a broad, acoustically buffered stabilizing platform that rested on the beach. The inertance for the oscillator was provided by the magnets and frame hardware.

The bed-of-nails foot was a reinforced rectangle of epoxy foam measuring 15 cm$\times$20 cm with a corrugated surface, from which thirty-eight 4-cm carpentry nails protruded in staggered rows, to embed in the beach. The corrugation and nail matrix served to distribute transducer stresses so that high power could be delivered to the beach surface without plastically deforming the sand matrix, to improve signal repeatability. A schematic of the transducer is given in Fig. 1.

The seismic output signal from the transducer, used for both *in situ* medium characterization and object detection, was chosen to be a Gaussian-enveloped cosine of center frequency 100 Hz. The full width at half-maximum of the signal envelope was chosen to equal one cycle period. Reasons for this choice of waveform and frequency band will be discussed below. The electrical input to produce the desired waveform was constructed in two steps, in an effort to remove as much of the transducer and medium response function as possible at the source. A broadband electrical test input was first applied, and the seismic waveform on a sensor 0.5 m from the source was recorded. The loaded response function of the transducer was obtained by dividing the Fourier transform of the measured seismic signal by that of the input. The Fourier transform of the desired output was then divided by the response function, and the result was inverse-Fourier transformed to obtain a time series for the desired electrical input.

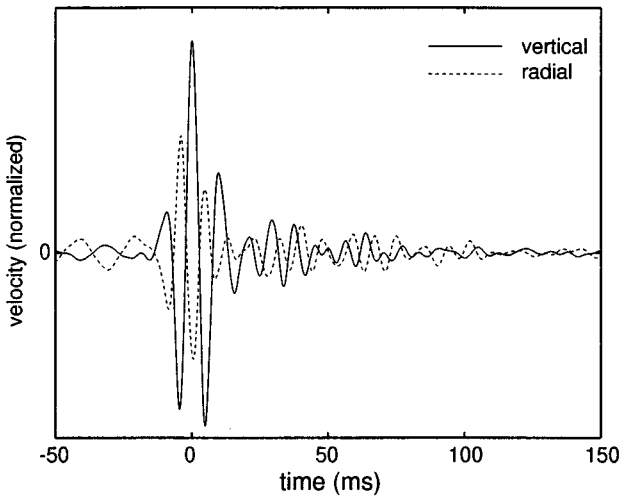


FIG. 2. A typical signal received on a geophone 0.5 m from the transducer foot. Solid (dashed) lines are vertical (radial) components of particle velocity.

In the ideal case, an electrical input produced with this method would pre-correct for the transducer response function, and reproduce the desired output. In realistic applications, the deconvolution of the response function tends to be unstable against noise. In this experiment, the seismic signal produced by pre-correction was found to acceptably approximate the desired output. The measured vertical and horizontal velocity signals produced at short range from the pre-corrected input are shown in Fig. 2.

The seismic sensors used in all measurements consisted of three 6-Hz geophones packaged in a balanced cylinder of epoxy foam to form a triaxial velocimeter. The sensor package had specific gravity 2.0 (comparable to that of saturated quartz sand[10]), height 5.6 cm and diameter 9.4 cm. When buried in the sand, the receiver packages appeared to be acoustically transparent to all of the signals used. Measurements with single sensors were used to obtain dispersion, time-of-flight and propagation loss measurements for *in situ* acoustic environment characterization, and the sensors were deployed in a linear array to form the receiver of the active sonar.

II. SITE CHARACTERIZATION AND SIGNAL PROPAGATION

The experimental site was the seaward beach of Mustang Island, a natural barrier island between Corpus Christi

bay and the Gulf of Mexico. Geologic studies have been conducted in the area, consisting of cores and detailed surface-sediment analysis.[11–13] The experiment described here was performed on the foreshore, defined in Ref. 14 as the region over which the surf travels between high and low tides. This region was very smooth and uniform, with a slope of roughly 3–6 degrees and width of 25–50 m. All measurements were carried out in a region of wet, but not wave-washed sediment, at times when the tide was below maximum.

Cores taken in the region[11] indicate that the top 30 m of the beach are composed of clean, well-sorted, and predominantly quartz sand, with some shell hash and other minerals. A grain size distribution of $\phi = 2.8 \pm 0.35$ was measured, where $\phi = -\log_2 D$ and D is the grain diameter in mm. While the sand is well sorted on small scales, cores showed some sedimentary layering, with layers of finer sand than that at the surface, from one to three meters thick, lying anywhere from 3 to 5 m subsurface. Layer depths and thicknesses varied among cores, though, and surface topography of the beach has been measured to vary by heights of roughly 1 m relative to mean sea level in the course of a year.[12] Thus the depths, thickness, or compositions of layers at a particular site are not precisely specified by existing records of geologic ground truth. The water table was estimated to lie roughly one meter below the beach surface while the current measurements were being taken.

In situ estimates of the effects of beach layering or sediment compaction with depth on acoustic propagation were obtained by measuring the dispersion curves of signals, created with a manual impact source, that propagated some 100 m along the surf line. The source was a cylindrical handle with a flat plate impactor welded to the end. The goal of these measurements was to determine what modes were created by sources at the beach surface, to assess their polarization and dispersion characteristics, and from these to identify optimal frequency ranges or signal processing methods for object detection.

Figure 3 shows Gabor matrix contour plots of the signals received on the horizontal and vertical geophones in the vertical plane looking from the receiver toward the source. A Gabor matrix is a time-frequency decomposition of the spectral density in a signal, plotted with frequency on the abscissa and time-of-flight on the ordinate. Figure 3 was produced by windowing the signal in the time domain with a sequence of Hanning windows, and plotting the spectra of the windowed segments, obtained as the absolute values of the fast Fourier transforms (FFT), as rows stacked according to the arrival time at the center of each window.

The group velocities of arriving pulses may be obtained from the Gabor plot by tracing along the maxima of the contours in regions of contiguous spectral density, and the phase velocities may be computed from this data.[1] The shear speed profile as a function of depth in the medium is then inferred by fitting the group velocity curve from a model profile to that measured from the Gabor plots. A simple two-layer shear speed model is shown as Fig. 3(c), and its associated group and phase velocities are plotted as the dashed and solid lines superimposed on the Gabor matrix plots. This

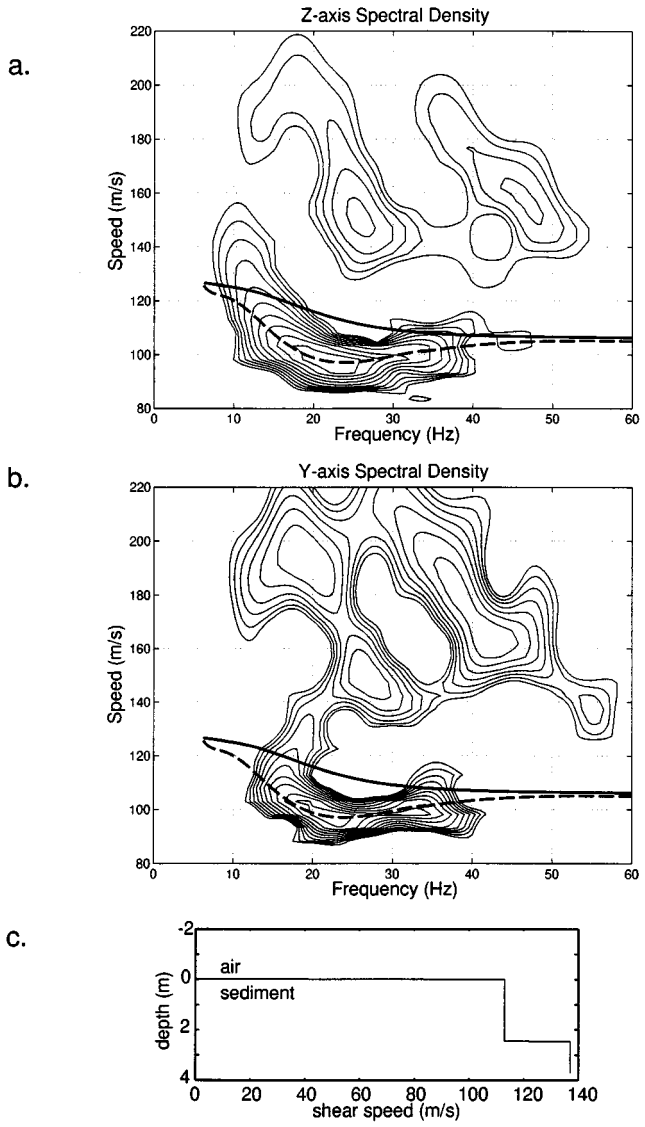


FIG. 3. Velocity-frequency decompositions (Gabor Matrix contour plots) from impacts, at 100-m source–receiver range along surf line. (a) radial; (b) vertical channels of particle velocity at receiver. Thick solid (dashed) lines are fits to phase (group) velocity from a model with an elastic layer over an elastic half-space. Shear speed profile for the model is shown in (c). Layer speed is 113 m/s, half-space speed 137 m/s, layer thickness 2.5 m. Compressional speeds are 1600 m/s in layer and half-space; model is insensitive to variations of compressional velocities as long as they are several times faster than shear speeds.

simple model of a homogeneous upper layer over a homogeneous half-space can be seen to match the measured group velocities within the uncertainty of the Gabor matrix contours above 15 Hz, which is the frequency range over which the Rayleigh wave propagation is controlled by the top 3 m or less of sediment.

The sequence of windows used in Fig. 3 was chosen so that the sum of all window values at any point in the time domain is unity. This makes it possible to decompose the Gabor matrix by ''cutting'' along boundaries between the modes, resumming the FFTs for all segments within a given cutting boundary and inverse Fourier transforming, to produce time series for the individual modes separately. When the modes are well separated, as they are in this case, this

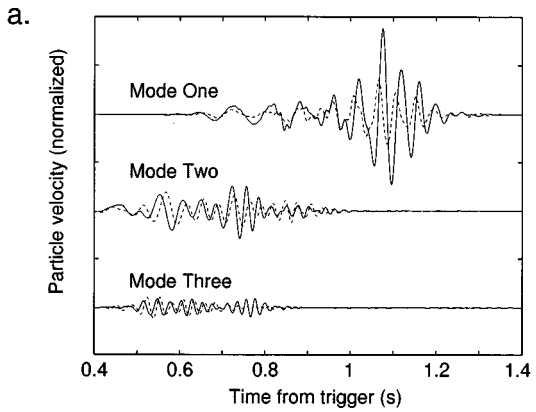


FIG. 4. Decomposition of the time series leading to Fig. 3 into modes, performed by summing FFT segments from contiguous regions of large spectral density in Gabor matrix. Solid (dashed) lines are vertical (radial) particle velocity components.

decomposition makes it possible to recover the individual mode time series with minimal corruption by either the window functions employed in the time domain, or those used to define the cutting prescription along rows in the Gabor matrix. The resulting modal time series for the three modes of Fig. 3 are shown in Fig. 4.

Plots of the horizontal signal versus the vertical signal yield hodograms (representations of actual particle velocity as the wave passes the receiver), which are shown for the three modes respectively in Fig. 5. The slowest mode, labeled ''Mode One,'' is vertical in aspect ratio and retrograde in sense of particle motion (polarization), which is consistent with the identification of this mode as the free-surface Rayleigh wave.[15] Both higher modes are prograde in polarization and more horizontal than the Rayleigh mode in aspect ratio.

Frequencies higher than 50 Hz were not visible above noise in the long-range dispersion analysis, so pulses from the transducer were used to examine higher-frequency propagation at closer ranges. Figure 6 shows received vertical time series and pulse envelopes for pulses propagated distances of 1, 5, and 10 m. The envelopes were obtained as the absolute values of the quadrature completions of the received time series. Center frequency of the pulses was 100 Hz. Time-of-flight estimates were made for these pulses by fitting a straight line to the maxima of the envelopes as a function of range. The mean pulse speed inferred by this method is 115 ± 5 m/s, consistent with the high-frequency asymptotic behavior shown in Fig. 3.

The *in situ* data were interpreted as indicating that an acoustically significant interface is present roughly 2 m below the beach surface (inferred from the layer model), forming a duct of nearly constant, slower shear speed between the beach surface and the lower interface. The cutoff frequency for shear modes in the duct coincides with the frequency of the group speed minimum for the Rayleigh wave, equal to roughly 25 Hz. Rayleigh waves of higher frequencies, being evanescent from the beach surface, involve particle motions almost entirely confined within the homogeneous upper layer and propagate effectively without dispersion.

To determine the geometric and dissipative losses characterizing high-frequency signals, a sequence of low energy,

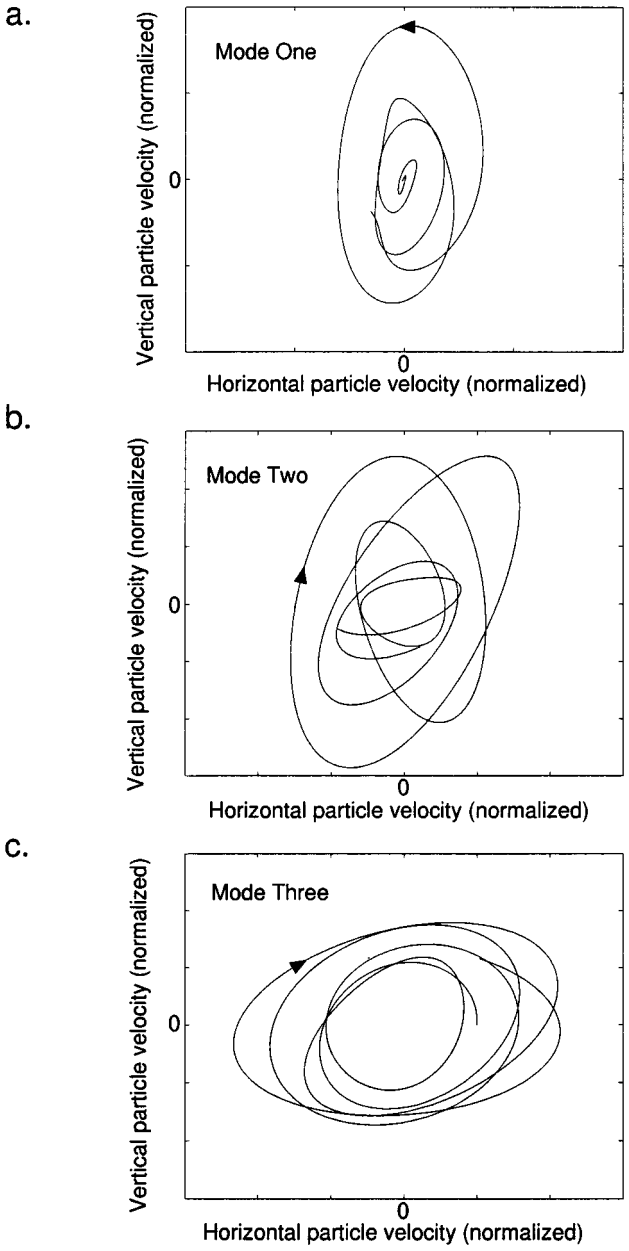


FIG. 5. Hodograms, or plots of particle velocity in the radial-vertical plane, for each of the time series shown in Fig. 4. Arrows indicate directions in which sediment frame traces the curves. Motion is retrograde for mode one, prograde for modes two and three.

calibrated impacts was used to produce broadband signals at ranges from 50 cm to 30 m from the receiver. The impacts were made by dropping an onyx sphere of diameter 8 cm from heights ranging from 10 cm to 2 m, to embed without bouncing in a region of wet but not saturated sand. The signals produced were found to scale in amplitude with the momentum of the drops, and were otherwise nearly independent of drop height. (This made it possible to vary drop height with range to optimize signal quality relative to noise, without overdriving the sensors. The spherical impactor provides very repeatable signals, unlike the plate impactor previously described.)

The received signals were band passed in 30-Hz bins and fitted to functions of the form $Ae^{-\kappa r}/r^{\alpha}$ in each bin.

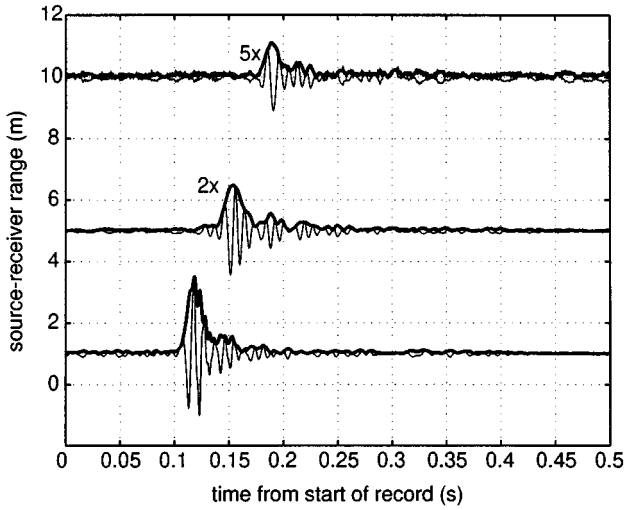


FIG. 6. Time series of pulses used in time-of-flight velocity measurement. Thin lines are vertical-channel geophone signals, and thick lines are rms amplitudes used to identify corresponding points in the group envelope at different ranges. Best-fit velocity for this population is 115 ± 5 m/s. Signals at 5-m and 10-m range have been multiplied by the indicated scale factors relative to the signal at 1 m, to aid visibility.

Data and fit for the bin from 90 to 120 Hz are shown in Fig. 7. The best-fit value of the power law for geometric spreading is $\alpha = 0.55$, and the best-fit dissipation factor is $\kappa = 0.912$ dB/m. The population shown represents the dominant frequency content for these impacts; results were similar, but scatter in the data was larger, in the bins at both higher and lower frequencies. These data confirm that the signal is carried by a cylindrically spreading wave, and the dissipation is consistent with values previously reported for clean sand.[10]

The results of these measurements were used to identify an optimal frequency range for the object-detection experiments. The best range resolution possible was desired, which recommended single-cycle pulses. A resolution of order 1 m could be obtained from pulses of center frequency 100 Hz. Such pulses were sufficiently bandlimited that most of the signal energy propagated in the high-frequency asymptotic part of the Rayleigh wave dispersion curve indicated in Fig. 3. Significantly higher frequencies were found to damp out too quickly to provide useful signal strengths over propagation paths of even 5–10 m, while lower frequencies than 100 Hz produced wavelengths longer than 1 m and could not be enveloped in single-cycle pulses without leading to significant dispersion of their lower-frequency components.

III. SIGNAL PROCESSING METHODS

Extraction and localization of seismic interface wave signals can employ a number of signal processing methods that are not available for simpler pressure-wave acoustic signals. This is because interface waves are vector excitations, containing a dominant shear component. In the case of Rayleigh waves, the radial component of the particle motion in the wave is collinear with its direction of travel. Further, interface waves, by their nature as evanescent waves that decay with distance from an interface, generically have phase offsets between the horizontal and vertical components of their particle motion. These phase offsets cause the particle motion to be elliptical, or *polarized*, in the radial-vertical plane, as appears in the hodograms of Fig. 5. It is possible to filter for this polarization, preferentially selecting interface-wave components of a signal, over in-phase components associated with bulk motion in the solid or the air. Both techniques were combined with traditional beamforming methods to create a signal-processing algorithm for an active seismic sonar, as described below.

The sonar source and targets were deployed along a line parallel to the surf line. Targets were placed at several ranges along this line, and a linear receiving array of three triaxial geophones was buried, with their top surfaces flush with the beach surface, and with the center element 0.5 m in front of the source. The receiving array was offset in angle, so that the source–target line was 22.6-deg off-broadside, and the element spacing used was 1.08 m. The nine directional signals from the array elements were received and stored, along with copies of the output signal that were used to specify a zero-time for pings.

Processing of the array signals to yield a PPI sonar plot was performed in three stages. First, for each position in the search field, a look direction was defined from each element. Because the distances between source, target and array elements were all comparable to the wavelengths used, exact look directions, rather than far-field approximations, were used. To define the planes of interest for imaging a cylindrically spreading Rayleigh wave, the vertical signal and a steered radial signal,

$$r_i(t,\mathbf{x}) \equiv x_i(t)\cos\,\theta_i(\mathbf{x}) + y_i(t)\sin\,\theta_i(\mathbf{x}), \tag{1}$$

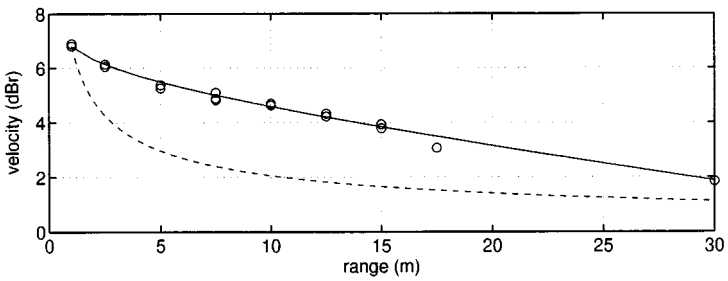


FIG. 7. Fit of bandpass-filtered packet envelopes, for a population of spherical impactor drops, to the function $Ae^{-\kappa r}/r^{\alpha}$, with r in meters. Pass band is 90–120 Hz. A is a normalization and fit parameters shown are $\alpha = 0.55$, $\kappa = 0.912$ dB/m. (Dashed line shows cylindrical spreading without dissipation, for reference.)

were retained. Here, $\mathbf{x}$ is the position in the search field, $x_i(t)$ and $y_i(t)$ are the two received time series on the horizontal geophones, at element i in the array, and $\theta_i(\mathbf{x})$ is the angle to position $\mathbf{x}$ from the ith element relative to the x sensor direction.

The vertical and radial signals from all three elements were multiplied by element shadings σ_i between 0 and 1, and the steered signals for each search position were summed with appropriate time-of-flight delays for that position to produce a two-channel array output:

$$v_{\text{array}}(t,\mathbf{x}) \equiv \sum_i \sigma_i v_i(t - t_i(\mathbf{x}),\mathbf{x});$$
$$r_{\text{array}}(t,\mathbf{x}) \equiv \sum_i \sigma_i r_i(t - t_i(\mathbf{x}),\mathbf{x}). \tag{2}$$

Here, $t_i(x)$ is the time delay between the central and ith element for search position $\mathbf{x}$, and v_i is the vertical signal measured on element i. The two components v_{array} and r_{array} sum constructively for the outgoing vertical and radial components that would be produced by a cylindrically spreading Rayleigh wave reflected from a target at position $\mathbf{x}$.

Polarization filtering for this application is the process of extracting, from a radial-vertical signal pair, the component of the crossed signal power that comes from the elliptical particle motion of the wave. It was performed on the radial and vertical array signals defined in Eq. (2) as follows: Each real-valued time series was Fourier transformed to produce a complex spectrum. Negative-frequency components of this spectrum were zeroed; positive-frequency components were multiplied by two; the resulting spectrum was inverse-Fourier transformed. This produces a quadrature completion of the real time series, the positive-frequency complex signal whose real part is the measured signal. The completed signals r and v will be denoted r_C and v_C (for ''Complex'').

A complex, crossed power function in the time domain, denoted P_{rv}, was then formed from r_C and v_C as

$$P_{rv}(t,\mathbf{x}) \equiv r^*_{\text{array } C}(t,\mathbf{x}) \times v_{\text{array } C}(t,\mathbf{x}), \tag{3}$$

where * denotes complex conjugation. P_{rv} contains the in-phase and out-of-phase power components that can be extracted from the signal pair, as its real and imaginary parts. If r_C and v_C happen to be two copies of the same signal, then P_{rv} is a purely real quantity, proportional to the rms power in the signal pulse. Other in-phase motions differ only by an overall scale factor, determined by the relative projections of the wave into the radial and vertical axes.

For an interface-wave pulse arriving at the receiver, the two quadrature-completed signals are again copies of the same time series, but the phase delay in the particle motions appears as a relative *complex* phase in the complex waveforms. P_{rv} is again proportional to the rms pulse power, but deflects in the complex plane along an axis defined by the phase angle.

Therefore, the imaginary component of P_{rv} projects out purely in-phase arriving signals, and deflects positively or negatively for polarized signals, according to whether the polarization is prograde or retrograde. An example of the process of polarization filtering is shown in Fig. 8.

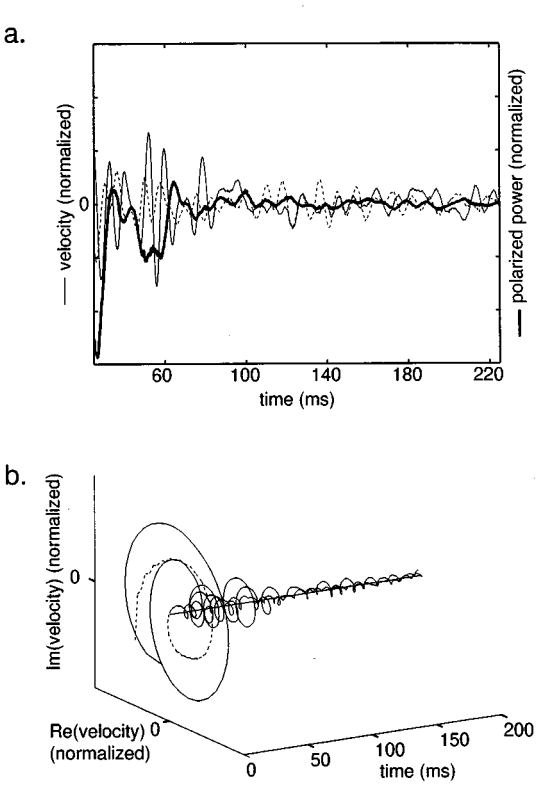


FIG. 8. Polarization filtering and crossed power in the time domain. (a) Vertical (solid-thin) and radial (dashed-thin) time series obtained by subtracting a 10-signal average of events with no target from a 10-signal average with a target present, received on the central array geophone of the active sonar apparatus. (b) The quadrature completions obtained from the real-valued time series in part (a). Using r to denote the complex radial time series (dashed), and v the complex vertical (solid), the crossed radial-vertical power is defined as the product $P_{rv} \equiv r^* \times v$ in the time domain. Its polarized component, $\text{Im}(P_{rv})$, is plotted as the heavy line superimposed on the time series in part (a).

This imaginary part of P_{rv}, representing the polarized power component, is plotted versus position $\mathbf{x}$ in the search field to produce the PPI output. Because the signal can deflect either positively or negatively, a linear scale is used, and the sign of the deflection gives an additional form of visual information about the arrival. For instance, if the arriving signal is known by some other means to be a Rayleigh wave, the sign of the polarized component resolves the forward-backward ambiguity of the linear array. Alternatively, if the direction of arrival is known, the polarization provides information about the type of mode that carried the signal.

An example of the sonar plot produced by this method is shown in Fig. 9. The inputs to the algorithm were the raw geophone signals measured with a target buried 3 m in front of the source, at bearing -22.6 deg from array broadside. A population average over ten ''pings'' was used to reduce random noise. The target used was a titanium cylinder of height 20 cm and diameter 22.5 cm, weighing 35.4 kg. It was buried with the cylinder axis vertical and its top surface flush with the beach surface. (For wavelengths larger than the typical target dimensions, this orientation is not expected to affect reflection signatures.[2]) The scale used for $\text{Im}(P_{rv})$ was chosen to normalize the polarized power of the outgoing

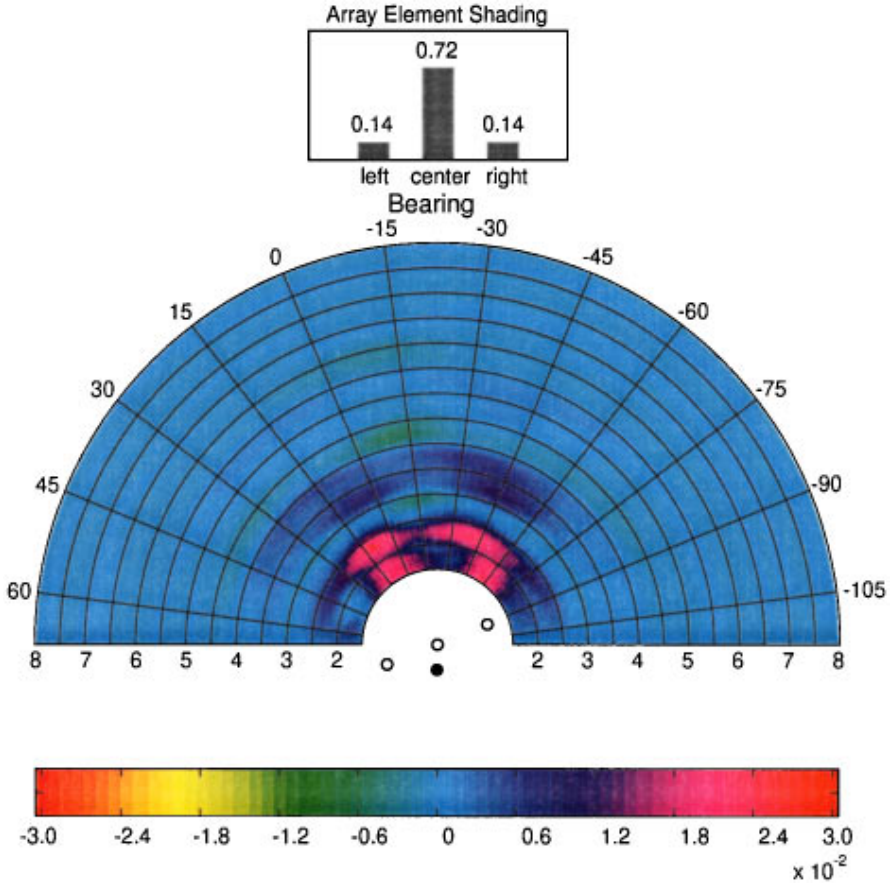


FIG. 9. Example sonar plot of a signal from a buried target before subtraction of reverberant backgrounds. Polarized component of the crossed radial-vertical power, $\mathrm{Im}(P_{rv})$, shown with linear color scale. Blank region in the center of the display shows source and receiver locations. Solid circle is the source position, and outlined circles are the array element positions. Plot produced with element shadings depicted, and mode velocity 113 m/s.

direct arrival to unity at 1 m (i.e., source level of 0 dB). No evidence of a target reflection is visible at the expected location, but numerous artifacts of order $\pm 2\%$ are present.

IV. ANALYSIS OF BACKGROUND AND NOISE FEATURES

In Fig. 2, deviation from the desired output is largest immediately following the main signal pulse. To understand the nature of this deviation, as well as the origin of the features in Fig. 9 and their relation to statistical and systematic components of noise, the signal creating a slice through the PPI plot, looking along the chosen source–target line, was analyzed in the following ways.

Figure 10(a) shows the autocorrelation of the vertical (quadrature completed) signal, defined as

$$\left| \int dt v^{*}_{\text{array } C}(t) \times v_{\text{array } C}(t-\tau) \right| \Big/ \int dt |v_{\text{array } C}(t)|^2,$$

as a function of offset time τ. Because of the form of the signal, shown in Fig. 2, the large direct arrival dominates the correlation, acting as a matched filter to extract copies of itself appearing later in the time series. Because the full complex signal is used, these can appear with any phase offset, relative to the direct arrival. The magnitude of peaks gives, to linear order, the *amplitude* of the reflected copies, because the main weight of the filter function is invariant.

A complicated series of peaks can be seen in the autocorrelation. Once the constructive interference of the direct arrival with itself is past, though, the maxima of the peaks fall nearly on a line, with most peaks in the series forming a regular sequence (indicated by the arrows). If these are interpreted as a sequence of reflections from a subsurface interface, the associated ray diagram for the four indicated peaks is shown in Fig. 10(b). Using the reference distance of 0.5 m from the source to the central receiver element, and the measured value of the shear speed, the inferred depth of the reflecting interface is 2 m, consistent with the value obtained from the dispersion analyses of Fig. 3. It is therefore concluded that a large component of the signal background is due to reverberation within the beach, with an amplitude reflection coefficient of roughly 0.5 indicated by the straight line in Fig. 10(a).

Because of the pre-correction process used to produce these signals, it is likely that reverberation signatures are present both directly and indirectly via the response-function deconvolution. This does not compromise the signal processing methods applied here, and it may explain why the sequence of peaks seen in Fig. 10(a) is more complex than the simplest single-layer model predicts.

To identify whether the form of these time series was stable or slowly variable, autocorrelations were computed for all shots in a ten-shot population, and their average graphed as a function of separation within the population in Fig. 11.

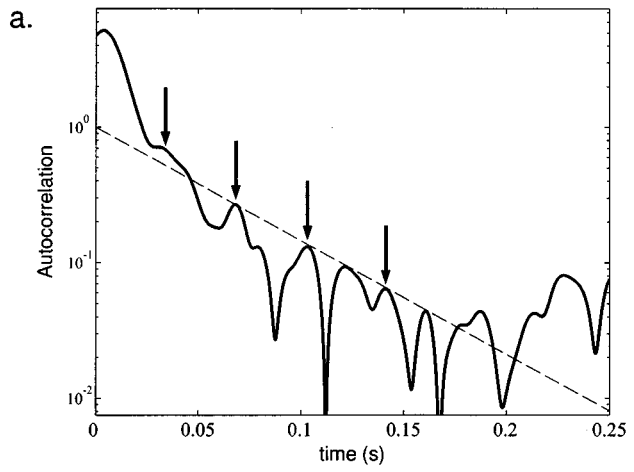


FIG. 10. Reverberation identification from the signal autocorrelation. (a) Absolute value (solid line) of the autocorrelation, from the quadrature completion of a typical vertical geophone signal with no target, as a function of time offset. Signal was measured on a geophone 0.5 m from the source. Dashed line is the exponential decay corresponding to a series of reflections with constant reflection coefficient 0.4. Arrows indicate first four reflected arrivals. (b) Ray diagram leading to the reflections marked in (a). Layer thickness implied for shear reflections at 113-m/s shear speed is 2.0 m.

Roughly, each shot in the population was recorded 2 min after the one preceding it. It should be noted that, for a population of ten shots, there are nine autocorrelation values contributing to the population average at separation 1, and only one value at separation 9, so the autocorrelations at large separation are more sensitive to individual deviant shots, or outliers, than those at small separations.

Autocorrelations were computed over three time intervals, to compare the repeatability of the direct arrival with that of longer time series containing more reverberation. For the shorter two intervals, centered on the direct arrival, there is no convincing evidence for a signal drift with time. Autocorrelations for separation intervals 1–7 and 9 are comparable. While the autocorrelation at separation 8 is lower, it may not be significant because it represents only a two-event average. For the longer time period, comprising the whole signal shown, there appears to be a small systematic decay of the autocorrelation, representing a change in medium response over the period of 20 min sampled. It is therefore concluded that the reverberation may be subject to a systematic drift over these time scales, which is larger than that for the direct arrival.

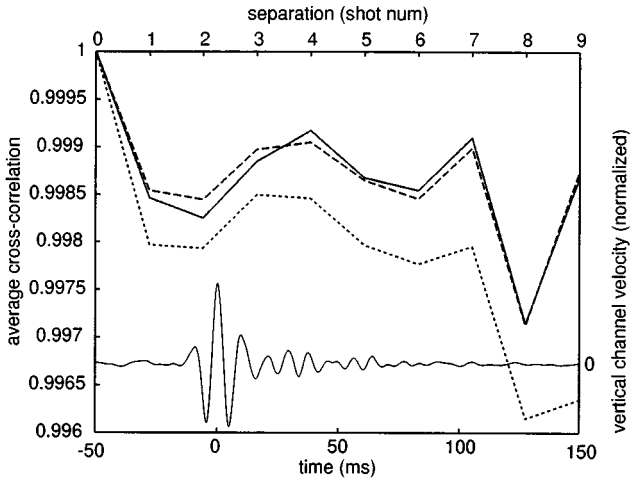


FIG. 11. Averaged crossed correlations for pairs of vertical geophone velocity signals from a 10-shot population, as a function of separation in shot numbers. Separation between events in the population is roughly 1–2 min per event number. Zero separation denotes the autocorrelation, normalized to unity. Nine correlation values are averaged to produce the 1-index result, and one value for the 9-index separation. Time series (thin line) plotted for reference. Heavy solid line is the correlation average for time-interval $(-20,30)$ ms, dashed is $(-20,80)$ ms, and dotted $(-50,150)$ ms, indicating a greater relative decay of correlation with time among shots including more of the reverberation with the direct arrival.

A final measure, of the statistical rather than systematic variation of shots, is a point-by-point signal-to-noise ratio (SNR) in the time domain, shown in Fig. 12. It is computed as $-20\log_{10}$ of the standard deviation obtained from the absolute values of differences among the quadrature-completed signals in the population. This measure incorporates both magnitude and phase errors in a mutually consistent way. The value of roughly 25 dB in the main pulse of the signal time series corresponds to an expected two-shot coherence of 0.997, which is comparable to the values given

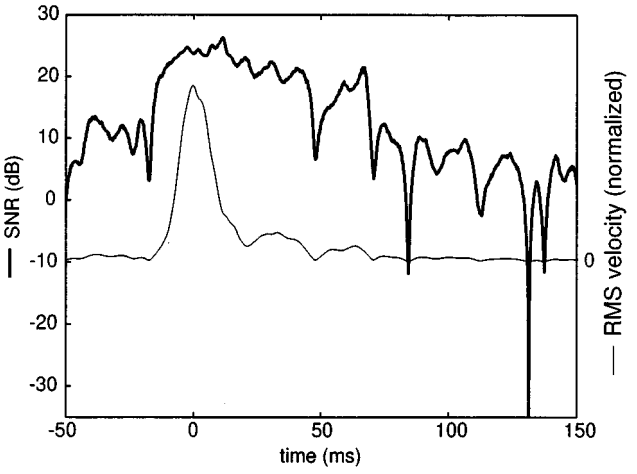


FIG. 12. Pointwise signal/noise ratio (SNR) as a function of time for the vertical channel of the central geophone, from a population of ten recorded pulses with no target present. rms signal amplitude (thin line, shown for reference), is the absolute value of the mean of the quadrature completion of the ten signals. SNR is computed from the absolute mean square deviation of the quadrature-completed signals, thus incorporating both phase and magnitude errors.

by the autocorrelation in Fig. 11, further suggesting that most of the loss of coherence in the autocorrelation of the direct arrival is statistical in nature, and not systematic.

The absolute signal level and the SNR at 55–65 ms, where the target reflection is expected, define the noise floor, relative to which subtraction of backgrounds should make a reflected signal visible. The central limit theorem relates the SNR of 10 dB for two individual pings, from Fig. 12, to an expected SNR of 15 dB for the difference of two ten-shot population averages, which will be used below. It then follows, from the signal level of roughly -24 dB in this range, *re*: the direct arrival at 1 m, that the noise floor is -39 dB. Thus a target reflection larger than -39 dB relative to source level can be expected to be recovered by subtraction of backgrounds from signals with the target present.

V. COHERENT SUBTRACTION AND TARGET LOCALIZATION

Both incoherent and coherent subtractions of backgrounds are possible in computations of the polarized power component, as for any quadratic function of the signal. Figure 13 shows the output of both differencing methods. The raw input containing reverberation, with the target present, is the same ten-shot average used to produce Fig. 9. Backgrounds alone are given by the average of ten shots produced with the same sonar deployment, immediately prior to burying the target in the beach. Figure 13(a) shows the vertical components v_{array} for the two populations, and the difference amplitude between them.

Figure 13(b) shows the corresponding polarized power values, $\text{Im}(P_{rv})$, for the signal and background, and also the difference of the power values (incoherent subtraction) and the power value produced from the signal difference (coherent subtraction). The polarized power produced by coherent subtraction (multiplied by 10 for visibility in the figure) is much smaller than the incoherently subtracted value. Also, whereas the incoherent subtraction leaves numerous features in the time series, coincident with large reverberant signatures in the background, the coherently subtracted power has only one large, retrograde feature at 3-m range, corresponding to the known position of the target.

The PPI output from the coherently subtracted signal is shown in Fig. 14. A clear retrograde signal is visible at the target location, with magnitude -34 dB. Array sidelobes of the same sign are also visible at bearings 30 deg and -75 deg, with strengths of -38 dB.

The array shadings used in Fig. 14 were chosen to produce the least cluttered sonar plot when three elements are included, and also the most consistent two-element outputs when any single element is zeroed. The output of the array with these indicated shadings, using real data, was similar to that predicted from *uniform* shadings for an idealized reflection. The origin of this deviation in array performance from the idealized case is not known at this time, but we conjecture that it is related to the close proximity of the large static load created by the transducer foot, which is centered 0.5 m from the center receiving element.

Because it was of interest to compare the reflected signal strength to predictions, the reflection was also measured using the central array element alone, with unit shading. The single-element reflected power was found to be within 0.5 dB of the value for the array shown in Fig. 14, where the array weights sum to unity. Deviation within the main body of the pulse was roughly 0.5 dB for the single-element signal, so the two values are taken to agree within measurement error.

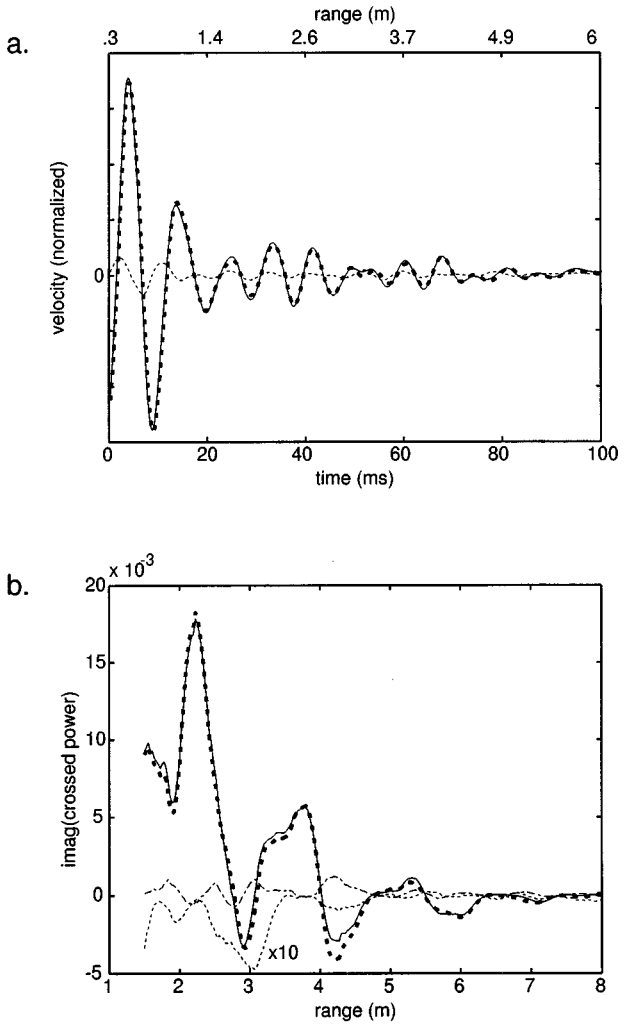


FIG. 13. Typical signal and polarized crossed power component for coherent and incoherent subtractions. (a) Center-geophone vertical component for 10-event population averages of signal with target present (solid), reverberation measured without target (heavy, dashed) and target signal with background subtracted (dotted). (Range label on top axis provided for reference.) (b) Sections through the PPI plot, looking at the target, for different input signals. 10-event average with target (solid), reverberation only (dashed), the difference of the power values, (incoherent subtraction, dash-dot), and the polarized power component obtained from the amplitude differences (coherent subtraction, dotted). The coherent subtraction is multiplied by 10 to aid visibility. All other values are shown with the normalization used in the PPI plots.

VI. COMPARISON TO PREDICTIONS

A perturbative method for approximating the scattering of interface waves from small buried objects is presented in a companion article.[2] Prior to now it has not been possible to test the predictions of this method. The perturbative method is analytic rather than numerical, and the first-order perturbative result, which is expected to be valid for weak scatter-

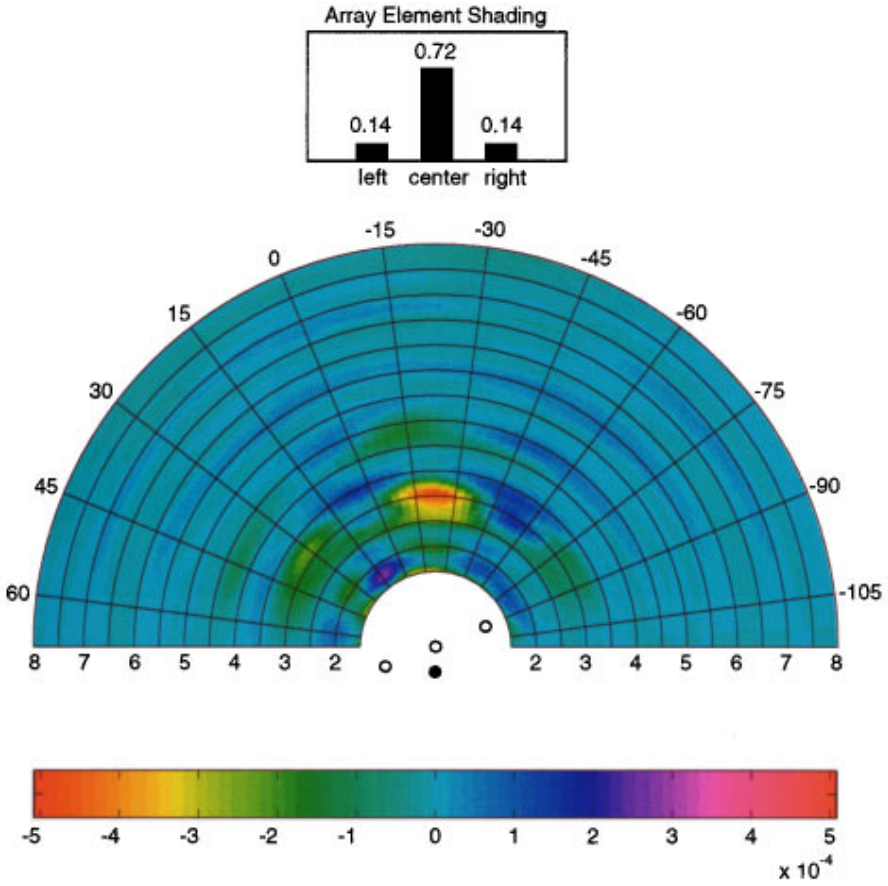


FIG. 14. Example of a sonar plot produced from measurements taken with target buried at 2.5-m range and -22.5 deg bearing, after coherent subtraction of backgrounds measured without target. Element shadings and speed are the same as those of Fig. 9.

ing from inclusions much smaller than the interface-wave wavelength, is computed algebraically. Therefore this method, when it can be used, is much less costly than finite-difference computation, or other numerical methods for treating the scatterer boundary conditions explicitly.

The perturbative computation predicts that scattering from the target used here, in an unconsolidated sediment medium, should be dominated by the discontinuity created by the target mass. The pointlike-scatterer approximation also should be valid for this target, because all of its characteristic dimensions are smaller than than $1/2\pi$ times the wavelength at 100 Hz. Using a density of 2 g/cm^3 for the sand, first-order perturbation theory predicts a target strength of -21.6 dB for the conversion of a plane-fronted Rayleigh wave incident on the target to a cylindrical outgoing Rayleigh wave at 1 m.

The transmission loss (TL) associated with the propagation path for this system is drawn from the acoustic medium characterization. Using the observed geometric spreading as $r^{-0.55}$ and the dissipation of 0.912 dB/m, the combined TL for the two paths, 1 m from the source to the target, and one meter from the target to the central receiver, is -11 dB. The resulting predicted reflection strength is -32.6 dB, which compares favorably with the measured value of -34 dB.

VII. CONCLUSIONS

The experimental results presented here show that it is possible to identify the reflection from a passive buried target insonified with an active, seismic interface-wave sonar, and to process the reflection to recover the position of the target.

Seismic-array signal processing was performed by combining dipole steering, time-delay beamforming, and polarization filtering techniques in a single algorithm. This type of processing enhances interface-wave components relative to in-phase seismic excitations, and provides polarization as well as amplitude information in the resulting PPI sonar plot.

The signals produced for echo location were found to contain a significant component of medium-induced reverberation, and it was necessary to coherently subtract the reverberant background from sonar pings with a target present, in order to isolate the target-induced reflection. Because the source produced repeatable signals, and statistical variation of backgrounds was found to be much smaller than the reverberation level, coherent subtraction was effective, rendering a reflection of -34 dB at the receiver clearly visible. Developing reliable differential signal processing methods that can be used under less idealized circumstances remains an important outstanding problem.

Simple algebraic methods based on perturbation theory can be used to predict interface-wave scattering strengths under the experimental conditions studied here. Prior to now, though, it has not been possible to test the perturbative predictions directly with measured data. The reflection strengths measured in this experiment, combined with transmission loss data for the site, were found to yield a target strength

that agrees closely with the value predicted by first-order perturbation theory.

ACKNOWLEDGMENTS

This work was supported by the Applied Research Laboratories Independent Research and Development Program, the U.S. Navy Office of Naval Research, and the Advanced Research Projects Agency, Marine Systems Technology Office.

[1] J. A. TenCate, T. G. Muir, A. Caiti, A. Kristensen, J. F. Manning, J. A. Shooter, R. A. Koch, and E. Michelozzi, ''Beamforming on seismic interface waves with an array of geophones on the shallow sea floor,'' IEEE J. Ocean Eng. **20**, 1–11 (1995).
[2] E. Smith, ''Scattering of interface waves from pointlike obstacles,'' J. Acoust. Soc. Am. **103**, 2344–2353 (1998).
[3] Å. Kristensen and R. Mittet, ''Using surface waves to detect a point diffractor buried at the sea bottom,'' STATOIL report N-7005, IKU report N-7034, Trondheim, Norway (1996).
[4] P. Gerstoft and H. Schmidt, ''A boundary element approach to ocean seismoacoustic facet reverberation,'' J. Acoust. Soc. Am. **89**, 1629 (1991).
[5] R. A. Stephen, ''Bottom penetration at subcritical grazing angles by scattering,'' J. Acoust. Soc. Am. **99**, 2475(A) (1996).
[6] R. A. Stephen and S. T. Bolmer, ''Interface waves in seafloor propagation and scattering problems,'' J. Acoust. Soc. Am. **99**, 2552(A) (1996).
[7] D. A. Sachs, ''Feasibility of acoustic landmine detection final technical report,'' BBN Technical Report No. 7677, 19 May 1992, submitted to U.S. Army Belvoir Research Development and Engineering Center.
[8] P. A. Krumhansl, B. G. Watters, and D. A. Sachs, ''Detection of buried landmines by imaging with surface waves,'' J. Acoust. Soc. Am. **93**, 2341 (1993).
[9] P. A. Krumhansl, ''Buried ordnance detection using seismic waves,'' Paper presented at Unexploded Ordnance Detection and Range Remediation Conference, Golden, Colorado, May 17–19, 1994.
[10] E. Hamilton, ''Elastic properties of marine sediments,'' J. Geophys. Res. **76**, 579 (1971).
[11] G. L. Shideler, ''Stratigraphic studies of a late quaternary barrier-type coastal complex, Mustang Island-Corpus Christi Bay area, South Texas Gulf Coast,'' U.S. Geological Survey Professional Paper 1328 (1986).
[12] E. W. Behrens and R. L. Watson, ''Corpus Christi Water Exchange Pass: A Case History of Sedimentation and Hydraulics During its First Year,'' Unpublished Report, U. T. Austin Dept. of Geol. (1974).
[13] L. F. Brown, Jr. and J. W. Macon *et al.*, ''Environmental geologic atlas of the Texas coastal zone,'' Bureau of Economic Geology, University of Texas at Austin (1976), p. 65.
[14] M. E. Milling and E. W. Behrens, ''Sedimentary Structures of Beach and Dune Deposits: Mustang Island, Texas,'' Institute of Marine Sciences, The University of Texas, internal report.
[15] M. Yu. Brekhovskikh, *Waves in Layered Media* (Academic, New York, 1980), 2nd ed.

Scattering of interface waves from pointlike obstacles

Eric Smith
Applied Research Laboratories, The University of Texas at Austin, P.O. Box 8029, Austin, Texas 78713-8029

(Received 10 October 1996; revised 1 August 1997; accepted 31 December 1997)

A method is presented for efficiently computing or estimating low-order elastic wave scattering, from a pointlike inhomogeneity in a bounded medium, into interface as well as bulk modes. The particular system considered is a half-space of fluid overlying a half-space of a linear elastic solid, with the obstacle placed in the solid very near the interface. Boundary constraints are enforced as an improvement to the bulk scattering vertex, so that the stratified-medium Green's function is used to propagate scattered waves at each perturbative order. Analytic approximation of the first-order scattering amplitude, based on the symmetries of the scattering vertex and the interface-wave pole structure, makes it easy to identify the qualitatively different components of the scattered wave, and their dependence on medium properties. © *1998 Acoustical Society of America.* [S0001-4966(98)03404-3]

PACS numbers: 43.30.Ma, 43.20.Gp [JHM]

INTRODUCTION

Relatively few phenomena involving interface-wave excitation can be treated efficiently with current computational models of acoustic propagation and conversion. Yet problems occur for which efficient computational techniques would be useful. For instance, because interface waves confine energy entirely within a small spatial region around a particular interface, their scattering can be used to identify and characterize defects or inhomogeneities at the interface itself, while suppressing sensitivity to features of the surrounding bulk media. In designing effective search methods to make use of interface-wave reflection, it is desirable to be able to search a large parameter region to identify what effects are possible, and in many cases very precise solution is not as important as the ability to qualitatively identify the origins and parameter dependence of signal features.

Another application involving interface-wave conversion concerns its presence as a loss mechanism in otherwise range-independent environments. While global-matrix codes[1] automatically account for the excitation of interface waves in response to placed sources, they necessarily omit secondary scattering, which occurs at isolated defects on the interface itself. Range-dependent enhancements of these methods[2] incorporate interface-wave production at explicit points of change, but are not efficient for modeling the presence of many such inclusions in an otherwise homogeneous medium.

Totally incoherent scattering into interface modes can be treated phenomenologically,[3] but scattering at isolated obstacles, even simple ones, requires augmentation of range-independent algorithms, preferably without adding great computational complexity.

This last application is of interest because ocean-bottom sensors have been used to monitor waterborne noise sources,[4,5] and the ability to estimate the importance of isolated obstacle-induced conversion, under the wide variety of ocean-bottom conditions that occur, is necessary to correctly interpret the results of such measurements. A finite-difference analysis of this problem has recently been presented, which can serve as a benchmark for analytic solutions based on the approach given here.[6,7]

It is therefore useful to have available a simple representation of the scattering amplitude into interface as well as bulk modes, from obstacles located near or on the interface itself. The representation that will be derived here is a perturbative multipole scattering expansion for a pointlike discontinuity of density, compressibility, and rigidity, located near a boundary in a homogeneous, isotropic elastic solid. The important feature of the calculation is the use of the stratified-medium Green's function to propagate scattered waves at intermediate orders of perturbation theory. For a dilute collection of obstacles, the discrete states induced by boundary conditions may be expected to dominate multiple as well as single scattering near the interface, so it is efficient to incorporate the boundary constraints as correction terms to the simple bulk scattering vertex. Because interface waves exist at a only a few allowed values of horizontal slowness (for practical purposes, often only one), the multipole expansion at lowest order can be reduced to an algebraic rather than integral form. This can be used as the basis for an algebraic perturbation expansion in interface-wave scattering, rather than one requiring wave number integrals. The closed-form first-order scattering amplitude reduces to a simple set of constants, which depend on medium parameters and give the proportionality between the relative strengths of radiated monopole, dipole and quadrupole interface mode excitation and the material inhomogeneity represented by the point defect. This approach is intended to be complementary to numerical techniques such as finite-element calculations that have been performed for equivalent physical systems.[8]

The primary motivation to perform the calculation in this way comes from an application of interface wave scattering to detect buried obstacles in unconsolidated sediments.[9] Experimental results on measurement and localization of such reflections are presented in a companion article.[10] The following two sections present the definition and construction of the scattering amplitude, and then its analysis and evaluation in a particular instance. The last sec-

tion provides a conversion from the the language of the scattering vertex to the notion of a target strength, and comments.

I. CONSTRUCTION OF THE SCATTERING THEORY

Point-obstacle perturbation theory, for a number of realistic problems, gives the ability to trade refinement of the solution, either of resolution or magnitude approximation, for computational efficiency. The basic perturbative building block is an amplitude for homogeneous waves, incident on a pointlike obstacle, to scatter into a basis of radiated continuum (bulk) and discrete (interface) modes. Because the scatterer is represented as a point, the elementary scattering vertex can be computed analytically, which makes it relatively easy to identify contributions from discrete and continuum modes, which can have very different propagation characteristics, to the scattered density of states. Complex or distributed scatterers may be modeled as collections of points at whatever resolution is appropriate to a given application, by means of a Huygens' construction of the scattered wave. Because the presence of the interface is represented as part of the elementary scattering vertex, any truncated perturbative order may be used to approximate the exact solution, and resummation techniques traditional in perturbation theory remain viable if coherent effects, such as resonant scattering, are important. This section contains the computation of the perturbative scattering amplitude in a simple system, chosen because it demonstrates the techniques and results of interest, but otherwise restricts attention to a relatively simple scattered spectrum.

A. General definitions and notation

Acoustic disturbances in a layered system of linear elastic solids or fluids are characterized by the displacement from equilibrium of the points of each of the media, denoted in each layer by a vector field **u**. It is convenient to decompose **u** as a sum of a gradient and a curl by writing

$$\mathbf{u}=\boldsymbol{\nabla}\phi+\boldsymbol{\nabla}\times\mathbf{A}, \tag{1}$$

where ϕ is a complex scalar field and **A** a complex vector field. The ambiguity of **A**, up to addition of an arbitrary gradient, can be removed by imposing the transversality condition

$$\boldsymbol{\nabla}\cdot\mathbf{A}=0. \tag{2}$$

If **A** has components A_i, and the gradient $\boldsymbol{\nabla}$ components $\boldsymbol{\nabla}_i$, with spatial index $i\in\{1,2,3\}$, the strain tensor $\boldsymbol{\epsilon}$ of the medium has components $\boldsymbol{\epsilon}_{ij}$ given by

$$\boldsymbol{\epsilon}_{ij}=\boldsymbol{\nabla}_i\boldsymbol{\nabla}_j\phi+\tfrac{1}{2}[\boldsymbol{\nabla}_i(\boldsymbol{\nabla}\times\mathbf{A})_j+\boldsymbol{\nabla}_j(\boldsymbol{\nabla}\times\mathbf{A})_i]. \tag{3}$$

The wave equation for the medium may be written in terms of the trace of the strain tensor $\mathrm{Tr}(\boldsymbol{\epsilon})\equiv\Sigma_i\boldsymbol{\epsilon}_{ii}$, and its transverse part $\boldsymbol{\epsilon}^T$, with components ${\boldsymbol{\epsilon}^T}_{ij}\equiv\boldsymbol{\epsilon}_{ij}-\delta_{ij}\,\mathrm{Tr}(\boldsymbol{\epsilon})$, as

$$\partial_t{}^2(\rho\mathbf{u})=\boldsymbol{\nabla}\cdot[2\mu\boldsymbol{\epsilon}^T]+\boldsymbol{\nabla}[(\lambda+2\mu)\mathrm{Tr}(\boldsymbol{\epsilon})]. \tag{4}$$

Here, λ and μ are the Lamé parameters, and ρ the density of the medium. The density and moduli are assumed to be time independent but spatially varying.

The particular case considered here will be that of a uniform isotropic half-space of fluid overlying a uniform isotropic half-space of elastic solid, with a scatterer located in the solid at ignorable distance below the interface. Because interface excitations are to leading order independent of system characteristics away from the interface that supports them, many of the important features which occur in a general case also appear in this very simple system. As a discontinuity of the solid, the scatterer can couple fully to its shear as well as to its compressional modulus and density. A rich structure of interface modes arises from the boundary conditions at the fluid–solid interface, which is sufficient to contain the features of interest in the companion application,[10] but simpler to compute than the most general case of layers of fluid and adjoining elastic solids. The description of this system will thus exclude phenomena such as scattering into Love modes[11] or more complicated normal modes trapped among multiple layers, but because these are simply additional discrete states of possible systems, they can be readily included either analytically at the price of more algebra, or with wave number integration[1] or mode codes,[12] using the vertex from the pointlike scatterer to provide the strengths of an effective source located in one of the media.

The fluid layer will be described by a homogeneous density ρ_f and zero shear modulus. All fluid parameters will be expressed in terms of ρ_f and the fluid sound speed c_f. The solid layer will be described by homogeneous and isotropic parameters except at a point, which must then be parametrized to represent the scatterer. While it may be natural to think of a small massive buried object as a point perturbation of the density with respect to wavelengths longer than the size of the object, it is less clear that pointlike perturbations in the moduli are meaningful. Physically, changes in moduli occur by finite amounts in macroscopic domains of the medium. Therefore, the perturbations of the moduli, while treated as pointlike, will be labeled with strengths which denote a finite change in modulus over some characteristic volume labeled R^3. The scaling with frequency of the resulting scattering amplitude will be used to argue that this representation is valid for perturbations smaller in extent than the relevant wavelengths. For weak scattering (i.e., in the Born approximation), it should then also be possible to represent larger regions of inhomogeneity with a distribution of pointlike scatterers.

The forms thus taken to describe the lower half-space with a single scatterer are

$$\rho(\mathbf{x})=\rho_0+M\,\delta^3(\mathbf{x}-\mathbf{x}_0), \tag{5}$$

$$\lambda(\mathbf{x})=\lambda_0+\Delta\lambda R^3\delta^3(\mathbf{x}-\mathbf{x}_0), \tag{6}$$

$$\mu(\mathbf{x})=\mu_0+\Delta\mu R^3\delta^3(\mathbf{x}-\mathbf{x}_0). \tag{7}$$

The compressional speed in the lower medium will be denoted by c and is given in terms of the background parameters by $c^2=(\lambda_0+2\mu_0)/\rho_0$, and the shear speed s by $s^2=\mu_0/\rho_0$. Because buried obstacles provide a physically natural application for low-order perturbative calculations, the density defect has been expressed in terms of a mass M of such an obstacle.

B. Solution of the wave equation

It is notationally convenient to treat the scalar and vector displacement potentials in each medium as elements of a four-component column vector ψ defined as

$$\psi \equiv \begin{bmatrix} \phi \\ \mathbf{A} \end{bmatrix}. \tag{8}$$

ψ will be taken to have a single frequency component and time dependence $e^{i\omega t}$ in all that follows. The Helmholtz equations for the scalar and vector potentials in a homogeneous medium are obtained by acting, respectively, with the gradient and curl operators on wave equation (4). The pair of operations can be represented as a matrix acting on the components of ψ which, after separation of the homogeneous from the inhomogeneous terms, takes the form

$$(\boldsymbol{\nabla}^2)\begin{bmatrix} (c^2\boldsymbol{\nabla}^2+\omega^2) & \\ & (s^2\boldsymbol{\nabla}^2+\omega^2)\mathbf{1} \end{bmatrix}\psi$$

$$= \sum_{\alpha=1,2} \mathbf{P}_1^{(\alpha)}(\boldsymbol{\nabla})[\mathbf{P}_2^{(\alpha)}\delta^3(\mathbf{x}-\mathbf{x}_0)(\mathbf{P}_3^{(\alpha)}(\boldsymbol{\nabla})\psi)]. \tag{9}$$

In this expression the $\mathbf{P}_i^{(\alpha)}$ are matrices with polynomial coefficients in components of the gradient operator and the perturbation strengths. The values $\alpha=1,2$ correspond, respectively, to perturbations in density or the moduli, which involve different orders of spatial derivatives. Explicit expressions for the $\mathbf{P}_i^{(\alpha)}$ are given in Appendix A, and $\mathbf{1}$ is the unit diagonal. This form for the equations of motion lends itself to a perturbative treatment in which the inverse of the operator on the left-hand side is the Green's function for the homogeneous medium and the polynomial expression on the right-hand side is the vertex associated with the perturbation.

To solve Eq. (9) perturbatively, one divides the wave ψ into parts $\psi=\psi_0+\delta\psi$, requiring that

$$(\boldsymbol{\nabla}^2)\begin{bmatrix} (c^2\boldsymbol{\nabla}^2+\omega^2) & \\ & (s^2\boldsymbol{\nabla}^2+\omega^2)\mathbf{1} \end{bmatrix}\psi_0 \equiv 0. \tag{10}$$

The homogeneous wave operator is invertible in a wave number basis, defined by the relation

$$\psi_{\mathbf{x}} \equiv \int d^3k e^{-i\mathbf{k}\cdot\mathbf{x}}\psi_{\mathbf{k}}. \tag{11}$$

The solution of the full equations of motion may be generated in two steps. The first is to compute an expression for the wave that would be scattered from the same pointlike perturbation in the lower medium in the *absence* of a boundary constraint. Denoted $\delta\psi_{\mathbf{k}_{\text{n.b.}}}$, this part of the scattered wave follows from inversion of Eq. (9), using Eq. (10),

$$\delta\psi_{\mathbf{k}_{\text{n.b.}}} = \frac{-1}{(2\pi)^3}\begin{bmatrix} \dfrac{1}{(\omega^2-c^2\mathbf{k}^2)} & \\ & \dfrac{1}{(\omega^2-s^2\mathbf{k}^2)}\mathbf{1} \end{bmatrix}\frac{1}{\mathbf{k}^2}\int d^3k'$$

$$\times \sum_{\alpha=1,2} \mathbf{P}_1^{(\alpha)}(-i\mathbf{k})\mathbf{P}_2^{(\alpha)}\mathbf{P}_3^{(\alpha)}(-i\mathbf{k}')e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}_0}\psi_{\mathbf{k}'}. \tag{12}$$

The function ψ appearing in the integrand is the solution to the complete equations of motion, including the constraint from the boundary. The scalar product $\mathbf{k}\cdot\mathbf{k}$ is denoted $\mathbf{k}^2$ to distinguish it from the scalar product of the two-component horizontal wave vector, which will be used explicitly in later steps.

The vertex strength is parametrized in terms of three ratios with the dimensions of volume, defined, respectively, for compression, density, and shear as

$$\mathbf{C} \equiv \frac{(\Delta\lambda)R^3}{(\lambda_0+2\mu_0)}, \tag{13}$$

$$\mathbf{D} \equiv \frac{M}{\rho_0}, \tag{14}$$

$$\mathbf{S} \equiv \frac{\Delta\mu R^3}{\mu_0}. \tag{15}$$

The inhomogeneous term in Eq. (12) may be written in terms of these, as

$$\sum_{\alpha=1,2} \mathbf{P}_1^{(\alpha)}(-i\mathbf{k})\mathbf{P}_2^{(\alpha)}\mathbf{P}_3^{(\alpha)}(-i\mathbf{k}')$$

$$= -c^2\mathbf{C}\begin{bmatrix} \mathbf{k}^2\mathbf{k}'^2 & 0 \\ 0 & 0 \end{bmatrix} + (\omega^2\mathbf{D} - 2s^2\mathbf{S}(\mathbf{k}\cdot\mathbf{k}'))$$

$$\times\begin{bmatrix} \mathbf{k}\cdot\mathbf{k}' & \mathbf{k}\times\mathbf{k}' \\ -\mathbf{k}\times\mathbf{k}' & [(\mathbf{k}\cdot\mathbf{k}')\mathbf{1}-\mathbf{k}'\mathbf{k}] \end{bmatrix}$$

$$+ s^2\mathbf{S}\begin{bmatrix} 0 & 0 \\ 0 & [(\mathbf{k}\times\mathbf{k}')(\mathbf{k}\times\mathbf{k}')+(\mathbf{k}\cdot\mathbf{k}')((\mathbf{k}\cdot\mathbf{k}')\mathbf{1}-\mathbf{k}'\mathbf{k})] \end{bmatrix}, \tag{16}$$

in which the products $(\mathbf{k}\times\mathbf{k}')(\mathbf{k}\times\mathbf{k}')$ and $\mathbf{k}'\mathbf{k}$ are to be understood as dyadics. The symmetry decomposition of this bulk scattering vertex is performed in Appendix B. It is shown that from arbitrary incoming waves satisfying the transversality condition (2), the vertex produces scattered monopole, dipole and quadrupole components, with expansions in terms of respectively zeroth, first, and second level spherical harmonics in a basis defined by the direction and polarization of the incoming wave. The symmetries of the bulk vertex are slightly different from those that will result from projection onto interface waves, because dipole components normal to the interface appear isotropic in the horizontal plane.

The boundary constraint is next implemented by solving for the wave that must be added to $\delta\psi_{\text{n.b.}}$ for the sum to obey the boundary conditions. The boundary-induced correction will be expressed as an improvement to the bulk vertex, thus producing a perturbation series for the solution which satisfies the boundary constraint order by order. The scatterer will be taken to lie at a small finite distance z_0 below the interface, with $|z_0|$ larger than any infinitesimals ϵ needed to define contour prescriptions in wave vector integrals. z_0 will then be taken to zero at the end of the calculation.

The constraint is represented as follows. A general incident wave satisfying the homogeneous wave equation at a given frequency is decomposed into a sum of plane waves

indexed by the horizontal wave number $\vec{k}$ and a polarization for the vector shear wave in the solid. (The corresponding vertical wave numbers are fixed by the wave equation.) The transversality condition $\nabla\cdot\mathbf{A}=0$ is satisfied without loss of generality in this system if the vector potential is canonically chosen to have polarization $\mathbf{A}\propto\hat{\mathbf{z}}\times\mathbf{k}$, where $\hat{\mathbf{z}}$ is the unit normal to the interface and $\mathbf{k}$ is the shear wave vector with horizontal component $\vec{k}$.

A set of arbitrarily normalized basis functions $\{\varphi^f_{\vec{k}},\varphi_{\vec{k}},\mathbf{a}_{\vec{k}}\}_{\rm in}$, $\{\varphi^f_{\vec{k}},\varphi_{\vec{k}},\mathbf{a}_{\vec{k}}\}_{\rm out}$ is introduced for respectively incoming and outgoing scalar components in the two layers, and the vector component in the solid. Square roots for the $\hat{\mathbf{z}}$ wave numbers are chosen so that outgoing radiative roots continue to evanescent roots bounded away from the interface, and incoming radiative roots to unbounded evanescent roots. This is the correct prescription because, with the scatterer in the bulk, the scattered wave incident on the interface is characterized by unbounded or in-radiative roots, whereas the physical Green's function propagating solutions from the interface selects entirely out-radiative or bounded roots. The no-boundary solution $\delta\psi_{\mathbf{k}_{\rm n.b.}}$ of Eq. (12) provides incoming components $\phi^f_{\rm in}$, $\phi_{\rm in}$, $\mathbf{A}_{\rm in}$ for an expansion of the full scattered wave in terms of the basis functions:

$$\begin{aligned}
\phi^f_{\vec{k};\rm in,out}&=\gamma^f_{\rm in,out}\varphi^f_{\vec{k};\rm in,out},\\
\phi_{\vec{k};\rm in,out}&=\gamma_{\rm in,out}\varphi_{\vec{k};\rm in,out},\\
\mathbf{A}_{\vec{k};\rm in,out}&=\sigma_{\rm in,out}\mathbf{a}_{\vec{k};\rm in,out}.
\end{aligned}\tag{17}$$

The outgoing components will be fixed by the boundary constraints.

The differential fluid–solid interface conditions, continuity of vertical motion and the vertical components of the stress tensor, may be written by means of the basis expansion (17) as algebraic constraints on the coefficients γ,σ. Expressed collectively as matrices acting on vectors of the coefficients, these may be arranged to read

$$\begin{aligned}
&\begin{bmatrix}\sqrt{\omega^2/c_f^2-k^2} & \sqrt{\omega^2/c^2-k^2} & -k^2\\ -\dfrac{\rho_f}{\rho}\left(\dfrac{\omega}{s}\right)^2 & -(2k^2-\omega^2/s^2) & -2k^2\sqrt{\omega^2/s^2-k^2}\\ 0 & 2\sqrt{\omega^2/c^2-k^2} & -(2k^2-\omega^2/s^2)\end{bmatrix}\\
&\quad\times\begin{bmatrix}\gamma^f_{\rm out}\\ \gamma_{\rm out}\\ \sigma_{\rm out}\end{bmatrix}\\
&=\begin{bmatrix}\sqrt{\omega^2/c_f^2-k^2} & \sqrt{\omega^2/c^2-k^2} & k^2\\ \dfrac{\rho_f}{\rho}\left(\dfrac{\omega}{s}\right)^2 & (2k^2-\omega^2/s^2) & -2k^2\sqrt{\omega^2/s^2-k^2}\\ 0 & 2\sqrt{\omega^2/c^2-k^2} & (2k^2-\omega^2/s^2)\end{bmatrix}\\
&\quad\times\begin{bmatrix}\gamma^f_{\rm in}\\ \gamma_{\rm in}\\ \sigma_{\rm in}\end{bmatrix}.
\end{aligned}\tag{18}$$

Square roots denote the out-radiative/bounded branch (which has opposite signs in the fluid and solid), and k^2 is the scalar product $\vec{k}\cdot\vec{k}$ of the horizontal wave vector.

Values of the transverse wave vector for which the first constraint matrix in Eq. (18) is degenerate admit outgoing waves which may exist with nonzero amplitude even if the incoming waves vanish. These, of course, are the interface wave states. With correct treatment of such degenerate cases, the first constraint matrix may be inverted to express the outgoing coefficients in terms of the incoming, in which case a solution, having the same incoming components as $\delta\psi_{\mathbf{k}_{\rm n.b.}}$, but satisfying the boundary conditions, is given by

$$\begin{aligned}
\phi^f_{\vec{k};\rm full}&=\gamma^f_{\rm in}\varphi^f_{\vec{k};\rm in}+\gamma^f_{\rm out}\varphi^f_{\vec{k};\rm out},\\
\phi_{\vec{k};\rm full}&=\gamma_{\rm in}\varphi_{\vec{k};\rm in}+\gamma_{\rm out}\varphi_{\vec{k};\rm out},\\
\mathbf{A}_{\vec{k};\rm full}&=\sigma_{\rm in}\mathbf{a}_{\vec{k};\rm in}+\sigma_{\rm out}\mathbf{a}_{\vec{k};\rm out}.
\end{aligned}\tag{19}$$

The coefficients are related by a product of the constraint matrices which may be treated as a Jacobian from the incoming to the outgoing basis:

$$\begin{aligned}
&\begin{bmatrix}\gamma^f_{\rm out}\\ \gamma_{\rm out}\\ \sigma_{\rm out}\end{bmatrix}\\
&=\begin{bmatrix}\sqrt{\omega^2/c_f^2-k^2} & \sqrt{\omega^2/c^2-k^2} & -k^2\\ -\dfrac{\rho_f}{\rho}\left(\dfrac{\omega}{s}\right)^2 & -(2k^2-\omega^2/s^2) & -2k^2\sqrt{\omega^2/s^2-k^2}\\ 0 & 2\sqrt{\omega^2/c^2-k^2} & -(2k^2-\omega^2/s^2)\end{bmatrix}^{-1}\\
&\quad\times\begin{bmatrix}\sqrt{\omega^2/c_f^2-k^2} & \sqrt{\omega^2/c^2-k^2} & k^2\\ \dfrac{\rho_f}{\rho}\left(\dfrac{\omega}{s}\right)^2 & (2k^2-\omega^2/s^2) & -2k^2\sqrt{\omega^2/s^2-k^2}\\ 0 & 2\sqrt{\omega^2/c^2-k^2} & (2k^2-\omega^2/s^2)\end{bmatrix}\\
&\quad\times\begin{bmatrix}\gamma^f_{\rm in}\\ \gamma_{\rm in}\\ \sigma_{\rm in}\end{bmatrix}\equiv\left[\frac{\partial(\gamma^f_{\rm out},\gamma_{\rm out},\sigma_{\rm out})}{\partial(\gamma^f_{\rm in},\gamma_{\rm in},\sigma_{\rm in})}\right]\begin{bmatrix}\gamma^f_{\rm in}\\ \gamma_{\rm in}\\ \sigma_{\rm in}\end{bmatrix}.
\end{aligned}\tag{20}$$

For application to the companion experiment,[10] as well as to ocean-bottom seismometry, the receiver as well as the obstacle will be taken to be anchored in the solid at the interface, so the motion of the fluid is not directly of interest. Since the incident wave from the scatterer consists entirely of components in the solid layer, it is sufficient to consider only the Jacobian $[\partial(\gamma_{\rm out},\sigma_{\rm out})/\partial(\gamma_{\rm in}\sigma_{\rm in})]$ which is the projection of the matrix defined in Eq. (20) onto the coefficients of the basis functions that describe the solid.

The full scattered wave in Eq. (19) may now be expressed in terms of the incoming component generated by Eq. (12). Fourier transforming solutions (17) of the wave equation back to a full three-dimensional plane wave representation, taking care with the Fourier transforms that relate incoming to outgoing $\hat{\mathbf{z}}$ wave numbers in the basis functions, gives the representation

$$\delta\psi_{\vec{k},\lambda}=\frac{-1}{(2\pi)^3}\begin{bmatrix}\dfrac{1}{(\omega^2/c^2-k^2-\lambda^2)} & \\ & \dfrac{1}{(\omega^2/s^2-k^2-\lambda^2)}\,\mathbf{1}\end{bmatrix}\int d\lambda'\left\{\delta(\lambda-\lambda')+\frac{i}{\pi}\begin{bmatrix}\sqrt{\omega^2/c^2-k^2} & \\ & \sqrt{\omega^2/s^2-k^2}\,\mathbf{1}\end{bmatrix}\right.$$
$$\left.\times\left[\frac{\partial(\gamma_{\rm out},\sigma_{\rm out})}{\partial(\gamma_{\rm in},\sigma_{\rm in})}\right]\begin{bmatrix}\dfrac{1}{(\omega^2/c^2-k^2-\lambda'^2)} & \\ & \dfrac{1}{(\omega^2/s^2-k^2-\lambda'^2)}\,\mathbf{1}\end{bmatrix}\right\}\begin{bmatrix}\dfrac{1}{c^2} & \\ & \dfrac{1}{s^2}\,\mathbf{1}\end{bmatrix}$$
$$\times\frac{e^{i\lambda'\epsilon}}{\mathbf{k}'^2}\int d^3k''\sum_{\alpha=1,2}\mathbf{P}_1^{(\alpha)}(-i\mathbf{k}')\mathbf{P}_2^{(\alpha)}\mathbf{P}_3^{(\alpha)}(-i\mathbf{k}'')e^{i(\mathbf{k}'-\mathbf{k}'')\cdot\mathbf{x}_0}\psi_{\mathbf{k}''}. \tag{21}$$

In expression (21) the transverse and vertical components of the wave vector are denoted explicitly. Three-component wave vectors are defined as $\mathbf{k}=(\vec{k},\lambda)$ and $\mathbf{k}'=(\vec{k},\lambda')$. The integral over λ' relates the incoming to the outgoing basis states used to implement the boundary condition in Eq. (19), and the term $e^{i\lambda'\epsilon}$ defines the prescription for selecting poles of the argument $1/\mathbf{k}'^2$. This additional integral may be taken as the improvement of the vertex term to one which automatically produces an outgoing state respecting the boundary conditions.

The physics of solution (21) may be better understood by decomposing the Jacobian into two very simple terms. Specifically, it is possible to write

$$\left[\frac{\partial(\gamma_{\rm out},\sigma_{\rm out})}{\partial(\gamma_{\rm in},\sigma_{\rm in})}\right]$$
$$=\begin{bmatrix}1 & 0\\ 0 & -\mathbf{1}\end{bmatrix}-\frac{2}{\rm Det}\begin{bmatrix}(2k^2-\omega^2/s^2)\\ 2\sqrt{\omega^2/c^2-k^2}\hat{\mathbf{z}}\times\mathbf{k}\end{bmatrix}[(2k^2-\omega^2/s^2)$$
$$-2\sqrt{\omega^2/s^2-k^2}\hat{\mathbf{z}}\times\mathbf{k}], \tag{22}$$

where

$$\text{Det}\equiv\frac{\rho_f}{\rho}\left(\frac{\omega}{s}\right)^4\frac{\sqrt{\omega^2/c^2-k^2}}{\sqrt{\omega^2/c_f^2-k^2}}+(2k^2-\omega^2/s^2)^2$$
$$+4k^2\sqrt{\omega^2/c^2-k^2}\sqrt{\omega^2/s^2-k^2} \tag{23}$$

is the appropriately scaled determinant of the constraint matrix acting on the outgoing coefficients in Eq. (18). Thus the values of the wave vector for which the outgoing constraint becomes degenerate are the zeros of the determinant, which produce poles in the density of states that make possible scattering of finite energy into discretely indexed modes.

The first term in decomposition (22), together with the incident wave preserved by the $\delta(\lambda-\lambda')$ in Eq. (21), produces a symmetric sum of in-going and out-going modes which has identically zero vertical motion (and vertically directed shear stress). The second term may be checked to be an out-going wave which satisfies the constraint of zero vertically directed shear, multiplied by a projector which acts on the state produced by the scatterer. All sensitivity to properties of the upper fluid layer is contained in the first summand in the determinant, which is a ratio of the mechanical impedances of the upper and lower compressional modes times a ratio of the vertical components of their velocities. Taking the impedance of the fluid to infinity effectively removes the term proportional to 1/Det, leaving only the zero-vertical motion state consistent with an immovable upper fluid layer. Taking the same impedance to zero produces the limiting Rayleigh solution. For finite impedance of the upper layer, corrections to the symmetric sum provided by the first term in Eq. (22) come from the wave vector integrals around the branch cuts induced by square roots in the term containing the determinant, and the strength of scattering into any bound states is given by the residues of its poles, because all other terms are explicitly regular.

It is desirable to separate various terms from the scattering integral according to the range dependence of the scattered wave components they contribute. Physically, one expects modes radiating into the bulk to die off as a larger power of distance from the scatterer than modes trapped at the interface. To leading order at long range, the scattered wave should therefore include only the trapped Scholte pole on the physical sheet of the complex wave vector plane (for realistic values of Poisson's ratio in granular sediments[13]). For systems with additional layers or gradients that refract toward the interface, however, the situation may be more complicated. Normally ignorable poles representing ''Leaky Scholte modes'' may be brought from unphysical to physical sheets as frequency is varied. Yet this emergence of an additional pole from what was formerly a branch integral results from a continuous deformation of parameters. The qualitative separation of modelike from bulk-radiative effects is thus not clearly defined. In what follows, branch integrals will not be evaluated explicitly. However, it is important to note that in addition to spherically spreading waves which are less important at long range, those integrals may contain resonances arising from ''partial contour integrals'' around nearby leaky poles, in which case the strength of scattering into such poles may be estimated analytically from complete contour integrals enclosing them, even when they are not located at physical wave number values.

II. EVALUATION OF THE SCATTERING MATRIX

The first step in evaluating Eq. (21) is to perform the λ' integration, so the fields satisfy the homogeneous wave equation in the bulk. The Green's function appearing in the full solution (21) contains more poles than would be expected from the physical degrees of freedom of the theory. These are created by the term $1/\mathbf{k}'^2$, and arise because the fields being propagated are not the physical degrees of freedom, but rather their scalar and vector potentials. The only such integral which will be evaluated here is the one involving the determinant, because the other, which leads to a factor of $1/\mathbf{k}^2$ in the outgoing state, affects only the symmetric, bulk-radiative term. The determinant itself is a function only of the horizontal wave vector, so in fact the only effect of the integration over λ' is to induce appropriate projections of the the terms involving $\mathbf{k}'$ in the vertex integral over $\mathbf{k}''$.

At this point a great simplification is obtained by using the fact that the perturbation series defined by Eq. (21) satisfies the boundary conditions at each order. This feature makes it possible to restrict attention to terms $\psi_{\mathbf{k}''}$ in the integrand which satisfy the boundary conditions, without concern for whether these are intended to represent the full solution or intermediate orders of perturbation theory in an iterative approximation. The general projection of such a wave onto the components of the lower half-space is given in the position representation by

$$\psi_{\mathbf{x}_{\text{b.c.}}} = \int d^2k\ e^{-i\vec{k}\cdot\vec{x}} \begin{bmatrix} (2k^2-\omega^2/s^2)e^{i\sqrt{\omega^2/c^2-k^2}z} \\ 2\sqrt{\omega^2/c^2-k^2}\hat{\mathbf{z}}\times\mathbf{k}e^{i\sqrt{\omega^2/s^2-k^2}z} \end{bmatrix} \psi_{\vec{k}}. \tag{24}$$

The three-vector $\mathbf{k}$ appearing in the cross product is defined to have as its horizontal component $\vec{k}$, and the remaining density $\psi_{\vec{k}}$ need not be further restricted at this point.

Evaluating the two integrals (over λ' and λ'', the vertical component of $\mathbf{k}''$), reduces solution (21) to a form

$$\begin{aligned}\delta\psi_{\vec{k},\lambda} &= \delta\psi_{\vec{k},\lambda,\text{symmetric bulk}} \\ &+ \frac{1}{(2\pi)^3} \begin{bmatrix} \dfrac{\sqrt{\omega^2/c^2-k^2}}{(\omega^2/c^2-k^2-\lambda^2)} & \\ & \dfrac{\sqrt{\omega^2/s^2-k^2}}{(\omega^2/s^2-k^2-\lambda^2)}\mathbf{1} \end{bmatrix} \\ &\times \begin{bmatrix} (2k^2-\omega^2/s^2) \\ 2\sqrt{\omega^2/c^2-k^2}\hat{\mathbf{z}}\times\mathbf{k} \end{bmatrix} \left(\frac{2}{\text{Det}}\right) \\ &\times \frac{1}{\sqrt{\omega^2/c^2-k^2}} \int d^2k''\ e^{i(\vec{k}-\vec{k}'')\cdot\vec{x}_0} [\mathbf{C}\ \mathbf{D}\ \mathbf{S}] \\ &\times \left[\frac{\partial(\mathscr{M},\mathscr{D},\mathscr{Q})}{\partial(\mathbf{C},\mathbf{D},\mathbf{S})}\right] \begin{bmatrix} 1 \\ (\vec{k}\cdot\vec{k}'') \\ (\vec{k}\cdot\vec{k}'')^2 \end{bmatrix} \psi_{\vec{k}''}. \end{aligned} \tag{25}$$

The first (diagonal) matrix in Eq. (25) may be recognized as the full three-vector Green's function for the compressional and shear Helmholtz equations in the solid. The column vector is the polarization vector for waves satisfying the boundary conditions, and the matrix $[\partial(\mathscr{M},\mathscr{D},\mathscr{Q})/\partial(\mathbf{C},\mathbf{D},\mathbf{S})]$ is the transformation that relates the perturbation strengths to the amplitudes of scattered components proportional to the zeroth, first and second powers of in-plane scattering angle.[14] The only nonzero components of this matrix are given in terms of the incoming and outgoing horizontal wave numbers as

$$\frac{\partial\mathscr{M}}{\partial\mathbf{C}} = -\frac{\omega^2}{c^2}(2k^2-\omega^2/s^2)(2k''^2-\omega^2/s^2),$$

$$\begin{aligned}\frac{\partial\mathscr{M}}{\partial\mathbf{D}} &= -2\frac{\omega^2k^2}{s^2}\left(1-\frac{i\sqrt{\omega^2/s^2-k^2}}{|k|}\right) \\ &\times\sqrt{\omega^2/c^2-k^2}\sqrt{\omega^2/c^2-k''^2},\end{aligned}$$

$$\begin{aligned}\frac{\partial\mathscr{M}}{\partial\mathbf{S}} &= 2\frac{s^2k^2}{\omega^2}\left(1-\frac{i\sqrt{\omega^2/c^2-k^2}}{|k|}\right)\left(\frac{i\sqrt{\omega^2/c^2-k^2}}{|k|}\right) \\ &\times(2k^2-\omega^2/s^2)\{(2k''^2-\omega^2/s^2)(k''^2-\omega^2/c^2) \\ &+2k''^2\sqrt{\omega^2/c^2-k''^2}\sqrt{\omega^2/s^2-k''^2}\},\end{aligned} \tag{26}$$

$$\begin{aligned}\frac{\partial\mathscr{D}}{\partial\mathbf{D}} &= \left(1-\frac{i\sqrt{\omega^2/c^2-k^2}}{|k|}\right)(2k^2-\omega^2/s^2) \\ &\times[(2k''^2-\omega^2/s^2)+2\sqrt{\omega^2/c^2-k''^2}\sqrt{\omega^2/s^2-k''^2}],\end{aligned}$$

$$\begin{aligned}\frac{\partial\mathscr{Q}}{\partial\mathbf{S}} &= -2\frac{s^2}{\omega^2}\left(1-\frac{i\sqrt{\omega^2/c^2-k^2}}{|k|}\right)(2k^2-\omega^2/s^2) \\ &\times[(2k''^2-\omega^2/s^2)+2\sqrt{\omega^2/c^2-k''^2}\sqrt{\omega^2/s^2-k''^2}].\end{aligned}$$

Density perturbation contributes a component isotropic in the plane, because of the large vertical component of interface wave motion, and the fact that dipole scattering oriented normal to the surface looks monopolar from the point of view of interface waves. Otherwise the symmetry of in-plane scattering respects that of the three-vector vertex (16) in the bulk.

In a dilute-scatterer approximation, where interface modes are expected to dominate, the appropriate expansion of the scattered wave is in terms of outgoing radial functions. Further, for weak scattering, where self-interaction of the scattered wave is expected to be small, it is sufficient to consider the effects of a plane incident wave satisfying the boundary conditions at a single wave number, in which case $\psi_{0,\mathbf{x}}$ takes the form (24), with

$$\psi_{0,\vec{k}} = \varphi_0\delta^2(\vec{k}-\vec{k}_0). \tag{27}$$

The wave number integral for the incident part of the full wave in this case becomes trivial.

Performing the wave number integrations, evaluating the residues at any poles of the determinant by Cauchy's theorem, and keeping explicit only leading lowest order terms in the far field gives the result

$$\delta\psi_{\mathbf{x}}=(\delta\psi_{\mathbf{x},\text{bulk}}+\delta\psi_{\mathbf{x},\text{higher order}})+\sum_{\text{poles}} e^{i(k\hat{x}-\vec{k}_0)\cdot\vec{x}_0}$$
$$\times\begin{bmatrix}(2k^2-\omega^2/s^2)e^{i\sqrt{\omega^2/c^2-k^2}z}\mathrm{J}_0^+(kx)\\ 2\sqrt{\omega^2/c^2-k^2}e^{i\sqrt{\omega^2/s^2-k^2}z}i\mathrm{J}_1^+(kx)k\hat{\mathbf{z}}\times\hat{\mathbf{x}}\end{bmatrix}\varphi_0$$
$$\times\left(2\sqrt{\omega^2/c^2-k^2}\,\frac{\partial\ \mathrm{Det}}{\partial k^2}\right)^{-1}[\mathbf{C}\ \ \mathbf{D}\ \ \mathbf{S}]$$
$$\times\left[\frac{\partial(\mathscr{M},\mathscr{D},\mathscr{Q})}{\partial(\mathbf{C},\mathbf{D},\mathbf{S})}\right]\begin{bmatrix}1\\ (k\hat{x}\cdot\vec{k}_0)\\ (k\hat{x}\cdot\vec{k}_0)^2\end{bmatrix}. \tag{28}$$

The three-vector $\mathbf{x}\equiv(\vec{x},z)$, with $x\equiv|\vec{x}|$ and $\hat{\mathbf{x}}\equiv(\vec{x}/x,0)\equiv(\hat{x},0)$. Similarly, $k\equiv|\vec{k}|$. In the expansion in leading terms, it has been assumed that the distance from the origin of coordinates to the scatterer $|\vec{x}_0|\ll x$. J_i^+ denotes the outgoing complex component of the Bessel function of order i, and as required the vector component in the scattered wave takes value zero at the origin, removing the orientation singularity of the vector polarization. The magnitude k is to be evaluated at the position of the pole in each term in the sum, and in the final factor, the wave vectors $\vec{k}$ and $\vec{k}''$ of Eq. (25) are evaluated as $k\hat{x}$ and $\vec{k}_0$, respectively.

Recalling the form of input wave (24), the first term is simply the outgoing radial wave with the same normalization at the origin as the incident plane wave, so the coefficients that follow give the relative magnitudes of the incident and scattered waves as a function of the wave numbers of each. The phase assigned to the component from each pole is simply that obtained by tracing along rays in a geometric-optics description, so when the scattered amplitude is perturbatively small, a representation of a distributed scatterer by Huygens' construction is consistent. The scattering strengths may be nondimensionalized by grouping a factor of $(\omega/s)^3=(2\pi/\lambda_{\text{shear}})^3$ with the perturbation strengths $[\mathbf{C}\ \mathbf{D}\ \mathbf{S}]$ to give the small-parameter vector $(2\pi R/\lambda_{\text{shear}})^3[\Delta\lambda/(\lambda_0+2\mu_0)\ M/(R^3\rho_0)\ \Delta\mu/\mu_0]$, and separating the norms k and $|\vec{k}_0|$ from their respective powers of the direction cosine $\hat{k}_0\cdot\hat{x}$. The vertex is thus expressed in terms of a matrix of dimensionless scattering strengths relating equivalently normalized planar incident and radial scattered waves at first order. Denoted $\mathbf{V}^{(1)}$, this matrix is given by

$$\mathbf{V}^{(1)}\equiv\left(2\sqrt{\omega^2/c^2-k^2}\,\frac{\partial\ \mathrm{Det}}{\partial k^2}\right)^{-1}\left(\frac{s}{\omega}\right)^3\left[\frac{\partial(\mathscr{M},\mathscr{D},\mathscr{Q})}{\partial(\mathbf{C},\mathbf{D},\mathbf{S})}\right]$$
$$\times\begin{bmatrix}1 & & \\ & (kk_0) & \\ & & (kk_0)^2\end{bmatrix}. \tag{29}$$

The definition of a pointlike scatterer is that the combination $(2\pi R/\lambda_{\text{shear}})^3$ be less than or comparable to unity, and the remaining specifications of the perturbation strengths are then expressed as fractional changes relative to the background values.

The detailed behavior of the first-order vertex depends on the choice of poles used for incident and scattered waves, and on parameter values. However, it has several features, including the general order of magnitude and the relative dependence of scattering on density and modular perturbations, which are qualitatively parameter independent. These may be represented by giving the values of $\mathbf{V}^{(1)}$ for the scattering of Rayleigh→Rayleigh waves, in the absence of dissipation, when the compressional speed is assumed much larger than the shear speed. For this case, the position and residue of the pole is asymptotically independent of the compressional speed. The prefactor in the scattering amplitude term of Eq. (28), including the residue from the determinant, contributes a nearly parameter-independent suppression of magnitude $\sim(0.1)$ (true also for scattering into the leaky Rayleigh pole). All vertex elements in Eq. (26) are bounded above on the physical sheet at order unity, but for this range of cases all except the coefficient $\partial\mathscr{M}/\partial D$ are suppressed by a factor of s^2/c^2. Supposing for example a shear speed of 200 m/s and a compressional speed of 1800 m/s yields a first-order vertex

$$\mathbf{V}^{(1)}_{\text{Rayleigh}\to\text{Rayleigh}}=\begin{bmatrix}-0.0019i & & \\ 0.1795i & 0.0004i & \\ 0.0009i & & -0.0010i\end{bmatrix}. \tag{30}$$

Thus increases in density by a factor of 2 or more across substantial fractions of a shear wavelength are expected to be visible at the 10% level in *amplitude* near the target. While changes in shear modulus much larger than unity are realistic, e.g., for hard inclusions in unconsolidated sediments, the difficulty of coupling such discontinuities to the sediment frame may make realizable scattering strengths smaller than this vertex would indicate. Also, contributions from realizable variations in compressibilities, as well as density-driven in-plane dipole scattering, are suppressed by powers of the compressional speed, corresponding to the smaller magnitude of compressional motion relative to that of shearing.

III. APPLICATION AND COMMENTS

An application for low-order perturbative scattering amplitudes is the location and characterization of objects buried in unconsolidated sediments.[10] Since this is essentially a sonar application, it conventionally employs a description of the scattering amplitude in terms of a target strength, defined as[15]

$$\mathrm{TS}\equiv 10\log_{10}\left(\frac{I_r(1\ \mathrm{m})}{I_i}\right), \tag{31}$$

where $I_r(1\ \mathrm{m})$ is the scattered (radiated) intensity at a distance of 1 m from the scatterer, and I_i in the incident intensity of the insonifying plane wave. Because the scatterer is a point source, the acoustic center from which distances are measured is simply its location.

For small scattered amplitude, the first-order perturbative result should be a useful approximation to the full scattered amplitude. In the respective limits of long and short wavelength, this approximation defines a first-order bistatic target strength as a function of the wavelengths of the inci-

dent and scattered waves and the angle θ between the direction of incidence and the direction of reception from the target, with the forms:

$$\mathrm{TS}^{(1)}(\lambda_i,\lambda_r,\theta) = \left\{ \begin{array}{ll} \lambda_r \gg \pi^2(\mathrm{m}); & -20\log_{10}(2) \\ \lambda_r \ll \pi^2(\mathrm{m}); & -20\log_{10}(2\pi)+10\log_{10}\left(\dfrac{\lambda_r}{1m}\right) \end{array} \right\} + 20\log_{10}\left(\left(\frac{\omega}{s}\right)^3 [\mathbf{C}\ \ \mathbf{D}\ \ \mathbf{S}]\ \ \mathbf{V}^{(1)}\Big|_{\lambda_i,\lambda_r,\theta} \begin{bmatrix} 1 \\ \cos\theta \\ \cos^2\theta \end{bmatrix} \right). \tag{32}$$

The scattering amplitude in the logarithm scales as ω^3, which may be recognized as the scaling of the Rayleigh cross section ($\sim\omega^{d+1}$) for $d=2$. Thus though the obstacle was introduced formally as a feature of the bulk, projection onto the interface-wave pole yields the characteristics of wave scattering from pointlike obstacles two dimensions. The combination of this scaling, for all perturbations, with the angular dependence of the scattered wave that distinguishes compressional, shear, and density coupling, shows how small finite regions may be consistently treated as point perturbations.

As an example, scattering of 1-m Rayleigh waves from a density mismatch of 100% occupying a $(1\ \mathrm{m}/2\pi)^3$ volume gives a monopole contribution at -31 dB. In unconsolidated sand of density 2 g/cm^3, this would describe a 16-kg buried object. When propagation characteristics are included, these numbers compare favorably to experimental results for density inclusions.[10]

A similar computation for a shear-only diffractor has been performed with finite-element methods in the point-scatterer limit.[8] The parameters for that system were 250-m/s background shear speed and 30-Hz signal, with a diffractor of volume roughly 1.5 m^3 buried 1 m below a solid–fluid interface in a model resembling a surf zone. The water overlayer at the position of the diffractor was very thin, and will be ignored here. The first-order perturbative calculation predicts almost identical monopole and quadrupole scattered Rayleigh wave components at one wavelength, of magnitude -63 dB relative to the wave impinging on the target. When this value is corrected for propagation loss terms, it appears to under-predict scattering by roughly 6 dB relative to the numerical result. This discrepancy is consistent, at one estimated standard deviation, with the approximations made in computing propagation loss and the difficulty of specifying the shape and effective radius of the inclusion in the coarse-grained, higher order finite-element scheme.[16] The perturbative prediction of comparable monopole and quadrupole components in the scattered wave also appears consistent with angular dependence shown in the numerical result.

Nominally, the target strength (32) is direction dependent. But, as noted, for a wide range of realistic parameter values the scattered amplitude is dominated by the monopole term coupled to density defects, so for practical purposes any directionality in the bistatic target strength is likely to be dominated by the spatial distribution of the target itself.

ACKNOWLEDGMENTS

This work has been supported by the Office of Naval Research, and by the Advanced Research Projects Agency, Maritime Systems Technology Office. The author also thanks T. G. Muir for support of this and the associated experimental work.

APPENDIX A: MATRIX EXPANSION OF THE INTERACTION TERMS

The inhomogeneous terms appearing in wave equation (9) may be written as a matrix of scattering vertex operators of the form

$$\sum_{\alpha=1,2} \mathbf{P}_1^{(\alpha)}(\boldsymbol{\nabla})[\mathbf{P}_2^{(\alpha)}\delta^3(\mathbf{x}-\mathbf{x}_0)(\mathbf{P}_3^{(\alpha)}(\boldsymbol{\nabla})\psi)] = \begin{bmatrix} V^{\phi\phi} & V^{\phi\mathbf{A}} \\ V^{\mathbf{A}\phi} & V^{\mathbf{AA}} \end{bmatrix} \psi. \tag{A1}$$

The action of these elements on the scalar and vector potential components will be indexed as

$$\begin{bmatrix} \delta\phi \\ \delta\mathbf{A}_l \end{bmatrix} = \begin{bmatrix} V^{\phi\phi} & V_k^{\phi\mathbf{A}} \\ V_l^{\mathbf{A}\phi} & V_{lk}^{\mathbf{AA}} \end{bmatrix} \begin{bmatrix} \phi \\ \mathbf{A}_k \end{bmatrix}, \tag{A2}$$

where spatial vector indices $k,l\in\{1,2,3\}$, and repeated indices are summed. The matrix product (A1) implicitly involves inner products over internal spatial indices, which will be denoted explicitly in the factors given below. In this Appendix only it is necessary to use this index convention to write the factorization of the operator in a manageable form. The unit diagonal matrix is implied wherever **P** operators are denoted as scalars. ϵ_{ijk} is the totally antisymmetric symbol, and δ_{ij} is the Kronecker delta.

The $\alpha=1$ term in Eq. (A1) involves only density perturbations, and is quadratic in spatial derivatives. Its factors are

$$\begin{aligned} [\mathbf{P}_1^{(1)}\ (\boldsymbol{\nabla})]_i &= \boldsymbol{\nabla}_i, \\ [\mathbf{P}_2^{(1)}]_{ip} &= \left(\frac{-M\omega^2}{\rho_0}\right) \begin{bmatrix} \delta_{ip} & \epsilon_{ipk} \\ -\epsilon_{ipl} & \epsilon_{mil}\epsilon_{mpk} \end{bmatrix}, \\ [\mathbf{P}_3^{(1)}\ (\boldsymbol{\nabla})]_p &= \boldsymbol{\nabla}_p. \end{aligned} \tag{A3}$$

The $\alpha=2$ term is quartic in derivatives, and comprises perturbations in both compressional and shear moduli. Its factors are

$$\begin{aligned} [\mathbf{P}_1^{(2)}\ (\boldsymbol{\nabla})]_{ij} &= \boldsymbol{\nabla}_i\boldsymbol{\nabla}_j, \\ [\mathbf{P}_2^{(2)}]_{ijpq} &= \left\{ \left(\frac{-\Delta\lambda R^3}{\rho_0}\right) \begin{bmatrix} \delta_{ij}\delta_{pq} & 0 \\ 0 & 0 \end{bmatrix} - \left(\frac{2\Delta\mu R^3}{\rho_0}\right) \times \begin{bmatrix} \delta_{ip}\delta_{jq} & \epsilon_{ipk}\delta_{jq} \\ -\epsilon_{ipl}\delta_{jq} & \tfrac{1}{2}\epsilon_{mil}[\epsilon_{mpk}\delta_{jq}+\epsilon_{jpk}\delta_{mq}] \end{bmatrix} \right\}, \\ [\mathbf{P}_3^{(1)}\ (\boldsymbol{\nabla})]_{pq} &= \boldsymbol{\nabla}_p\boldsymbol{\nabla}_q. \end{aligned} \tag{A4}$$

Contraction of these terms, after transformation to a wave number basis, gives the expansion of Eq. (16).

APPENDIX B: SYMMETRIES OF THE SCATTERING VERTEX

The symmetry decomposition of the scattering vertex (16) is performed by fixing the wave vector $\mathbf{k}'$ of input fields ϕ, $\mathbf{A}$, and identifying the dependence of the various components of the scattered wave,

$$\sum_{\alpha=1,2} [\mathbf{P}_1^{(\alpha)}(-i\mathbf{k})\mathbf{P}_2^{(\alpha)}\mathbf{P}_3^{(\alpha)}(-i\mathbf{k}')]\cdot\begin{bmatrix}\phi\\ \mathbf{A}\end{bmatrix}_{\mathbf{k}'},$$

on the angle of the output wave vector $\mathbf{k}$.

For $\mathbf{A}$ satisfying the transversality condition (2), the polarization of the shear wave defines a unit vector $\hat{\mathbf{p}}$ as $\mathbf{A} \equiv |\mathbf{A}|(\hat{\mathbf{p}}\times\mathbf{k}')/|\mathbf{k}'| \equiv |\mathbf{A}|(\hat{\mathbf{p}}\times\hat{\mathbf{k}}')$. Without loss of generality, $\hat{\mathbf{p}}$ may be taken to satisfy $\hat{\mathbf{p}}\cdot\hat{\mathbf{k}}'=0$. An orthonormal basis is therefore defined by the three unit vectors

$$\hat{\mathbf{e}}_1=\hat{\mathbf{p}}\times\hat{\mathbf{k}}', \quad \hat{\mathbf{e}}_2\equiv\hat{\mathbf{p}}, \quad \hat{\mathbf{e}}_3\equiv\hat{\mathbf{k}}'. \tag{B1}$$

Denoting $\hat{\mathbf{k}}\equiv\mathbf{k}/|\mathbf{k}|$, and introducing angles θ, φ for an expansion in spherical harmonics $Y_{mn}^{e,0}(\theta,\varphi)$,[17] such that

$$\begin{aligned}
&\hat{\mathbf{k}}\cdot\hat{\mathbf{k}}' \equiv \cos\,\theta = Y_{01}^e,\\
&\hat{\mathbf{k}}\cdot(\hat{\mathbf{p}}\times\hat{\mathbf{k}}') \equiv \cos\,\varphi\,\sin\,\theta = Y_{11}^e,\\
&\hat{\mathbf{k}}\cdot\hat{\mathbf{p}} \equiv \sin\,\varphi\,\sin\,\theta = Y_{11}^0,
\end{aligned} \tag{B2}$$

the components of the scattered wave may be grouped according to the level of their representation of the rotation algebra.

Identities useful in computing the action of the matrix proportional to $\mathbf{D}$ are

$$\begin{aligned}
&(\hat{\mathbf{k}}\times\hat{\mathbf{k}}')\cdot(\hat{\mathbf{p}}\times\hat{\mathbf{k}}') = Y_{11}^0,\\
&-(\hat{\mathbf{k}}\times\hat{\mathbf{k}}') = -\hat{\mathbf{e}}_1 Y_{11}^0 + \hat{\mathbf{e}}_2 Y_{11}^e,\\
&[(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}}')\mathbf{1}-\hat{\mathbf{k}}'\hat{\mathbf{k}}]\cdot(\hat{\mathbf{p}}\times\hat{\mathbf{k}}') = \hat{\mathbf{e}}_1 Y_{01}^e - \hat{\mathbf{e}}_3 Y_{11}^e.
\end{aligned} \tag{B3}$$

Similar useful identities for the terms proportional to $\mathbf{S}$ are

$$\begin{aligned}
&(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}}')^2 = \tfrac{1}{3}Y_{00}^e + \tfrac{2}{3}Y_{02}^e,\\
&(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}}')(\hat{\mathbf{k}}\times\hat{\mathbf{k}}')\cdot(\hat{\mathbf{p}}\times\hat{\mathbf{k}}') = \tfrac{1}{3}Y_{12}^0,\\
&-(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}}')(\hat{\mathbf{k}}\times\hat{\mathbf{k}}') = -\hat{\mathbf{e}}_1\tfrac{1}{3}Y_{12}^0 + \hat{\mathbf{e}}_2\tfrac{1}{3}Y_{12}^e,\\
&(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}}')[(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}}')\mathbf{1}-\hat{\mathbf{k}}'\hat{\mathbf{k}}]\cdot(\hat{\mathbf{p}}\times\hat{\mathbf{k}}') = \hat{\mathbf{e}}_1(\tfrac{1}{3}Y_{00}^e + \tfrac{2}{3}Y_{02}^e) - \tfrac{1}{3}\hat{\mathbf{e}}_3 Y_{12}^e,\\
&(\hat{\mathbf{k}}\times\hat{\mathbf{k}}')(\hat{\mathbf{k}}\times\hat{\mathbf{k}}')\cdot(\hat{\mathbf{p}}\times\hat{\mathbf{k}}') = \hat{\mathbf{e}}_1(\tfrac{1}{3}Y_{00}^e - \tfrac{1}{3}Y_{02}^e - \tfrac{1}{6}Y_{22}^e) - \tfrac{1}{6}\hat{\mathbf{e}}_2 Y_{22}^0.
\end{aligned} \tag{B4}$$

Substituting these expansions into Eq. (16) gives an expression for the scattered wave in terms of the magnitudes of the inputs ϕ and $\mathbf{A}$,

$$\begin{aligned}
&\sum_{\alpha=1,2} [\mathbf{P}_1^{(\alpha)}(-i\mathbf{k})\mathbf{P}_2^{(\alpha)}\mathbf{P}_3^{(\alpha)}(-i\mathbf{k}')]\cdot\begin{bmatrix}\phi\\ \mathbf{A}\end{bmatrix}_{\mathbf{k}'}\\
&= \left\{ -\left(c^2\mathbf{C} + \frac{2s^2\mathbf{S}}{3}\right)\mathbf{k}^2\mathbf{k}'^2\begin{bmatrix} Y_{00}^e & 0\\ 0 & 0\end{bmatrix} + \omega^2\mathbf{D}|\mathbf{k}||\mathbf{k}'|\begin{bmatrix} Y_{01}^e & Y_{11}^0\\ -\hat{\mathbf{e}}_1 Y_{11}^0 + \hat{\mathbf{e}}_2 Y_{11}^e & \hat{\mathbf{e}}_1 Y_{01}^e - \hat{\mathbf{e}}_3 Y_{11}^e\end{bmatrix}\right.\\
&\left. - \frac{s^2\mathbf{S}}{6}\,\mathbf{k}^2\mathbf{k}'^2\begin{bmatrix} 8Y_{02}^e & 4Y_{12}^0\\ -4\hat{\mathbf{e}}_1 Y_{12}^0 + 4\hat{\mathbf{e}}_2 Y_{12}^e & \hat{\mathbf{e}}_1(6Y_{02}^e + Y_{22}^e) + \hat{\mathbf{e}}_2 Y_{22}^0 - 2\hat{\mathbf{e}}_3 Y_{12}^e\end{bmatrix}\right\}\begin{bmatrix}\phi\\ |\mathbf{A}|\end{bmatrix}_{\mathbf{k}'}.
\end{aligned} \tag{B5}$$

Both $\mathbf{C}$ and $\mathbf{S}$ contribute to isotropic, or monopole scattering, with constant angular dependence $Y_{00}^e \equiv 1$. $\mathbf{D}$ multiplies only components $Y_{m1}^{e,0}$, and thus gives rise to pure dipole scattering. Finally, $\mathbf{S}$ contributes additional terms in $Y_{m2}^{e,0}$, representing quadrupole scattering, as would be expected from the symmetry of shear deformations.

[1] H. Schmidt, ''SAFARI: Seismo-acoustic Fast Field Algorithm for Range Independent Environments. User's Guide,'' SR 113, SACLANT ASW Research Centre, La Spezia, Italy (1987).

[2] J. T. Goh and H. Schmidt, ''Benchmark cases for range-dependent seismoacoustic propagation codes,'' J. Acoust. Soc. Am. **98**, 2970(A) (1995).

[3] See, for example, L. M. Brekhovskikh and Yu. Lysanov, *Fundamentals of Ocean Acoustics* (Springer-Verlag, Berlin, 1982), or J. A. Ogilvy, ''Wave scattering from rough surfaces,'' Rep. Prog. Phys. **50**, 1553 (1987). A tutorial introduction to the effective medium approach is given in J. M. Hovem, ''Scattering from Particles and Inclusions in Solids,'' Lecture notes for Norfa Course in Acoustical Scattering, Visby, Sweden (1995).

[4] B. Schmalfelt, ''A comparison of seismic and hydroacoustic measurements at very low Frequencies in different shallow water areas,'' in *Ocean Seismo-Acoustics*, edited by T. Akal and J. Berkson (Plenum, New York, 1986) (Volume of the NATO Conference Series: Marine Science, SACLANT ASW Research Centre, La Spezia, Italy).

[5] B. Schmalfelt and D. Rauch, ''Ambient and ship-induced low-frequency noise in shallow water,'' in *Bottom-Interacting Ocean Acoustics*, edited by W. A. Kuperman and F. B. Jensen (Plenum, New York, 1986) (Volume of the NATO Conference Series: Marine Science, SACLANT ASW Research Centre, La Spezia, Italy).

[6] R. A. Stephen, ''Bottom penetration at subcritical grazing angles by scattering,'' J. Acoust. Soc. Am. **99**, 2475(A) (1996).

[7] R. A. Stephen and S. T. Bolmer, ''Interface waves in seafloor propagation and scattering problems,'' J. Acoust. Soc. Am. **99**, 2552(A) (1996).

[8] Å. Kristensen and R. Mittet, ''Using surface waves to detect a point diffractor buried at the sea bottom,'' STATOIL report N-7005, IKU report N-7034, Trondheim, Norway (1996).

[9] E. Smith, J. Behrens, P. Wilson, B. Wolz, T. Muir, and A. Atchley, ''Scattering of seismic interface waves from pointlike obstacles in unconsolidated sediments,'' J. Acoust. Soc. Am. **96**, 3330(A) (1994).

[10] E. Smith, P. S. Wilson, F. W. Bacon, J. F. Manning, J. A. Behrens, and T. G. Muir, ''Measurement and localization of interface wave reflections from a buried target,'' J. Acoust. Soc. Am. **103**, 2333–2343 (1998).

[11] A. E. H. Love, *Some Problems of Geodynamics* (Cambridge U.P., Cambridge, 1926).

[12] E. K. Westwood, C. T. Tindle, and N. R. Chapman, ''A Normal Mode

Model for Acousto-Elastic Ocean Environments,'' J. Acoust. Soc. Am. (submitted).
[13] E. Hamilton, ''Elastic Properties of Marine Sediments,'' J. Geophys. Res. **76**, 579 (1971).
[14] Strictly, the column represented by $\mathcal{C}$ contains monopole as well as quadrupole components, but an expansion in powers of the scattering angle is convenient for this application.
[15] R. J. Urick, *Principles of Underwater Sound* (McGraw-Hill, New York, 1983), 3rd ed., p. 291.
[16] R. Mittet, private communication.
[17] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), pp. 1264, 1325.

A fast integral equation approach to acoustic scattering from three-dimensional objects using a natural basis set

D. T. Moroney and P. J. Cullen
Department of Electronic and Electrical Engineering, Trinity College Dublin, Dublin 2, Ireland

(Received 1 February 1996; revised 29 December 1996; accepted 31 December 1997)

Boundary value integral equations and the moment method solution technique provide a succinct and elegant strategy for the solution of acoustic surface scattering problems. This paper explores the application of a subdomain basis set, based on the phase of the incident radiation, which is used to solve for the unknown normal pressure gradient on the surface of the scatterer using the moment method. The advantage of this basis set lies in the fact that it can accurately represent the fast phase character of the pressure gradient on the surface of the scatterer. This allows for far fewer discretizations/basis elements in the moment method formulation and therefore significantly reduces the density of unknowns that must be numerically computed. The main contribution of this paper is to show how to efficiently employ this natural basis and to explore the physical justification for using this approach. The method is compared with other techniques including the Kirchhoff solution technique, high frequency single scattering techniques, the Nussenzveig–Fock technique, and the parabolic equation approximation. Unlike these methods no approximation to the physical system need be introduced. © *1998 Acoustical Society of America.* [S0001-4966(98)02904-X]

PACS numbers: 43.30.Gv [JHM]

INTRODUCTION

The problem we shall address is the efficient computation of the scattered field caused by the presence of an irregular, three-dimensional, pressure-release scatterer, embedded in free space and irradiated by a source with wave number k. The starting point of our analysis will be the Helmholtz integral equation. If we consider a two-dimensional surface S with Dirichlet boundary conditions the total acoustic pressure $p(\boldsymbol{\rho})$ that results from an incident field $p_{\text{inc}}(\boldsymbol{\rho})$ interacting with S is given by

$$p(\boldsymbol{\rho}) = p_{\text{inc}}(\boldsymbol{\rho}) - \int_S \Psi(\boldsymbol{\rho},\boldsymbol{\rho}') \frac{\partial p(\boldsymbol{\rho}')}{\partial n'} dS'. \tag{1}$$

Here, $\partial p(\boldsymbol{\rho}')/\partial n'$ is the unknown normal pressure gradient at the surface which we will hereafter denote as J_n, $\Psi(s,s',t,t')$ (s and t parametrize the boundary surface) is the free-space Green's function

$$\frac{\exp(-ikR)}{4\pi R} \tag{2}$$

with R the distance between points on the surface and integration points, and dS' is a surface element (see Fig. 1). Using the boundary condition $p(\boldsymbol{\rho}_s)=0$, where $\boldsymbol{\rho}_s$ is a point on the surface, we arrive at an integral equation for $J_n(\boldsymbol{\rho}_s)$

$$p_{\text{inc}}(s,t) = \int_S \Psi(s,s',t,t') J_n(s',t') dS', \tag{3}$$

where s and t parametrize the boundary surface.

The moment method solves boundary value problems (BVPs) described as surface integral equations over finite domains by converting continuous integral equations into systems of algebraic equations. In the case where N Dirac testing functions are used (point collocation) such systems can be written as

$$\mathcal{Z}\mathbf{a} = \mathbf{P}_{\text{inc}}, \tag{4}$$

where $\mathbf{P}_{\text{inc}}$ is the vector consisting of the incident field at the collocation points, $\mathcal{Z}$ is the $N\times N$ kernel of the discretized integral equation (which we will refer to as the impedance matrix), and $\mathbf{a}$ is the unknown vector containing the coefficients a_n of the expansion of the pressure gradient over the basis $b_n(s,t)$

$$\hat{J}_n(s,t) = \sum_{n=1}^{N} a_n b_n(s,t). \tag{5}$$

Once the a_n are determined Eq. (5) provides an approximation to the exact solution of the integral equation.

The moment method is usually considered a low frequency method[1,2] because the size N of the moment matrix increases with frequency. To accurately model the fast moving phase we must discretize the surface pressure gradient at a rate of at least four surface patches (triangular patches are most commonly used) per wavelength in any direction on the surface. Moreover the impedance matrix $\mathcal{Z}$, where N is the total number of surface elements, is full and has complex entries.

Over the past few years a number of methods have been proposed to reduce the number of unknowns in the formulation of integral equation problems. Recent approaches to speed up iterative solutions include the fast multipole method and its variants[3–5] as well as the matrix decomposition algorithm.[6] The impedance matrix localization method[7] can reduce solution complexity by creating a very sparse impedance matrix. The factoring out of either the physical optics pressure gradient (or current in the electrical case) or its phase factor have been suggested by James.[8] The use of the phase factor of the physical optics current to construct a basis set for the unknown surface current density has been suggested in Refs. 9 and 10. It is this technique which we

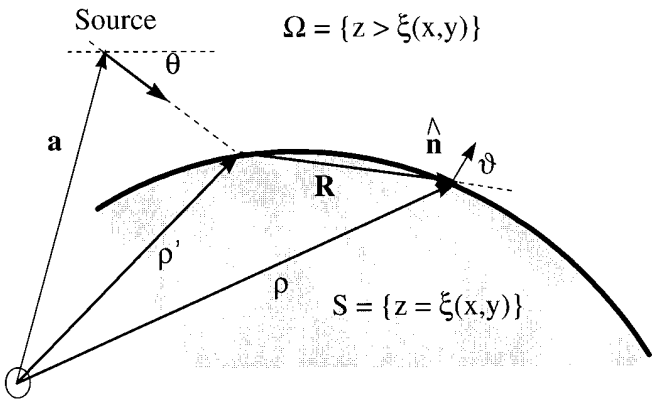


FIG. 1. Scattering geometry.

shall explore in this paper. Incidentally, the relationship between this method and some of the aforementioned techniques is considered in Ref. 11 and more recently in Ref. 12. In the latter citation very high speed iterative methods which are more robust than the natural basis method, for the case of multiple scattering, are outlined for the two-dimensional electromagnetic case.

Any physical knowledge we may have of the solution of a scattering problem can be used to inform the construction of a good basis set. The more we can inform the basis, the less numerically intensive will be our task of finding the exact solution. This paper is concerned with a particular, and rather natural, case of this general approach. The optimum basis function for the unknown pressure gradient in a moment method formulation of a scattering problem is of course, the trivial one that tends towards the unknown gradient itself.

We have mentioned James[8] who constructed a modified integral equation called the High Frequency Panel Method, for uniform plane wave excitation by dividing the surface into linear segments (panels) and assuming that the phase of the surface current on the panels could be expressed in terms of the phase of the physical optics current. The resulting integral equation where a piecewise constant basis function was assumed was solved for the residual function and a reduction in the density by a factor of between 20 and 100 was claimed as possible and a reduction nearer 20 was actually demonstrated. The essence of our approach is the same, we use the normalized incident field (the field divided by its magnitude, i.e., the physical optics phase-factor) as a basis.

Moroney and Cullen in Ref. 9 introduced the idea of a natural basis for two-dimensional terrain scattering. This basis is a set of linearly independent subdomain complex valued functions defined to be the phase of the incident radiation as a function of arclength (they consider the 2-D case) on the surface boundary of the scatterer. This basis set can greatly reduce the number of unknowns N and thus provide a method suitable for many very large scattering problems. A reduction of N by a factor of 1000 has been demonstrated for a practical terrain propagation problem where the scatterer was of the order $10^5\lambda$ in total arclength, allowing the scattered field to be calculated in a few seconds on a desktop computer.

Abregg and Peterson,[10] taking their lead from James, describe a method where the phase of the physical optics current is used to construct a moment method basis for the unknown current. They also employ an edge basis function for some of the cases they consider and their surface elements are considered to be parabolic. Aberegg and Peterson call their method the integral equation asymptotic phase (IE-AP) technique. A reduction in the density of unknowns to less than one sample per wavelength was achieved.

In this paper we consider the application of the natural basis to three-dimensional problems. The main contribution of this paper is the solution methodology (in particular, the construction of the matrix operator) and the exploration of the physical justification of this approach. This is important because it is required to understand under what conditions we can expect significant computational savings to be had using the natural basis.

Computationally, the numerical scattering problem can be divided into four elementary steps: (1) Preprocessing—surface discretization, physical parameter input. (2) Matrix fill. (3) Inversion/pseudo-inversion to calculate the unknown $\hat{J}_n$. (4) Calculation of total field using $\hat{J}_n$. It will be seen that using the natural basis set can speed up each of the steps with the exception of preprocessing; and although the analysis will rely on the Helmholtz integral[13] with pressure-release boundary condition, i.e., $p=0$ on the surface of the scatterer, the extension to the condition $\partial p/\partial n=0$ on the surface of the scatterer will be evident. It appears likely that the method can also be applied to lossy homogeneous isotropic scatterers, this is certainly the case for two-dimensional scattering by dielectrics which is considered in Ref. 14.

I. A NATURAL BASIS SET IN THREE DIMENSIONS

The natural basis set is extracted from the high frequency complex function $\phi(\boldsymbol{\rho}_s)=p_{\text{inc}}(\boldsymbol{\rho}_s)/|p_{\text{inc}}(\boldsymbol{\rho}_s)|$. Writing J_n as $J_n=\overline{J_n}\cdot\phi$ we argue that, for a significant class of problems, ϕ contains the fast moving phase of the true surface pressure gradient J_n. Whenever this is true we can significantly reduce the size of the impedance matrix $\mathcal{Z}$. Generally if $\overline{J_n}$ is in any way slower moving than J_n then, of course, a computational speedup can be achieved.

Expressing $\overline{J_n}$ in terms of the incident and scattered field gradients

$$\overline{J_n}=\hat{n}\cdot\nabla(p_{\text{inc}}+p_{\text{sca}})/\phi$$

and assuming that the incident field magnitude will typically be slowly varying in practical applications, we see that if $\overline{J_n}$ is to be slowly varying on the surface then

$$F=[\nabla\Phi(p_{\text{sca}})A(\boldsymbol{\rho})-i\nabla A(\boldsymbol{\rho})]$$

must be also be slowly varying where we have expressed the scattered pressure $p_{\text{sca}}=A(\boldsymbol{\rho})\exp(-i\Phi(p_{\text{sca}}(\boldsymbol{\rho})))$ and we use the function $\Phi(x)=x/|x|$ to extract its phase factor. Computational gain requires that F be piecewise slowly varying in amplitude and phase over regions that are large compared to a wavelength.

For computational purposes we rely on the weaker assumption that $\overline{J_n}$ can be represented as a set of low order

polynomials over disjoint subdomains. We assume that there exist suitable conditions such that $\Phi(J_n)$ can be well approximated, in local areas, with possibly some constant phase shift as $\Phi(p_{\text{inc}}(s))$. This occurs when the phases and amplitudes of the function F, restricted to the surface of the scatterer, are quasiconstant in local basis areas. Specifically the complex basis set is

$$b_n(s,t)=\begin{cases} e^{-i\Phi(p_{\text{inc}}(s,t))}, & s,t\in\Delta S_n \\ 0, & s,t\notin\Delta S_n \end{cases} \tag{6}$$

defined over *large* subdomains $s,t\in\Delta S_n$ of the surface. None of the elements $b_n(s,t)$ lie in the space spanned by $\{b_m(s,t)\}_{m=1,m\neq n}^{N}$ thus $\{b_n(s,t)\}_{n=1}^{N}$ is a linearly independent set. The restriction on the set is that it forms a complete set in the domain of the operator $\mathscr{Z}$ (Ref. 15) i.e., $\{b_n(s,t)\}_{n=1}^{N}$ should satisfy the boundary conditions and any differentiability condition of the problem.

We may supplement the complex basis set with a real set $\{r_n\}_{n=1}^{N}$ to better approximate the change in amplitude and phase over a basis element on ΔS_n giving a basis set $\{b_n r_n\}_{n=1}^{N}$ over the same subdomains ΔS_n. The subdomains ΔS_n over which the basis functions are defined are, of course, variable in length for each n.

II. RELATION OF THE NATURAL BASIS TO OTHER SOLUTION TECHNIQUES

We will now produce evidence to suggest that $\overline{J_n}$ varies slowly relative to ϕ, for a certain class of problems. We will first consider the formulation of our problem in terms of a second kind integral equation exploring the contribution of the integral term in the case of large smooth scatterers. We then examine the behavior of the first few terms of the Born series solution to justify our thesis that $\overline{J_n}$ can be a slowly varying function. We will then consider in turn the Nussenveig–Fock correction, the Green's function perturbation method (GFPM), and the parabolic wave approximation in each case to support our thesis that when we remove ϕ that $\overline{J_n}$ is slowly varying.

A. Second kind integral equation

Consider first (in the two-dimensional case), the modified Helmholtz equation (second kind integral equation) with Dirichlet boundary condition

$$J_n=2J_n^i-\frac{1}{2i}\int_S(\hat{n}\cdot\hat{R})G'(s,s')J_n ds', \tag{7}$$

to help gauge the validity of our representation of ϕ. J_n^i is the incident pressure derivative on the surface and we can identify $2J_n^i$ as the Kirchhoff approximation J_K for the pressure gradient. The geometry is illustrated in Fig. 1. For ease of use we rewrite Eq. (7) as $J_n=J_K+I$. These two right hand terms can be interpreted as the Kirchhoff–tangent plane approximation J_K and all corrections I to this approximation. In Ref. 16 Poggio and Miller make the observation that due to geometrical factors in the kernel of I, namely $\hat{n}\cdot\hat{R}=\cos\vartheta$, for large smooth conductors the contribution from I is only of second order importance. The phase $\Phi(J_K)$ of the Kirchhoff term J_K is directly calculable and so we postulate that J_n can be written as $\overline{J_n}\Phi(J_K)$. This is trivially true when S is a flat plate because $\cos\vartheta$, and hence I, is zero everywhere. Thus our thesis is supported for large smooth scatterers.

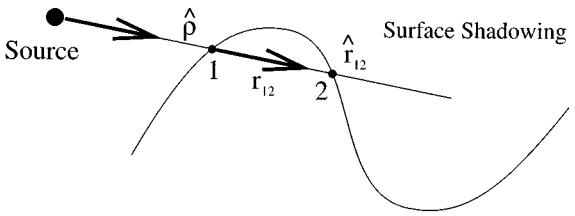


FIG. 2. Geometrical interpretation of a surface shadowing event.

B. Born series: Multiple scattering

Writing the Born (the Born or Neumann series is sometimes referred to as the Meecham series for the specific case of the acoustic analog of the magnetic field integral equation) series for Eq. (7)[17]

$$J_n=J_K-\mathscr{Z}J_K+\mathscr{Z}^2J_K-\cdots+\mathscr{Z}^NJ_K-\cdots \tag{8}$$

we can explore the contribution of terms after the tangent plane pressure gradient approximation J_K to the induced pressure gradient J_n. An evaluation of the contribution to J_n of the term $\mathscr{Z}J_K$ in the limit $k\to\infty$ is studied using the method of stationary phase by Liszka and McCoy[18] for the two-dimensional case where the Green's function $G(s,s')$ is given by the Hankel function $H_0^{(2)}$. However, the analysis extends, quite naturally, to the case where $G(s,s',t,t')=\exp(-ikR)/(4\pi R)$. They found that when the addition to the integral (7) from the local region $kR\leqslant 1$ is negligible, two stationary points exist. The domain of integration is divided into local $kR\leqslant 1$ and outer $kR>1$ so that asymptotic forms of the Hankel function $H_0^{(2)}$, for small and large arguments, can be used in the analysis. Referring to the geometry of Figs. 2 and 3 stationary points are given by

$$\hat{\boldsymbol{\rho}}=-\hat{r}_{12}, \tag{9}$$

where $\hat{\boldsymbol{\rho}}$ is the unit vector from the radiating source to an observation point 1, on the surface, and $\hat{r}_{12}$ is a unit vector, equivalent to $\hat{R}$, along the vector joining the observation point 1 and any other point 2 indicating surface shadowing and

$$\hat{\boldsymbol{\rho}}\cdot\hat{t}_2=-\hat{r}_{12}\cdot\hat{t}_2, \tag{10}$$

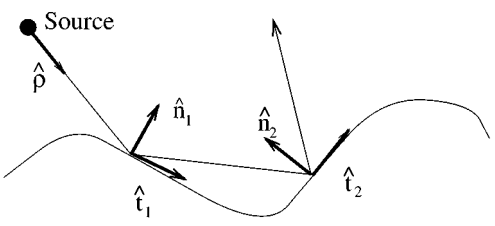


FIG. 3. Geometrical interpretation of a single scattering event.

where $\hat{t}_2$ is the unit tangent vector at a point 2, indicating a double reflection. What is most important about these two instances is that their contribution to J_n is a piecewise scalar multiple of J_K, thus the phase of J_n is equal to the phase of J_K. Again we find $J_n = \overline{J_n}\Phi(J_K)$ and our requirement is satisfied for surfaces with up to double scattering events (see Fig. 3) where the dimensions of the surface are large compared to the dimension of the incident wavelength.

C. Born series: Second term

The contribution from local regions, where $kR \leqslant 1$, is considered by Voronovich.[19] He limits himself to surfaces that can be represented, in the local area $kR \leqslant 1$ about the observation point $\boldsymbol{\rho}$, by $\xi(x,y) \approx x^2/2\rho_x + y^2/2\rho_y$ where ρ_x and ρ_y are the radii of curvature in the xz and yz planes at (x,y). He shows, in the case of an incident plane wave, that the addition, from local areas, to J_n from the second term in the Born series is

$$J_K\left(\frac{i}{2k\rho_x \sin^3\theta} + \frac{i}{2k\rho_y \sin\theta}\right), \tag{11}$$

where θ is the angle of incidence. Thus the contribution to the phase from local areas, other than that from J_K, in the second term of the Born series, is from the purely complex function $i(1/(2k\rho_x \sin^3\theta) + 1/(2k\rho_y \sin\theta))$ whose phase angle is zero. Once again this lends support to our thesis.

D. Green's function perturbation method

Alternative evidence can be had by considering the most general form of convolutional approximate solution to the boundary operator problem is introduced as a Green's function perturbation method (GFPM) in Refs. 20 and 21.

The difference in phase of the normal derivative of the pressure calculated by convolution $\Phi(P_c)$, using the GFPM, and $\Phi(J_K)$ is

$$\Phi\left(\mathscr{F}^{-1}\left[\frac{\mathscr{F}[p_{\text{inc}}(s,t)]}{\mathscr{F}[G(g(s,s',t,t'))]}\right]\right) - \frac{2J_n(s,t)}{|2J_n(s,t)|}, \tag{12}$$

where $s = \int_0^x \sqrt{1 + (dy/dx)^2 (dz/dx)^2}\, dx$ and $g(s,t)$ is judiciously chosen. It is a straightforward task to show that again, in this case, $\overline{J_n}$ is slowly varying. As an example, $g(s,t)$ may be defined by $g(s,t) = \langle \mathbf{d}(s',t',s'-s,t'-t)\rangle_{s,t}$ where $\langle\cdot\rangle_{s,t}$ defines the average function $1/L\int_S \cdot\, ds\, dt$ and $\mathbf{d}(\cdot,\cdot)$ is the Euclidean distance function.

E. Nussenveig–Fock correction

Fock[22] considered another general set of surfaces, namely the set of paraboloids of revolution, in order to investigate diffraction into shadowed regions. (A point is *shadowed* if it is not geometrically visible to the source.) He found that he could express the current (pressure gradient) J_n as

$$J_n = J_K G(\zeta,0), \tag{13}$$

where

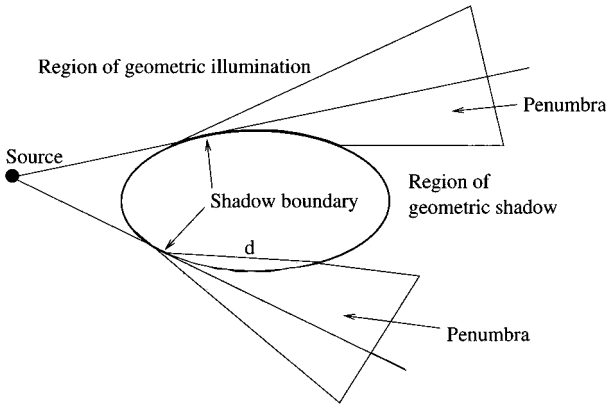


FIG. 4. Geometry for the Nussenzveig–Fock correction.

$$G(\zeta,0) = \exp\left(i\frac{\zeta^3}{3}\right)\frac{1}{\sqrt{\pi}}\int_c \frac{\exp(i\zeta t)}{w'(t)}\, dt \tag{14}$$

with $\zeta = \pm ld$, where l is the distance from the geometrical shadow boundary and the sign convention is $+$ in the geometrical shadow and $-$ otherwise, d is the width of the transition region from illumination to shadow (see Fig. 4), w is the complex Airy function and c is a contour in the complex t plane. He called his method the principle of the local field in the penumbra region (the *penumbra* is that region which, heuristically, can be defined as the finite transition region from illumination, according to geometrical optics, to total shadow). The function G tends to a limit in amplitude of 2, for ζ negative and large; and exponentially to zero, for ζ positive; thus showing the physical requirements of diffraction and reflection. For our purposes, the most important trait is that G, for the paraboloids of revolution, was found to be a slowly varying function of position. This agrees with our postulation exactly. The functions G are difficult to evaluate and so Fock envisaged a lookup table with G evaluated asymptotically over individual scatterers. This approach is often used as a post-process in improving the Kirchhoff approximation and is referred to as the Nussenzveig-Fock correction.[23] We, on the other hand, evaluate the Fock functions implicitly in the moment method.

F. Parabolic equation

The parabolic equation (PE) approach to solving an approximate form of the Helmholtz equation was first introduced by Leontovich and can be found in Ref. 22. Leontovich was primarily concerned with propagation around the earth, first considered as a plane and then as a sphere. Tappert[24] is regarded as the first to use the PE method for complex scattering problems. Physically the PE requires that energy propagates in a forward direction away from the source and that an insignificant amount is backscattered from any obstacle along the path of propagation. In terms of the pressure p, satisfying $\nabla^2 p + k^2 p = 0$, we write p as $W\exp(ik_0(R\cdot\hat{x}))/\sqrt{(R\cdot\hat{x})}$, where W is a slowly varying envelope about the fast phase component $\exp(ik_0 x)/\sqrt{x}$. The PE is given by[24]

TABLE I. Subdomain integration types.

d_1 ↓	d_2→	Monotonic increasing	Monotonic decreasing
monotonic increasing		*A*	*B*
monotonic decreasing		*B*	*A*

$$\frac{\partial^2 W}{\partial z^2}+\frac{1}{r^2}\frac{\partial^2 W}{\partial \theta^2}+2ik_0\frac{\partial W}{\partial x}+k_0(n^2-1)W=0. \tag{15}$$

Note that k_0 is a nominal wave number but is usually taken to be $2\pi/\lambda$ where λ is the mean wavelength of the illuminating radiation source. The symbol n is the refractive index of the propagating medium which may be a function of position. Within the bounds of the PE approximation n is considered close to constant. The conditions necessary for the validity of the PE are

$$\frac{\partial^2 W}{\partial x^2}\ll 2ik_0\frac{\partial W}{\partial x} \tag{16}$$

indicating that the change in the envelope W is small over a wavelength and

$$k_0 x\gg 1 \tag{17}$$

demanding observations far from the source. Tappert continues by indicating that the rapidity of changes in W depend on n and the boundary conditions and that for variations that are much more gradual in the horizontal, compared to the vertical, the step size needed to solve Eq. (15) is much greater than a wavelength. Methods that are often used to solve the PE are finite difference[25,26] and finite element.[27] It is evident that with $W\exp(ikR\cdot\hat{x})/\sqrt{R\cdot\hat{x}}$ on the surface on the scatterer

$$J_n=W\varphi\phi. \tag{18}$$

If φ is a low frequency function then problems that can be solved by the PE method can be solved using the natural basis set.

III. COMPUTATIONALLY EFFECTIVE ADAPTIVE MATRIX FILL

An extra advantage of the complex basis set in Eq. (6) arises in the calculation of the subintegrals required to fill the moment matrix. A naïve approach,[28] using small basis functions with compact support, would sample the subdomains for each matrix slot by at least four samples per wavelength along the arclength of the scatterer and then apply some numerical integration scheme. James in Ref. 29 considers a range of asymptotic methods for speeding the calculation of matrix elements.

Here we follow a different approach which is physically inspired and particularly well suited to grazing incidence scattering from undulating surfaces. Subintegrals of collocation type in the moment method with the derived basis set (6) take the following form

$$\int_{\Delta S} G(s,s',t,t')\exp(-i\Phi(p_{\text{inc}}(s',t'))dS'. \tag{19}$$

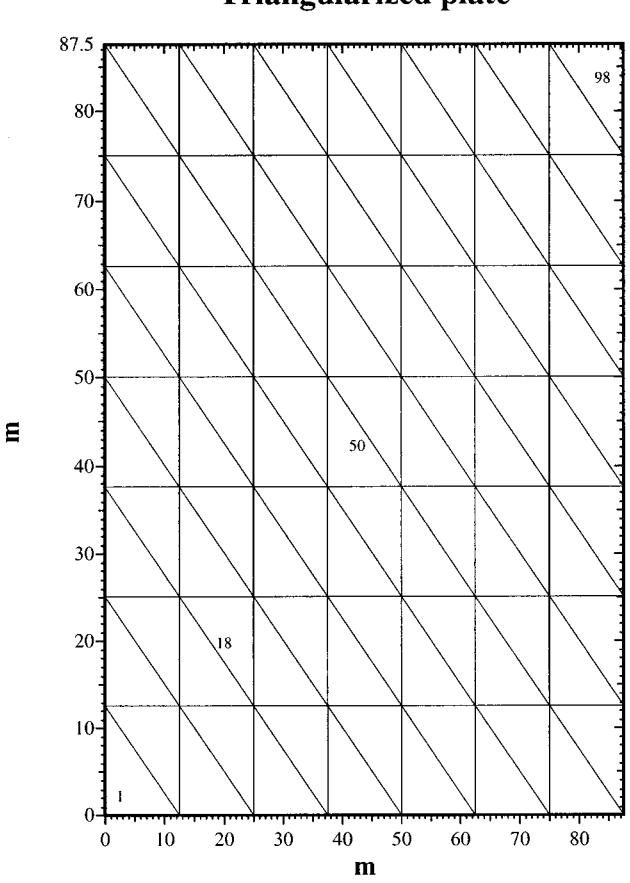


FIG. 5. Triangularization of flat plate.

Equation (19) is a generalized Fourier integral dependent on the two distance functions $d_1=|\boldsymbol{\rho}-\boldsymbol{\rho}'|$, via the Green's function, and $d_2=|\mathbf{a}-\boldsymbol{\rho}'|$, via the basis, as illustrated in Fig. 1, where **a** is a vector describing the position of the point source.

Over any surface S the oscillation in the subintegral is characterized by the function $D(s',t')=d_1+d_2$ and the distance functions can be divided into segments where d_1 and d_2 are, respectively, monotonic in the variables s,t. The four possible interactions of these functions shown in Table I produce two different integral types. Integrations of type *A* have integrands that oscillate more rapidly than the Green's function $\exp(-jkR)/(4\pi R)$ and integrands of type *B* oscillate more slowly than the Green's function. The specific oscillation rates over subdomains are calculated by the respective partial derivatives in s and t of D.

As an example consider the scattering of the field due to a point source by a flat plate: a thin screen of dimensions 87.5×87.5 m with leftmost position at the origin, see Fig. 5. The (x,y,z) coordinates of the source (the wavelength is $\lambda=1$ m) are $(-10,-10,4)$m. The plate is triangulated as in Fig. 5. We will consider the contributions from basis subdomains

$$\Delta S_1=\{(0,0,0),(12.5,0,0),(0,12.5,0)\}, \tag{20}$$

$$\Delta S_{18}=\{(25,12,5,0),(25,25,0),(12.5,25,0)\}, \tag{21}$$

$$\Delta S_{98}=\{(87.5,75,0),(87.5,87,5,0),(75,87.5,0)\}, \tag{22}$$

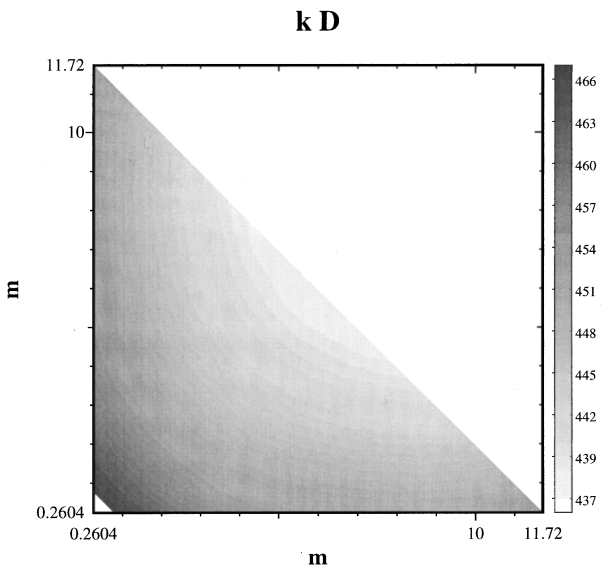


FIG. 6. The function kD in basis ΔS_1.

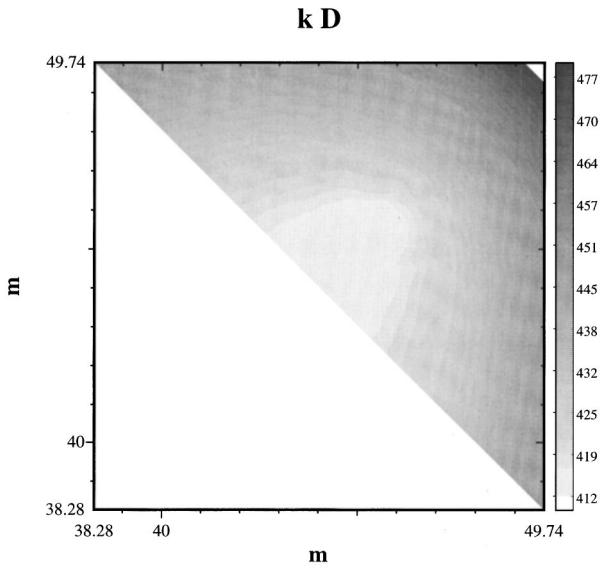


FIG. 8. The function kD in basis ΔS_{50}.

and the basis self-term

$$\Delta S_{50}=\{(50,37.5,0),(50,50,0),(37.5,50,0)\} \tag{23}$$

to the integration point $P=(45.83,45.83,0)$ indicated in Fig. 5. We note that the subscript i in subdomains ΔS_i is a label to a triangle where labeling is left to right and then bottom to top. Each ΔS_i is characterized by its three corners. The basis subdomains are each 12.5×12.5 m in area (i.e., greater than $156\lambda^2$).

The first case, where the phase of the contribution to integration point P is from ΔS_1, has the distance functions d_1 and d_2, where d_1 is monotonically decreasing and d_2 is monotonically increasing. The integration over this element is therefore of type B with kD, the sum of distance functions times wave number is shown in Fig. 6. kD is quasiconstant indicating that the integrand behaves very much like $|G(|\boldsymbol{\rho}-\boldsymbol{\rho}'|)|$. A second similar case is noted for ΔS_{18} in Fig. 7. When Eq. (19) represents a slowly oscillating function a low order Gaussian[30] or Simpson rule can be used to calculate the integration.

The third case, where the phase contribution to integration point P is from ΔS_{98}, has the distance functions d_1 and d_2 whose sum times wave number is shown in Fig. 9. In this case both d_1 and d_2 are monotonically increasing and the integration over this element is of type A. The accumulated distance function D has a very large slope indicating that the integrand is oscillating at a fast rate.

Figure 8 shows the case of the phase of the contribution to P from the self-term ΔS_{50}. The transition from integration of type A to type B at the point P is evident. When Eq. (19) represents a very quickly oscillating function an asymptotic technique[31] can be used in a fixed number of operations to a particular order once a truncation criterion for the asymptotic series is fixed. A specific and simple method for expanding a rapidly oscillating integral in terms of wave number k is repeated integration by parts.

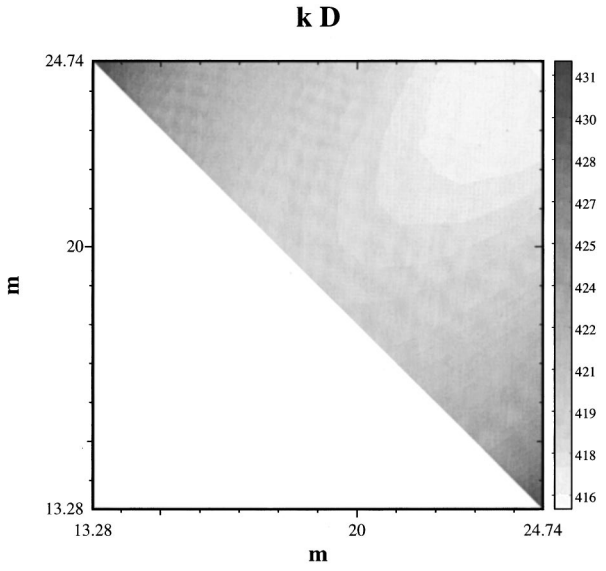


FIG. 7. The function kD in basis ΔS_{18}.

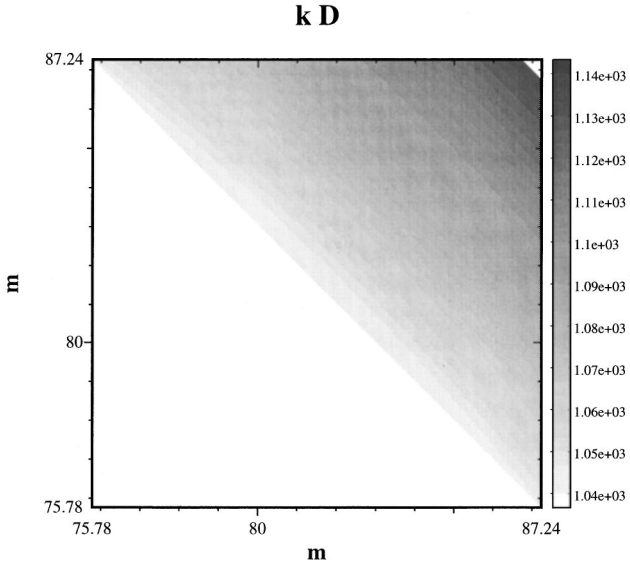


FIG. 9. The function kD in basis ΔS_{98}.

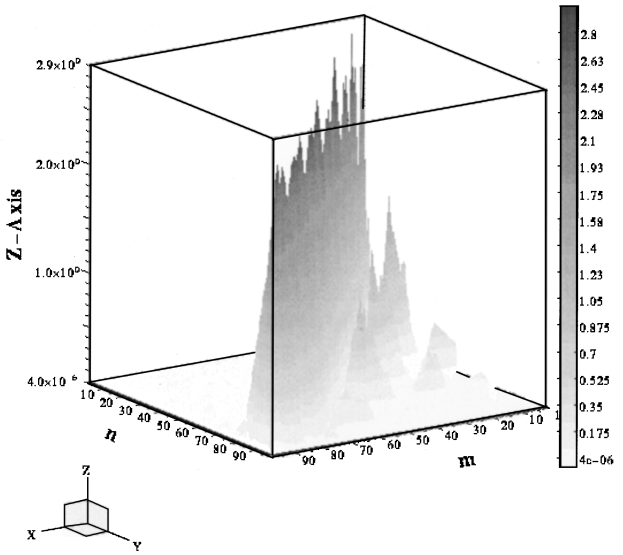


FIG. 10. Absolute value of matrix entries for the double wedge with a source at (100,45,16) with basis size (12.5×12.5).

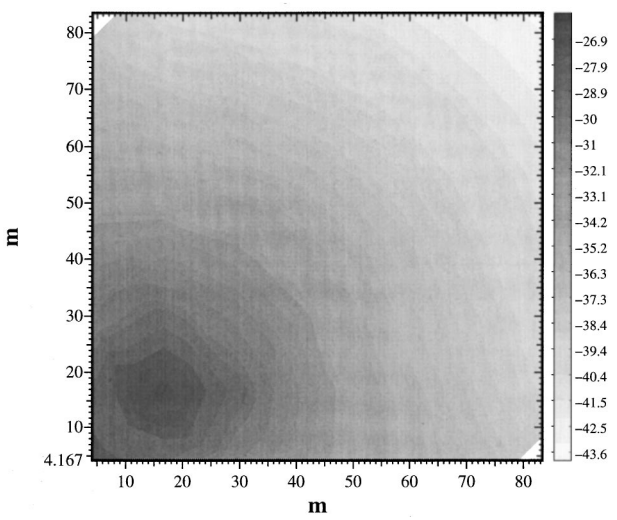


FIG. 11. Total field in dB 2.4 m above a 87.5×87.5 m flat plate.

The generalized Fourier integral can be represented as

$$I(k)=\int_{a_t}^{b_t}\int_{a_s}^{b_s} f(s',t')\exp(ikh(s',t'))ds'dt' \tag{24}$$

$$=\int_{a_t}^{b_t}\left[\int_{a_s}^{b_s}\frac{f(s',t')}{ik(d/ds')\,h(s',t')}\,\frac{d}{ds'}e^{ikh(s',t')}ds'\right]dt'. \tag{25}$$

Thus the integral (19) can be written as Eq. (25) with

$$f=\frac{1}{k\pi d_2} \tag{26}$$

and

$$h=-(d_1+d_2) \tag{27}$$

giving

$$I(k)=-\int_{\Delta t}\left[\int_{\Delta s} f\left[ik\frac{d}{ds'}(d_1+d_2)\right]^{-1}\times e^{-ik(d_1+d_2)}ds'\right]dt'. \tag{28}$$

Integration by parts yields,

$$I(k)=-\int_{\Delta t}\left[\left[\frac{\sqrt{2/k\pi d_2}[ik(d/ds')(d_1+d_2)]^{-1}}{ik(d/ds')(d_1+d_2)}e^{ikh(s')}\right]_{a_s}^{b_s}-\int_{a_s}^{b_s}\frac{d}{ds'}\left[\frac{\sqrt{2/k\pi d_2}[ik(d/ds')(d_1+d_2)]^{-1}}{ik(d/ds')(d_1+d_2)}\right]\times e^{-ik(d_1+d_2)}ds'\right]dt'. \tag{29}$$

Clearly, continued integration creates a series in inverse powers of k for each integration variable. We note the restriction that $\nexists\xi\in[a,b]$ s.t. $h'(\xi)=0$. When $h'(\xi)=0$ the method of stationary phase[32] should be used. This is in no way prohibitive since, if $h'(\xi)<\varepsilon$ (i.e., when d_1 and d_2 are almost collinear in the specific or coplanar in the general), where ε is small, an integration of type B would be in order and a quadrature rule may be used. The functions $h'(s')$ and $f(s')$ give the information needed to pick the order. The error produced in the asymptotic integration is[32] $\mathcal{O}(k^{-(n+1)})$ when n terms of the asymptotic series are retained.

In the case of terrain propagation, where the angle of incidence of the propagating wave is small and the terrain undulating (i.e., with hills and mountains) the integrands relating to forward propagation are very slowly oscillating and may even be thought of as pseudo-linear allowing low order Gaussian rule to be used to evaluate the integrals, while the integrands relating to backward propagation are very quickly oscillating (of the order twice that of the forcing frequency) and are ideal candidates for asymptotic integration. Diagonal terms, as illustrated in Fig. 8, exhibit both forward and backward characteristics about the match point.

The natural basis set (6) has the property that integrands over domains that produce a significant contribution to the kernel using (6) are low frequency (slowly oscillating) and integrands over domains that are less significant are high frequency (quickly oscillating).

An interaction is defined by the contribution from a basis element subdomain to a point

$$\int_{\Delta S_n} G(s,t,s',t')(\beta|\boldsymbol{\rho}-\boldsymbol{\rho}'|)b_n(s',t')dS'. \tag{30}$$

In particular, each entry $\mathcal{Z}_{mn}$ in the moment matrix represents the contribution from the domain of the nth basis function to the mth match point for collocation (or mth basis domain for Galerkin's method). These interactions over basis elements explicitly show, by means of amplitude $|\mathcal{Z}_{mn}|$, the areas of significant physical interaction. Thus in a system where we expect forward scattering to be a dominant factor the lower triangular section of the moment matrix dominates the operator. This triangular matrix fits the criteria for suitable preconditioners. It is easily invertible, its inverse is

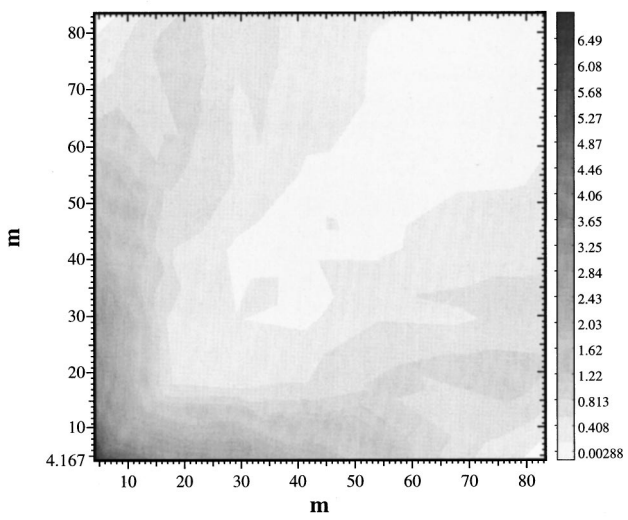


FIG. 12. Absolute difference in dB between the natural basis method and the image method.

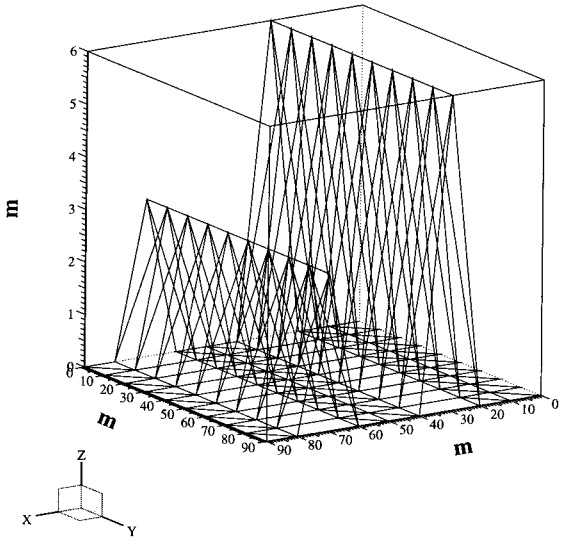


FIG. 14. Geometry of three dimensional double wedge.

close to $\mathscr{Z}^{-1}$ and, importantly, it is easily found, both by means of physical understanding (using a forward propagation assumption for example) and by a simple numerical scheme that removes values that fail a certain tolerance criterion. Figure 10 shows the amplitudes of the matrix kernel, for the double wedge problem illustrated in Fig. 14 with a radiating source of $\lambda = 0.5$ m situated at (100,45,16). It clearly shows that the operator discriminates in the forward scattering direction (by having large values of $|\mathscr{Z}_{mn}|$ when $m \leqslant n$ and small values of $|\mathscr{Z}_{mn}|$ for $m > n$). It also shows us that truncation to lower triangular matrix will produce a preconditioner $\mathscr{P}^{-1}$. For operators that are not solely forward scattering the same heuristic can be used to drop values in a well conditioned operator that fail a tolerance criterion.

IV. ANALYSIS AND RESULTS OVER CANONICAL BOUNDARIES

In this section we will discuss the results found, using the complex basis set (6), over some canonical boundaries. A canonical boundary is one that is considered a standard structured boundary. The canonical boundaries that we investigate are the flat plate and the wedge. Series type solutions for these simple shapes can be found in Bowman *et al.*[33]

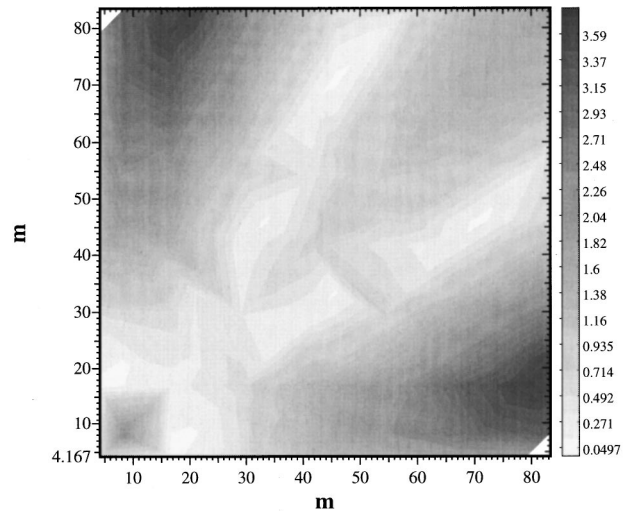


FIG. 13. Absolute difference in dB between the natural basis method and the physical optics method.

A. The flat plate

Figure 11 shows the total field in dB 2.4 m above a 87.5 $\times$87.5 m flat plate irradiated by a line source at $\lambda = 1$ m, positioned at (0,0,10) m calculated by three different methods; the method of images, which is our reference solution (it can be considered quite accurate for the example chosen) and the Helmholtz integral using the natural basis set as described in Eq. (6). The basis dimensions were 12.5$\times$12.5 m. It should be noted that, neglecting edge effects, the natural basis set is an exact representation of the phase of the induced pressure gradient for any orientation of the plate. The size of the moment matrix used is 98$\times$98 whereas the accepted published norm[34] with discretizations of at least four samples per wavelength would create a matrix of order $1.225\times10^5\times1.225\times10^5$. The core storage saving is then greater than 1.5625×10^8%. An algorithm consistent answer is assured by using the conjugate gradient squared (CGS) algorithm and its condition number is 5.8345, using the $\|\cdot\|_2$ norm, indicating fast convergence even without preconditioning.

Figure 12 shows the absolute dB difference between the natural basis method and the method of images. Figure 13 shows the absolute dB difference between the natural basis method and the Kirchhoff/physical optics method. In both cases the differences are small. The edge effect due to the location of the source is clearly visible in the comparison with the image method.

B. The double wedge

One of the most studied canonical boundaries is the wedge. Solutions in the form of infinite series (see Bowman[33] and Balanis[35] where the wedge is considered infinite in one coordinate direction) are well known as are high frequency asymptotic solutions.[36] Series solutions suffer,

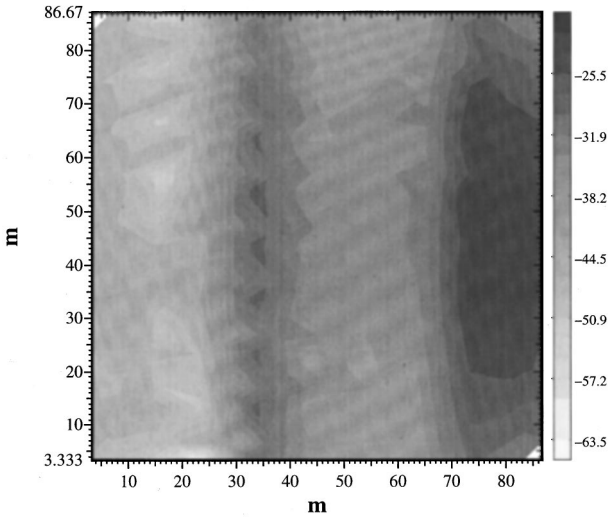


FIG. 15. Total field in dB 2.4 m above the three-dimensional double wedge.

computationally, from poor convergence characteristics and high frequency asymptotic methods such as the geometrical theory of diffraction,[36] which is used for terrain scattering problems,[37] have convergence difficulties when multiple wedges have apexes in shadow regions.[38] In this section we will consider an example of a double wedge configuration as illustrated in Fig. 14. The incident radiation is from a point source positioned at (100,45,16), as indicated in the figure and the radiating frequency is $\lambda = 0.5$ m.

Figure 15 illustrates the total field, in dB, 2.4 m above the surface of the scatterer. Figure 16 shows the absolute difference of the total fields, in *dB*, between the natural basis and the Kirchhoff method. Typically, the Kirchhoff method overestimates the field strength in the shadowed region. The absolute value of the kernel for this problem is illustrated in Fig. 17 and shows how many of the entries are close to zero. Its condition number, using the $\|\cdot\|_2$ norm, is 14.7082.

C. A box on a plane

Another interesting scattering problem is that of a box on a plane as illustrated in Fig. 18. In this example the source

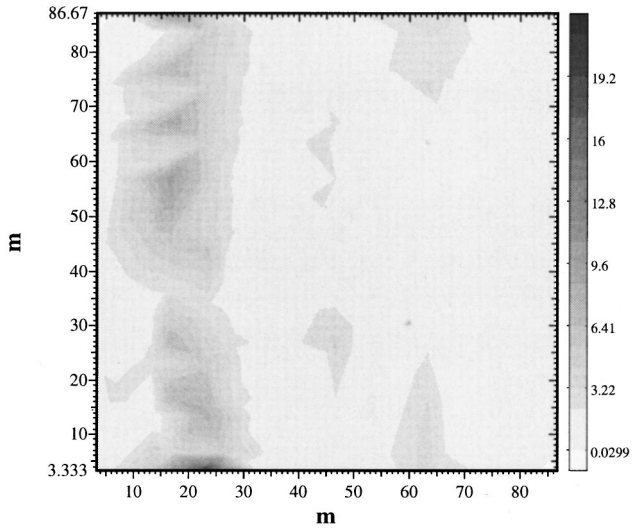


FIG. 16. Absolute difference, in dB, between the natural basis and Kirchhoff method for the three-dimensional double wedge.

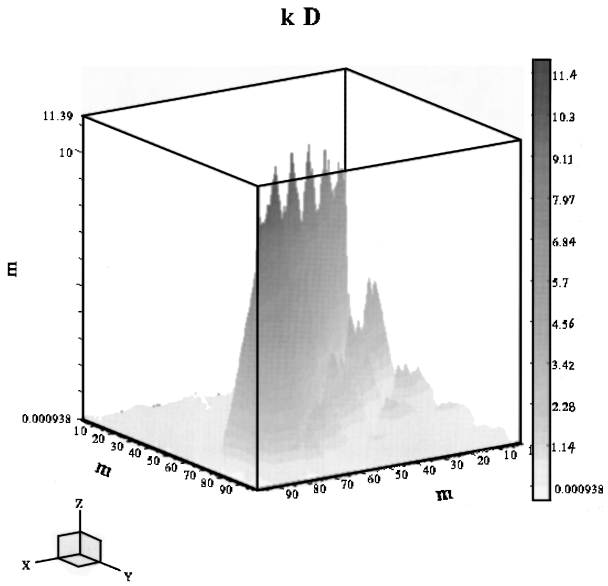


FIG. 17. Absolute values of the kernel $\mathcal{Z}$ for the three dimensional double wedge.

is situated at (0,0,10) and the frequency of the incident radiation is $\lambda = 0.33$ m. Figure 19 shows the magnitude, in dB, of the total field 5.0 m above the scatterer. A comparison with the Kirchhoff method is made in Fig. 20 where the absolute difference, in dB, between the Kirchhoff method and the natural basis method is given. Again, the overestimation, by the Kirchhoff method, of the field strength in the shadowed region is evident.

V. CONCLUSIONS

In this paper we have explored the application of a natural basis moment method to acoustically large objects achieving a large reduction in computational complexity. A reduction of three orders of magnitude in the density of unknowns over the direct approach has been demonstrated for grazing propagation over undulating surfaces, which is of significance for rough surface scattering.

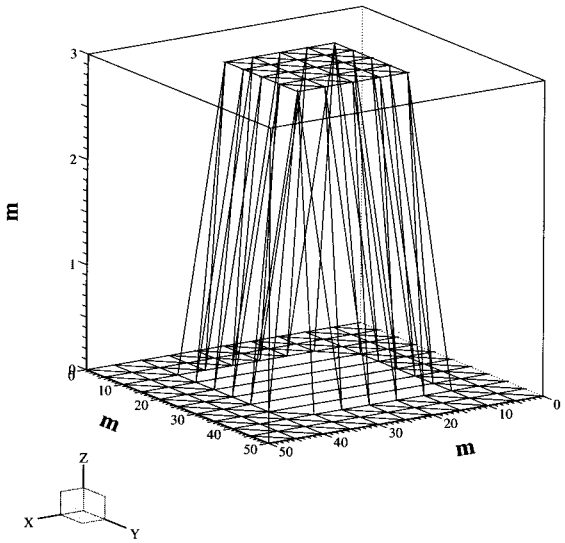


FIG. 18. Triangularization of box on plate.

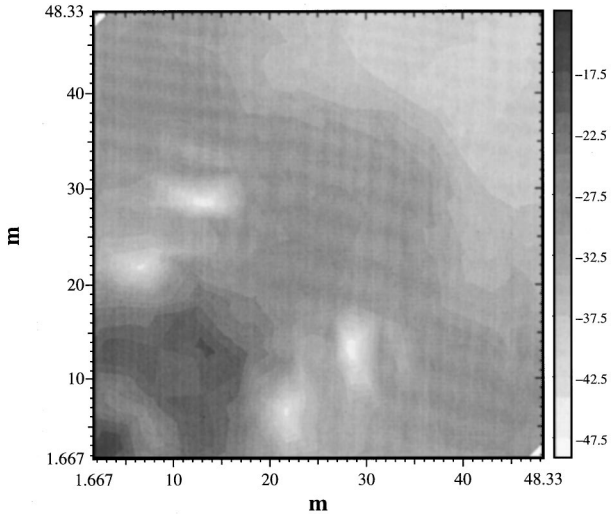


FIG. 19. Total field, in dB, 4.0 m above the surface given in Fig. 18.

We have demonstrated how to substantially increase the computational efficiency of the matrix fill integrations using a low order method for the forward scattering terms and an asymptotic method to solve the generalized Fourier integral for the backscattering terms. We have also suggested a heuristic for the construction of a preconditioner.

We have carefully compared the natural basis method to a number of other completely different methods including: (i) Born or Neumann series, (ii) The Green's function perturbation method, (iii) The Nussenzveig–Fock correction to the Kirchhoff approximation, and (iv) The parabolic equation method. In making these comparisons we have provided guidelines on the applicability of the method and provided a physical justification for the choice of basis function.

The method described can be applied to any integral equation where we have to calculate the unknown pressure gradient. It is directly applicable to the Dirichlet and Von Neumann boundary condition cases. It is probably applicable to scattering from lossy penetrable media although to date, we have only verified this in the lossy electrical dielectric case in two dimensions.

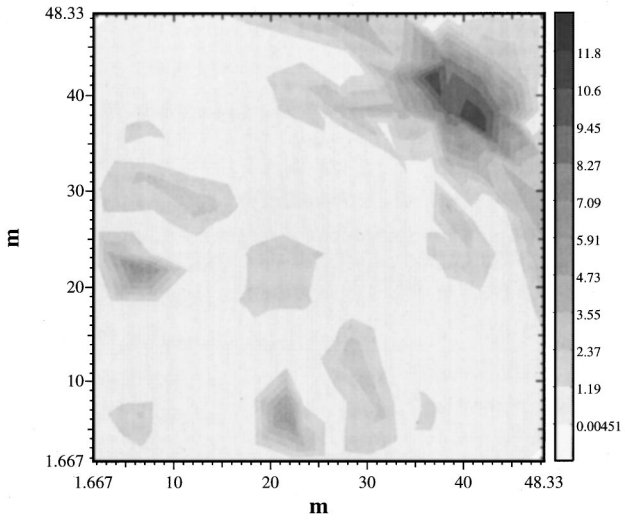


FIG. 20. Absolute difference, in dB, between natural basis and Kirchhoff method.

[1] K. Umashankar and A. Taflove, *Computational Electromagnetics* (Artech House, London, 1993).
[2] R. E. Kleinman, *The Review of Radio Science 1990–1992* (Oxford U. P., New York, 1993), Chap. 3.
[3] V. Rokhlin, ''Rapid Solution of Integral Equations of Scattering Theory in Two Dimensions,'' J. Comput. Phys. **86**, 414–439 (1990).
[4] C. C. Lu and W. C. Chew, ''A multilevel algorithm for solving boundary integral equations,'' Microw. Opt. Technol. Lett. **7**, 466–470 (1994).
[5] R. L. Wagner and W. C. Chew, ''A ray propagation fast multipole algorithm,'' Microw. Opt. Technol. Lett. **7**, 435–438 (1994).
[6] E. Michielssen and A. Boag, ''A multilevel matrix decomposition algorithm for analysing scattering from large structures,'' IEEE Trans. Antennas Propag. **44** (1996).
[7] F. X. Canning, ''Interaction matrix localization (IML) permits solution of larger scattering problems,'' IEEE. Trans. Magn. **27**, No. 5 (1991).
[8] R. M. James, ''A contribution to scattering calculation for small wavelengths–the high frequency panel method,'' IEEE Trans. Antennas Propag. **38**, 1625–1630 (1990).
[9] D. T. Moroney and P. J. Cullen, ''A Fast Integral Equation Approach to UHF Coverage Estimation in Mobile and Personal Communications,'' *Proceedings of the 2nd Joint COST 227/231Workshop on Mobile and Personal Communications, Florence 1995*, edited by E. Del Re (Elsevier, New York, 1995), pp. 343–350.
[10] K. R. Aberegg and A. F. Peterson, ''Application of the integral equation-asymptotic phase method to two-dimensional scattering,'' IEEE Trans. Antennas Propag. **43**, 534–537 (1995).
[11] C. Brennan and P. J. Cullen, ''A method to speed up iterative solutions of large-scale scattering problems,'' *IEE International Conference on Antennas and Propagation*, ICAP'97, Edinburgh, 1997 (IEE, London, 1997).
[12] P. J. Cullen and C. Brennan, ''Efficient methods for the computation of UHF grazing incidence terrain propagation,'' *International Conference on Electromagnetics in Advanced Applications* (Invited Paper), ICEAA'97 Torino, 1997.
[13] P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces* (Pergamon, New York, 1963), Appendix A.
[14] E. Kenny and P. J. Cullen, ''Fast numerical calculation of UHF propagation over two dimensional dielectric irregular terrain using an integral equation method,'' *IEEE Vehicular Technology Conference*, VTC'97 Phoenix, 1997 (IEEE, New York, 1997).
[15] T. K. Sarkar, ''A Note on the variational method (Rayleigh-Ritz), Galerkin's Method and the method of least squares,'' Radio Sci. **18**, 1207–1224 (1983).
[16] A. J. Poggio and E. K. Miller, *Integral Equation Solutions of Three Dimensional Scattering Problems in Computer Techniques for Electromagnetics* (Pergamon, New York, 1973), p. 168.
[17] W. C. Meecham, ''On the use of the Kirchhoff approximation for the solution of reflection problems,'' J. Rational Mech. Anal. **5**, 323–333 (1956).
[18] E. G. Liskza and J. J. McCoy, ''Scattering at a rough boundary – Extensions of the Kirchhoff approximation,'' J. Acoust. Soc. Am. **71**, 1093–1100 (1982).
[19] A. G. Voronovich, *Wave Scattering from Rough Surfaces* (Springer-Verlag, Berlin, 1994), pp. 127–130.
[20] D. T. Moroney, ''Computational methods for the calculation of electromagnetic scattering from large-scale perfect electrical conductors,'' Ph.D. thesis, University of Dublin, Trinity College, 1995.
[21] D. T. Moroney and P. J. Cullen, ''An integral equation approach to UHF coverage estimation,'' *IEE Ninth International Conference on Antennas and Propagation* (IEE, London, 1995), Pub. No. 407, Vol. 2, pp. 367–372.
[22] V. A. Fock, *Electromagnetic Diffraction and Propagation Problems* (Pergamon, New York, 1965).
[23] H. M. Nussenzveig, ''High-frequency scattering by an impenetrable sphere,'' Ann. Phys. **34**, 23–95 (1965).
[24] F. D. Tappert, *The Parabolic Approximation Method in Wave Propagation and Underwater Acoustics* (Springer-Verlag, Berlin, 1977).
[25] D. Lee and S. T. McDaniel, ''Ocean acoustic propagation by finite difference methods,'' J. Comp. Math. Appl. **14**, 305–423 (1987).
[26] G. D. Smith, *Numerical Solution of Partial Differential Equations* (Oxford U. P., New York, 1969).

[27] J. Jianming, *The Finite Element Method in Electromagnetics* (Wiley, New York, 1993).
[28] C. A. Balanis, *Advanced Engineering Electromagnetics* (Wiley, New York, 1989), Chap. 12.
[29] R. James, ''A three part analysis of scattering calculations at small wavelengths,'' BOEING Technical report D180-31586-1, 1989.
[30] A. H. Stroud and D. Secrest, *Gaussian Quadrature Formulas* (Prentice Hall, Englewood Cliffs, NJ, 1966).
[31] R. Wong, *Asymptotic Approximations of Integrals* (Academic, New York, 1989).
[32] C. M. Bender and S. A. Orszeg, *Advanced Mathematical Methods for Scientists* (McGraw-Hill, New York, 1978).
[33] J. J. Bowman, T. B. A. Senior, and P. L. E. Uslenghi, *Electromagnetic and Acoustic Scattering by Simple Shapes* (North Holland, Amsterdam, 1969).
[34] *Moment Methods in Antennas and Scattering*, edited by R. C. Hanson (Artech, Boston, 1990).
[35] C. A. Balanis, *Advanced Engineering Electromagnetics* (Wiley, New York, 1989), Chap. 11.
[36] G. L. Jones, *Geometrical Theory of Diffraction* (Peregrinus, London, 1986).
[37] R. J. Lubbers, ''Propagation prediction for hilly terrain using G.T.D. wedge diffraction,'' IEEE Trans. Antennas Propag. **AP-32**, no. 9 (1984).
[38] M. Schneider and R. J. Luebbers, ''A general, uniform double wedge diffraction coefficient,'' IEEE Trans. Antennas Propag. **AP-39**, no. 9 (1991).

Acoustical properties of end-initiated explosive line charges

William J. Marshall
BBN Technologies, Union Station, New London, Connecticut 06320

(Received 15 July 1997; accepted for publication 6 February 1998)

Acoustic radiation produced by end-fired explosive line charges is examined experimentally and theoretically. Three explosive compositions are studied: PETN detonating cord, RDX detonating cord, and thin ribbons of HLX. Specific output levels are formulated for each of the three explosive materials over the four 1-oct bands spanning 50–800 Hz. Bubble period relationships are derived for each material and compared to previous studies. In terms of directivity effects, long line charges are found to behave like beam-steered continuous arrays over moderately wide bands. A math model is developed which predicts acoustic levels and spectrum shapes for the shock wave component over a wide range of view angles and linear charge densities. This model is driven by two parameters which are extracted from the experimental data for the three compositions studied. The result is a practical means of designing explosive line arrays to desired source levels, bandwidths, and beam patterns. © *1998 Acoustical Society of America.* [S0001-4966(98)05305-3]

PACS numbers: 43.30.Lz, 43.30.Xm [DLB]

INTRODUCTION

Explosives have a long history of use in underwater acoustics, their main advantages over modulatable sources being short pulse lengths/wide bandwidths (for good time resolution) and high source levels (for long-range propagation and deep bottom penetration), all at modest expense. Fairly accurate performance predictions are possible for a number of explosive compositions,[1–5] but only when arranged in compact, omnidirectional configurations. Spatially extended configurations, particularly line charges, are also of interest because of their directivity and ease of fabrication. The main benefit of directivity is that sound is focused in desired directions and away from undesired directions. An important secondary benefit is that impressive far-field source levels may be achieved with comparatively small near-field pressures (that is, less environmental impact). For these reasons it is desirable to have data and techniques available for designing this type of explosive source.

The library of papers on acoustical properties of line charges is fairly sparse. Slifko[6] devised procedures for calculating spectra and beam patterns for the shock wave component. Since his aim was to find scaling laws for end-initiated lines, the actual composition of his charges was not identified and his data are of limited usefulness in designing new sources. Wentzell *et al.*[7] studied the bubble periods from point and line charges made of Composition A3 (waxed RDX). Sari and Nelson[8] reported on measurements of long vertical lines of PETN detonating cord, finding that the interplay of shock wave and bubble pulse characteristics caused tilting and splitting of the beam patterns. Their experiments explored a variety of cord charge densities and depths, and observed a bubble period coefficient significantly different from Wentzell's.

The present work describes a series of at-sea measurements of vertical top-fired line charges and develops a model describing shock wave radiation for this class of sources. The composition, charge density, length, and depth of the test objects were varied in order to investigate variations in parameters across these four dimensions. Three explosive compositions were used: HLX ribbon (an RDX-based flexible explosive equivalent to SX-2), PETN detonating cord, and RDX detonating cord. Acoustic measurements were made using a semivertical array of 13 hydrophones at a nominal test distance of 400 m.

Experimental results in the following categories are presented: maximum energy source level (ESL) and beamwidth, by band; energy spectral density (ESD) versus vertical view angle; bubble period versus depth; and impulse and shock wave decay time on the maximum response axis (MRA).

Definitions of ESL and ESD can be found in Ref. 9.

Empirical parameters required for the math model are extracted from the database. The model is generally consistent with Slifko's findings, but offers three advantages over his method: (1) closer agreement with measured data, particularly for modern flexible explosives, (2) a gradual rather than abrupt transition between view angle regimes, and (3) a time domain description. Since all measurements were made at a single range, direct verification of range-dependent line charge parameters is not possible with these data.

I. MEASUREMENTS

A. Test description

1. Geometry, technique, and environment

The test configuration consisted of a receiver platform suspending a bottom weighted 13-element vertical array of hydrophones, tethered downwind from a moored source deployment platform. Water depth at the test site was approximately 520 m. In addition to deploying the explosives, the source ship suspended three shallow hydrophones to monitor waveforms above the source. Since the test took place in a high current area (Straits of Florida), considerable tilting and bowing of both source and hydrophone arrays was inevitable. To account for this effect each explosive line array was augmented with small charges placed near its upper and lower end. These charges (called marker pops) were initiated

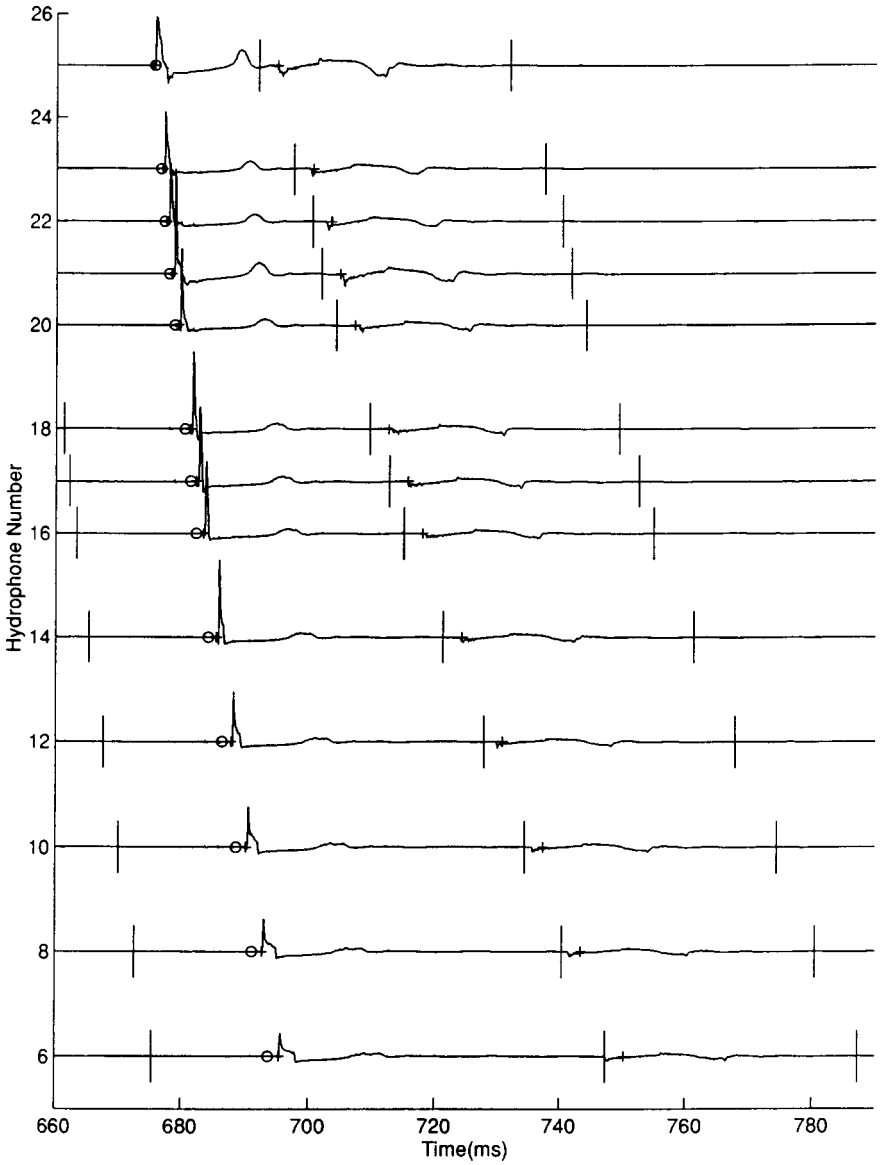


FIG. 1. Display used to determine source tilt and hydrophone locations by matching predicted and actual marker pop signals. Vertical spacing of traces is proportional to hydrophone spacing (Event C3-10).

a fraction of a second before the main charge, thus providing timing marks to determine range and relative angle from each end of the source to each hydrophone. This technique compensated well for tilts of both arrays and for hydrophone array bowing, but not for source bowing which remains an uncertainty in these measurements.

The hydrophone array included depth gauges at top and bottom. During data reduction the test geometry for each event was reconstructed by plotting a cascade of hydrophone signals (Fig. 1) and superimposing on each trace predicted marker pop arrivals, via direct (open circle) and surface reflected (plus sign) paths, for pops at known depths (calculated from their bubble periods) but displaced laterally in accordance with a variable source tilt angle. Varying source tilt in the simulation caused changes in perceived hydrophone locations and simulated marker pop arrival times. Comparing top and bottom hydrophone depths against the two depth gauge readings while monitoring simulated marker pop arrival times led to the optimum tilt value. The three vertical bars on each trace in Fig. 1 are analysis region boundaries for direct and surface-reflected portions of the signal. Figure 2 shows typical geometries observed for each of the three source lengths.

Data were recorded on 16-channel digital audio tape recorders with a useable bandwidth of 5 kHz (12-kHz sampling rate per channel). To synchronize recordings, the composite detonating signal for main charge and marker pops was telemetered from source to receiver ship and recorded at both locations.

The sound velocity profile (SVP) at the receiver ship featured a sharp thermocline but remained steady over the 2 days of the test. A typical SVP from this period, with lengths and depths of sources added, is given in Fig. 3. A constant sound speed of 1530 m/s was used for all data reduction and modeling purposes.

2. Objects tested

Table I gives a list of sources tested and shows the parameter variation plan. C1–C4 were compared to study the

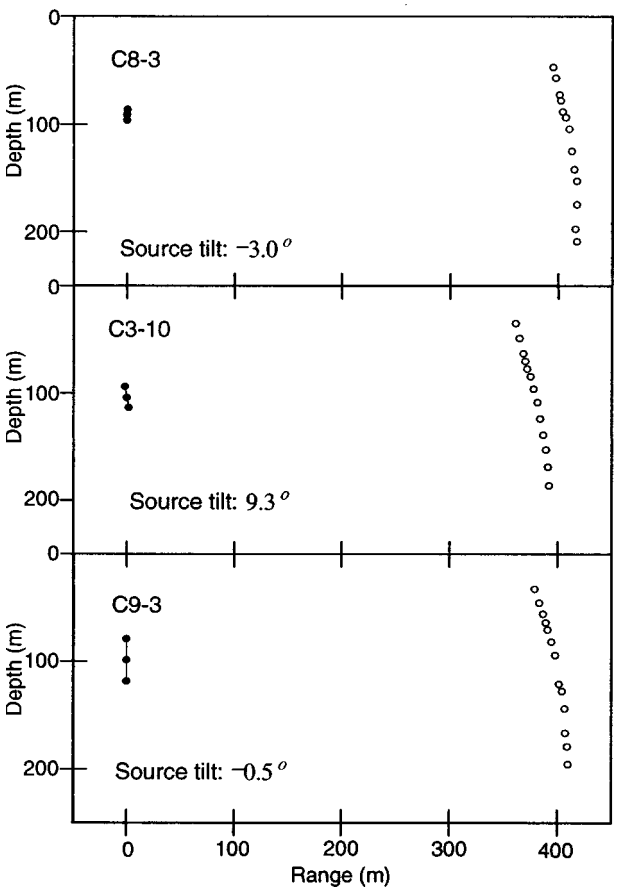


FIG. 2. Typical source/receiver geometries for three lengths of line source. Filled circles indicate source position (upper and lower marker pops, midpoint of interconnecting line); open circles indicate hydrophone positions.

effect of varying linear charge density, C5–C7 plus C3 to study depth variation, C8 and C9 plus C3 to study length variation, and so forth. C3, the standard configuration, was a 20-m (65-ft) length of HLX ribbon, top initiated at a mid-

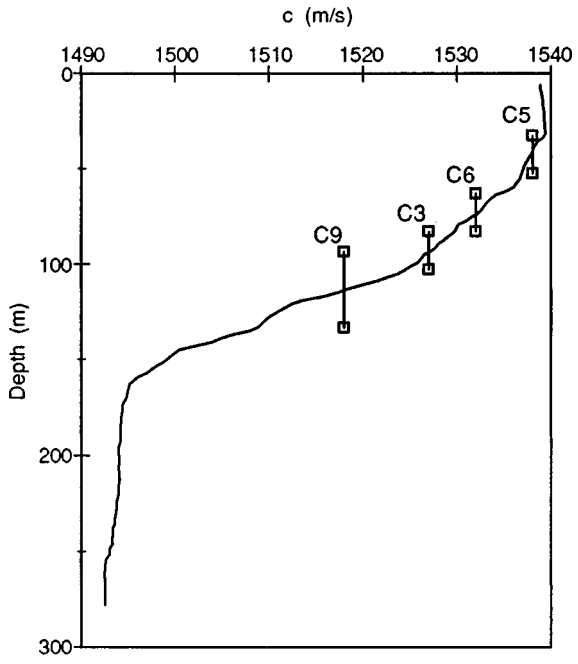


FIG. 3. Sound velocity profile during the test.

TABLE I. Test object catalog.

Name	Material	Density (grain/ft)	Length (m)	Mid-depth (m)	Number tested
C1	HLX	50	20	92.5	4
C2	HLX	100	20	92.5	6
C3	HLX	200	20	92.8	9
C4	HLX	400	20	103.5	4
C5	HLX	200	20	42.5	3
C6	HLX	200	20	73	5
C7	HLX	200	20	134	3
C8	HLX	200	10	98.5	4
C9	HLX	200	40	113.4	4
C10	PETN	50	20	103.5	3
C11	PETN	100	20	103.5	3
C12	PETN	200	20	103.5	3
C13	PETN	400	20	103.5	3
C15	RDX	80	20	101.3	3
C16	RDX	160	20	101.3	3

depth of 92.8 m. The HLX material is extruded initially into a flat noodle having a thickness of 3.2 mm (0.125 in.) and a width of 18 mm (0.72 in.) in which shape it has an effective linear charge density (excluding nonexplosive binders) of 400 grain/ft (0.085 kg/m). Smaller sizes are then formed by slitting this ribbon in halves, quarters, and eighths to achieve charge densities of 200, 100, and 50 grain/ft. The smallest of these was near the critical lateral dimension for this material and did not propagate reliably, so C1 results were not analyzed.

Detonating cord samples for the smaller charge densities were single strands of cord; multiple strands were placed side by side and taped at 1 ft intervals to form the larger sizes, 2×100, 4×100, and 2×80 grain/ft. All charges were initiated by duPont SSS blasting caps augmented with 6 g slip-on boosters.

B. Test results

Consider the typical set of waveforms shown in Fig. 1. The first disturbance in each trace is the initial shock wave. At view angles away from the shock axis this has the form of a triangular or rectangular impulse with a peaked leading edge. Nearer the MRA it narrows and grows in height until, exactly on the MRA, it becomes a brief sharp spike. The shock wave phase is followed by a period of slightly negative pressure and then a broad positive hump associated with collapse of the gas bubble formed by the explosion. (Occasionally a second bubble rebound can be seen in traces near the MRA.) The shape and relative position of the bubble hump is fairly consistent from phone to phone, indicating that this phenomenon has little directionality for these source sizes and depths. The last features are long-duration negative polarity impulses which follow the previous groups at ever longer delays on deeper hydrophones. These are surface-reflected echoes of the shock wave and usually display the low plateau seen far from the shock axis, but inverted due to having been reflected from a pressure release surface. Sometimes this feature overlaps the bubble pulse on shallower hydrophones.

Analysis of data for the 60 events in Table I is divided into three categories:

(1) angular variation of spectra and levels of total acoustic energy;
(2) dependence of bubble rebound period on depth; and
(3) derivation of MRA signal parameters used in the math model.

When designing a line source for given acoustical requirements the information in items (1) and (2) may be used for initial top level choices, then the math model described in Sec. II can be used to obtain more detailed spectra and band level energy predictions for the shock wave portion of the signal.

1. Energy spectra and beam patterns

Data reduction methods will be described using the C3 series as an example. For each hydrophone the shock wave and first bubble rebound portions of the waveforms are spectrum analyzed and the spectra summed over four 1-oct bands and one 4-oct band to obtain bandlimited ESLs. Plotting these values against hydrophone view angle (defined from hydrophone to source array center, with zero at broadside and positive angles toward the surface) produces the energy beam patterns in Fig. 4(a)–(e). Overlaying beam patterns from several events provides an improved picture of source behavior since the different measurement geometries sample different angles and angular regions. Note that one shot in Fig. 4 was at an extreme tilt angle so that only the sector from $-34°$ to $-60°$ was captured, thus providing useful information about the lower edge of the pattern. Beams become narrower in successively higher 1-oct bands, as expected, but higher bands also have higher peak levels because the on-axis spectrum rolls off at less than 3 dB per octave so there is a net energy gain in the higher bands. The 4-oct band has even higher levels [note the scale change in Fig. 4(e)].

Patterns in the higher bands consistently peak near $-10°$, while the lowest band, though it has little directivity and therefore an indistinct peak, seems to be spread fairly evenly between $+5°$ and $-10°$. Like the lowest band, the 100–200 Hz band shows considerable shot-to-shot variability, however, the sample having the highest output has a well-shaped peak aimed between $0°$ and $-5°$. Including the bubble rebound in the analyzed waveform generally adds 2 to 3 dB to wideband ESLs below about 500 Hz, and less at higher frequencies. Including the bubble also makes the source less directive, but only in the band containing the bubble frequency fundamental. The last panel in Fig. 4 demonstrates this by comparing spectra of signals measured on the hydrophone nearest the MRA, with and without the bubble portion of the waveform. Energy levels in narrow bands surrounding peaks of the bubble spectrum can exceed those of the shock wave-only case by as much as 6 dB.

Beam pattern results may be condensed and related to source properties by characterizing the maximum of each pattern by three metrics: angle of maximum response, maximum ESL value, and beamwidth. An MRA curve fitting procedure, described in the Appendix, is used to derive these three parameters for every test group in Table I, then these values are used to explore each family of sources.

The MRA direction for the early, shock wave portion of the emission is normal to the shock front. Calling v the detonation velocity and c the small-amplitude speed of sound, the Mach number, $M=v/c$, determines ϕ_s, the MRA angle measured from broadside, through

$$\sin\ \phi_s = 1/M. \tag{1}$$

ϕ_s ranges between $11.0°$ and $11.7°$ for these materials and test conditions.

Directivity effects are an antenna property (i.e., related to the spatial distribution of radiating elements) and should be the same for each of the three source materials. Let L be line charge length and λ the wavelength in water. For continuous line arrays the theoretical narrow-band 3-dB beamwidth of the broadside beam is $0.886\lambda/L$, and this relationship continues to hold for moderately wide bands if the spectrum is flat and λ is interpreted as the wavelength at the arithmetic band center. For beams steered away from broadside the theoretical relationship[10] changes to

$$B_{-3} = \sin^{-1}(0.443\lambda/L + \sin\ \phi_s) + \sin^{-1}(0.443\lambda/L - \sin\ \phi_s). \tag{2}$$

Fitted beamwidths are plotted against λ/L in Fig. 5. These charts show that measured beamwidths, in both 1- and 4-oct bands, follow the curve given by (2). Greater shot-to-shot variations and departures from theory are seen for the broader beams (shorter L's, lower frequencies) because these have little directivity which makes the beamwidth more difficult to measure.

Acoustic output in a given band is determined by three factors: composition of the energetic material, quantity of that material, and array directivity. Directivity is accounted for by recognizing that the effect of concentrating an omnidirectional energy level E_0 into an axisymmetric fan beam B radians wide is to raise the apparent level seen on the MRA by the factor $1/\sin(B/2)$. For small B this factor can be approximated by $2/B$, with the approximation error less than 0.1 dB for $B<40°$. If the spectrum is fairly flat over a band, then the acoustical energy in that band will be directly proportional to the net explosive weight, W,

$$E_{\max} = E_0 W(2/B) \tag{3}$$

or, expressed in decibels,

$$\mathrm{ESL_{MRA}} = \mathrm{ESL_0} + 10\log\ W + 3 - 10\log\ B, \tag{4}$$

where $\mathrm{ESL_0}$, interpreted as the omnidirectional source level per kg of active material, varies with explosive composition and band. By normalizing output by explosive weight and beamwidth, (4) condenses the measured source level data to a set of $\mathrm{ESL_0}$ values which are presented in Table II.

Constants in Table II were obtained by solving (4) for $\mathrm{ESL_0}$ and averaging over the set of measurements for each material. Because of the large scatter in measured beamwidths seen in Fig. 5, lower $\mathrm{ESL_0}$ variances were obtained using theoretical beamwidths from (2) rather than measured

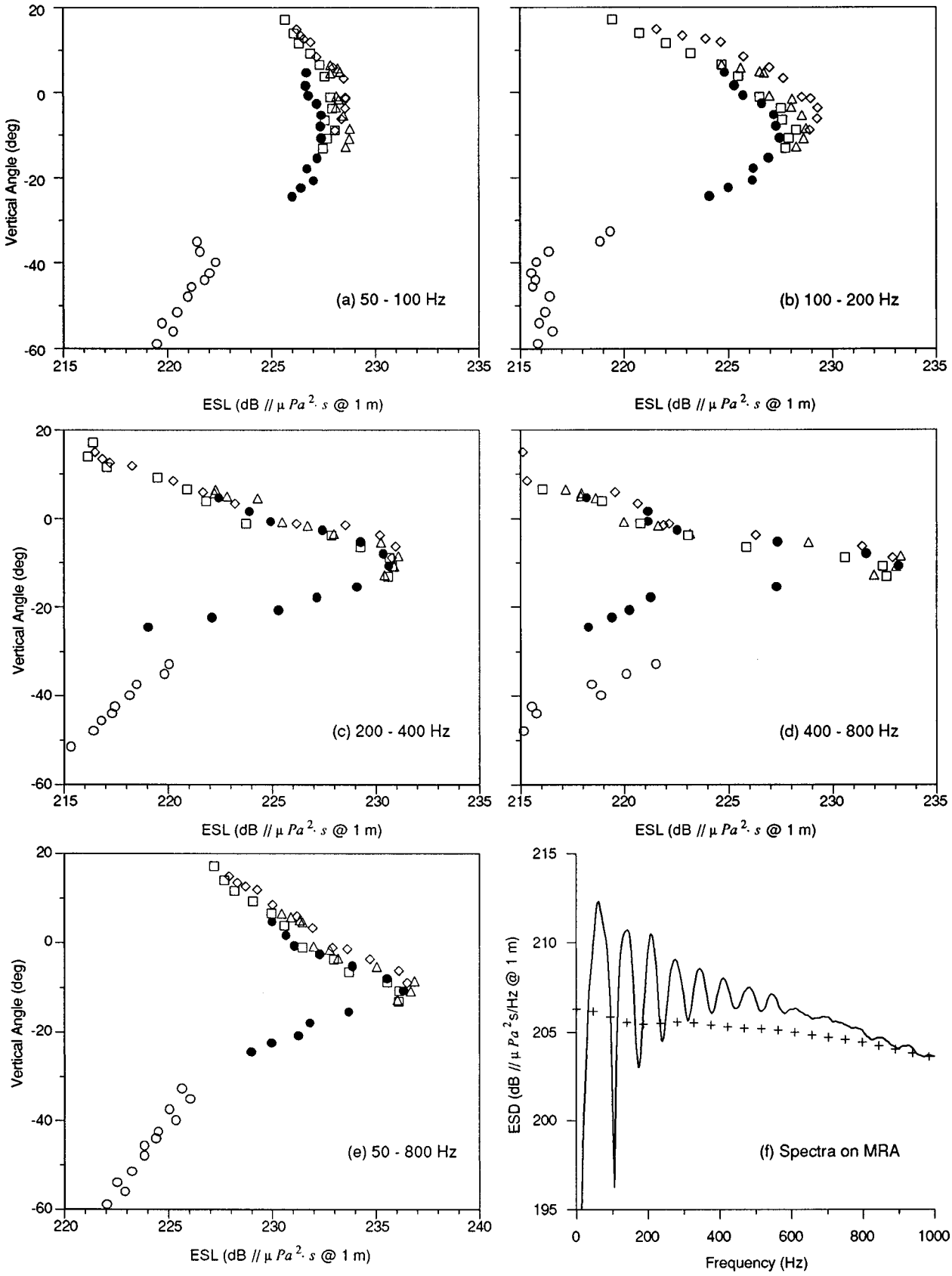


FIG. 4. (a)–(e) Superposition of ESL beam patterns for five C3 events. (f) Spectra from hydrophone nearest the MRA for C3-2 (solid line: complete direct path waveform, +: shock wave portion only). C3-2 is plotted with □ symbols in (a)–(e).

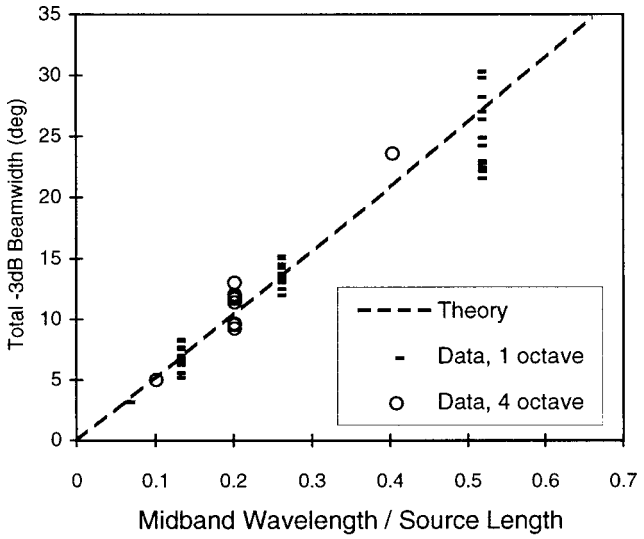


FIG. 5. Measured and theoretical beamwidths. Dashed line is the theoretical narrow-band curve for a continuous line steered 11° off broadside [Eq. (2)].

ones. By connecting source level to predicted directivity, this choice enhances the use of Table II and Eqs. (2) and (4) as a design tool.

Two trends are seen in Table II. First, frequency dependence over the four 1-oct bands is weak except for HLX which shows significantly higher output in the lowest band. This is due to the bubble effect which adds energy to the band containing the bubble frequency. For HLX samples this frequency remains in the lowest band consistently for all test conditions, thus raising the average level in that band. For the two detonating cord samples, depending on cord density and depth, the bubble frequency can fall into either of the lowest two bands, which lowers its impact on either one. The second trend noted in Table II is that, comparing the three materials, PETN cord has the largest acoustic output per active weight, about 2 dB greater than HLX ribbon.

2. Bubble period

For the depths and charge sizes used in these tests, oscillation of the gas bubble generally produces nearly omnidirectional radiation (the wavelength associated with the bubble period is usually longer than the charge length), and the bubble period can be measured most conveniently on a near-MRA waveform where a clear shock wave peak is present. Observed bubble rebounds vary from consistent, sharp, well-defined peaks to weak, broad, and sometimes double humped features. No bubble data was taken for the C5 series (the shallowest) because the surface reflection of the shock wave obscured the bubble rebound.

TABLE II. Average values of ESL_0, by band, in dB//μPa^2 s/kg at 1 m. The uncertainty span represents the standard deviation for HLX and PETN (8 and 4 samples, respectively) and the half range to extrema for RDX (2 samples).

Band	HLX	PETN	RDX
50–100 Hz	225.1±1.5	225.9±1.3	225.2±1.6
100–200 Hz	222.8±0.8	225.0±0.7	225.1±0.7
200–400 Hz	222.3±0.5	224.0±0.2	223.1±0.1
400–800 Hz	221.5±0.8	223.3±0.2	222.6±0.0
50–800 Hz	229.2±0.6	231.2±0.4	230.5±0.1

Wentzell[7] has shown that the first bubble period for line charges is directly proportional to the square root of the linear charge density and inversely proportional to the hydrostatic pressure,

$$T_B = K_1 (W/L)^{1/2}/(H+10), \tag{5}$$

where H is the center depth of the charge, in m, and K_1 is a material-specific constant to be determined experimentally. With the limited range of depths available in this data set, the best method for finding K_1 is to rearrange (5) into

$$H+10 = K_1 \frac{(W/L)^{1/2}}{T_B} \tag{6}$$

and form a scatter plot of hydrostatic depth against the reciprocal of reduced bubble period. K_1 is the slope of the best straight line from the origin through the resulting cluster of points.

Seven sets of HLX points and a straight line fit through these data are plotted in Fig. 6(a). The C9 data points have the greatest deviation, presumably because it is the longest array. For the RDX and PETN data in Fig. 6(b), two curve fits are shown. The RDX slope matches that for HLX, while most of the PETN data shows a higher K_1 (longer period, lower bubble frequency) than HLX and RDX. The smallest charge density PETN cord, C10, has lower bubble periods than the remaining three cases. These results are tabulated and compared to prior measurements in Table III. The present measurements agree with Sari and Nelson's[8] value for PETN cord, but show an RDX value some 40% larger than that obtained by Wentzell for Composition A-3 pressed into long thin-walled aluminum cylinders. (Composition A-3 is 91% RDX by weight.) Some of this difference can be attributed to the inertial effect of the aluminum casing, but most of it remains unexplained.

3. Shock wave decay time constant and impulse

The model to be developed in Sec. II requires two empirical parameters related to the MRA waveform, shock wave decay rate and either peak pressure or impulse. Ideally, the far-field pressure waveform of a line charge resembles that of a point charge when viewed on the shock axis: a nearly instantaneous rise to a sharp peak followed by a steady, and initially exponential, decay. In the present experiment, bowing of the source line prevents the MRA waveform from achieving ideal form, and this complicates finding the peak parameters.

Figure 7 shows a one-sided impulse with exponential decay, sampled at evenly spaced time steps. Call M_n $(n=-1,0,1,...)$ the time series for this waveform. Assume the ideal form for the impulse and let it reach its peak value at time t_p. Then the nth measurement during the decaying portion of the signal is

$$M_n = p_0 \exp(-(t_n - t_p)/\tau). \tag{7}$$

Taking the ratio of any two successive readings eliminates both p_0 and the start time t_p yielding an estimate of τ alone

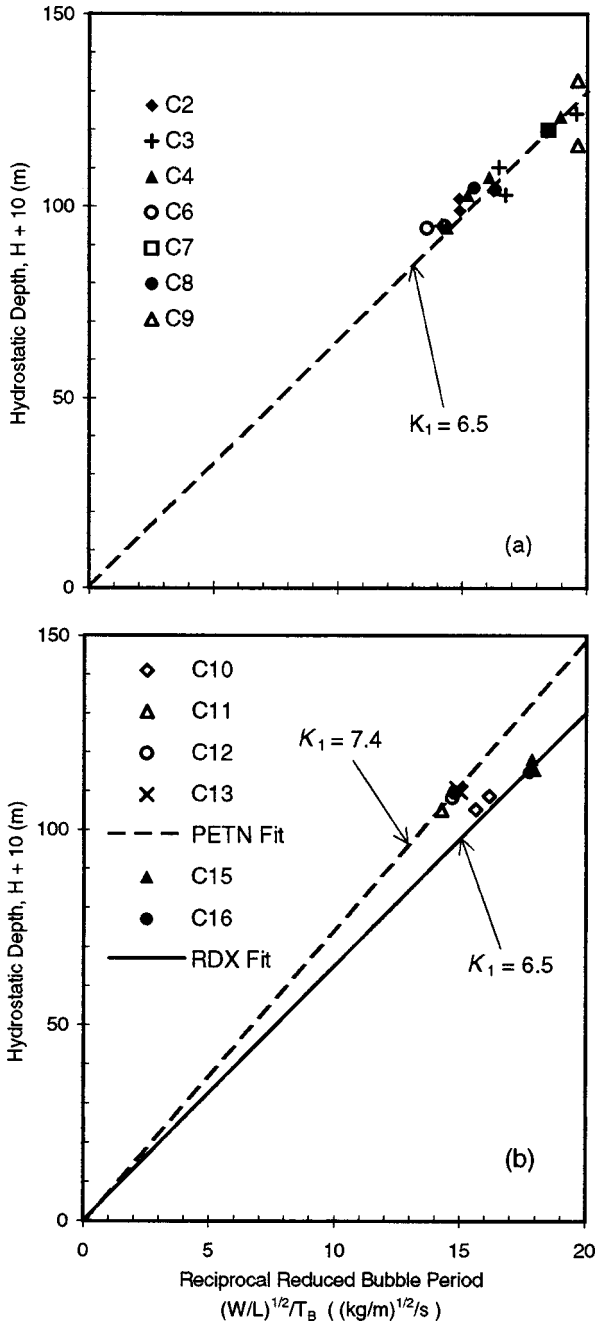


FIG. 6. Bubble period data with lines from Eq. (6) drawn through clusters. (a) HLX ribbon, (b) PETN and RDX cord.

$$\tau_n = \frac{\Delta t}{\ln(M_n/M_{n+1})}, \tag{8}$$

where $\Delta t = t_{n+1} - t_n$ is the sampling interval. Usually the estimate formed from the first two large values in the sequence is used.

It follows from the principle of similarity that τ for spherical charges is proportional to $W^{1/3}(W^{1/3}/R)^{\alpha}$, where R is the measurement distance and α an experimentally determined parameter given by Arons[2] as -0.22 (for TNT and Pentolite) and, more recently, by Chapman[11] as -0.14 (for Cyclotol). In the near field of a line charge the type of wave changes as it propagates, and there is no similitude equation for time constant in this case;[12] however, transition to spherical propagation occurs at about $R \approx L/2$ so at our test distances ($10L$ for the longest specimen) a simple scaling law can be found for τ as observed on the MRA. The scaling variable, however, is not explosive weight but linear charge density, W/L.

TABLE III. First bubble period constants (MKS units: $m^{3/2}\,kg^{-1/2}\,s$, English units: $ft^{3/2}\,lb^{-1/2}\,s$).

Material	Data source	K_1 (MKS)	K_1 (English)
Comp A-3	Wentzell *et al.*	4.5	18
RDX	This paper	6.5	26
HLX	This paper	6.5	26
PETN	Sari/Nelson	7.4	30
PETN	This paper	7.4	30

For the remainder of the paper let τ without a subscript denote the decay time constant of a line charge, as observed on the MRA. To find the form of $\tau(W/L)$, consider the line charge to be a series of independent bits of explosive strung end to end. A sensible choice for the length of these bits is something proportional to the distance the detonation moves along the line in time τ, say $Av\tau$. Taking Chapman's τ formula as the most accurate recent measurement for an underwater point charge we have, using MKS units,

$$\begin{aligned}\tau &= 8.12(10)^{-5} W^{(1-0.14)/3} R^{0.14} \\ &= 8.12(10)^{-5} ((W/L)Av\tau)^{0.86/3} R^{0.14},\end{aligned} \tag{9}$$

which reduces to

$$\tau = 1.84(10)^{-6} (Av)^{0.402} (W/L)^{0.402} R^{0.196}, \tag{10}$$

suggesting that an exponent of about 0.4 on charge density might describe line charge τ data.

Figure 8 displays near-MRA τ measurements at constant R plotted against W/L. The large τ variation at each W/L arises from two causes: not being exactly on the shock axis, and curvature of the source in the plane containing the hydrophone. Both of these factors cause observed values to exceed true values, so the smallest of a set is that closest to the desired value. The power law that best fits the minimal PETN τ's in Fig. 8(a) is $\tau = 15.2(W/L)^{0.4}$ with τ in microseconds and W/L in the traditional linear charge density units, grains/ft. The good match of the slope of this line to that of the locus of data minima justifies treating the line

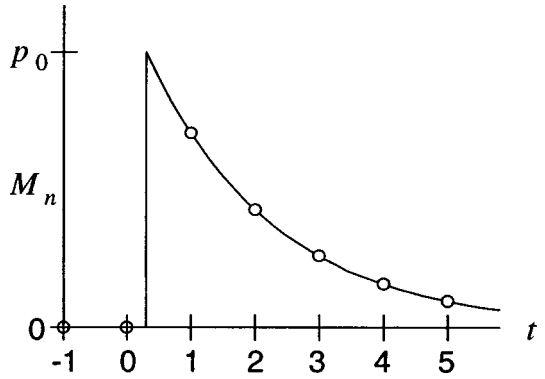


FIG. 7. Shock wave front pressure samples taken at regular time intervals.

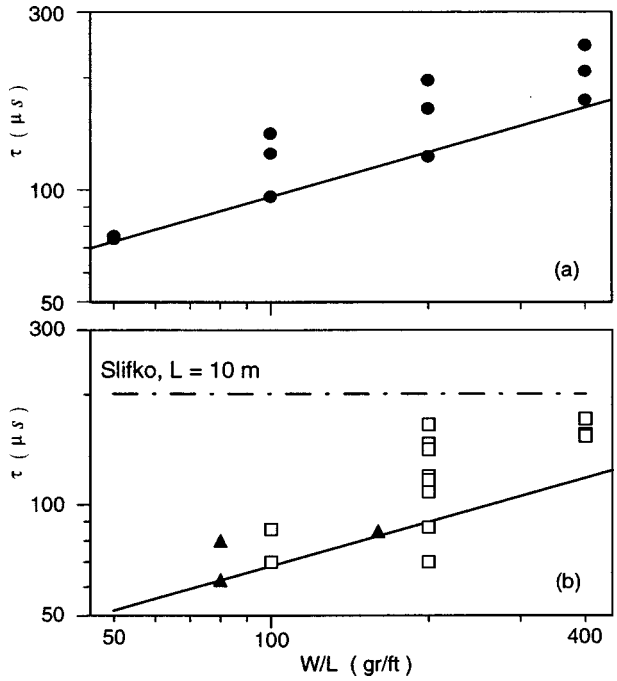


FIG. 8. Shock wave decay time constant near MRA, measured data and curves of the form $K(W/L)^{0.4}$, with W/L in grains/ft. (a) circles: PETN cord, line: $K=15.2$. (b) squares: HLX ribbon, triangles: RDX cord, solid line: $K=10.8$.

charge as a series of independent bits. Finding the appropriate value of A will be addressed after the impulse measurements are evaluated.

The τ data for the other two materials are given in Fig. 8(b). The tall spread of data points for 200 gr/ft HLX samples can be divided into three groups: the lowest cluster for the shortest arrays (C8), the middle group for the standard 10-m lengths (C7), and the highest group for the longest lengths (C9). Although a definite trend toward longer τ for heavier ribbons can be seen, the locus of minima does not form a straight line, so an empirical line is not given for the HLX data. The very limited amount of RDX data, however, does seem consistent with a 0.4 power law, and the line which passes through the lowest RDX point is given by $\tau = 10.8(W/L)^{0.4}$.

Slifko's experiments[6] on short horizontal lengths of cord (composition unspecified) measured decay rate as a function of deviation from the shock axis. His on-axis result, at $R = 8L$ and for two very different cord densities, 120 and 1400 gr/ft, was that τ did not obey a power law like $(W/L)^{\alpha}$ but, rather, was independent of W and directly proportional to L: $\tau = 0.031(L/c)$. That result, plotted in Fig. 8(b) for a 10-m line length, is completely incompatible with our data. One might think having τ proportional to L is a reasonable result: the time for the detonation to travel the length of cord is L/v, and that certainly influences waveform duration when away from the shock axis, and even on axis when in the near field. However, the shock front from the entire line arrives at one instant when viewed in the far field on the MRA, so L cannot have a strong influence on τ in this case.

Impulse is defined as $\int p(t)dt$ with the integration extending over the positive portion of the pressure waveform. The impulse of an ideal shock wave (7) is simply $p_0\tau$, therefore p_0 for an exponentially decaying transient can be inferred from a measurement of its impulse. Impulse is computed from time series data by

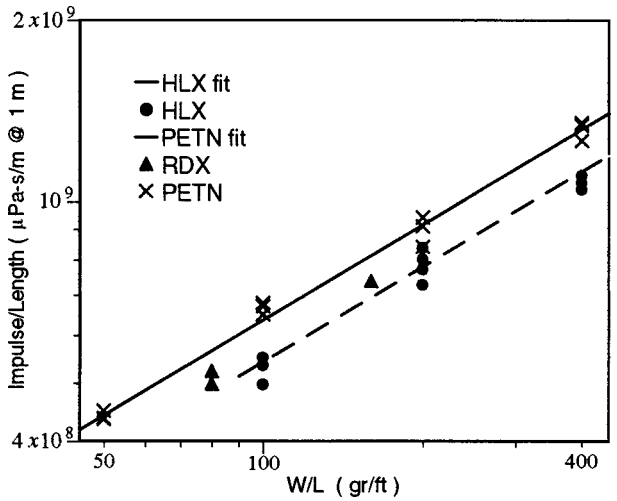


FIG. 9. Shock wave impulse per unit length, near MRA. Solid line: $5.65 \times 10^7\ (W/L)^{0.526}$, dashed line: $4.81 \times 10^7\ (W/L)^{0.526}$ (W/L in grains/ft).

$$I = \Delta t \sum M_n \tag{11}$$

with the summation taken over contiguous positive M_n surrounding the peak.

To determine the dependence of line charge impulse on source parameters, start with Chapman's formula for the peak pressure of a single isolated charge of weight W/N

$$p_0 = C_0(W/N)^{0.377}/R^{1.13} \tag{12}$$

and, as before, regard the line as $N = L/Av\tau$ independent bits whose shock waves add coherently to a peak pressure of Np_0 on the MRA. The impulse on the MRA is then

$$I = \tau N p_0 = \frac{LC_0}{Av}\left(\frac{W}{L}Av\tau\right)^{0.377} R^{-1.13}$$

$$= C_0 L (Av)^{0.377-1}\left(\frac{W}{L}\tau\right)^{0.377} R^{-1.13}. \tag{13}$$

Replacing τ in (13) by (10) results in

$$I = C_1 L (C_2 (W/L)^{1.402} R^{0.196})^{0.377} R^{-1.13}$$

$$= C_3 L (W/L)^{0.529} R^{-1.056}, \tag{14}$$

where C_0, C_1, C_2, and C_3 are various proportionality constants.

Experimental values of impulse per unit length are plotted against W/L in Fig. 9. A least-squares fit of logarithms of the PETN data (solid line) gives a power law exponent of 0.526, in excellent agreement with (14). The dashed line in Fig. 9 is that curve with the same exponent which best fits the HLX data. The coefficient for the HLX line is 0.85 times that for the PETN line, or -0.7 dB. It is clear from the graph that the sparse RDX data would fall near a third line between and parallel to the first two. Experimental determination of the R dependence of I is not possible with constant distance measurements, but it is interesting to note that the exponent

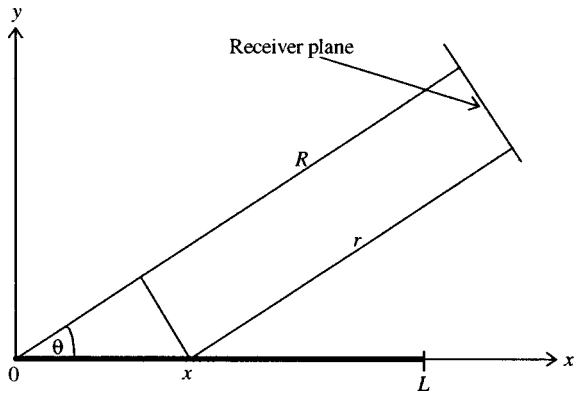


FIG. 10. Notation for model of radiation from an end-fired line charge.

of R reported by Cole[13] for the impulse of Pentolite point charges is -1.05, in close agreement with (14).

Preserving the A dependence of C_3 in (14) and using $v=8000$ m/s one finds that (14) matches the experimental I/L curve for PETN when A is close to 1, as hypothesized. Inserting this A in (10) predicts τ's about half as large as those from the minimal τ curve in Fig. 8(a). We conclude, therefore, that τ's predicted by (10) are a lower bound for those realized on long, curved cords.

II. MODEL

A. Derivation

Consider the geometry of Fig. 10 where a line of length L is detonated at the origin at time $t'=0$, and the detonation travels along the positive x axis with velocity v, reaching position $x=vt'$ at t'. Assume the shock wave pressure observed at a distant field point (R,θ) can be found by adding the pressures developed by successive sections of a cylinder of radius a and infinitesimal length dx. Then

$$p(t')=\frac{\rho}{4\pi}\int_{x_1}^{x_2}\frac{2\pi a\ddot{a}(x,t'-r/c)}{r}\,dx, \tag{15}$$

where ρ is the undisturbed water density, $\ddot{a}$ the radial acceleration of the cylinder's surface at the moment of detonation, and the limits of integration are functions of t' to be determined presently. Assume that each cord element emits a pressure pulse like that of a small point charge,

$$\ddot{a}(x,t')=\begin{cases}\alpha\,\exp(-(t'-x/v)/\tau), & 0<x<vt'\\ 0, & x\leqslant 0,\ x\geqslant vt'\end{cases}, \tag{16}$$

where α and τ are experimentally determined properties of the charge. Using (16) and $r=R-x\cos\theta$ from Fig. 10, (15) becomes

$$p(t')=\frac{\rho a\alpha}{2}\int_{x_1}^{x_2}\frac{\exp(-[t'-(R/c)+(x/c)\cos\theta-x/v]/\tau)}{R-x\cos\theta}\times dx. \tag{17}$$

Now make the usual far-field simplification that when $R\gg L$, variations in the argument of the exponential function have a strong influence on the integral, while variations in the denominator do not. Then the denominator can be replaced by R and (17) reduces to

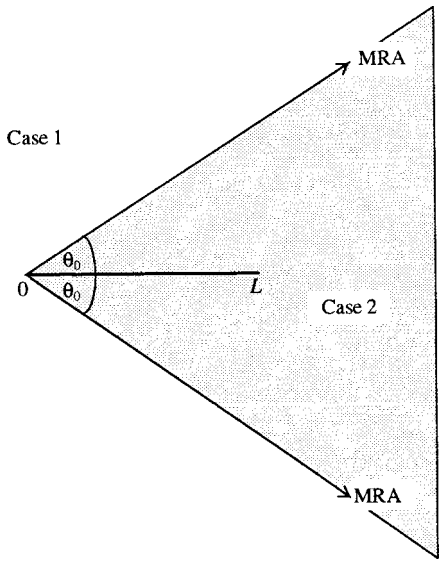


FIG. 11. Two regions are distinguished: behind (case 1) and in front of (case 2) the MRA.

$$p(t')=\frac{\rho a\alpha}{2R}\,\Phi(t'-t_0)\times\exp\left(-\frac{(t'-t_0)}{\tau}\right)\int_{x_1}^{x_2}\exp\left(\frac{\beta}{\tau}\frac{x}{v}\right)dx, \tag{18}$$

where Φ is the Heaviside step function, $t_0=R/c$ is the travel time from the origin to R, and $\beta=1-M\cos\theta$. Finally, shifting the time variable to $t=t'-t_0$ (so time is measured from the instant sound from the detonator reaches R) and evaluating the definite integral, (18) becomes

$$p(t)=\frac{\rho a\alpha}{2R}\frac{v\tau}{\beta}\,\Phi(t)\exp\left(-\frac{t}{\tau}\right)\times[\exp(\beta x_2/v\tau)-\exp(\beta x_1/v\tau)]. \tag{19}$$

Two cases are evident (see Fig. 11). Case 1 is when $\beta>0$ or $\theta>\cos^{-1}(1/M)$, the region behind the shock axis, and case 2, with $\beta<0$, the smaller region forward of the shock axis.

In case 1, $x<v(t'-r/c)=v(t'-t_0)+xM\cos\theta$ so that $x(1-M\cos\theta)<v(t'-t_0)$ or $x<vt/\beta$. The first sound to reach R arrives from the detonator end of the line at $t=0$ and $x_1=0$. During the interval $0<t<\beta L/v$ the upper limit follows $x_2=vt/\beta$ as the detonation progresses down the line. After the detonation reaches L, x_2 remains fixed at L for $t\geqslant\beta L/v$.

In case 2, $x<v(t_0-t')/(-\beta)=vt/\beta$ with both β and t negative. The first sound to reach R arrives from the end of the line opposite the detonator at $t=\beta L/v$ and $x_2=L$. During the interval $\beta L/v<t<0$ sound from earlier positions on the line arrives and the lower limit follows $x_1=vt/\beta$ until after $t\geqslant 0$ when it stays fixed at $x_1=0$.

Evaluating (19) with these limits produces the general waveform

$$p_{1,2}(t)=\frac{\rho a\alpha}{2R}\frac{v\tau}{\beta}G_{1,2}(t)\exp(-t/\tau) \tag{20}$$

with the subscript indicating case 1 or 2 and

$$G_1(t)=\begin{cases}0, & t\leqslant 0\\ 1-e^{t/\tau}, & 0<t<L\beta/v\\ 1-e^{L\beta/v\tau}, & t\geqslant 0\end{cases}$$
$$G_2(t)=\begin{cases}0, & t\leqslant L\beta/v\\ e^{t/\tau}-e^{-L\beta/v\tau}, & L\beta/v<t<0\\ 1-e^{-L\beta/v\tau}, & t\geqslant 0.\end{cases} \tag{21}$$

Exactly on the MRA neither form can be used, however, direct evaluation of (18) with $\beta=0$, $x_1=0$, and $x_2=L$ yields

$$p_{\mathrm{MRA}}=\frac{\rho a\alpha}{2R}L\Phi(t)\exp(-t/\tau), \tag{22}$$

which has the form of a point charge shock wave, as expected.

The impulse calculated from this model is independent of θ,

$$I_0=\frac{\rho a\alpha}{2R}L\tau \tag{23}$$

and this is confirmed experimentally. (For instance Slifko reports[6] that, at ranges beyond $1.5L$, measured line charge impulse varies little with emission angle.) Therefore data measured on the MRA may be used to establish the two parameters needed in the model: time constant, directly from the empirical equations in Fig. 8, and amplitude, through (23) and the impulse equations in Fig. 9. For example, for PETN detonating cord we have

$$\tau_{\mathrm{PETN}}=15.2(10)^{-6}(W/L)^{0.4} \tag{24}$$

and

$$\left.\frac{\rho a\alpha}{2R}\right|_{\mathrm{PETN}}=\frac{(I_0/L)}{\tau}=\frac{5.65(10)^7(W/L)^{0.526}}{15.2(10)^{-6}(W/L)^{0.4}}$$
$$=3.72(10)^{12}(W/L)^{0.26}. \tag{25}$$

Typical waveforms produced by this model are shown in Fig. 12. Away from the shock axis the impulse is long and flat-topped. Nearer the axis it becomes shorter and grows in height, maintaining the area under the curve (impulse) constant. All examples in Fig. 12 are from the region behind the MRA (case 1); curves for positive deviations (case 2) have the same size and shape but are displaced to the right and begin at $t=0$.

B. Model/data comparisons

The model has been validated by comparing predictions to data for much of the data reported in Sec. I. This is illustrated for a subset of view angles for one example, C8-2, in Fig. 13. Characteristics of unit C8-2 used in the model were

$$L=9.9\ \mathrm{m},\quad M=5.33,\quad \tau=80\ \mu\mathrm{s},$$

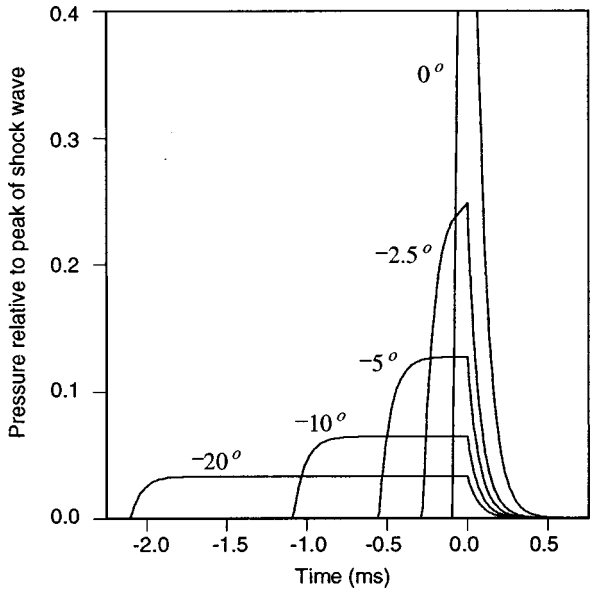


FIG. 12. Modeled line charge waveforms at indicated angles relative to the shock axis, normalized by peak pressure exactly on the shock axis ($M=5.33$).

$$I_0=8.31\times 10^9\ \ \mu\mathrm{Pa\ s\ at\ 1\ m}$$

and model time was discretized at 12 kHz, the same rate used in the measurement. All spectra were cut off at the Nyquist frequency, 6 kHz.

Of the five angles depicted in Fig. 13, (b) is the closest to the theoretical shock axis angle of 79.2° for this M. The other angles span the measurement extremes for this unit. It was found that the match improved if the angle used in the model was tweaked slightly from the value derived from marker pop analysis. This adjustment varied from a few tenths of a degree near the MRA to 2° at the maximum view angle, and is believed to be due to uncertainty in M, angle measurement errors, and source curvature.

Overall the agreement is quite good, and major features of the measurement are predicted well in both time and frequency domains. The negative phase following the shock wave is not modeled, but this feature affects only the very low-frequency end of the spectrum.

III. CONCLUSIONS

(1) Directivity effects for explosive line arrays can be approximated over moderately wide bands by continuous array methods.

(2) Ranking the three materials by specific acoustic output per active weight, PETN cord is the strongest, RDX cord intermediate, and HLX ribbon weakest. The difference between PETN and HLX is about 2 dB.

(3) Bubble period data in this experiment agreed with Sari and Nelson's result for PETN cord, but deviated significantly from Wentzell's result for RDX.

(4) A model of shock wave acoustics for end-initiated line charges has been developed and validated. Scaling laws which relate τ and I, the two parameters controlling the model, to linear charge density are based on the hypothesis that a line charge may be treated as a string of independent

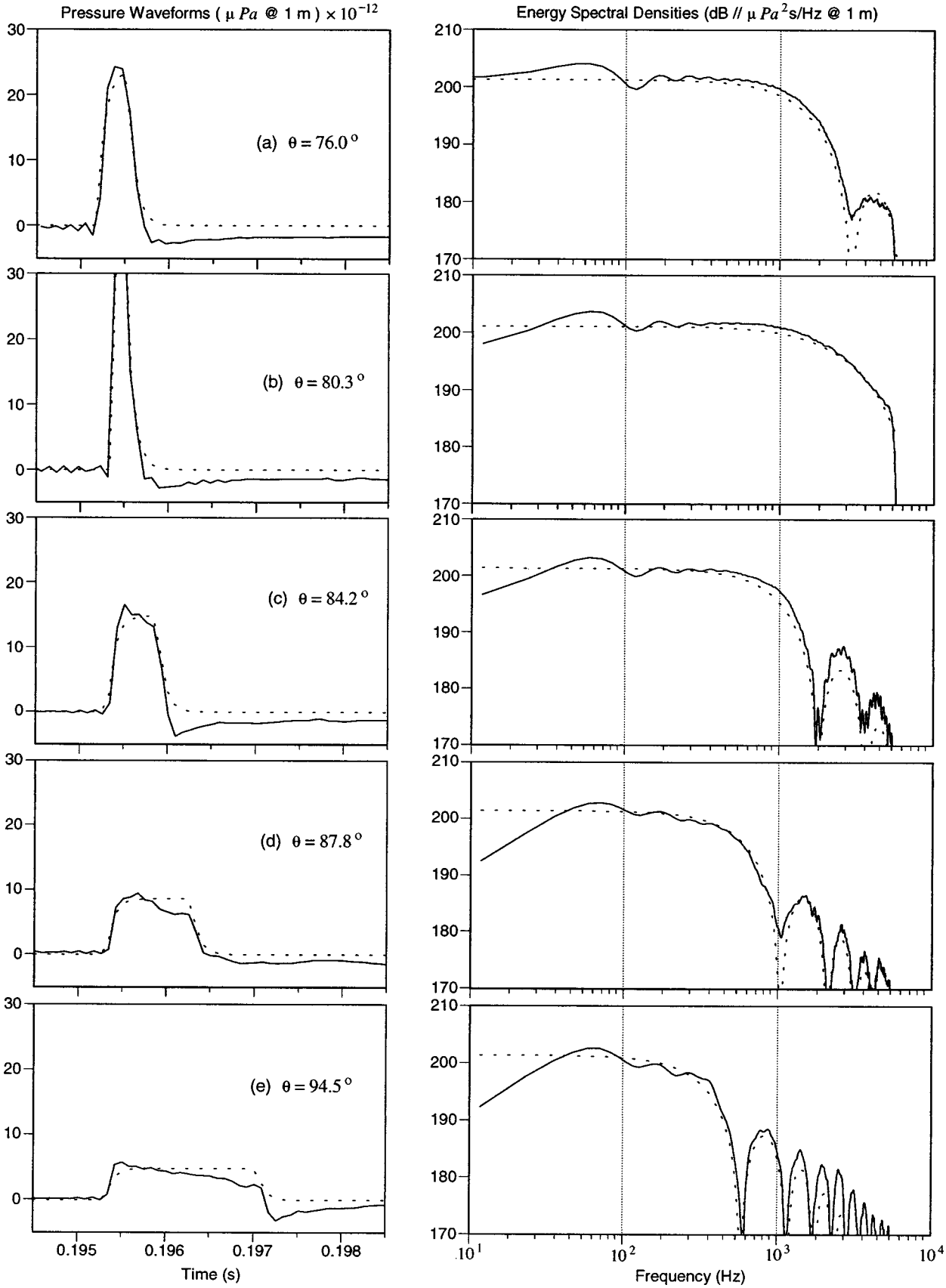


FIG. 13. Model results (dashed lines) compared to measured data (solid lines) for unit C8-2 at indicated view angles θ. Model parameters used are given in the text.

discrete charges if the length of each discrete bit is the distance the detonation travels in time τ. Finding that measured τ and I follow predicted trends confirms the hypothesis.

(5) Of the two parameters driving the model, impulse is easier to measure because it is substantially independent of view angle and, therefore, relatively unaffected by source curvature. The on-axis decay time constant is hard to measure on long cords because small amounts of curvature can increase apparent τ by a factor of 2 or more. The best model predictions result from using actual τ's.

ACKNOWLEDGMENTS

The author wishes to thank J. Barger for suggesting the mathematical method employed in Sec. II and, in particular, the idea of discretizing the line into bits of length $v\tau$. T. Farrell designed and wrote the data reduction procedures used in locating the hydrophones from marker pop arrival times. This work was sponsored by the Tactical Technology Office (TTO) of the Defense Advanced Research Projects Agency (DARPA).

APPENDIX: BEAM PATTERN CURVE FITTING

Assume the energy beam pattern of a directive source is the same as that of a continuous line,

$$\begin{aligned} \mathrm{ESL}(\theta) &= \mathrm{ESL}_0 + 20\log\left|\frac{\sin x}{x}\right|, \\ x &= \pi(L/\lambda)(\cos\theta - \cos\theta_0), \end{aligned} \tag{A1}$$

where θ_0 is the MRA direction, ESL_0 the maximum value on the peak of the pattern, and angles are measured from end-fire. Express L/λ in terms of the beamwidth,

$$\pi(L/\lambda) = 50.4\pi/\beta = 158/\beta, \tag{A2}$$

where β is the total beamwidth between the -3-dB points, in degrees. On the main lobe of the pattern (θ near θ_0) the x-dependent term can be expanded in a Taylor's series,

$$20\log\left|\frac{\sin x}{x}\right| = -1.45\left(1+\frac{x^2}{30}\right)x^2 + \cdots . \tag{A3}$$

Dropping the fourth power term in x, the peak of the pattern reduces to a parabola,

$$\mathrm{ESL}(\theta) = \mathrm{ESL}_0 - 1.45x^2, \tag{A4}$$

where

$$x = \frac{158}{\beta}(\cos\theta - \cos\theta_0). \tag{A5}$$

Pairs of (ESL,θ) near the MRA are least-squares fitted to this equation to obtain the three parameters, ESL_0, β, and θ_0, which best describe the peak. The nonlinear fit is done using a Mathcad solve block.

[1] R. H. Cole, *Underwater Explosions* (Princeton U. P., Princeton, NJ, 1948).
[2] A. B. Arons, ''Underwater explosion shock wave parameters at large distances from the charge,'' J. Acoust. Soc. Am. **26**, 343–346 (1954).
[3] D. E. Weston, ''Underwater explosions as acoustic sources,'' Proc. Phys. Soc. London **B76**, 233–249 (1960).
[4] M. Blaik and E. A. Christian, ''Near-surface measurements of deep explosions. I. Pressure pulses from small charges,'' J. Acoust. Soc. Am. **38**, 50–56 (1965).
[5] N. R. Chapman, ''Source levels of shallow explosive charges,'' J. Acoust. Soc. Am. **84**, 697–702 (1988).
[6] J. P. Slifko, ''Approximate methods of estimating energy spectra of line charges fired underwater,'' J. Acoust. Soc. Am. **41**, 182–188 (1967).
[7] R. A. Wentzell, R. H. Adlington, and J. C. Moldon, ''Depth dependence of bubble pulse periods of point and end-fired line charges,'' J. Acoust. Soc. Am. **48**, 1283–1286 (1970).
[8] J. W. Sari and D. E. Nelson, ''Shock and bubble pulse characteristics of long-line charges (*U*),'' *Proceedings of the 35th Navy Symposium on Underwater Acoustics, 1-259 to 1-274* (Office of Naval Research, Arlington, VA, 1983) (CONF).
[9] W. J. Marshall, ''Descriptors of impulsive signal levels commonly used in underwater acoustics,'' IEEE J. Ocean Eng. **20**, 108–110 (1996).
[10] *Handbook of Array Design Technology, Vol. I* (Applied Hydro-Acoustics Research, Inc., Rockville, MD, 1976).
[11] N. R. Chapman, ''Measurement of the waveform parameters of shallow explosive charges,'' J. Acoust. Soc. Am. **78**, 672–681 (1985).
[12] See Ref. 1, Sec. 4.4.
[13] See Ref. 1, Table 7.3.

Nonlinear acoustic method for bubble density measurements in water

A. M. Sutin,[a] S. W. Yoon, E. J. Kim, and I. N. Didenkulov[a]
Acoustics Research Laboratory, Department of Physics, Sung Kyun Kwan University, Suwon 440-746, Republic of Korea

(Received 11 January 1996; revised 29 January 1997; accepted 24 November 1997)

Nonlinear scattering due to bubble oscillations was theoretically and experimentally investigated for bubble size distribution measurements in water. If two primary acoustic waves of different frequencies are incident on a bubble, the difference frequency component of primary waves is produced in the scattered field. The incoherent scattering at the difference frequency from bubbles was theoretically studied. It was shown that the difference frequency amplitude is proportional to the density of bubbles having resonance frequencies close to the difference frequency. It allows the use of difference frequency measurements to determine the bubble density in water. For experimental investigations two focused transducers were used to increase the levels of primary waves in the interaction zone. One of the primary frequencies was kept constant 2.25 MHz while another was changed from 2.22 to 1.93 MHz, so that the observed difference frequency was varied from 30 to 320 kHz. This frequency band corresponds to the resonant bubble radii from 109 to 10 μm. For the bubble clouds produced in a laboratory tank by an electrolysis-type and a slit-type bubble maker the bubble densities were well estimated with the present nonlinear acoustic method. [S0001-4966(98)02603-4]

PACS numbers: 43.30.Pc, 43.30.Qd, 43.30.Lz [JHM]

INTRODUCTION

It is well known that the ocean surface layer is very inhomogeneous and full of entrapped air bubbles. In the ocean there are various bubble-generating mechanisms. The breaking of surface waves is one of the most likely generating mechanisms. The bubble clouds can be formed in the upper layer and be transported down to depths of tens of meters in the ocean by Langmuir circulation, turbulence, and so on.[1] Such bubble clouds influence the gas flux between atmosphere and ocean[2] and also play significant roles in sound propagation as well as sound generation.[3–6] Bubble clouds have been also experimentally confirmed as one of the most likely sources of underwater ambient noise in the ocean over wide frequency ranges.[7,8]

Measurements of bubble size and density in a sound propagating medium require understanding of the acoustic characteristics of the medium. Even the presence of a small number of bubbles makes enormous changes to acoustic properties in water. Since bubbles are good resonant scatterers and absorbers of sounds in the ocean, sound scattering and attenuation data are used in acoustic bubble sizing methods and for bubble cloud observations.[1,3,5]

For bubble diagnostics in the ocean nonlinear acoustic methods have been lately attracting increasing attention. Since a bubble is a highly nonlinear scatterer, the acoustic scattered fields from a bubble also show highly nonlinear properties. Such scattered fields from a bubble are easily observed at the second or higher harmonics as well as at the fundamental frequency. If two primary acoustic waves of different frequencies are incident on a bubble, the scattered fields at the sum and the difference frequencies of the primary waves also arise. These nonlinear acoustic responses make nonlinear acoustic diagnostics possible for bubble sizing in water. The usage of the nonlinear acoustic techniques easily allows one to distinguish a bubble from the other scatterers since nonlinear scattering from a bubble is much stronger than the other scatterers such as solid particles or any other inhomogeneities in fluid.

One of the early experimental investigations with a method of second harmonic measurements was an observation of gas emboli in divers' tissues during decompression.[9] A method based on second harmonic observation was also developed for estimating bubble distribution in the ocean[10,11] and tested for bubble detection in pipelines.[12] A difference frequency method was developed with dual primary transmitting waves.[13,14] For this difference frequency observation the primary frequencies are taken as those higher than bubble resonance frequencies. The difference frequency is chosen around bubble resonance. Another nonlinear acoustic technique is a modulation method based on modulation of a high-frequency probe wave by a low-frequency primary wave.[15,16] The modulation level increases if the primary frequency becomes close to the bubble resonance frequency. Another modulation technique is a method based on the modulation of a probe wave by the subharmonic frequency of primary wave.[17] In the above-mentioned experiments the nonlinear scattering from a single bubble was mainly studied.

When sound waves propagate in the ocean, acoustic scattering from many bubbles usually takes place. Such scattering consists of coherent and incoherent parts. The coherent component arises when there is spatial temporal correla-

[a]On leave from Institute of Applied Physics, Russian Academy of Sciences, Nizhny Novgorod 603600, Russia.

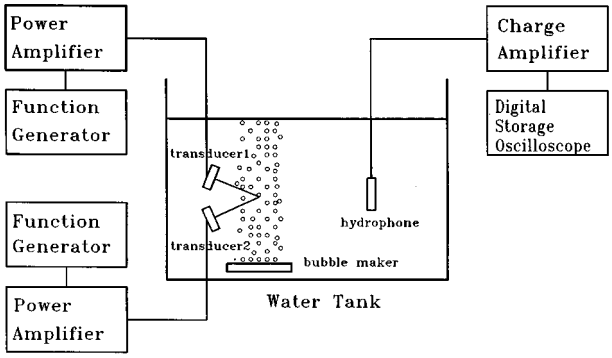


FIG. 1. Schematic diagram of experimental setup.

tion between scattered and incoming waves. It takes place, for example, for forward scattering when the coherent scattering yields distortion in sound velocity and sound attenuation. Measurements of these parameters are used in linear acoustic bubble sizing methods. The coherent technique based on sound velocity measurements was applied for total void fraction estimation,[3] and the sound attenuation measurement technique was used for bubble density determination.[18,19] The incoherent components in the scattered signals exist when phases of scattered signals from scatterers are stochastically distributed. Such a case usually occurs in a subsurface bubble layer in the ocean when the backscattering signal is investigated. The incoherent methods of bubble sizing[6,20,21] are much more widely used than the coherent methods.

The nonlinear scattering signal consists not only of coherent but also of incoherent components. Nonlinear coherent scattering can take place for the waves scattered in a forward direction or as a result of nonlinear reflection from the ocean surface layer. Such nonlinear coherent scattering from bubble layers has been observed in laboratory measurements.[22,23] Nonlinear incoherent scattering can take place in a backward direction. Such scattering was investigated when the bubble resonance occurred at the primary frequencies.[10,11,21] In Ref. 21 such nonlinear scattering was considered as a disturbance factor for conventional linear measurements of bubble cloud spatial distribution.

In this paper nonlinear incoherent scattering of the difference frequency for a wide function of bubble size distribution is theoretically investigated for the case that individual bubble resonance frequencies are within the variation range of difference frequencies. The parameters of nonlinear scattering are related to a bubble size distribution that permits measurement of the bubble density. Experimental verification of the present method was carried out in a laboratory facility.

I. EXPERIMENTAL SETUP

The experimental measurements were carried out in a laboratory tank of dimensions 0.30 m×0.91 m×0.45 m. A schematic diagram of the experimental setup is shown in Fig. 1. Anechoic lining rubber plates were placed in the tank opposite to the transmitters to reduce the reflection from tank walls.

Two different bubble makers were used in the measurements. An electrolysis-type bubble maker was composed of nets of wires for a positive and a negative electrode. A positive constant electric voltage about 20 V was applied at the positive electrode. Hydrogen bubbles were generated from the positive electrode to the water surface. The negative electrode was placed near the tank wall to avoid the influence of oxygen bubbles produced by itself in an acoustical field. Much higher bubble concentrations were created with a slit-type bubble maker. The slit-type bubble maker consisted of a porous plate and a compressed air supply system. Bubbles can be generated through the plate made from pressed cardboards with compressed air. The compressed air from an air tank was fed to the bubble maker through a flow control system. The size of the plate was 9 cm×14 cm. The flow rate Q of gas volume fed to the bubble maker was measured with a flow meter, and the bubble rising speed V_r was also determined by direct measurements of time and distance of bubble rising. The total void fraction β of the bubble layer produced by the bubble maker was estimated from the gas volume flow rate Q and the average speed V_r of rising bubbles as follows:

$$\beta = \frac{Q}{SV_r},$$

where S is the effective cross-sectional area for bubble generation of the slit-type bubble maker.

Two high-frequency primary beams were generated by two separate focused transducers (Panametrics type A304S) to make more interactions of the primary waves in the beam-crossing region. Each transducer has a resonance frequency around 2.25 MHz, a radius $a=12.7$ mm, and a focal length $F=101.6$ mm. Each transducer was calibrated with a reference source by the comparison method.[4] It was observed that the amplitude of the acoustic wave near the transmitter was about 6.6×10^4 Pa near a resonance frequency. For this pressure field measurement a tone burst signal was generated and its reflection signal from a water surface was used. The primary acoustic power radiated by each transducer was about 1.5 W. The radiation amplitude decreased when the frequency was off from the resonance. The size of the interaction zone can be approximated to the diameter of the focal spot. In our case this diameter d can be estimated by a formula[24]

$$d = 1.21\lambda F/a,$$

where λ is the wavelength of the pump wave. In the present case the diameter of the focal spot was about 6.4 mm. More details of the acoustic pressure distribution near the focal point are described in Sec. II.

One of the primary frequencies was kept constant at 2.25 MHz, and another was varied from 2.22 to 1.93 MHz to get difference frequencies from 30 to 320 kHz. Such difference frequency variation allowed measurements of bubble size distribution in the range from 109 to 10 μm of bubble radius. The two primary beams intersected in the region near their focal spots. The positions of both primary pump radiators were tuned with a spherical steel ball of 3 mm in diameter placed at the center of interaction zone. A tone burst

signal was radiated by one transducer as a transmitter and was received by another transducer as a receiver. The locations of both transducers were adjusted to reach the maximum amplitude of signals reflected from the steel ball at the intersection point. A bubble maker was placed under the beam-crossing interaction zone. As bubbles rose and passed through the interaction zone, the difference frequency wave was generated in the interaction zone. The signal was received by a hydrophone situated at approximately 130 mm from the focal point. It was amplified with a preamplifier (Panametrics model 5678) and analyzed by a digital storage oscilloscope (LeCroy 9310M).

The measurements of the difference frequency radiation signal were carried out by using continuous wave (cw) primary waves for different frequencies. Since the interaction zone of the primary waves was very small, bubbles randomly passed through the interaction zone and fluctuations of the received signal were observed. To get their mean value the averaging of the received signal intensities for a time period of 60 s at each difference frequency was used. Even though anechoic lining rubber plates were used, all the influence of tank resonance could not be avoided in the case of the electrolysis-type bubble maker, where the attenuation of the acoustic wave was not enough to suppress such reverberation due to tank resonance. Since the attenuation was rather strong for the slit-type bubble maker, the effects of reverberation in the tank were not observable in this case. To reduce tank resonance effects in the measurements the following two methods were applied to the data processing:

(1) The frequency dependence of received signals reveals tank resonance when the frequency is swept. To avoid such interference due to the tank resonance a frequency averaging method was used. In our experiments the levels of five different frequency signals near every chosen frequency were determined and these intensities were averaged. The step between the main frequency points was 10 kHz and a step of 1 kHz was used to measure the levels around the main point for averaging.
(2) Beside such an averaging method the tank resonance effect was also estimated by using a small spherical omnidirectional radiator of which volume was close to the volume of interaction zone. This radiator was centered in the interaction zone, so that it can be considered as a reference source for absolute measurements of the sound produced by bubbles in the interaction zone. The transfer function $K(\Omega)$ from a reference radiator to a hydrophone is defined as the ratio of the electric input voltage amplitude U_1 into the radiator to the electric received voltage amplitude U_2 from the hydrophone, i.e., as $K(\Omega)=U_1/U_2$ for an observed difference frequency.

In our experiments the electric output amplitude U of the hydrophone at a difference frequency produced by bubbles was measured. To recalculate that amplitude for the free field condition the transfer function $K(\Omega)$ was used. Since the transmitting sensitivity S of the radiator is known, the acoustic pressure P_Ω of the difference frequency radiated by bubbles in a free field at the distance 1 m from the interaction zone was estimated by

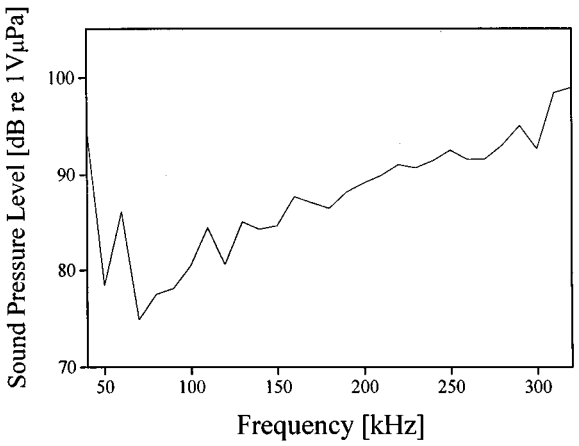


FIG. 2. Frequency dependence of the difference frequency field generated through the bubble layer from the slit-type bubble maker.

$$P_\Omega = S(\Omega)K(\Omega)U. \tag{1}$$

The frequency dependence of the received difference frequency signals is presented in Figs. 2 and 3 for bubble clouds from the respective bubble makers, respectively.

To calculate bubble density function from such measurements we also need to take account of the attenuation of the primary high-frequency waves in a bubbly medium. The attenuation coefficients for the primary pump waves were determined by the insertion loss measurements with a bubble layer. A hydrophone was placed behind the bubble layer as a receiver. The received amplitudes of primary frequency waves through the bubble layer were measured and compared with the amplitudes without the bubble layer. The attenuation coefficient α in dB per m can be estimated from such measurements as follows:

$$\alpha = \frac{\ln(U_0/U_b)}{l},$$

where U_0 is the received amplitude of the primary wave without the bubble layer, U_b is the received amplitude with the bubble layer, and l is the length of the acoustic path inside the bubble layer.

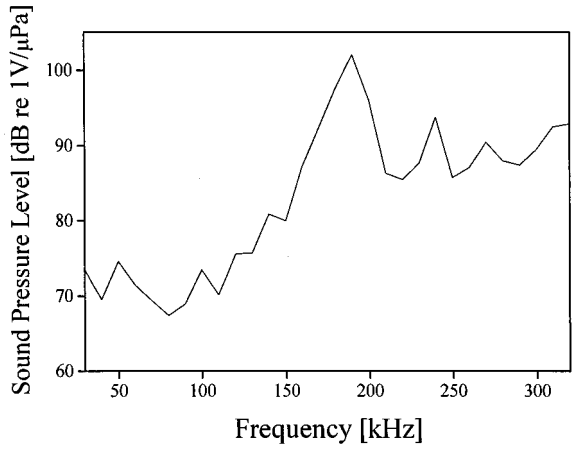


FIG. 3. Frequency dependence of the difference frequency field generated through the bubble layer from the electrolysis-type bubble maker.

II. THEORETICAL CONSIDERATION

For a theoretical description of the nonlinear incoherent scattering it is necessary to start from nonlinear scattering by a single bubble. The formula for a nonlinear scattered field at the combination frequencies can be derived by two typical approaches in the literature. One is the volume displacement approach[25] and another is the radial displacement approach.[12,15,26] In Refs. 26 and 27 both approaches were compared and their authors had different conclusions. In Ref. 26 it was concluded that the volume displacement approach was not correct, and in Ref. 27 it was pointed out the errors taken in the Ref. 26. The opinion of the authors of the present paper is that the approach of Zabolotskaya and Soluyan[25] is valid.

The derivation of the formulas will be given below following the concept of Zabolotskaya and Soluyan[25] but from more fundamental equations.

The first equation of movement, i.e., the Euler equation which has a form for spherically symmetric motion, is described as follows:

$$\frac{\partial U}{\partial t}+U\frac{\partial U}{\partial r}=-\frac{1}{\rho_o}\frac{\partial p}{\partial r}, \tag{2}$$

where U is the fluid velocity at the distance r from the center of a bubble, p is the pressure, and ρ_o is the ambient equilibrium fluid density.

The second equation is the linear wave equation in terms of a velocity potential φ:[28]

$$\nabla^2\varphi-\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2}=0, \tag{3}$$

where c is the sound velocity. The fluid velocity is given by the velocity potential as follows:

$$U=\nabla\varphi. \tag{4}$$

Since the nonlinear disturbances due to bubble oscillations are much higher than the distortion of acoustic waves in a pure fluid surrounding a bubble, the higher-order terms of acoustic nonlinearity of fluid can be neglected. The Rayleigh equation for bubble oscillations can be derived from Eqs. (2)–(4). The solution of the wave equation for radially symmetric motion is written in the form

$$\varphi=\frac{A(t-r/c)}{r}, \tag{5}$$

where A is an arbitrary function. Since only small bubbles are under consideration, that is, the radii R are much smaller than the acoustic wavelength λ, the parameter $R/\lambda\ll 1$ for the asymptotic expansion of the sound wave equation can be used and the following assumption is valid:

$$\frac{|(1/c^2)(\partial^2\varphi/\partial t^2)|}{|\nabla^2\varphi|}\sim\left(\frac{R}{\lambda}\right)^2\ll 1.$$

Therefore, the sound wave equation can be written in the following formula;

$$\nabla^2\varphi=0, \tag{6}$$

that is, the equation of continuity for an incompressible liquid in terms of the velocity potential. Equation (6) has a solution of

$$\varphi=\frac{A(t)}{r}, \tag{7}$$

where A is the arbitrary function of time.

The same result can be obtained directly from Eq. (5) for the distance $r\ll\lambda$. With the velocity potential φ for radial motion

$$U=\frac{\partial\varphi}{\partial r},$$

the Euler equation, Eq. (2), becomes

$$\frac{\partial^2\varphi}{\partial t\partial r}+\frac{1}{2}\frac{\partial}{\partial r}\left(\frac{\partial\varphi}{\partial r}\right)^2=-\frac{1}{\rho_o}\frac{\partial p}{\partial r}. \tag{8}$$

Let us integrate the last equation from r to r_1, where $R\ll r_1\ll\lambda$. It gives

$$\frac{\partial\varphi}{\partial t}+\frac{1}{2}\left(\frac{\partial\varphi}{\partial r}\right)^2=-\frac{p-P_o-p_a}{\rho_o}, \tag{9}$$

where p_a is the applied acoustic pressure produced by external sources, and P_o is the ambient equilibrium fluid pressure. It is assumed that at the distance $r_1\gg R$ the acoustical pressure disturbance produced by a bubble is small. The boundary conditions at the bubble surface are given as follows:

$$\dot{R}=\left.\frac{\partial\varphi}{\partial r}\right|_{r=R},\quad p_g=p|_{r=R}, \tag{10}$$

where p_g is the gas pressure inside the bubble. After substitution of Eq. (7) to the first boundary condition of Eq. (10), the velocity potential can be described as

$$\varphi=\frac{A}{r}=-\frac{R^2\dot{R}}{r}. \tag{11}$$

Substituting Eq. (11) in Eq. (9) one can obtain the well-known Rayleigh equation:

$$R\ddot{R}+\frac{3\dot{R}^2}{2}=\frac{p_g-P_o-p_a}{\rho_o}. \tag{12}$$

Since the surface tension effect is neglected, Eq. (12) might be good for the bubbles larger than several microns in radius. Neglecting the heat exchange between the bubble and the medium, the adiabatic gas law can be used as the equation of state:

$$p_g=P_o\left(\frac{V_o}{V_o+V}\right)^{\gamma}, \tag{13}$$

where V is the volume variation of single bubble, $\gamma=1.4$ is the specific heat ratio of the gas in a bubble, and P_o is the static pressure in a bubble. The subscript ''o'' denotes the quantity in a state of equilibrium. For small bubble oscillations, i.e., $V\ll V_o$, substituting Eq. (13) into Eq. (12) and keeping all the terms up to $(V/V_o)^2$ order, the Rayleigh equation, Eq. (12), becomes[25]

$$\ddot{V}+\omega_o^2 V+\nu\dot{V}-\frac{1}{8\pi R_o^3}\left[3(\gamma+1)\omega_o^2 V^2+2\dot{V}V+(\dot{V})^2\right]$$

$$=\frac{-4\pi R_o}{\rho_o}\,p_a, \tag{14}$$

where R_o is the equilibrium bubble radius and ω_o is the bubble resonance frequency known by the Minnaert formula as

$$\omega_o=\frac{1}{R_o}\sqrt{\frac{3\gamma P_o}{\rho_o}}. \tag{15}$$

The energy dissipation term is included phenomenologically in Eq. (14) and ν is the loss factor. For a time harmonic wave the loss factor ν can be written as $\nu=\omega/Q$ or as $\nu=\omega\delta$, where Q is the quality factor and δ is the bubble damping constant. The damping constant is made up of the damping constants due to reradiation or scatter, plus shear viscosity and thermal conductivity.[3,29] Equation (14) describing the forced bubble oscillations under acoustic pressure p_a up to the quadratic nonlinearity terms was first derived and investigated by Zabolotskaya and Soluyan in Ref. 25.

The substitution of Eq. (11) to Eq. (9) allows us to find the pressure around a bubble:

$$p(r)=P_o+p_a+\frac{\rho_o[R^2(t_r)\ddot{R}(t_r)+2\dot{R}^2(t_r)R(t_r)]}{r}$$

$$-\frac{\rho_o}{2}\frac{\dot{R}^2(t_r)R^4(t_r)}{r^4}, \tag{16}$$

where a retarded time $t_r=t-r/c$ is used instead of t to take into account that the pressure disturbances produced by a bubble propagate with a finite velocity, i.e., sound velocity. The pressure field consists of two parts: the far-field component which depends on r as $1/r$, and the near-field component which decreases with r as $1/r^4$. For the distance $r\gg R$ the near-field component becomes negligibly small, and the acoustic pressure scattered by a bubble, p_s, is defined by the third term:

$$p_s=\frac{\rho_o[R^2(t_r)\ddot{R}(t_r)+2\dot{R}^2(t_r)R(t_r)]}{r}=\frac{\rho_o\ddot{V}(t_r)}{4\pi r}, \tag{17}$$

where $V=(4\pi/3)(R^3-R_o^3)$ is the bubble volume variation. Equations (14) and (17) are the basic equations to determine the processes of nonlinear scattering by bubbles. These equations were the same as those in Ref. 25 and the derivation made in the present paper shows that the volume displacement approach is correct. It is interesting to consider why another solution has been obtained in the paper of Du and Wu.[26] The point of disagreement with the radial displacement approach[12,15,26] is in the form of pressure pulsation near a small bubble. Du and Wu used the acoustic pressure field in the form

$$p=\frac{B(t-r/c)}{r}, \tag{18}$$

but the correct formula, Eq. (16), contains also the near-field component, which is very essential for the nonlinear bubble oscillations, and the near-field component cannot be neglected.

Let us investigate the bubble radiation field at the difference frequency. In the case of biharmonic waves, which consist of two frequency components ω_1 and ω_2 with the pressure amplitudes P_1 and P_2, respectively, the pressure field of scattered wave at the difference frequency $\Omega=\omega_1-\omega_2$ can be determined by the bubble volume variation at this frequency. The acoustic radiation field is expressed with the bubble volume variation V_Ω as follows:

$$P_\Omega=\frac{\rho_o\Omega^2 V_\Omega}{4\pi r}, \tag{19}$$

where r is the distance from a bubble to an observation point. Relation between the volume variation at the difference frequency and the pressure amplitude in an incident wave can be obtained from Eq. (14). It leads to an easy solution for the bubble volume variation V_Ω by the perturbation method for a given pressure of an external wave and small additions due to nonlinearity (see Ref. 25):

$$V_\Omega=\frac{2\pi[3(\gamma+1)\omega_o^2-\omega_1^2-\omega_2^2+\omega_1\omega_2)]P_1P_2}{\rho_o^2R_o[\omega_o^2-\omega_1^2+i\delta\omega_1^2][\omega_o^2-\omega_2^2-i\delta\omega_2^2][\omega_o^2-\Omega^2+i\delta\Omega^2]}, \tag{20}$$

where $i=\sqrt{-1}$. Assume that there is no bubble with radius less than 1.4 μm, which has a resonance at the primary frequency around 2.3 MHz. It allows one to simplify Eq. (20) and for a very low difference frequency compared with the primaries, i.e., $\Omega\ll\omega_1,\omega_2$, the amplitude of the bubble volume variation at the difference frequency can be written as

$$V_\Omega=\frac{2\pi P_1P_2}{\rho_o^2R_o\omega_1^2[\omega_o^2-\Omega^2+i\delta\Omega^2]}. \tag{21}$$

It is common procedure to describe the scattering strength by an effective scattering cross section. Let us introduce the nonlinear scattering cross section σ_n, which is defined as the ratio of the total intensity of nonlinearly scattered waves to that of the incident wave in the free space:

$$P_\Omega^2=\sigma_n P_1P_2/4\pi r^2. \tag{22}$$

From Eqs. (19), (21), and (22) the nonlinear scattering cross section σ_n at the difference frequency can be expressed as

$$\sigma_n=\frac{\pi\Omega^4P_1P_2}{\rho_o^2R_o^2\omega_1^4[(\omega_o^2-\Omega^2)^2+\delta^2\Omega^4]}. \tag{23}$$

The nonlinear scattering cross section depends on the amplitudes of incident waves, while the linear scattering cross sec-

tion does not. It seems noteworthy that the above relations were obtained with the perturbation method assuming the amplitude of nonlinear scattered signal is much smaller than that of the linear one. Equations (22) and (23) give the information on the difference frequency signal produced by a single bubble. In our experiments the time-averaged intensity of received signals was used as described in Sec. I.

Bubbles randomly pass through the interaction zone. Since the interaction zone is small, the probability that more than one bubble exists in this zone at the same moment is also small. The time-averaged signal from the small size interaction zone can be considered as a model for the larger size interaction zone where many bubbles can exist at the same moment. Such a situation can take place in ocean bubble measurements with transducers of larger sizes than those in the laboratory experiments. Let us consider a small volume element dv in the interaction zone. In the processing of the received signal the time-averaged level of the power spectrum was measured, that is, the time-averaged intensity of the signal produced by bubbles. The mean intensity can be found by using the function of bubble size distribution $n(r)$. This function determines the probability W that the bubble having radii from R to $R+dR$ is located in a volume dv as $W=n(r)dRdv$. Thus, the averaged signal produced by a volume element dv is

$$dP_{\Omega}^2=\frac{S_n P_1 P_2\, dv}{4\pi r^2}, \tag{24}$$

where

$$S_n=\int_{R_{\min}}^{R_{\max}}\sigma_n(r)n(r)dR. \tag{25}$$

Here S_n is the nonlinear scattering coefficient, which is equal to the effective nonlinear scattering cross section per unit volume of a medium. A similar linear backscattering coefficient is well known in underwater acoustics.[3,4] The nonlinear backscattering coefficient can be found in the same manner as the linear one by integrating the scattering cross section of an individual bubble over the radius with the bubble density function. Equation (25) is mainly contributed by the resonant bubbles and can be calculated with the same approximation method, which is usually used for determining the absorption coefficient and the linear scattering coefficient:[3]

$$S_n=\frac{\pi^2 n(R_\Omega)P_1P_2}{2\rho_o^2\delta\omega_1^4 R_\Omega}, \tag{26}$$

where R_Ω is the bubble radius having the resonance frequency of Ω. Equations (24) and (26) give the information on the acoustical signal produced by the bubbles passing through the volume element dv. Any bubbles passing through the interaction zone Δv make a contribution to the received signal. For this contribution Eq. (25) should be integrated through the beam interaction zone, where the nonlinear scattering from bubbles originates. When the dimension of the interaction volume element Δv is much smaller than the distance r from the interaction zone to the hydrophone, the nonlinear scattering field can be described as

$$P_\Omega^2=\frac{\pi n(R_\Omega)}{8\rho_o^2\delta\omega_1^4 R_\Omega r^2}\int_{\Delta v}P_1^2P_2^2\,dv. \tag{27}$$

Equation (27) relates the intensity of the difference frequency field with the concentration of bubbles in the scattering volume which are resonant at the difference frequency. This equation allows the measurements of bubble density and size distribution for various difference frequencies. The intensity distributions of the primary beams in the scattering volume are considered to be known.

In a laboratory experiment the focused transducers were used as the primary beam transmitters. For such a case the integration in Eq. (27) can be analytically solved by considering the acoustic pressure distribution near the focal point of the transmitter for the primary pressure fields.[24] The transmitter axis is chosen with the x axis, and then the distance from the transmitter axis is given as $\sqrt{y^2+z^2}$. For this transmitter whose focal distance F is much larger than its radius a, i.e., $F\gg a$, it is possible to neglect the pressure variation along the x axis for the distance from the focal point much less than $\lambda F^2/a^2$, where λ is the wavelength of the primary wave. In the interaction zone the primary pressure field is described as the following formula:

$$P_1(x,y,z)=\frac{A_1\pi a^2 J_1(2X)}{\lambda F X}\,e^{-\alpha l}, \tag{28}$$

where A_1 is the acoustic pressure near the transmitter, α is the attenuation coefficient for the primary wave, l is the length that the primary wave passes inside of the bubble cloud, and

$$X=\frac{\pi a\sqrt{y^2+z^2}}{\lambda F}. \tag{29}$$

If two primary beams are perpendicular to each other, we can choose the y axis for the second transmitter axis and the field near focal point P_2 is determined by a similar equation with A_2 the acoustical pressure amplitude near the second transducer. With the primary beams of Eq. (28), Eq. (27) for the difference frequency can be easily calculated:

$$P_\Omega^2=\frac{\pi^2 a^5\chi A_1^2A_2^2 n(R_\Omega)}{8\rho_o^2\lambda F\delta\omega_1^4R_\Omega r^2}\,e^{-4\alpha l}, \tag{30}$$

where

$$\chi=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\frac{J_1(2\pi\sqrt{X^2+Z^2})}{\pi\sqrt{X^2+Z^2}}\right]^2\left[\frac{J_1(2\pi\sqrt{Y^2+Z^2})}{\pi\sqrt{Y^2+Z^2}}\right]^2 \times dX\,dY\,dZ\approx 0.11. \tag{31}$$

It is seen that the nonlinear scattering field is proportional to the bubble density function with the bubble resonance radius at the difference frequency. From Eqs. (1) and (30) the bubble density function $n(R_\Omega)$ can be described as follows:

$$n(R_\Omega)=\frac{8\rho_o^2K^2S^2U^2\delta\omega_1^4R_\Omega\lambda F e^{4\alpha l}}{\pi^2a^5\chi A_1^2A_2^2}. \tag{32}$$

This makes it possible to determine bubble density from direct measurements.

Let us derive the conditions for Eqs. (32) and (27) to be valid. It is evident that Eq. (32) is applicable for the case that the probability of more than one resonant bubble existing in the interaction zone is small. Such a condition can be written in the form:

$$n(r)\,\delta R\Delta v \ll 1. \tag{33}$$

In the present paper Eq. (32) can be applicable even for higher bubble density when Eq. (33) is not satisfied and many bubbles can be situated in the scattering volume. Since the bubbles in a liquid move due to floating force and turbulence, the scattered signals fluctuate with time. Therefore, the averaged values of such an ensemble scattering only have physical meanings. There are two components in the resulting signal after averaging. One is the coherent component that appears when the averaged signal amplitude does not become zero. It can be if there is spatial–temporal correlation in bubble distribution such as bubble layers. Forward scattering also gives the coherent component. Another is the incoherent part that has a meaning of pure stochastic scattering. The intensity of incoherent field is equal to the total intensity of the fields scattered by all the bubbles. The incoherent scattering field is described by Eqs. (27) and (32). The incoherent scattering from many bubbles in the interaction zone takes place if the angle between two incident beams is not small because the scale of the phase variation is about the primary pump wavelength. In other words, the phases of nonlinear scattered signals depend on the phases of primary waves at the scatterer positions. Different bubbles which stay in the interaction zone not closer from each other than the wavelength of pump waves have very different phases in the scattered field at the difference frequency. It means that the distance between bubbles must be more than the wavelength of pump waves and the condition of applicability of difference frequency incoherent scattering becomes smoother:

$$n(r)\,\delta R\lambda^3 \ll 1. \tag{34}$$

If this condition is not satisfied in a system of scatterers, it is necessary to take into account the processes of multiple scattering[30–33] and mutual interference.[34] However, the multiple scattering effect is essential mainly for the coherent scattering and very high bubble concentrations with void fraction of order 10^{-2}.[33] Such a high void fraction can hardly exist in the ocean. It is a reason why the processes of multiple scattering are not taken into account in the linear backscattering theory.[3,4] The interference effect was considered in Ref. 34 only for the case of planar distribution of scatterers.

III. RESULTS OF MEASUREMENTS

The experimentally determined bubble densities are presented in Fig. 4 for the bubble clouds produced by the slit-type bubble maker and in Fig. 5 for the bubble clouds produced by the electrolysis-type bubble maker. The slit-type bubble maker can produce a much higher bubble concentration than the electrolysis-type bubble maker. The peak bubble radius in the bubble size distribution for the electrolysis-type bubble maker was near 20 μm. For the bubble clouds produced by the slit-type bubble maker the bubble density increases with bubble radius. Bubble densities presented in Figs. 4 and 5 were calculated by using Eq. (32). Equation (32) includes the parameters of the transducers such as the focal distance F, the radius a of the transducer, the pressure A_1, A_2 at the transducer surface, the frequencies $\omega_{1,2}$ of the primary beams, the transfer function $K(\Omega)$, and the transmitting sensitivity $S(\Omega)$ of the reference radiator at the difference frequency. In Eq. (32) R_Ω is the radius of a bubble having a resonance at frequency Ω defined by the Minnaert formula, Eq. (15). Here $U=U(\Omega)$ is an experimentally measured output voltage at the hydrophone at the difference frequency Ω. The total gas void fraction β can be easily calculated from the bubble density function:

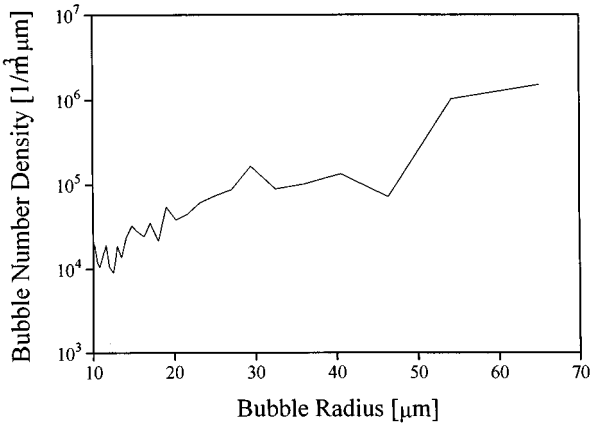


FIG. 4. Bubble density as a function of bubble radius for the bubble layer of the void fraction 7.2×10^{-3} produced by the slit-type bubble maker.

$$\beta = \int_{R_{\min}}^{R_{\max}} \frac{4\pi R^3 n(r)}{3}\, dR. \tag{35}$$

For numerical calculations $\delta \approx 0.1$ was used for the damping constant.[29] The estimated void fraction for the bubble clouds produced by the slit-type bubble maker was $\beta \approx 7.2\times10^{-3}$. It agreed very well with the directly measured void fraction $\beta \approx 7.1\times10^{-3}$. For the bubble clouds produced by the electrolysis-type bubble maker the void fraction was estimated as $\beta \approx 6.0\times10^{-5}$.

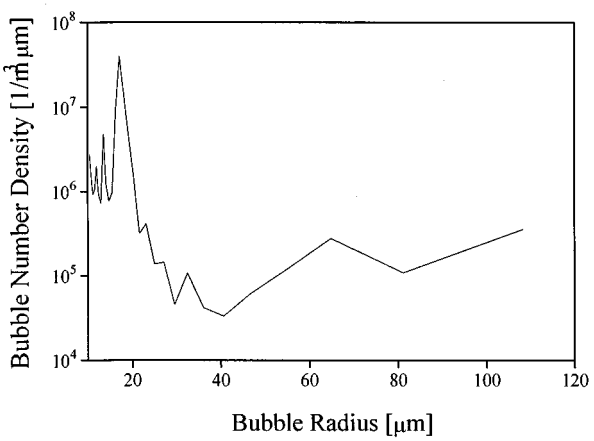


FIG. 5. Bubble density as a function of bubble radius for the bubble layer of the void fraction 6.0×10^{-5} produced by the electrolysis-type bubble maker.

IV. CONCLUSION

In this paper a nonlinear acoustic bubble sizing method based on the incoherent scattering of the difference frequency by bubbles from the intersecting zone of two high-frequency primary acoustic beams was theoretically investigated. This nonlinear acoustic method allows the measurements of the very low bubble density case, which is almost impossible to be measured by a linear acoustic method. The present nonlinear method was verified in the laboratory measurements for bubbles produced by two types of bubble makers. This nonlinear method seems very attractive to the ocean application because it provides high selectivity of bubbles from the other types of ocean scatterers which do not have such large nonlinear response as bubbles. It also provides a rather wide frequency band by relatively small changes of the frequency for one of the primary beams.

ACKNOWLEDGMENTS

The authors appreciate an anonymous referee's comments which helped to greatly clarify this paper. This work was supported by the Korea–Russia collaboration program through the STEPI, Republic of Korea. One of the authors (SWY) was supported in part by the Korea Science and Engineering Foundation. Two of them (AMS and IND) were supported in part by International Science Foundation and Russian Foundation of Fundamental Research.

[1] S. A. Thorpe, ''On the clouds of bubbles formed by breaking waves in deep water and their role in air-sea gas transfer,'' Philos. Trans. R. Soc. London, Ser. A **304**, 155–210 (1982).
[2] D. K. Woolf and S. A. Thorpe, ''Bubbles and the air-sea exchange of gases in near saturation conditions,'' J. Mater. Chem. **49**, 435–466 (1991).
[3] C. S. Clay and H. Medwin, *Acoustical Oceanography: Principles and Applications* (Wiley, New York, 1977).
[4] R. J. Urick, *Principles of Underwater Sound* (McGraw–Hill, New York, 1983), 3rd ed.
[5] D. M. Farmer and S. Vagle, ''Waveguide propagation of ambient sound the ocean-surface bubble layer,'' J. Acoust. Soc. Am. **86**, 1897–1908 (1989).
[6] S. O. McConnell, ''Acoustic measurements of bubble densities at 15–50 kHz,'' in *Natural Physical Sources of Underwater Sound*, edited by B. R. Kerman (Kluwer Academic, Dordrecht, 1993), pp. 237–252.
[7] N. Q. Lu, A. Prosperetti, and S. W. Yoon, ''Underwater noise emission from bubble clouds,'' IEEE J. Ocean Eng. **15**, 275–281 (1990).
[8] S. W. Yoon, L. A. Crum, A. Prosperetti, and N. Q. Lu, ''An investigation of the collective oscillations of a bubble cloud,'' J. Acoust. Soc. Am. **89**, 700–706 (1991).
[9] F. N. Fenlon and J. W. Wohn, ''On the Amplification of Modulated Acoustic Waves in Gas-Liquid Mixtures,'' in *Cavitation and Inhomogeneities in Underwater Acoustics*, edited by W. Lauterborn (Springer-Verlag, Heidelberg, 1980), pp. 141–150.
[10] L. A. Ostrovsky and A. M. Sutin, ''Nonlinear acoustic diagnostics of discrete inhomogeneities in liquids and solids,'' in *Proceedings of the 11th International Congress on Acoustics* (Paris, 1983), Vol. 2, pp. 137–140.
[11] L. A. Ostrovsky and A. M. Sutin, ''Nonlinear sound scattering from subsurface bubble layer,'' in *Natural Physical Sources of Underwater Sound*, edited by B. R. Kerman (Kluwer Academic, Dordrecht, 1993), pp. 363–370.
[12] D. L. Miller, ''Ultrasonic detection of resonance cavitation bubbles in a flow tube by their second-harmonic emissions,'' Ultrasonics **19**, 217–224 (1981).
[13] G. Gimenez, M. Chamant, and J. P. Farnand, ''Non-linear response of a single bubble driven by a two-components exciting wave,'' in *Proc. 10th Int. Symp. on Nonlinear Acoustics*, Kobe, Japan (1984), pp. 83–87.
[14] B. M. Sandler, D. A. Selivanovsky, and A. Y. Sokolov, ''New results of bubble concentration with radii from 6 to 20 m at sea,'' Sov. Phys. Tech. Phys. **27**, 1038–1039 (1982).
[15] V. L. Newhouse and P. M. Shankar, ''Bubble size measurement using the nonlinear mixing of two frequencies,'' J. Acoust. Soc. Am. **75**, 1473–1477 (1984).
[16] J. Y. Chapelon, P. M. Shankar, and V. L. Newhouse, ''Ultrasonic measurement of bubble cloud size profiles,'' J. Acoust. Soc. Am. **78**, 196–201 (1985).
[17] A. D. Phelps and T. G. Leighton, ''Investigations into the use of two frequency excitation to accurately determine bubble sizes,'' in *Bubble Dynamics and Interface Phenomena*, edited by J. R. Blake *et al.* (Kluwer Academic, Dordrecht, 1994), pp. 475–484.
[18] H. Medwin, ''Counting bubbles acoustically. Review,'' Ultrasonics **15**, 7–13 (1977).
[19] H. Medwin, ''*In situ* acoustic measurements of microbubbles at sea,'' J. Geophys. Res. **82**, 971–976 (1977).
[20] I. R. Schippers, ''Density of air-bubbles below the sea surface, theory and experiments,'' in *Cavitation and Inhomogeneities in Underwater Acoustics*, edited by Lauterborn (Springer-Verlag, Heidelberg, 1980), pp. 205–210.
[21] M. Gensane, ''Bubble population measurements with a parametric array,'' J. Acoust. Soc. Am. **95**, 3183–3190 (1994).
[22] D. M. Donskoi, S. V. Zamolin, L. M. Kustov, and A. M. Sutin, ''Nonlinear backscattering of acoustic waves in a bubble layer,'' Acoust. Lett. **7**, 134–135 (1984).
[23] L. M. Kustov, V. E. Nazarov, and A. M. Sutin, ''Nonlinear sound scattering by a bubble layer,'' Sov. Phys. Acoust. **30**, 500–503 (1984).
[24] G. S. Kino, *Acoustic Waves: Devices, Imaging, and Analog Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1987), p. 185.
[25] E. A. Zabolotskaya and S. I. Soluyan, ''Emission of harmonic and combination frequency waves by air bubbles,'' Sov. Phys. Acoust. **18**, 396–398 (1972).
[26] G. Du and J. Wu, ''Comparison between two approaches for solving nonlinear radiations from a bubble in a liquid,'' J. Acoust. Soc. Am. **87**, 1965–1967 (1990).
[27] Yu. A. Ilinskii and E. A. Zabolotskaya, ''Cooperative radiation and scattering of acoustic waves by gas bubbles in liquids,'' J. Acoust. Soc. Am. **92**, 2837–2841 (1992).
[28] L. Landau and E. Lifshits, *Hydrodynamics* (Nauka, Moscow, 1986), 3rd ed.
[29] C. Devin, Jr., ''Survey of thermal radiation and viscous damping of pulsation air bubbles in water,'' J. Acoust. Soc. Am. **31**, 1654–1657 (1959).
[30] L. O. Foldy, ''The multiple scattering of waves,'' Phys. Rev. B **67**, 107–119 (1945).
[31] A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, New York, 1978).
[32] I. Tolstoy, ''Superresonant system of scatterers I,'' J. Acoust. Soc. Am. **80**, 282–294 (1986).
[33] Z. Ye and L. Ding, ''Acoustic dispersion and attenuation relation in bubbly mixture,'' J. Acoust. Soc. Am. **93**, 1629–1636 (1995).
[34] Z. Sun and G. Gimenez, ''Evaluation of interference effect in the energy investigation of echoes scattered by an uncorrelated planar distribution of spherical targets,'' J. Acoust. Soc. Am. **92**, 3259–3270 (1992).

Radiation impedance of resonant ultrasound spectroscopy modes in fused silica

Hong Zhang, R. S. Sorbello, Carsten Hucho, Joseph Herro, Jeffrey R. Feller, D. E. Beck, and Moises Levy
Physics Department, University of Wisconsin—Milwaukee, Milwaukee, Wisconsin 53201

D. Isaak, J. D. Carnes, and O. Anderson
Institute of Geophysics and Planetary Physics, University of California at Los Angeles, Los Angeles, California 90095-1567

(Received 9 May 1997; accepted for publication 21 January 1998)

The pressure dependence of the resonance frequency of several resonant ultrasound spectroscopy modes in a sample of fused silica has been measured at UCLA in atmospheres of air, helium, and argon near ambient temperature. For both compressional and torsional modes, the radiation resistance is linearly dependent upon pressure and increases with the molecular mass of the surrounding gas. The effects are larger for breathing modes than for torsional modes. They also increase with the molecular mass of the gas. A radiation impedance model is presented which explains some of these data qualitatively and quantitatively. © *1998 Acoustical Society of America.* [S0001-4966(98)00505-0]

PACS numbers: 43.35.Cg [HEB]

INTRODUCTION

The adiabatic elastic moduli of solids, and the pressure and temperature dependencies of these moduli, are important for a wide range of problems relevant to solid state physics and geophysics. The pressure dependencies of the elastic moduli are related to the third-order elastic constants and are important in developing theoretical models of solid state properties since they provide information about the interatomic potentials.[1] We emphasize that information about the third-order elastic constants reveals much about the fundamental physics of a solid since the third-order constants are obtained by taking appropriate derivatives of the free energy.[2,3] Indeed, some theoretical attempts to predict solid phase transformations and/or solid–liquid phase changes involve consideration of large strains, i.e., a few percent, in the crystalline lattice.[4–9] Consideration of strains of this magnitude means that anharmonic effects, i.e., nonlinearity in the elastic constants, must be understood. The third-order elastic constants are also required for accurate equations of state used in the investigation of solid material under pressure.[1] Accurate equations of state at low pressure are critical when making extrapolations to the megabar range as is done in studies of planetary interiors. Of prime importance in these equations of state is a particularly simple third-order elastic constant, the pressure dependence of the adiabatic bulk modulus.

There is a long, fruitful history of using ultrasonic techniques in measuring elastic properties and their pressure and temperature dependencies. Two methods in particular have been used in these ultrasonic experimental studies on single crystals. The first is the pulse-echo method, and various versions of it, all based on principles of interferometry. The second is resonant ultrasound spectroscopy (RUS). The RUS method has been used extensively in the study of the temperature dependencies of elastic properties up to temperatures of 1500–1800 K.[10–12]

Virtually all ultrasonic experimental work on the pressure dependencies of elasticity, i.e., on some of the third-order elastic constants, has been done using interferometry. Although interferometry is a powerful experimental tool, there remain questions regarding the effects of bonds between the specimen and the transducers. One advantage of the RUS technique is that RUS requires no bonds between the transducers and the specimen. Even though advances during the past few years have been realized using interferometry techniques[13,14] the effect of bonds required to couple the acoustic energy to the specimen continues to be problematic. In describing their state-of-the-art GHz-ultrasonic interferometer apparatus, Spetzler *et al.*[14] provide a lengthy discussion of theoretical and experimental approaches to problems originating with the bonds in high precision ultrasonic experiments. They state, ''...conclusive results on the nature and the effect of the bonds are still lacking.'' In view of the experimental uncertainties noted above, it is appropriate to develop alternative ultrasonic methods by which the third-order elastic properties of materials are determined experimentally.

We propose using the RUS technique to determine the third-order elastic constant of solids. This will be done by measuring the pressure dependence of the resonance peaks of appropriately shaped samples. Isotropic pressure is exerted by a surrounding gas of helium, air, or argon. Different gases are used in order to experimentally determine the effect of the molecular mass of the gas on the resonance peaks, exclusive of the intrinsic effect of the pressure on the resonance peaks via the third-order elastic constants. Once experimental data were obtained on a fused silica sample in the shape of a rectangular parallelepiped, it became obvious that a theoretical model had to be developed in order to account for the effect of radiation impedance on the resonance frequency

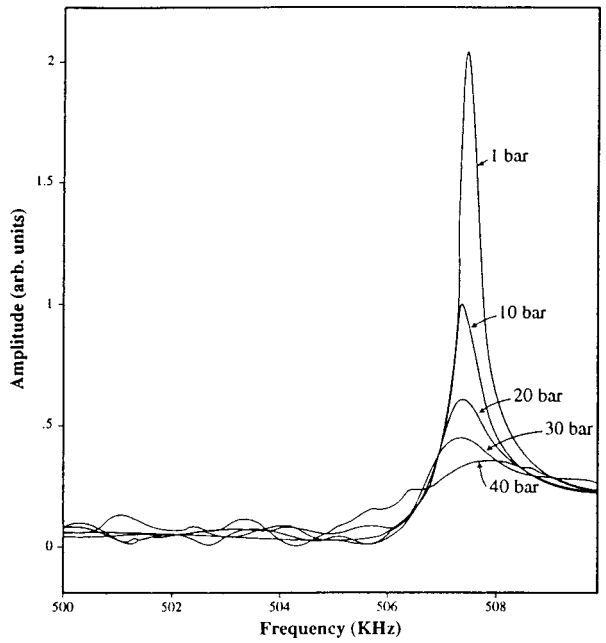


FIG. 1. Resonant peaks of torsional mode A_{1u} measured in fused silica in air at a temperature of 24 °C at various pressures.

of the sample. However, since the resonance frequency could be affected by both the third-order elastic constants and by the mass loading due to the surrounding gas, it was imperative to check the accuracy of the theoretical model by comparing the theoretical results with another experimentally obtained quantity which was negligibly dependent upon the pressure dependence of the intrinsic properties of the sample. Such a quantity is the quality factor of the resonance peaks, Q. It is assumed that the changes in Q being produced by the changes in the applied pressure are principally produced by the energy radiated into the surrounding gas, the intrinsic change in attenuation in the sample itself produced by the compression being negligible.

The RUS measurements reported here were performed on a rectangular crystal of fused silica. The resonant lines of two vibrating modes, identified as torsional mode (A_{1u}) and compressional mode (A_{1g}), were monitored when the sample was in three different surrounding gases, helium, air, and argon, and the pressure of the surrounding gas was changed from 1 bar up to 100 bar. For both modes, it was found that the quality factor of the resonance, Q_r, where the subscript r refers to energy loss due to radiation impedance, decreased when the pressure increased, and decreased with increasing molecular mass of the surrounding gas. For the torsional mode in all three gases, the resonant line was shifted to lower frequency when the pressure was increased, while for the compressional mode, it was shifted to lower frequency in helium but higher in argon.

In our theoretical analysis, we use a two-dimensional model to solve the resonance problem for a bar, and we calculate the sound radiation impedance for each face of the rectangular parallelepiped sample assuming a response similar to that for a vibrating piston mounted in an infinite baffle. The two-dimensional model is applicable because the sound wavelength in the surrounding gas is much smaller than the

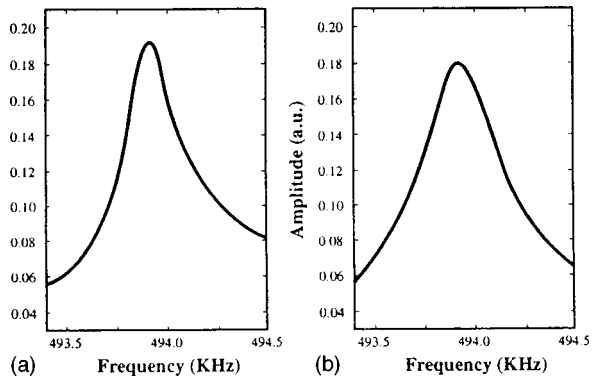


FIG. 2. Lorentzian resonant peaks obtained before (a) and after a background subtraction (b).

dimensions of the rectangular faces of the crystal, and, therefore, the directivity of the sound beams is mostly in the forward direction.

I. EXPERIMENTAL RESULTS

The resonance peaks of the torsional A_{1u} and the compressional A_{1g} modes of a fused silica crystal with dimensions of $0.3194\times0.3083\times0.2491$ cm^3 and a mass of 0.0545 g were measured under various pressures of surrounding gases of helium, air, or argon. A typical result for the A_{1u} mode in air at room temperature, 23.8 °C, is given in Fig. 1. It can be seen that, as the pressure is increased from 0 to 40 bar, the resonant peak becomes broader and shifts to lower frequencies at higher pressures. Since the resonant peaks do not always display a symmetric Lorentzian shape due to cross talk between the two transducers used in the experiment, a background needs to be subtracted in order to determine the resonant center frequency and the quality factor accurately. The original data resonant peak has the form

$$\left|\frac{a_1}{a_2^2-x^2-ia_3x}+a_4\right|^2,$$

where x is the frequency and the coefficients a_1, a_2, and a_3 are real. The coefficient a_4 is complex and arises from coherent background noise, including electromagnetic and mechanical pickup. After removal of the background term a_4, one obtains the Lorentzian-like peak of the form

$$\frac{a_1^2}{(a_2^2-x^2)^2+a_3^2x^2},$$

where a_1 is the amplitude, a_2 is the center frequency of the peak, and a_3 is the width of the peak. The computer code used to analyze the data was obtained from A. Migliori.[15] Figure 2 shows one of the resonant peaks before and after the background subtraction.[16]

It was assumed that the total quality factor Q_t obtained after the background subtraction was composed of two loss mechanisms, one due to radiation impedance Q_r and the other due to the sample support and electronic losses Q_0. These latter losses are assumed to remain constant as a function of pressure. Therefore,

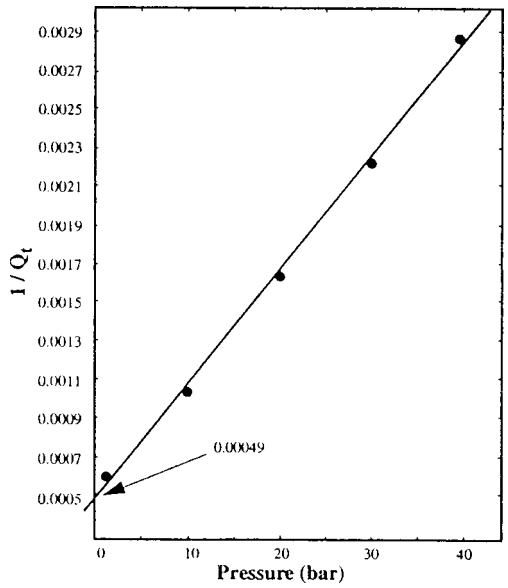


FIG. 3. Inverse total quality factor Q_t versus pressure for fused silica in argon (torsional mode).

$$\frac{1}{Q_t}=\frac{1}{Q_r}+\frac{1}{Q_0},$$

where Q_t is the Q factor measured from the corrected resonant peak using the background subtraction described above (using a_3), and $1/Q_0$ is a constant (y intercept) obtained by extrapolating the $1/Q_t$ versus pressure curve to zero pressure, as shown in Fig. 3, which plots $1/Q_t$ versus pressure for the torsional mode in argon. In Fig. 4, we have plotted the resulting values for $1/Q_r$ as a function of pressure for the torsional mode in helium, air, and argon. Similar values for the compressional mode in helium and argon are shown in Fig. 5. It is clearly seen that for both modes in the various gases, $1/Q_r$ is proportional to the pressure of the surrounding gases. In Fig. 6, we plot $1/Q_r$ at 1 atm as a function of molecular mass M of the surrounding gas on a log–log scale for both modes, and find that the slope is proportional to M^{γ}, with $\gamma=0.9$ for the torsional mode and $\gamma=0.5$ for the compressional mode.

The frequency of the resonant peak (a_2) obtained from the Lorentzian-like peak after the background subtraction

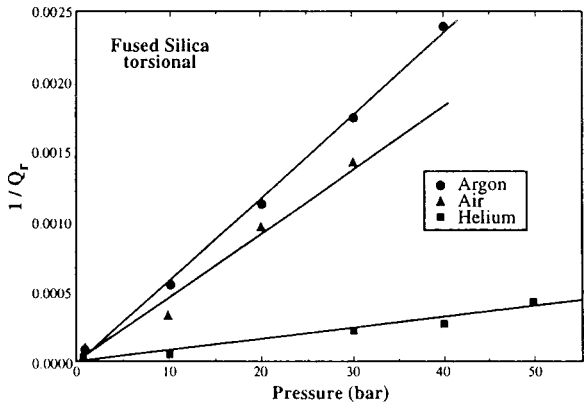


FIG. 4. Inverse radiation quality factor Q_r versus pressure for fused silica in helium, air, and argon (torsional mode).

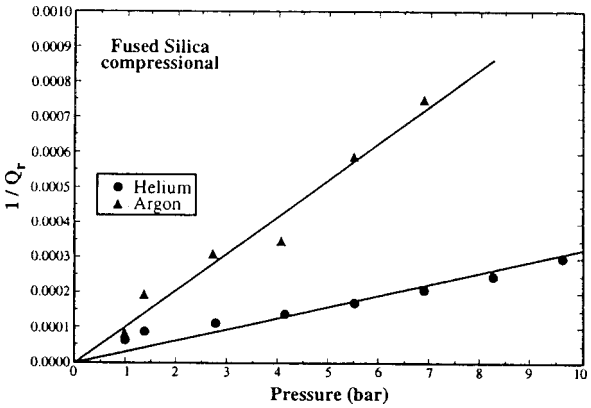


FIG. 5. Inverse radiation quality factor Q_r versus pressure for fused silica in helium and argon (compressional mode).

was plotted as a function of pressure in Fig. 7 for the torsional mode in helium, air, and argon, and in Fig. 8 for the compressional mode in helium and argon. The resonant peak shifts to lower frequencies at higher pressures in all cases except for the compressional mode in argon. Such a shift to lower frequencies would be expected as a result of the third-order elastic constants. However, before final conclusions can be drawn about the value of the third-order elastic constants, we need to estimate the mass loading effect due to the acoustic radiation reactance from the surrounding gas in order to determine its contribution to the frequency shift.

In the following two sections, we will carry out theoretical calculations of radiation impedance for a rectangular parallelepiped sample in two resonant modes, and compare the calculated Q_r factor and resonant frequency shift to the above experimental results.

II. THEORETICAL CALCULATION OF RADIATION IMPEDANCE

A. Two-dimensional model

Consider a solid bar of length l with a linear density of ρ_l, and with its two ends exposed to surrounding media, as

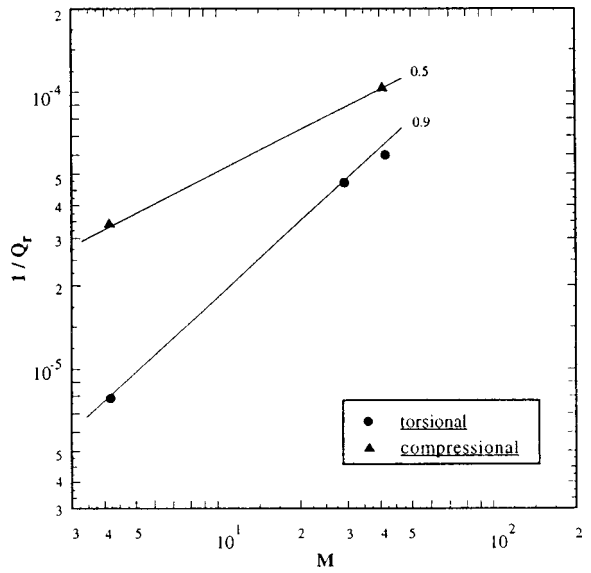


FIG. 6. Inverse radiation quality factor Q_r versus molecular mass of the surrounding gases for fused silica (torsional and compressional modes).

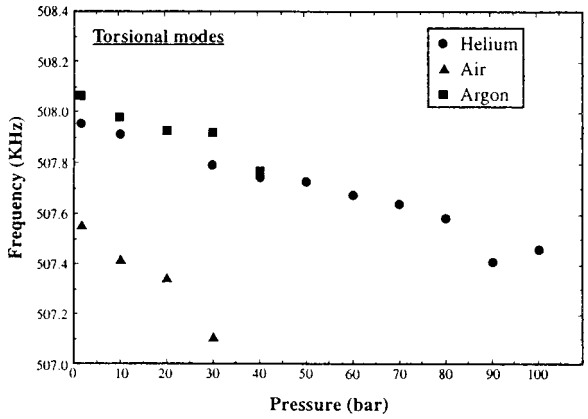


FIG. 7. Resonant frequency versus pressure for fused silica in helium, air, and argon (torsional mode).

shown in Fig. 9. The radiation impedances through the media at each end of the bar are Z_{m0} and Z_{ml}, respectively, and the sound velocity in the solid medium is v. The resonant frequencies of this bar can be obtained by solving the following transcendental equation:[17]

$$\tan Kl = \frac{i(Z_{m0}/\rho_l v + Z_{ml}/\rho_l v)}{1+(Z_{m0}/\rho_l v)(Z_{ml}/\rho_l v)}, \tag{1}$$

where $K=\omega/v$, with ω being the angular frequency of vibration.

For our case, the media at both ends are the same, and we may therefore take

$$Z_{m0}=Z_{ml}=Z_m=R_m+iX_m,$$

which allows Eq. (1) to be expressed as

$$\tan Kl = \frac{2iZ_m/\rho_l v}{1+(Z_m/\rho_l v)^2}. \tag{2}$$

It can be assumed that for our samples and the surrounding gases used, the following conditions hold: $R_m \ll \rho_l v$ and $X_m \ll \rho_l v$. Therefore, Eq. (2) simplifies to the following form:

$$\tan Kl = 2i\,\frac{Z_m}{\rho_l v}. \tag{3}$$

Now the remaining problem is to find Z_m for the three pairs of opposite faces of our rectangular parallelepiped sample.

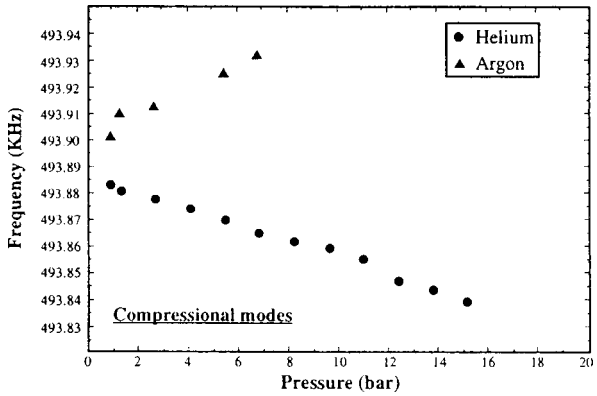


FIG. 8. Resonant frequency versus pressure for fused silica in helium and argon (compressional mode).

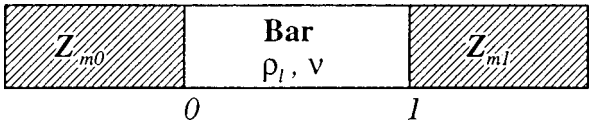


FIG. 9. A bar with a linear density of ρ_l, a length of l, sound velocity v, and radiation impedances of Z_{m0} and Z_{ml} at each end, respectively.

B. Radiation impedance of a vibrating piston

Before determining the radiation impedance of a rectangular piston with the surface moving uniformly or as a sine or cosine function, we shall review the procedure for evaluating the radiation impedance of a circular piston, moving uniformly in an infinite baffle.

1. Circular piston with uniform motion

When a piston is vibrating, it will set up acoustic pressure waves in a surrounding medium which will lead to reaction forces exerted by the medium on the driving surface (the piston itself). This causes an additional impedance for the vibration of the piston. In calculating the reaction force, let us consider an infinitesimal area dS of the surface of the piston as shown in Fig. 10. The increment in pressure that the motion of dS produces in the medium at a point adjacent to some other element of area of the piston, dS', can be written as[18]

$$dp = (i\rho ck/2\pi r)U_0\, dS\, \exp(i(\omega t - kr)), \tag{4}$$

where ρ is the density of the medium, c is the sound velocity in the medium, $k=\omega/c$ is the wave vector of the acoustic wave in the medium, r is the distance along the surface of the piston between dS and dS', U_0 is the amplitude of the piston velocity, and ω is the angular frequency of the vibration. The total acoustic pressure p in the medium adjacent to dS' can then be obtained by integrating expression (4) over the surface of the piston:

$$p \int\int \frac{i\rho ck}{2\pi r}\, U_0\, \exp(i(\omega t - kr))\, dS. \tag{5}$$

Now the total reaction force acting on the piston due to the medium can be obtained by integrating p over the piston:

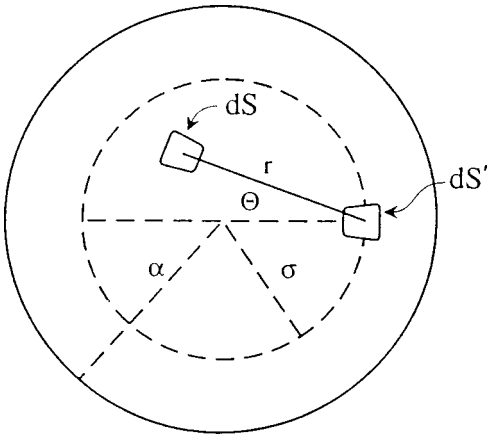


FIG. 10. A circular piston mounted on an infinite baffle with uniform motion in the z direction.

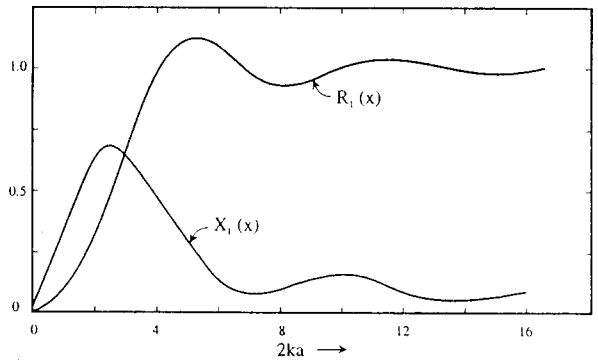


FIG. 11. Radiation resistance function $R_1(2ka)$ and radiation reactance function $X_1(2ka)$ for a cylindrical piston undergoing uniform motion.

$$f_m = -\int\int p\, dS' = -\frac{i\rho ck}{2\pi} U_0 \exp(i\omega t) \cdot \int\int dS' \int\int \frac{\exp(-ikr)}{r}\, dS. \tag{6}$$

Since the reaction force acting on the element dS' due to the motion of dS is the same as the force acting on dS due to the motion of dS', the result of the double integration in Eq. (6) will be twice as much as what would be obtained if the limits of integration were chosen to include the force due to each pair of elements only once. The choice of integration limits shown in Fig. 10 will considerably simplify the problem and therefore the resulting integral needs to be multiplied by a factor of 2:

$$f_m = -\frac{i\rho ck}{\pi} U_0 \exp(i\omega t) \int_0^a \sigma\, d\sigma \int_0^{2\pi} d\psi \int_{-\pi/2}^{\pi/2} d\theta \times \int_0^{2\sigma\cos\theta} \exp(-ikr)\, dr, \tag{7}$$

where a is the radius of the piston. The above integration results in

$$f_m = -\rho c\pi a^2 U_0 \exp(i\omega t)[R_1(2ka) + iX_1(2ka)], \tag{8}$$

where $R_1(x)$ and $X_1(x)$ are piston impedance functions defined as

$$R_1(x) = \frac{x^2}{2\cdot 4} - \frac{x^4}{2\cdot 4^2\cdot 6} + \frac{x^6}{2\cdot 4^2\cdot 6^2\cdot 8} - \cdots, \tag{9}$$

$$X_1(x) = \frac{4}{\pi}\left(\frac{x}{3} - \frac{x^3}{3^2\cdot 5} + \frac{x^5}{3^2\cdot 5^2\cdot 7} - \cdots\right). \tag{10}$$

A common way of defining the radiation impedance of the piston is to set it equal to the ratio of the force exerted by the piston on the medium to the velocity of the piston. When the force and velocity vary over the piston, however, the impedance should be evaluated from the (complex) power divided by the square of the velocity amplitude averaged over the piston area. For the special case of a uniformly moving piston, this approach yields the same impedance as the simple force-to-velocity ratio since the velocity can be factored out of the integrals. We shall first use the simple approach in what follows, and then use the power to evaluate Z_m for the square piston. Thus, for the uniformly moving cylindrical piston in an infinite baffle, we write

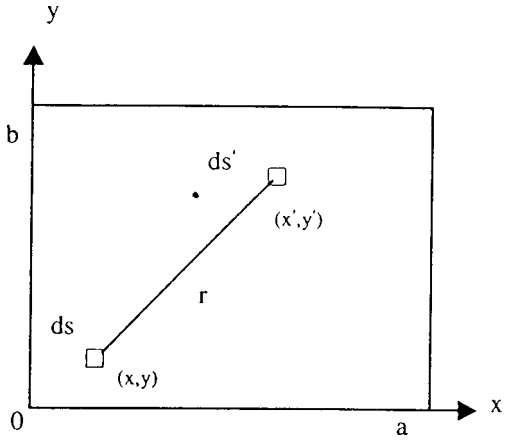


FIG. 12. A rectangular piston with dimensions of $a\times b$, and its surface vibrating in various modes in the z direction.

$$Z_m = \frac{-f_m}{U_0 \exp(i\omega t)} = \rho c\pi a^2[R_1(2ka) + iX_1(2ka)]. \tag{11}$$

The functions $R_1(2ka)$ and $X_1(2ka)$ are plotted in Fig. 11. The real part of the radiation impedance Z_m is called the radiation resistance R_m and the imaginary part the radiation reactance X_m. Therefore, for a circular piston mounted in an infinite baffle, we have

$$R_m = \rho c\pi a^2 R_1(2ka) \tag{12}$$

and

$$X_m = \rho c\pi a^2 X_1(2ka). \tag{13}$$

Notice that these are just equal to the piston impedance functions multiplied by the area of the pistons πa^2 and the acoustic impedance of the medium ρc.

Although the material that was presented in this section is well known, it was presented in order to acquaint the reader with the mathematical devices used to solve the circular baffle problem and which will be used for the rectangular piston in the following section. Other authors have used somewhat different approaches.[19,20]

2. Rectangular piston with motion in different modes

The radiation of sound from a vibrating rectangular piston has been addressed by several researchers.[21–25] However, the authors are not aware of any work which explicitly calculates the complex impedance for the nonuniform motion of a rectangular piston. Let us consider a rectangular piston which is vibrating in a particular mode in the z direction, as shown in Fig. 12. The velocity in the z direction of the area element dS located at position (x,y) is denoted by $U(x,y)\exp(i\omega t)$, and the distance between the element dS and dS' is

$$r = \sqrt{(x-x')^2 + (y-y')^2}. \tag{14}$$

Similar to the case of the circular piston, the increment in pressure that the motion of dS produces in the medium at dS' can be written as

$$dp = \frac{i\rho ck}{2\pi} U(x,y) \frac{\exp(i(\omega t - kr))}{r} dS. \tag{15}$$

Since the motion of the piston in the z direction is no longer uniform, the total pressure exerted by the medium at position (x',y') due to the vibration of the whole piston is

$$p(x',y') = \int\int \frac{i\rho ck}{2\pi} U(x,y) \frac{\exp(i(\omega t - kr))}{r} dS, \tag{16}$$

where, in contrast to Eq. (5), the velocity is not uniform but is allowed to vary with position.

As mentioned in the paragraph preceding Eq. (11), instead of calculating the total force exerted by the medium on the piston as in the case of the circular piston in uniform motion, one should consider the total power radiated by the vibrating rectangular piston in order to obtain the radiation impedance. The complex power radiated is given by the product of the complex pressure amplitude times the velocity amplitude, i.e.,

$$P = -\int\int U(x',y') p(x',y') dS', \tag{17}$$

where it is understood that the factor $\exp(i\omega t)$ is now dropped in $p(x',y')$. Substituting Eq. (16) into Eq. (17), one obtains

$$P = -\frac{i\rho ck}{2\pi} \int\int U(x',y') dS' \int\int U(x,y) \times \frac{\exp(-ikr)}{r} dS. \tag{18}$$

In evaluating the integral in Eq. (18), one needs the velocity of the area element on the piston located at (x,y) for various modes. For simplicity, here we consider only two modes: the torsional mode A_{1u} and the compressional mode A_{1g}, as illustrated in Fig. 13. For comparison, the integral is also evaluated for uniform motion of the piston. It can be seen from Fig. 13 that, for the torsional mode,

$$U(x,y) = U_0 \sin(\alpha x)\sin(\beta y) \tag{19}$$

is a good approximation[26] for the velocity of area element dS at (x,y), where U_0 is a constant and α and β are wave vectors in the piston with $\alpha = \pi/a$ and $\beta = \pi/b$. Similarly, for the compressional mode,

$$U(x,y) = U_0 \cos(\alpha x)\cos(\beta y); \tag{20}$$

and for uniform motion,

$$U(x,y) = U_0. \tag{21}$$

In the Cartesian coordinates shown in Fig. 12, the integral to be evaluated for a rectangular piston with dimensions of a and b $(a > b)$ has the following form:

$$I(k) = -ik \int_{-a/2}^{a/2} dx' \int_{-b/2}^{b/2} dy' \int_{-a/2}^{a/2} dx \int_{-b/2}^{b/2} dy\, U(x',y') U(x,y) \frac{\exp(-ik\sqrt{(x-x')^2+(y-y')^2})}{\sqrt{(x-x')^2+(y-y')^2}}. \tag{22}$$

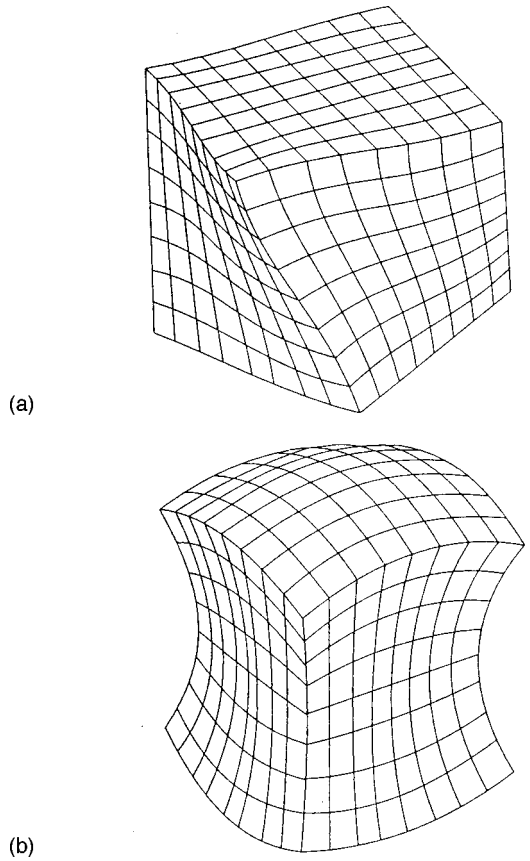


FIG. 13. Resonances of a cubic sample in (a) torsional mode A_{1u} and (b) compressional mode A_{1g}.

The analytical evaluation of the integral in Eq. (22) could be rather difficult, especially for the two nonuniform modes. In the following, we will show how to evaluate this integral numerically. It should be noticed that the integrand diverges at the origin of the Cartesian coordinates, and it is quite computer-time consuming to numerically carry out this fourfold integral. Therefore, a change of variables was undertaken in order to get around the divergence and to permit the analytical integrations which reduce the integral from fourfold to twofold. First, it can be shown that by letting

$$z = x \pm x', \quad z' = y \pm y',$$

and introducing

$$f(|z|,|z'|) = \frac{\exp(-ik\sqrt{z^2+z'^2})}{\sqrt{z^2+z'^2}},$$

we can rearrange the integral over x and x' and integrate analytically to obtain

$$\int_{-a/2}^{a/2} dx' \int_{-a/2}^{a/2} dx\; U(x,x') f(|x-x'|,|z'|) = U_0 \int_0^a dz\; f(|z|,|z'|) g_\alpha(a,z),$$

where

$$g_\alpha(a,z)=(a-z)\left(\cos(\alpha z)\mp\frac{\sin(\alpha(a-z))}{\alpha(a-z)}\right),$$

with the $-$ sign taken for the torsional mode, and the $+$ sign taken for the compressional mode. [For the uniform mode, $g_\alpha(a,z)\equiv 1$.] Thus, the fourfold integral in Eq. (22) is replaced by

$$I(k)=-ikU_0^2\int_0^b dz'\int_0^a dz\; g_\alpha(a,z)g_\beta(b,z')f(|z|,|z'|), \tag{23}$$

which is a twofold integral and requires much less computing time for numerical evaluation.

The straightforward numerical integration over z and z' in the above equation is frustrated by the $\rho=\sqrt{z^2+z'^2}$ factor in the denominator of the integrand which causes the integrand to diverge at the origin. For the circular piston this factor is cancelled by the same factor in the cleverly chosen differential surface element. We can take advantage of the same clever procedure by employing circular coordinates ρ and θ such that $z=\rho\cos\theta$ and $z'=\rho\sin\theta$. Using these variables, we can cast Eq. (23) into the form

$$\begin{aligned}I(k)=-ikU_0^2\Bigg(&\int_0^{\theta_0}d\theta\int_0^{a/(\cos\theta)}d\rho\\&\times\exp(-i\rho k)g_\alpha(a,\rho\cos\theta)g_\beta(b,\rho\sin\theta)\\&+\int_0^{\pi/2-\theta_0}d\phi\int_0^{b/(\cos\phi)}d\rho\\&\times\exp(-i\rho k)g_\alpha(a,\rho\sin\phi)g_\beta(b,\rho\cos\phi)\Bigg),\end{aligned} \tag{24}$$

where $\theta_0=\tan^{-1}(b/a)\leqslant\pi/4$ and $\phi=\pi/2-\theta$.

It is convenient to express $I(k)$ in terms of dimensionless impedance functions $R(ka)$ and $X(ka)$, as follows:

$$I(k)=2\pi CU_0^2ab(R(ka)+iX(ka)), \tag{25}$$

where we choose $C=1$ for the uniform mode [Eq. (21)] and $C=\frac{1}{4}$ for the sine and cosine modes [Eqs. (19) and (20)]. Although we have written R and X as explicit functions of ka, they also depend implicitly upon the aspect ratio a/b. The functions $R(ka)$ and $X(ka)$ for one face of the sample are shown in Figs. 14 and 15. Note that the introduction of the coefficient C in Eq. (25) results in $R(ka)$ having an asymptotic value equal to 1 for all modes.

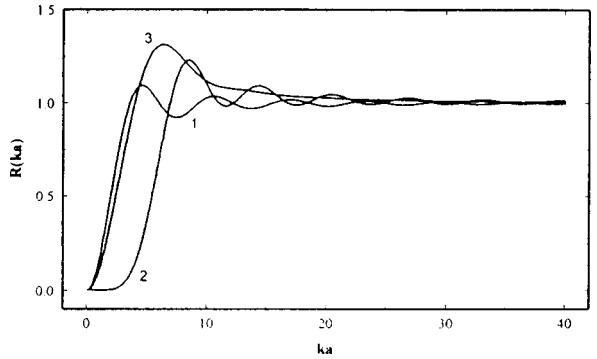


FIG. 14. Radiation resistance function $R(ka)$ for (1) $U(x,y)=U_0$, (2) $U(x,y)=U_0\sin(\alpha x)\sin(\beta y)$, and (3) $U(x,y)=U_0\cos(\alpha x)\cos(\beta y)$. As seen in Eqs. (30) and (31), $R(ka)$ is related to the energy radiated into the surrounding gas.

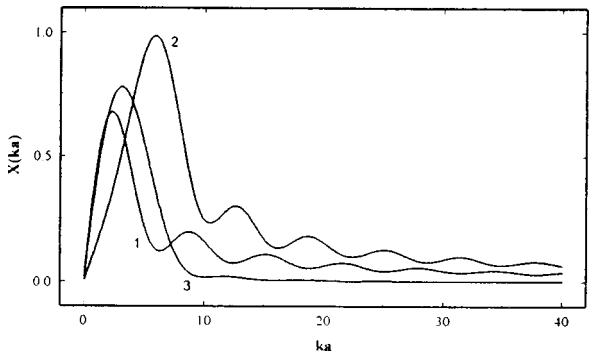


FIG. 15. Radiation reactance function $X(ka)$ for (1) $U(x,y)=U_0$, (2) $U(x,y)=U_0\sin(\alpha x)\sin(\beta y)$, and (3) $U(x,y)=U_0\cos(\alpha x)\cos(\beta y)$. As seen in Eq. (33), $X(ka)$ is related to the frequency shift due to mass loading by the surrounding gas.

The radiation impedance, Z_m, is related to $R(ka)$ and $X(ka)$ by the expression

$$Z_m=\rho cab(R(ka)+iX(ka)), \tag{26}$$

which follows directly from the prescription that the impedance is the complex power P divided by the square of the velocity amplitude $U(x,y)$ averaged over the area of the piston. This prescription is equivalent to calculating the impedance as the ratio F/U_0, where F is the amplitude of the force in mode pattern (α,β). For example, in the case of torsional modes, if we expand the pressure pattern as

$$p(x,y)=\sum_{n,m}p_{nm}\sin\left(\frac{n\pi x}{a}\right)\sin\left(\frac{m\pi y}{b}\right), \tag{27}$$

then $F=-p_{11}ab$ for the mode of interest in which $\alpha=\pi/a$ and $\beta=\pi/b$. From Eqs. (17) and (27) it is then easy to show that $P/(\int U(x,y)^2\,dxdy/ab)=F/U_0$. Note that Eq. (26) is similar to Eq. (11) for the case of the circular piston, and that the radiation impedance is proportional to ρc times the area of the piston, regardless of the mode of vibration.

The results shown in Figs. 14 and 15 are somewhat unexpected. In Fig. 14 the asymptotic value of $R(ka)$ for all three cases is 1. If we were to plot the values obtained for the integral $I(k)$ from Eq. (24), the asymptotic values would be four times larger for the uniform mode than for the sine and cosine modes. However, we are not interested in these values *per se*, but in the values of Z_m that are obtained by dividing the complex power by the square of the velocity averaged over the area of the piston. The complex power is equal to $\rho cI(k)/2\pi$, but the velocity average gives U_0^2 for the uniform mode and $\frac{1}{4}U_0^2$ for the sine and cosine modes. Thus, after the division, one obtains the same asymptotic value for all three cases.

There are features in the individual curves which are also interesting. For instance, we have the following:

(1) The R values for the sine curves are negligibly small up to values of $ka \simeq 2$, while the other curves increase as $(ka)^2$.
(2) The R and X values for both the uniform and sine modes exhibit gradually decaying oscillations for ka values larger than 15, while the cosine mode shows no perceptible oscillations past this value of ka.
(3) The period of the oscillation appears to be $ka = 2\pi$.

The first feature is related to the fact that the sine mode is antisymmetric and therefore the four quarters of the rectangular faces are effectively vibrating as a quadrupole source, which cancels the power radiation for large wavelengths or small values of ka, apparently up to $ka \simeq 2$. The last two features are interrelated. The piston faces are assumed to be vibrating in a semi-infinite baffle. Only the cosine mode does not exhibit a discontinuity in velocity and displacement at its boundaries with respect to the semi-infinite baffle. It is discontinuities that usually provoke oscillations in response functions. Therefore, the smooth transition of the cosine mode eliminates oscillations in the Fourier transform into frequency space. The integrals that are being performed are effectively Fourier transforms of the displacement functions (weighted by a $1/r$ factor). Following this line of reasoning, then, one may say that the oscillations in X and R for the sine and uniform modes are associated with the Fourier transform of the displacement of the faces and the discontinuities at the boundaries. The period of the modulation then should go as 2π, similar to the results for a cylindrical piston, which exhibit Bessel function dependence for large values of ka. Notice also that the piston impedance functions shown in Fig. 11 have the same period with respect to kd, where $2a = d$, the diameter of the piston.

The oscillations observed in Figs. 11, 14, and 15 may also be understood from an argument based on Babinet's principle. According to this principle, the pressure field above a vibrating piston that is contained in an infinite baffle is equal to the pressure field due to an entire vibrating plane minus the pressure field due to the (fictitious) vibrations of the infinite planar region consisting of only the baffle, with the piston regarded as stationary. Now, the pressure field from the entire vibrating plane equals $\rho c U(x,y)$ at the piston. This contribution by itself results in values of R and X appropriate to the asymptotic regime, namely $R(ka) = 1$ and $X(ka) = 0$. Departures from these values must therefore arise from the pressure field due to the (fictitious) vibrating infinite baffle. However, this pressure field in the region of the piston that is well away from the baffle is small because the distance to the baffle is much greater than the wavelength in the gas, and hence there is a great deal of cancellation due to the $r^{-1} \exp(-ikr)$ factor modulating the various contributions. [See, for example, Eq. (15).] On the other hand, the vibrating baffle creates a substantial pressure field near the edge of the piston because the $r^{-1} \exp(-ikr)$ factor will not effectively cancel-out contributions at positions within a few wavelengths, $2\pi/k$, from the edge. It follows that the significant departures from the asymptotic $R(ka)$ and $X(ka)$ values arise predominantly from regions within a few wavelengths from the edge of the piston. These contributions are relatively large when the piston (and hence, the fictitious baffle) has a relatively large displacement at the edge. This effect is further magnified because the $R(ka)$ and $X(ka)$ are proportional to the real and imaginary parts of the complex power, respectively, and the latter is found by integrating the product of pressure and velocity over the piston area. It follows that modes for which the displacement field is relatively large at the edge of the piston will have relatively large oscillations in $R(ka)$ and $X(ka)$. This is in qualitative agreement with the patterns observed in Figs. 11, 14, and 15.

In the next section we will use our calculated impedances for each face of the vibrating rectangular parallelepiped sample. In effect we are treating the six faces of the sample as independently vibrating pistons. This simple approach is justified because the wave vector α on the faces of the parallelepiped is about an order of magnitude smaller than the wave vector k in the gas. That is, the wavelength of the sound radiated in the gas is about an order of magnitude smaller than the dimensions of the faces of the parallelepiped since the sound velocity in the gas is about an order of magnitude smaller than in fused silica. Thus, we may think of the faces as independently vibrating pistons because their directivity pattern is mostly in the forward direction in the $ka \gg 1$ regime, which is the regime appropriate to the experimental measurements.

III. COMPARISON WITH EXPERIMENTAL RESULTS

In order to compare the above theoretical results to the experimental data, one needs to calculate the attenuation coefficient α, or the quality factor Q_r due to the radiation resistance, and the mass loading Δm, or the frequency shift Δf due to the radiation reactance. We start with the transcendental equation (3) obtained earlier in the two-dimensional model:

$$\tan Kl = 2i(Z_m/\rho_l v),$$

which, for $|Kl - n\pi| \ll 1$, becomes

$$\tan Kl \simeq Kl - n\pi = 2i(Z_m/\rho_l v),$$

and therefore

$$Kl = 2i(Z_m/\rho_l v) + n\pi. \tag{28}$$

Since K and Z_m are complex, one may let

$$K = K + i\alpha,$$

and with

$$Z_m = R_m + iX_m$$

Eq. (28) becomes

$$Kl + i\alpha l = -2X_m/\rho_l v + n\pi + 2iR_m/\rho_l v. \tag{29}$$

From the imaginary part of Eq. (29), one obtains

$$\alpha = 2R_m/\rho_l l v = 2R_m/M_B v,$$

where M_B is the mass of the bar. Since α is the attenuation coefficient per unit length, we find that the attenuation coefficient per unit time is given by

$$\alpha' = 2R_m/M_B.$$

The quality factor Q_r can now be calculated as the ratio of the angular frequency ω over $2\alpha'$, which gives

$$Q_r = \omega M_B / 4R_m.$$

Now, from Eq. (26) we have for the rectangular piston,

$$R_m = \rho c a b R(ka)$$

and (30)

$$X_m = \rho c a b X(ka).$$

Therefore,

$$Q_r = \frac{\omega M_B}{4R_m} = \frac{\omega M_B}{4\rho c a b R(ka)}, \tag{31}$$

and, since $R(ka)$ is approximately constant for values of $ka > 10$ (this is the range of values for ka in the experimental data presented in Sec. I), we see that

$$1/Q_r \propto \rho c, \tag{32}$$

where ρ is the density of the medium and c is the sound velocity in the medium. From the above relation, it is easy to see that the pressure dependence of the Q factor is

$$1/Q_r \propto p$$

since ρ is proportional to the pressure p of the surrounding gas, and c is independent of p. This is in agreement with the data shown in Figs. 4 and 5. In addition, one also finds that

$$1/Q_r \propto M^{1/2}$$

as ρ is proportional to the molecular mass of the surrounding gas M, and c is proportional to $\sqrt{1/M}$. This result is in agreement with the data shown for the compressional mode in Fig. 6.

From the real part of Eq. (29), one may calculate the mass loading due to the radiation reactance X_m as follows. The real part of Eq. (29) gives

$$Kl = -2X_m/\rho_l v + n\pi,$$

which results in a shift in the wave vector ΔK:

$$\Delta K l = -2X_m/\rho_l v.$$

Thus, we have

$$\Delta K = -\frac{2X_m}{\rho_l l v} = -\frac{2X_m}{M_B v} = -\frac{2X_m K}{M_B \omega}$$

or

$$\frac{\Delta K}{K} = -\frac{2X_m}{M_B \omega} = -\frac{2\Delta m}{M_B},$$

where $\Delta m = X_m/\omega$ is the mass loading. From the relation

TABLE I. Q_r in various surrounding gases at 1 atm (torsional mode).

Q_r	Theory	Experiment
Ar	1.04×10^4	1.67×10^4
Air	1.06×10^4	2.15×10^4
He	3.07×10^4	1.40×10^5

TABLE II. Q_r in various surrounding gases at 1 atm (compressional mode).

Q_r	Theory	Experiment
Ar	1.04×10^4	0.95×10^4
Air	1.04×10^4	⋯
He	2.98×10^4	3.14×10^4

$$\frac{\Delta K}{K} = \frac{\Delta\omega}{\omega} = \frac{\Delta f}{f},$$

we may calculate the resonant frequency shift Δf as

$$\frac{\Delta f}{f} = -\frac{2\Delta m}{M_B}$$

or

$$\Delta f = -\frac{X_m}{\pi M_B} = -\frac{\rho c a b X(ka)}{\pi M_B}. \tag{33}$$

To obtain the quality factor Q_r and the frequency shift Δf for the rectangular parallelepiped sample, we note that Eqs. (31) and (33) take into account the effects of a single pair of opposite faces of the sample. Therefore, to incorporate contributions due to all sets of faces, we average the Δf and $1/Q_r$ results over each of the three pairs of opposite faces. (An alternative approximation scheme would weigh the contributions from each pair of faces by the fraction of the total energy residing in the displacement field associated with each pair of faces. The final result would not be significantly different, however.) The reason that an average is required rather than a sum is that the vibration of each set of faces is associated with a different component of velocity. Therefore, the vibrations associated with each additional set of faces contributes both power radiated *and* energy stored. However, since $1/Q_r$ is proportional to the total power radiated divided by the total energy stored per cycle, it follows that an average of the contributions from each set of faces is the relevant quantity.[27] Similarly, the frequency shift can be understood as a mass loading on each face, but the net effect on the dynamics of the rectangular parallelepiped sample is approximated by a simple average within our two-dimensional model.

Now let us compare the quality factor of resonance Q_r and the frequency shift Δf measured experimentally to those calculated using Eq. (31) and Eq. (33) and the simple averaging process described in the previous paragraph. Table I gives the results of Q_r for the torsional mode for three different surrounding gases, and Table II shows the results for the compressional mode. It can be seen that the Q_r's calculated from Eq. (31) have the same order of magnitude as those of the experimental results for both modes, but the

TABLE III. Δf in various surrounding gases at 1 atm (torsional mode). (''Theory'' refers to radiation impedance contribution only.)

Δf (Hz)	Theory	Experiment
Ar	−1.69	−6.82
Air	−1.91	−23.6
He	−2.11	−5.34

TABLE IV. Δf in various surrounding gases at 1 atm (compressional mode). (''Theory'' refers to radiation impedance contribution only.)

Δf (Hz)	Theory	Experiment
Ar	-0.01	4.63
Air	-0.02	...
He	-0.15	-3.06

result of the torsional mode in helium agrees the best. Table III summarizes the results of Δf for the torsional mode for three different surrounding gases, while Table IV is the result for the compressional mode. For the compressional mode, the model predicts Q's that are within a few percent of the experimentally obtained ones and the changes in resonance frequency are one or two orders of magnitude lower than the measured ones. Therefore, for this mode it may be assumed that mass loading does not affect the values that would be deduced for the third-order elastic constants. However, the results are different for the torsional mode, where the theoretical Q's are as much as three and a half times lower than seen experimentally, and the theoretical changes in frequency are somewhere between a quarter and a twelfth of the experimental ones. Therefore, a better model has to be developed for these modes, since it appears that mass loading could make a significant contribution to the estimated third order elastic constant.

Since the typical ka values for the experiment are in the range $ka \approx 10$ to $ka \approx 35$, our calculated values of the frequency shifts are somewhat sensitive to the oscillations in the $X(ka)$ curves. It therefore follows from our discussion at the end of Sec. II concerning the sources of these oscillations that before definitive statements can be made about the effects of radiation impedance on the observed frequency shifts, calculations have to be performed for unbaffled pistons, and a better theoretical model has to be developed for the vibrating rectangular parallelepiped. However, it can be safely said that the present model accurately predicts the pressure dependence for the inverse of Q, and also predicts an increase in $1/Q$ with the increasing molecular mass of the surrounding gas.

At present, data are being collected on a spherical sample, which is more amenable to the actual calculation of the radiation impedance in three dimensions. The authors are also performing an *ab initio* calculation of the resonant modes of a sphere in a gas atmosphere and are developing a three-dimensional model for determining the radiation impedance of a rectangular parallelepiped.

ACKNOWLEDGMENTS

The authors would like to acknowledge interesting discussions with Albert Migliori, Jay Maynard, and Logan Hargrove. Research at UWM and UCLA was supported by the Office of Naval Research.

[1] See, for example, O. L. Anderson, *Equations of State for Geophysics and Ceramic Science* (Oxford U.P., Oxford, 1995), Chaps. 1, 8, and 9.
[2] R. N. Thurston and K. Brugger, ''Third-order elastic constants and the velocity of small amplitude elastic waves in homogeneously stressed media,'' Phys. Rev. **133**, A1604–A1610 (1964).
[3] K. Brugger, ''Thermodynamic definition of higher order elastic coefficients,'' Phys. Rev. **133**, A1611–A1612 (1964).
[4] L. L. Boyer, ''Theory of melting based on lattice instability,'' in *Phase Transitions* (Gordon and Breach, New York, and OPA Ltd., 1985), Vol. 5, pp. 1–48.
[5] L. L. Boyer, ''Magic strains in face-centered and body-centered cubic lattices,'' Acta Crystallogr., Sect. A: Found. Crystallogr. **45**, FC29–FC32 (1989).
[6] L. L. Boyer, ''Molecular-dynamics study of the melting of hexagonal and square lattices in two dimensions,'' Phys. Rev. B **53**, 3145–3149 (1996).
[7] L. L. Boyer and M. J. Mehl, ''Calculation of energy barriers for physically allowed lattice-invariant strains in aluminum and iridium,'' Phys. Rev. B **43**, 9498–9502 (1991).
[8] L. L. Boyer, E. Kaxiras, J. L. Feldman, J. Q. Broughton, and M. J. Mehl, ''New low-energy crystal structure for silicon,'' Phys. Rev. Lett. **67**, 715–718 (1991).
[9] E. Kaxiras and L. L. Boyer, ''Energetics of large lattice strains: application to silicon,'' Phys. Rev. B **50**, 1535 (1994).
[10] D. G. Isaak, O. L. Anderson, and T. Coto, ''Measured elastic moduli of single-crystal MgO up to 1800 K,'' Phys. Chem. Miner. **16**, 704–713 (1989).
[11] D. G. Isaak, ''High temperature elasticity of iron-bearing olivines,'' J. Geophys. Res. **97**, 1871–1885 (1992).
[12] O. L. Anderson and D. G. Isaak, ''Elastic constants of mantle minerals at high temperature,'' in *Mineral Physics and Crystallography: A Handbook of Physical Constants*, AGU Reference Shelf 2 (American Geophysical Union, Washington, DC, 1995), pp. 64–97.
[13] G. Chen, H. A. Spetzler, S. Whitehead, and I. C. Getting, ''Characterizing sub-millimeter crystals with a wide-band GHz interferometer,'' Ultrasonics International 1993 Proceedings 1993, pp. 363–366.
[14] H. A. Spetzler, G. Chen, S. Whitehead, and I. C. Getting, ''A new ultrasonic interferometer for the determination of equation of state parameters of sub-millimeter single crystals,'' Pure Appl. Geophys. **141**, 341–377 (1993).
[15] A. Migliori, private communication. See also, A. Migliori and J. L. Sarrao, *Resonant Ultrasound Spectroscopy* (Wiley, New York, 1997).
[16] The remaining asymmetry in the resulting peak in Fig. 2(b) might be removed by allowing the background coefficient a_4 to vary with frequency. Such refinements, however, are not warranted in view of the approximate nature of our theoretical model and the experimental uncertainties in the measurement.
[17] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics* (Wiley, New York, 1994), 3rd ed., p. 65.
[18] L. E. Kinsler and A. R. Frey, *Fundamentals of Acoustics* (Wiley, New York, 1962), 2nd ed., pp. 177–179.
[19] J. A. Archer-Hall, A. I. Bashter, and A. J. Hazelwood, ''A means of calculating the Kirchoff surface integral for a disk radiator as a single integral with fixed limits,'' J. Acoust. Soc. Am. **65**, 1568–1570 (1979).
[20] E. G. Williams and J. D. Maynard, ''Numerical evaluation of the Rayleigh integral for planar radiators using the FFT,'' J. Acoust. Soc. Am. **72**, 2020–2030 (1982).
[21] G. W. Swenson and W. E. Johnson, ''Radiation impedance of a rigid square piston in an infinite baffle,'' J. Acoust. Soc. Am. **24**, 84 (1952).
[22] C. E. Wallace, ''Radiation resistance of a rectangular panel,'' J. Acoust. Soc. Am. **51**, 946 (1972).
[23] R. J. Donato, ''Direct derivation of radiation resistance of a vibrating panel,'' J. Sound Vib. **28**, 87–92 (1973).
[24] C. L. Dym, ''A more direct derivation of the radiation resistance of a panel,'' J. Sound Vib. **32**, 279–282 (1974).
[25] M. C. Gomperts, ''Sound radiation from baffled thin rectangular plates,'' Acustica **37**, 93–102 (1977).
[26] The trigonometric function forms are used for the piston displacements in order to clearly demonstrate the consequences of the symmetries involved. These trigonometric functions are the leading terms of a series expansion of the exact solutions of the motion of the individual surfaces. However, it is felt that in view of the nature of the approximations used, the summation of the entire series is not warranted since the corrections produced by the absence of the summation would be comparable to those incurred by assuming a baffled solution.
[27] In an accurate calculation of Q, one would use the actual displacement field at each point within the sample to compute the total energy stored. Our more approximate approach avoids the necessity of dealing with the precise details of the displacement field inside the sample.

A new approximation method for thermoacoustic calculations

Richard Raspet, James Brewster, and Henry E. Bass
National Center for Physical Acoustics and Department of Physics and Astronomy, University of Mississippi, University, Mississippi 38677

(Received 27 November 1996; revised 21 January 1998; accepted 24 January 1998)

In many cases, it is convenient to have approximate relations to evaluate thermoacoustic gain and efficiency. G. W. Swift [J. Acoust. Soc. Am. **84**, 1145–1180 (1988)] has provided such relations for short stacks in a standing wave. This paper describes an approximation method based on a finite difference form of the thermoacoustic equations and provides results for thermoacoustic engines or refrigerators for both standing and traveling waves. The short stack assumption is retained. The method is limited at present to linear temperature gradients across the stack. The calculations are compared to measurements and in the appropriate limits to Rott's boundary conditions for a temperature discontinuity [N. Rott, Adv. Appl. Mech. **20**, 135–175 (1980)]. The approximation is applied to both thermoacoustic prime movers and thermoacoustic refrigerators. © *1998 Acoustical Society of America.* [S0001-4966(98)03605-4]

PACS numbers: 43.35.Ud [MAB]

INTRODUCTION

Swift's short stack approximation[1] has found wide application in explaining thermoacoustic phenomena in resonant tubes. In short, a uniform standing wave is assumed, and the work per unit volume is calculated using the gas properties at the center of the stack. This method has not been successfully applied to traveling wave devices or devices with both standing and traveling waves since there is no *a priori* way to apportion the work between the transmitted and reflected waves.

Atchley *et al.*[2] modeled the acoustic gain in a resonant thermoacoustic system by treating the second-order differential equation for acoustic pressure as a wave equation with constant coefficients. The free space wave number was used to calculate the complex compressibility. The results are readily expressed in terms of waves traveling in opposite directions. This approach leads to good predictions of quality factor and resonant frequency as a function of temperature difference across the stack. Atchley[3] reexamined the data from Ref. 2 using an energy-based standing wave short stack approximation which incorporated changes in gas properties through the stack, and improved the fit of theory to data.

Raspet *et al.*[4,5] developed an energy-based short stack traveling wave approximation and applied this theory to measurements on a traveling wave tube. Subsequent research demonstrated that this method failed in the Stirling regime.[6] We have attempted to apply and extend the calculational method of Ref. 2, only to find the method failed as the pore size was decreased. Other wave-based approximations we developed were valid for small stack pores, but failed for wider pore spaces. The goal of this paper is to develop an approximate method which will provide quick estimates of thermoacoustic performance while retaining the concept of counter-propagating waves necessary to develop an understanding of the relation between standing wave and traveling wave thermoacoustics. As such, this paper is a continuation of the research begun in Ref. 2 and for consistency will utilize the experimental results of that paper.

I. THEORY

The calculations of this paper are based on the formulation of Arnott *et al.*,[7] which express the thermoacoustic equations in terms of the thermoviscous function $F(\lambda)$, commonly used in the literature of sound propagation through porous media. This notation is chosen rather than Rott's since the notation and relevant approximations have been well developed in the analysis of sound propagation through a wide variety of porous media.[8] The translation between the porous media notation used and Rott's notation is presented in Appendix A. For parallel plate stacks with spacing y_0, $F(\lambda)=1-(2/\lambda\sqrt{-i})\tanh(\sqrt{-i}\lambda/2)$, where $\lambda=y_0(\rho\omega/\eta)^{1/2}$ is the shear wave number and $\lambda_T=y_0(\rho\omega c_p/\kappa)^{1/2}$ is the thermal wave number. The physical parameters of the gas are the density ρ, viscosity η, thermal conductivity κ, and specific heat per unit mass at constant pressure c_p. In small pores, the acoustic pressure is constant across the pore cross section and is a function only of axial distance along the pore. Also, the radial velocity is small with respect to the axial velocity. With these restrictions, an equation for the area-averaged complex axial particle velocity in terms of the pressure gradient is

$$\hat{v}(z)=\frac{d\hat{p}(z)/dz\ F(\lambda)}{i\omega\rho(z)}, \tag{1}$$

where $\rho(z)$ is the density, ω is the angular frequency, and $\hat{p}(z)$ is complex acoustic pressure amplitude.[7] The other equation necessary for our development relates the average axial velocity gradient to the average axial velocity and acoustic pressure. This equation can be developed by combining Eqs. (13) and (28) of Ref. 7 to give

$$\frac{d\hat{v}}{dz}+\frac{1}{T}\frac{dT}{dz}\left[\frac{F(\lambda_T)/F(\lambda)-1}{1-\sigma}\right]\hat{v}+\frac{F(\lambda)}{i\omega\rho(z)}k^2\hat{p}=0, \tag{2}$$

where we have substituted the ideal gas relationship $\beta=1/T$ for the coefficient of thermal expansion. Here k is the wave number in a pore and is given by

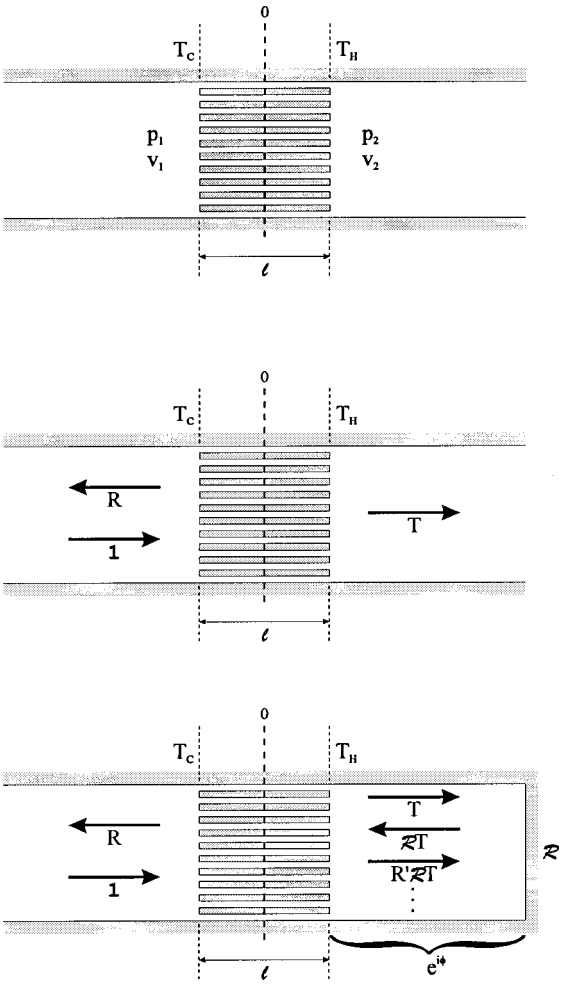


FIG. 1. Idealization of the stack in a thermoacoustic device with (a) the notation used for the finite difference scheme; (b) the notation used for reflection and transmission coefficients for traveling waves; and (c) the notation used for the standing wave gain calculation.

$$k^2=\frac{\omega^2}{c^2}\frac{[\gamma-(\gamma-1)F(\lambda_T)]}{F(\lambda)}, \tag{3}$$

where c is the adiabatic speed of sound in the free field, γ is the ratio of specific heats, and σ is the Prandtl number. Note that this wave number corresponds to the wave number in the pores of the stack and can be much larger than the wave number in free space. Equations (1) and (2) are convenient for an Euler finite difference solution for propagation in the stack. The stacks are typically very short with respect to a wavelength in the porous material, so a single-step Euler calculation can provide a reasonable estimate.

Consider the finite difference form of Eqs. (1) and (2) using average values at the center of the stack [see Fig. 1(a)]:

$$\frac{\hat{p}_2-\hat{p}_1}{l}=iA\frac{\hat{v}_1+\hat{v}_2}{2}, \tag{4}$$

$$iA\frac{(\hat{v}_2-\hat{v}_1)}{l}+iAB\frac{(\hat{v}_2+\hat{v}_1)}{2}+k^2\frac{\hat{p}_1+\hat{p}_2}{2}=0, \tag{5}$$

$$A=\omega\rho/F(\lambda), \tag{6}$$

and

$$B=\frac{1}{T}\frac{dT}{dz}\left[\frac{F(\lambda_T)/F(\lambda)-1}{1-\sigma}\right], \tag{7}$$

where l is the length of the stack. Here A, B, and k^2 are evaluated at the center of the stack, consistent with finite difference methods.

To adjust for the relative porosity of the stack, Ω, we replace $\hat{v}$ with $\hat{v}/\Omega$. Equations (4)–(7) can now be solved to form an impedance translation formula through the stack,

$$\frac{\hat{p}_1}{\hat{v}_1}=\frac{A}{k\Omega}\frac{\left[\frac{\hat{p}_2}{\hat{v}_2}\left(1-\frac{Bl}{2}-\frac{k^2l^2}{4}\right)-i\frac{A}{k\Omega}(kl)\right]}{\left[-i\frac{\hat{p}_2}{\hat{v}_2}kl+\frac{A}{k\Omega}\left(1-\frac{k^2l^2}{4}\right)+\frac{A}{k\Omega}\frac{B}{2}l\right]}. \tag{8}$$

Two ideal cases help us to interpret Eq. (8). The first is the case when $dT/dz=0$; there is no thermoacoustic effect. Then Eq. (8) becomes

$$\frac{\hat{p}_1}{\hat{v}_1}=\frac{A}{k\Omega}\frac{\left[\frac{\hat{p}_2}{\hat{v}_2}\left(1-\frac{k^2l^2}{4}\right)-i\frac{A}{k\Omega}(kl)\right]}{\left[-i\frac{\hat{p}_2}{\hat{v}_2}kl+\frac{A}{k\Omega}\left(1-\frac{k^2l^2}{4}\right)\right]}. \tag{9}$$

Comparison of Eq. (9) with the impedance translation theorem,[9]

$$Z_1=Z_{\text{int}}\frac{[Z_2\cos(kl)-iZ_{\text{int}}\sin(kl)]}{[Z_{\text{int}}\cos(kl)-iZ_2\sin(kl)]}, \tag{10}$$

demonstrates that the finite difference calculation is correct to first order in kl with $Z_{\text{int}}=A/k\Omega$, $Z_1=\hat{p}_1/\hat{v}_1$, and $Z_2=\hat{p}_2/\hat{v}_2$.

A second case is that of a vanishingly short stack with a fixed temperature difference. This corresponds to Rott's boundary condition for a temperature discontinuity.[10] In this limit

$$\lim_{l\to 0} Bl\to\frac{\Delta T}{T_m}\left[\frac{F(\lambda_T)/F(\lambda)-1}{(1-\sigma)}\right], \tag{11}$$

remains finite, and depends only on the total temperature change, while kl approaches zero.

Equation (8) then becomes

$$\frac{\hat{p}_1}{\hat{v}_1}=\frac{\hat{p}_2}{\hat{v}_2}\frac{(1-Bl/2)}{(1+Bl/2)}, \tag{12}$$

where Bl is given by Eq. (11). At the discontinuity, the pressure is continuous so that the boundary condition on particle velocity is

$$\hat{v}_2=\hat{v}_1\frac{(1-Bl/2)}{(1+Bl/2)}. \tag{13}$$

For positive ΔT, Bl has negative real and imaginary parts and therefore $\hat{v}_2$ is greater than $\hat{v}_1$; the wave is amplified in propagating from the cold side to the hot side of the stack.

Rott's boundary condition for a temperature discontinuity can be expressed as

$$\hat{p}_2=\hat{p}_1 \tag{14}$$

and

$$\hat{v}_2 = \hat{v}_1 e^{-B'l}, \tag{15}$$

where

$$B' = \frac{1}{l}\int_0^l \left[\frac{F(\lambda_T)/F(\lambda)-1}{1-\sigma}\right]\frac{1}{T}\frac{dT}{dz}\,dz. \tag{16}$$

The Rott impedance relation is then

$$\frac{\hat{p}_1}{\hat{v}_1} = \frac{\hat{p}_2}{\hat{v}_2}\,e^{-B'l}. \tag{17}$$

Here B' has negative real and imaginary parts for ΔT positive so that Eq. (15) produces a gain in particle velocity for propagation in the direction of positive temperature gradient. For thermoacoustic prime movers with the cold side near room temperature, the thermoviscous functions are slowly varying and the stack length is small with respect to a wavelength. For these conditions Eq. (16) can be approximated by

$$B' \cong \frac{1}{l}\,\frac{F(\lambda_T)/F(\lambda)-1}{1-\sigma}\cdot\frac{\Delta T}{T_m} = B, \tag{18}$$

where the gas properties are evaluated at the center of the stack.

The close agreement between the Euler form for the impedance change due to a short stack [Eq. (12)] and Rott's form [Eq. (17)] can be understood by comparing the series expansion of $(1-x/2)/(1+x/2)$ and e^{-x} with $x=Bl$:

$$\frac{1-x/2}{1+x/2} = 1-x+\frac{x^2}{2}-\frac{x^3}{4}\cdots, \tag{19a}$$

$$e^{-x} = 1-x+\frac{x^2}{2}-\frac{x^3}{6}\cdots. \tag{19b}$$

In the limit of very short stacks, the single-step finite difference agrees with Rott's boundary condition to second order in Bl, and the third-order term differs only by a factor of 1.5.

For Stirling devices, λ is small and

$$\frac{F(\lambda_T)/F(\lambda)-1}{1-\sigma} \approx -1. \tag{20}$$

Rott's integral becomes

$$B' = -\frac{1}{l}\int_0^l \frac{1}{T}\cdot\frac{dT}{dx}\,dx = -\frac{1}{l}\ln\left(\frac{T_H}{T_C}\right). \tag{21}$$

Equation (17) then yields

$$Z_1 = Z_2\,\frac{T_H}{T_C}. \tag{22}$$

An impedance relation similar to Eq. (22) has been derived by Ceperley for Stirling engines.[11]

The single-step finite difference method well approximates three analytic limits for short stacks and moderate $\Delta T/T_m$. In the thermoacoustic regime Bl is about 0.75 for $\Delta T/T_m = 1$; the difference between the expansions in Eqs. (19a) and (19b) is only 4%. In the Stirling limit Bl is close to 1 for $\Delta T/T_m = 1$, but the finite difference method reproduces the Stirling result exactly. The Euler method can be improved by subdividing the stack, and applying the finite difference calculation to the subsections. In the next section, we will compare results from a single-step Euler approximation to results using multiple steps for two realistic examples.

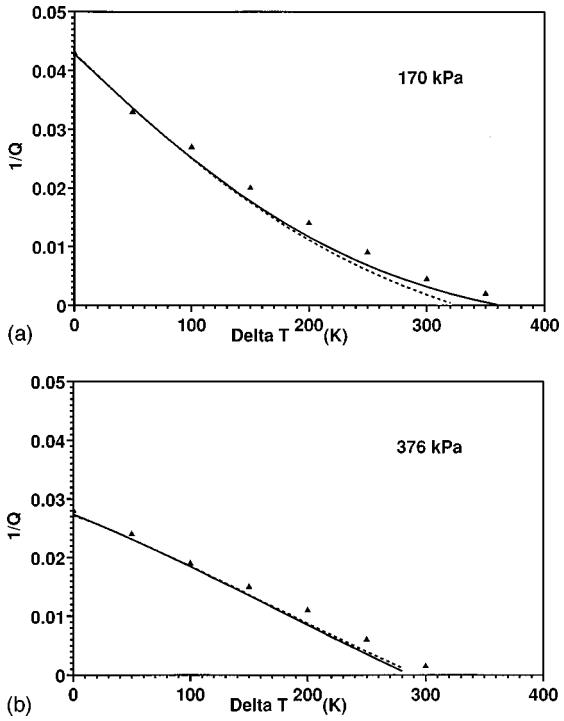


FIG. 2. Graph of $1/Q$ versus the temperature difference across the stack for a prime mover filled with helium. The symbols represent experimental data (Ref. 1). ——— Prediction using the single step Euler approximation for the stack. --- Prediction with two steps. (a) Mean pressure 170 kPa. (b) Mean pressure 376 kPa.

II. APPLICATIONS

A. Prime mover

We demonstrate the accuracy of the single-step Euler's method solution by comparison with a multiple step integration of the equations and previously published experimental data. Atchley *et al.*[2] measured the quality factor of a resonator containing a thermoacoustic prime mover as a function of the temperature gradient across the stack. The apparatus consisted of a cylindrical tube, 1.0 m long and 3.8 cm in diameter, filled with pressurized helium. The stack was 3.5 cm in length and sandwiched between heat exchangers with its hot end located 6.3 cm from the end of the tube.

We calculated the frequency response of this resonator in the following manner. In sections other than the stack, the temperature is uniform and the solution to Eqs. (1) and (2) is

$$p(x) = A^+ e^{+ikx} + A^- e^{-ikx}. \tag{23}$$

In the stack, Eqs. (4) and (5) are applied, either in a single step or in multiple iterations with the stack divided into sections of equal length. The boundary conditions between sections of the device were that the pressure and volume velocity be continuous. The impedance at the ends was taken to be to that of a solid wall.[9]

Figure 2(a) and (b) shows comparison of this calculation to experimental results taken from Ref. 2 in which the am-

bient pressures are 170 and 376 kPa, respectively. The calculation was performed using one step, and two steps to span the stack.

In general, the Euler method approximation predicts the behavior of the experimental data. The prediction overestimates the quality factor as the system approaches onset. The direction of this error is consistent with extra losses in the experiment or a nonzero temperature difference between the heat exchanger and the stack. The error is comparable for both cases displayed. The 170-kPa case [Fig. 2(a)] displays a significant but small difference between the one-step and two-step Euler methods. Calculations with three steps cannot be distinguished from the two-step calculation. As a diagnostic, the magnitude of kl and Bl were printed out. The one-step and two-step results differ noticeably only when $|kl| > 0.3$. The single-step approximation is still useful for initial design purposes when $|kl| \cong 0.5$.

In Atchley *et al.*,[2] the authors reduced the porosity of the stack from its true value of 0.76 to an effective value of 0.53. This was justified on the basis that the pores of the stack and its neighboring heat exchangers did not form continuous ducts but were offset. The value chosen for the porosity reflected the estimated extent of the blockage resulting from this offset. The reduced porosity provided better agreement between measurements and calculations, especially near onset. The calculations reported here agree with experimental results reasonably well without using a modified porosity.

In a later paper,[3] Atchley reanalyzed this data using an extension of Swift's short stack approximation involving integrals of the varying gas properties through the stack. The success of the single-step approximation reported here indicates that changing gas properties in the stack are a small perturbation.

B. Refrigeration

The finite difference approach to thermoacoustics can also be applied to the analysis of refrigeration stacks. The temperature gradient is typically less than that in a prime mover but the stack is often longer. Refrigeration technology is not limited to stacks driven by standing waves and we will demonstrate that the single-step approximation extends to cases where there is a significant traveling wave component present. Hofler has reported on a similar analysis using a different calculational method.[12]

Consider the general case of a thermoacoustic refrigeration stack. We wish to determine the coefficient of performance (CoP), defined by

$$\mathrm{CoP} = \frac{Q_C}{W_C - W_H}, \tag{24}$$

where Q_C is the heat flow at the cold end and W_H and W_C are the work flows at the hot and cold ends, respectively. In any particular device, the amplitude and phase of the pressure and velocity on each side of the stack are determined by the design of the resonator and driver in addition to the response of the stack. Here the problem is approached in a more general manner by taking the pressure and velocity to be known at one end. If the pressure and velocity are known, then the heat and work flows can be found by substitution into Eqs. (11), (47), and (48) of Ref. 7:

$$W = \frac{A_{\mathrm{res}}}{2} \mathrm{Re}(\hat{v}\hat{p}^*), \tag{25}$$

$$Q = \frac{A_{\mathrm{res}}}{2} \frac{1}{1+\sigma} \mathrm{Re}\left(\hat{v}\hat{p}^*\left(\frac{F^*(\lambda_T)}{F(\lambda)} - 1\right)\right) - \frac{A_{\mathrm{res}}}{2} \frac{dT}{dx} \frac{c_p}{\omega} \left(\left|\frac{\hat{v}}{F(\lambda)}\right|\right)^2 \frac{\rho}{1-\sigma^2} \mathrm{Im}(F^*(\lambda_T) + \sigma F(\lambda)), \tag{26}$$

where W is the work flow, Q is the heat flow, A_{res} is the cross-sectional area, and c_p the heat capacity at constant pressure. For the purpose of this demonstration the porosity is set equal to unity so the velocity can be assumed continuous between the stack and its surroundings. If we choose the cold end to be the end at which the pressure and velocity are known, then the only quantity in Eq. (24) dependent on the propagation model employed to span the stack is W_H.

Arguments based on similitude[13] show that for a given value of λ_T the results are independent of the frequency if the stack length is expressed in units of the wavelength at a fixed reference temperature. This temperature is chosen to be that of the hot end. The coefficient of performance is also independent of area.

The acoustic field on the cold side of the stack is described by the superposition of a unit traveling wave with a reflected wave of arbitrary real amplitude and phase:

$$p = e^{ikx} + Ae^{i\phi}e^{-ikx}. \tag{27}$$

The cold end of the stack is located at $x=0$. We define a ''standing wave ratio'' (SWR) in terms of A:

$$\mathrm{SWR} = \frac{1+A}{1-A}. \tag{28}$$

The amplitude A ranges from 0 for a pure traveling wave to the right to $+\infty$ for a pure traveling wave to the left. The inverse standing wave ratio then indicates the relative extent to which the wave is a standing wave; 1/SWR is equal to $+1$ for a traveling wave moving into the stack, 0 for a standing wave and -1 for a traveling wave moving away from the cold end of the stack.

To demonstrate the use of the finite difference method for refrigeration with arbitrary traveling and standing waves, we calculate the relative coefficient of performance of two example refrigerators as a function of the inverse of the standing wave ratio. The relative coefficient of performance is the coefficient of performance divided by the Carnot coefficient of performance $T_C/(T_H - T_C)$.

First, we calculate the relative coefficient for a helium-filled thermoacoustic refrigerator operating between 290 K and 270 K. The plate spacing is such that $\lambda_T = 2.9$ and the length of the stack is 0.02 wavelengths in free space. These are evaluated at 290 K. The phase is set to 0.225π, which gives the largest maximum relative coefficient of performance. Figure 3(a) displays the results of the calculation for

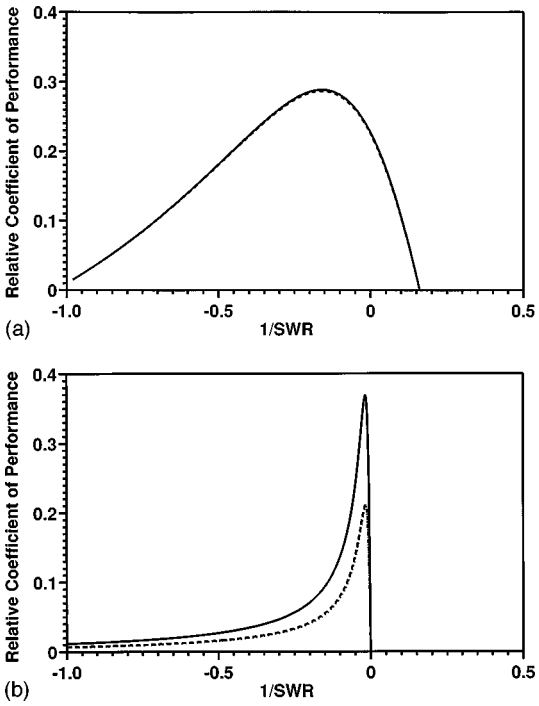


FIG. 3. Relative coefficient of performance for example refrigerators versus inverse standing wave ratio. (a) A thermoacoustic refrigerator for temperature span 290 K to 270 K, $\lambda_T = 2.9$. ——— one step, --- two steps. (b) A Stirling regime cryo-cooler for temperature span 270 K to 70 K, $\lambda_T = .035$. ——— one step, --- two steps.

one step and two steps. The results are nearly indistinguishable. For this case $|kl|$ and $|Bl|$ are 0.19 and 0.05, respectively, and the one-step method is sufficient.

The next case will use parameters more like a Stirling or pulse tube refrigerator. To investigate the limits of the discrete step method we choose a temperature span typical of cryo-coolers: the hot side is held at 270 K and the cold side at 70 K. The stack plate spacing is such that $\lambda_T = 0.035$ and the stack length is 0.002 wavelengths in free space evaluated in the hot end. This corresponds to a $k_0 l$ value of 0.013 rad, where k_0 is the free space wave number. The phase, ϕ, is set to 0.

The results are displayed in Fig. 3(b) for one step and two steps. The results for three steps agree with the two-step calculation to within the lines used in Fig. 3(b). Although the approximation is very poor for the one-step calculation, it quickly converges for two and three steps. The single-step calculation fails since both $|kl|$ and $|Bl|$ are greater than 1. Surprisingly, the two-step calculation is sufficient (agrees with the three-step calculation) although $|kl| = 0.62$ and $|Bl| = 0.455$.

Note that $|kl|$ is much larger than the ''free space'' value of 0.013 rad. The wave number in finely divided porous media can be orders of magnitude larger than the free space value. Calculations with half the stack length or half the temperature span used in Fig. 3(b) indicate that the restriction on $|kl|$ is the principle restriction on the single-step accuracy, not the value of $|Bl|$. For small λ_T, large $|kl|$ implies large acoustic losses since the real and imaginary parts of the wave number are equal. The losses in our example cryo-cooler may be larger than losses in an optimized realistic device.

As noted by Hofler,[12] the highest coefficient of performance for a thermoacoustic refrigerator is achieved when there is a significant traveling wave component in the refrigerator. Note also that the coefficient of performance of the Stirling device is zero for a pure standing wave.

III. A TRAVELING WAVE INTERPRETATION OF THERMOACOUSTIC GAIN

The finite difference model derived above can be used to develop a traveling wave interpretation of thermoacoustic gain. Referring to Fig. 1(b), a traveling wave of complex pressure amplitude 1 is incident on a stack comprised of a porous material supporting a linear temperature gradient. A portion of the wave is reflected and a portion is transmitted. We denote the complex reflection coefficient by R and the complex transmission coefficient by T when the wave is incident on the stack in the direction of increasing temperature and by R' and T' when the direction of the incident wave opposes the temperature gradient.

The impedance seen by the incident wave (Z_1) given a traveling wave on the right-hand side can be calculated using Eq. (9) and

$$Z_2 = \hat{p}_2 / \hat{v}_2 = \rho(T_H) c(T_H), \tag{29}$$

the traveling wave impedance. The reflection factor from the left-hand side of the stack is given by

$$R = e^{-ik_1 l}\left(\frac{z_1 - 1}{z_1 + 1}\right), \tag{30}$$

where z_1 is the impedance Z_1, normalized by $\rho(T_C)c(T_C)$. The factor $e^{-ik_1 l}$ results from choosing the center of the stack as the zero of the coordinate system when evaluating R. This means that k_1, the free space wave number, is evaluated at temperature T_C.

The transmission coefficient is evaluated by solving Eqs. (4) and (5) for $\hat{p}_2$ in terms of $\hat{p}_1$ and $\hat{v}_1$,

$$T = e^{-ik_3(l/2)} \frac{\left[\hat{p}_1\left(1 + \frac{Bl}{2} - \frac{k^2 l^2}{4}\right) + \hat{v}_1 \frac{iAl}{\Omega}\right]}{1 + \frac{Bl}{2} + \frac{k^2 l^2}{4}}, \tag{31}$$

where k_3 is the wave number at temperature T_H. The effect of stack porosity has been incorporated into Eq. (31). Figure 4 displays the transmission coefficient as a function of thermal wave number, λ_T, for a parallel plate stack 3.5 cm long with 0.77-mm spacing and porosity 0.76, with helium at 170 kPa as the working fluid. The cold end is held at 300 K and the hot end is 300, 400, 500, and 600 K for the curves displayed. The traveling wave gain for waves propagating from the cold side to the hot side is largest for λ_T small, the Stirling regime. In any resonant device, there are traveling waves propagating from the cold side of the stack to the hot side and vice versa. Also displayed in Fig. 4 is the transmitted intensity for a unit intensity traveling wave incident on the hot end for a 300-K temperature difference.

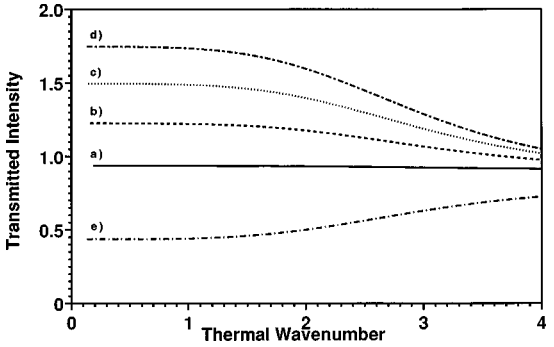


FIG. 4. Transmitted intensity relative to an incident intensity of one versus thermal wave number for the example stack. Transmitted intensity from the cold side at 300 K to the hot side at (a) 300 K ———, (b) 400 K – – –, (c) 500 K ···, and (d) 600 K ——–. (e) Transmitted intensity from the hot side at 600 K when the cold side is held at 300 K -·-·-.

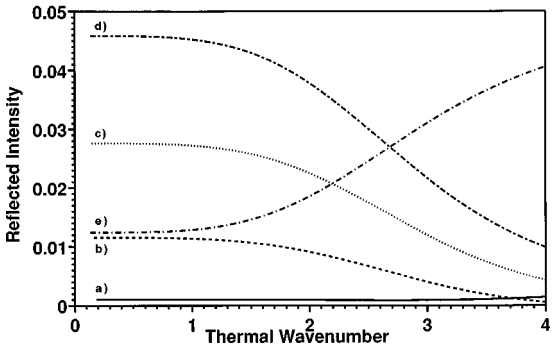


FIG. 6. Reflected intensity from the cold end of the stack at 300 K versus thermal wave number for a hot end temperature of (a) 300 K ———, (b) 400 K – – –, (c) 500 K ···, and (d) 600 K ——–. (e) Reflected intensity from the hot end of the stack at 600 K when the cold end is held at 300 K -·-·-.

A crude model for a resonant device is a traveling wave incident on the cold side of the stack transmitted to the hot side, reversed in direction and then transmitted from the hot side to the cold side without reflections off the stack. For this perfect reflectionless device, the total gain is the product of the intensity gain in propagating from the cold side of the stack to the hot side of the stack times the intensity loss in propagating from hot side to cold side. The calculated two-way intensity for our example stack as a function of λ_T is illustrated in Fig. 5. The net gain for the ideal device is less than 1 for all values of λ_T. There is a maximum near $\lambda_T = 2.7$, but this maximum still represents a net loss of 20% of the incident intensity. The product of the transmitted intensity increases with increasing temperature difference from cold to hot but still remains less than 1. Evidently, the energy reflected from the stack faces must also contribute to the intensity gain for thermoacoustic prime movers to operate.

The reflected intensities for the same temperature differences used in Fig. 4 are plotted in Fig. 6. These are small with respect to the transmitted intensities and are in fact smaller than would be calculated for a temperature discontinuity without a stack. For future use, the phase of the transmitted and reflected waves are displayed in Fig. 7. Note that the phase of the reflection coefficient from the cold side differs by approximately π from the phase of the reflection coefficient from the hot side of the stack.

A simple calculation which takes the reflected wave contributions into account will clarify the relationship between traveling wave gain and onset of standing wave devices. Consider the wave contribution in Fig. 1(c). A wave of amplitude 1 is incident on the cold side of the stack. A portion is reflected with complex coefficient R and a portion is transmitted with complex transmission coefficient T. The transmitted wave is reflected from the tube termination with reflection coefficient $\Re$. We will consider $\Re$ as real and incorporate any phase change on reflection into the phase factor ϕ. If the reflection factor of the termination is real and positive, ϕ is the free space phase change from the end of the stack to the termination at the hot end temperature. This is a completely general description of a passive termination of a resonance tube. A portion of this reflected wave is transmitted with coefficient T' and a portion reflected with coefficient R'. The net gain received at the left-hand side of the stack is expressed by the series

$$G = R + T\Re e^{i2\phi} + T\Re e^{i2\phi} R' \Re e^{i2\phi} T' + \cdots \tag{32}$$

or

$$G = R + T\Re T' e^{i2\phi}(1 + R'\Re e^{i2\phi} + (R'\Re e^{i2\phi})^2 + \cdots)$$

$$= R + \frac{T\Re T' e^{i2\phi}}{1 - R'\Re e^{i2\phi}}. \tag{33}$$

The series clearly converges since $\Re$ and R' are less than 1.

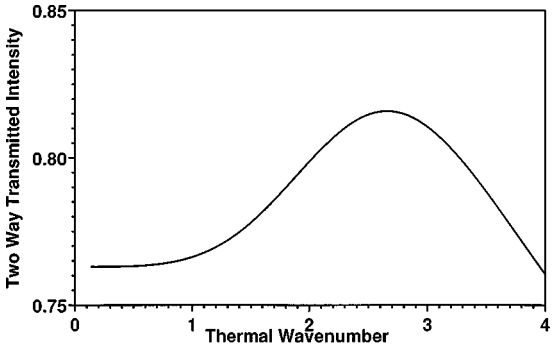


FIG. 5. Product of the transmitted intensities versus thermal wave number λ_T for the example stack with the cold end at 300 K and the hot end at 600 K.

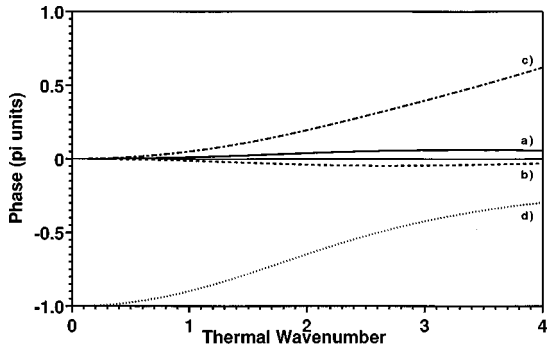


FIG. 7. The phase of the transmitted and reflected waves in fractions of π as a function of thermal wave number for (a) ——— transmitted wave for wave incident on cold side; (b) – – – transmitted wave for wave incident on the hot side; (c) -·-·- reflected wave for a wave incident on the cold side; and (d) ··· reflected wave on an incident wave on the hot side.

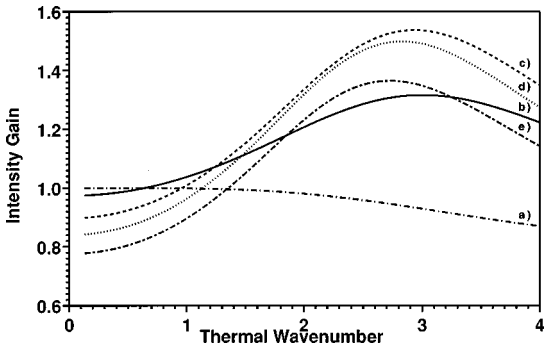


FIG. 8. Gain [Eq. (33)] as a function of thermal wave number for the example stack with ϕ given by (a) -·-·-, 0.0π; (b) ———, $0.10\ \pi$; (c) – – –, 0.20π; (d) ···, $0.25\ \pi$; and (e) ————, $0.30\ \pi$.

Note that the total gain produced by the stack plus resonance tube end is maximized when $\Re\, R' e^{i2\phi}$ is real and when the phase of R matches the phase of $e^{i2\phi}$. This is illustrated in Fig. 8 which displays the total gain of the example stack in a lossless resonator. The optimum phase change ($2\phi = 0.4\pi$) corresponds closely to the phase of the reflected wave from the cold end of the stack and also is close to the negative of the phase of the reflection from the hot end of the stack.

The total gain in the resonant system is strongly dependent on the phase and magnitude of the reflected waves. In the Stirling limit (small λ_T) the magnitude and phase of the reflection coefficients are small and the magnitude of the product TT' is always less than 1 so that it is not possible to construct a resonant device with net gain. The large traveling wave gain predicted in the Stirling limit cannot be exploited in a resonant device. The thermoacoustic engine ($\lambda_T \approx 3.0$) has smaller traveling wave gains but also has relatively smaller traveling waves losses against the temperature gradient. The reflected waves also contribute in phase to the gain in a properly designed resonance tube, so that instability can occur even when the product of the transmission coefficients is less than 1.

This analysis does not include annular resonators that are being studied by others,[14,15] but the methods described above can be adapted to such devices.

IV. CONCLUSIONS

The finite difference approximation method developed in this paper is useful for initial design estimates for thermoacoustic engines and refrigerators. This method may also be used for devices operated in the Stirling regime, with some care. All approximations in this calculation are clearly understood.

The approximation has been tested by use of a multistep extension of the finite difference method for three example systems; a thermoacoustic engine below onset and a refrigerator in the thermoacoustic and Stirling regimes. The accuracy of the calculations is determined by the size of $|kl|$ and $|Bl|$. We have found that the limiting factor for refrigerators is usually $|kl|$ exceeding 0.5, while the limiting factor for typical engines is $|Bl|$ exceeding 0.5. For Stirling devices, large wave numbers in the porous stacks are a limiting factor on the computational method. It is emphasized that this calculation is limited to a linear stack temperature gradient.

A traveling wave interpretation of thermoacoustic engines has been presented, which explains gain in thermoacoustic resonators in terms the interactions of traveling wave contributions, and demonstrates that the reflected components are a significant contributor to thermoacoustic gain. Optimum phasing is shown to be determined by the phase of the reflected waves relative to the incident waves.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research. We thank Tom Hofler for his assistance in modeling Stirling regime refrigeration and Greg Swift for his comments.

APPENDIX A: RELATION OF POROUS MEDIA NOTATION TO ROTT'S NOTATION FOR THERMOACOUSTICS

The relation between shear wave number, λ, and viscous penetration depth, δ_v is given by

$$\lambda = 2^{1/2} R/\delta_v, \tag{A1}$$

and the relation between thermal disturbance number, λ_T, and the thermal boundary layer thickness, δ_κ, is given by

$$\lambda_T = 2^{1/2} R/\delta_\kappa, \tag{A2}$$

where R is the hydrodynamic radius defined as twice the pore area divided by the pore perimeter.

The relations between $F(\lambda)$ and $F(\lambda_T)$ and the Rott functions f_ν and f_κ are given by

$$F^*(\lambda) = 1 - f_v, \tag{A3}$$

$$F^*(\lambda_T) = 1 - f_\kappa. \tag{A4}$$

The complex conjugate arises due to our use of $e^{-i\omega t}$ notation, rather than Rott's $e^{i\omega t}$ notation.

APPENDIX B: TRANSMISSION AND REFLECTION COEFFICIENTS CALCULATED USING ROTT'S BOUNDARY CONDITION

Analogous expressions to Eqs. (30) and (31) can be developed from the impedance of a short stack calculated from Rott's boundary condition for a temperature discontinuity.

The substitution of $Z_1 = \hat{p}_1/\hat{v}_1$ and $Z_2 = \hat{p}_2/\hat{v}_2$ into Eq. (17) yields

$$z_1 = z_2 \frac{\rho(T_H)c(T_H)}{\rho(T_C)c(T_C)}\, e^{-B'l}, \tag{B1}$$

where z_1 and z_2 are the normalized impedances.

The reflection coefficient calculated by assuming a traveling wave on the right-hand side is

$$R = \frac{z_1 - 1}{z_1 + 1} = \frac{\sqrt{T_C/T_H}\, e^{-B'l} - 1}{\sqrt{T_C/T_H}\, e^{-B'l} + 1}. \tag{B2}$$

The reflection coefficient for a wave incident on the left-hand side can be calculated by exchanging T_C and T_H in Eq. (B2) and yields $R' = -R$, so the reflection coefficient from the hot

side has a π phase shift with respect to the reflection coefficient from the cold side as observed approximately in Fig. 7.

The transmission coefficients are easily calculated since pressure is continuous across the discontinuity:

$$T = 1 + R \tag{B3}$$

and

$$T' = 1 + R' = 1 - R. \tag{B4}$$

These may be used in Eq. (33) to yield results for the gain of a short lossless stack.

[1] G. W. Swift, ''Thermoacoustic engines,'' J. Acoust. Soc. Am. **84**, 1145–1180 (1988).

[2] A. A. Atchley, H. E. Bass, T. J. Hofler, and H. T. Lin, ''Study of a thermoacoustic prime mover below the onset of oscillation,'' J. Acoust. Soc. Am. **91**, 734–743 (1992).

[3] A. A. Atchley, ''Standing wave analysis of a thermoacoustic prime mover below onset of self-oscillation,'' J. Acoust. Soc. Am. **92**, 2907–2914 (1992).

[4] R. Raspet, H. E. Bass, and J. Kordomenos, ''Thermoacoustics of traveling waves: Theoretical analysis for an inviscid ideal gas,'' J. Acoust. Soc. Am. **94**, 2232–2239 (1993).

[5] J. Kordomenos, A. A. Atchley, R. Raspet, and H. E. Bass, ''Experimental study of a thermoacoustic termination of a traveling-wave tube,'' J. Acoust. Soc. Am. **98**, 1623–1628 (1995).

[6] R. Raspet, ''Erratum: 'Thermoacoustics of traveling waves: Theoretical analysis for an inviscid ideal gas' [J. Acoust. Soc. Am. **94**, 2232–2239 (1993)] and 'Experimental study of a thermoacoustic termination of a traveling-wave tube' [J. Acoust. Soc. Am. **98**, 1623–1628 (1995)].'' J. Acoust. Soc. Am. **100**, 673 (E).

[7] W. P. Arnott, H. E. Bass, and R. Raspet, ''General formulation of thermoacoustics for stacks having arbitrarily shaped pore cross sections,'' J. Acoust. Soc. Am. **90**, 3228–3227 (1991).

[8] K. Attenborough, ''Acoustical characteristics of porous materials,'' Phys. Rep. **82**, 179–227 (1982).

[9] A. D. Pierce, *Acoustics, An Introduction to its Physical Principals and Applications* (Acoustical Society of America, Woodbury, NY, 1989).

[10] N. Rott, ''Thermoacoustics,'' Adv. Appl. Mech. **20**, 135–175 (1980).

[11] P. H. Ceperley, ''Gain and efficiency of a traveling wave heat engines,'' J. Acoust. Soc. Am. **72**, 1688–1694 (1982); ''Gain and efficiency of a short traveling wave heat engine,'' J. Acoust. Soc. Am. **77**, 1239–1244 (1985).

[12] T. J. Hofler, ''Performance of a short parallel-plate thermoacoustic stack with arbitrary plate spacing,'' J. Acoust. Soc. Am. **88**, 594 (1990).

[13] J. R. Olson and G. W. Swift, ''Similitude in thermoacoustics,'' J. Acoust. Soc. Am. **95**, 1405–1412 (1991).

[14] H. T. Lin, R. T. Muehleisen, and A. A. Atchley, ''Investigation of an annular thermoacoustic prime mover,'' J. Acoust. Soc. Am. **101**, 3021 (1997).

[15] R. T. Muehleisen, H.-T. Lin, and A. A. Atchley, ''Control system analysis of an annular prime mover,'' J. Acoust. Soc. Am. **101**, 3021 (1997).

Focusing and steering through absorbing and aberrating layers: Application to ultrasonic propagation through the skull

Mickaël Tanter, Jean-Louis Thomas, and Mathias Fink
Laboratoire Ondes et Acoustique, ESPCI, Université Paris VII, U.R.A. C.N.R.S. 1503, 10 rue Vauquelin, 75005 Paris, France

(Received 9 May 1997; accepted for publication 20 January 1998)

The time-reversal process is applied to focus pulsed ultrasonic waves through the human skull bone. The aim here is to treat brain tumors, which are difficult to reach with classical surgery means. Such a surgical application requires precise control of the size and location of the therapeutic focal beam. The severe ultrasonic attenuation in the skull reduces the efficiency of the time reversal process. Nevertheless, an improvement of the time reversal process in absorbing media has been investigated and applied to the focusing through the skull [J.-L. Thomas and M. Fink, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 1122–1129 (1996)]. Here an extension of this technique is presented in order to focus on a set of points surrounding an initial artificial source implanted in the tissue volume to treat. From the knowledge of the Green's function matched to this initial source location a new Green's function matched to various points of interest is deduced in order to treat the whole volume. In a homogeneous medium, conventional steering consists of tilting the wave front focused on the acoustical source. In a heterogeneous medium, this process is only valid for small angles or when aberrations are located in a layer close to the array. It is shown here how to extend this method to aberrating and absorbing layers, like the skull bone, located at any distance from the array of transducers. © *1998 Acoustical Society of America.* [S0001-4966(98)04904-2]

PACS numbers: 43.35.Wa, 43.80.Sh [HEB]

INTRODUCTION

A number of recent studies have demonstrated the feasibility of the practical clinical realization of noninvasive surgery using high intensity focal ultrasound.[1,2] Such surgical applications require precise control of the size and location of the focal of the therapy beam. However, most ultrasonic therapeutic systems do not take into account the inhomogeneities of the human tissues: The acoustic velocity is assumed to be constant, while steering and focusing the acoustic energy. In fact, the spatial variations of the speed of sound (from approximately 1440 m s^{-1} in fat to 1580 m s^{-1} in muscle) induce a phase and amplitude distortion of the wave front propagating in the medium. These aberrations degrade the spatial resolution, which can reduce the therapeutic performance. Furthermore, in the particular case of the skull, a large discrepancy in acoustic velocities between brain tissue and skull tissue (about 1500 m s^{-1} vs 3000 m s^{-1}, respectively) and a severe attenuation of ultrasound in the skull bone severely degrade the beam shape. This problem was first discovered by White *et al.* (1967)[3] during investigation into brain imaging.

There have been many investigations into the effects of biological tissue heterogeneities upon resolution and imaging quality. O'Donnell and Flax[4] have measured *in vivo* phase and amplitude distortions of the wave front in subjects producing low and high image quality. They have shown that the inhomogeneities can be modeled like a phase screen close to the array. Most investigations in correction of ultrasonic beam degradation have restricted their means of compensation to time shift in the receiving aperture, although it is not a complete solution of the aberration correction problem. Zhu and Steinberg[5,6] have shown that the wave front distortion due to breast tissue is not only described by a phase distortion, but also an amplitude distortion. Important improvements in breast imaging are made by phase aberration correction, but even when all phase errors are corrected, the sidelobes are higher than in homogeneous medium.

The inability of time shift compensation in the receiving aperture to correct perfectly for the effect of a phase screen that is not near the receiving aperture has been shown experimentally[7] and explained by our group.[8] Recently, Liu and Waag[9] have modeled the aberrating medium as a phase screen placed some distance from the receiving aperture. They removed amplitude and phase distortion by backpropagating the wave front before applying the time shift compensation. Optimal focusing in any heterogeneous nonabsorbing medium requires taking into account all the information received on the array from the point source. For a monochromatic signal, this information corresponds to the amplitude and phase modulation recorded on each element of the array. For a pulsed echo, each spectral component of the wave front is subject to different phase and amplitude distortions: Then, all the temporal waveforms need to be recorded.

Time reversal[10–12] represents an original way to compensate for distortion due to various inhomogeneous media. Time-reversal mirrors (TRM) take advantage of the invariance of the wave equation in a lossless medium under a time-reversal operation. This means that if focusing through any inhomogeneous medium is the objective, the distorted wave field coming from a source (active or passive) located at the desired focal point should be recorded and then transmitted in time reversed form. The time-reversed wave field back propagates through the inhomogeneities and optimally focuses on the source. Such an adaptive system is locked on

the pointlike reflector position or on the brightest point of an extended target. Usually, in a practical situation, such an acoustic source is not available in the region of interest. Nevertheless, for hyperthermia, invasive methods using implantable acoustic sensors inside tumors for acoustic feedback have been proposed.[13]

In this paper, we first show the limit of the time reversal process in focusing an implantable point source through an absorbing and aberrating layer, like the skull. Section I illustrates with a new experiment the main results previously shown in Ref. 14. Then, we review how to combine time reversal with an amplitude compensation procedure to improve the focusing, when the skull can be modeled as a thin absorbing layer close to the array of transducers. We discuss and justify this assumption of thin absorbing layer. We review also how a numerical backpropagation extends this technique, when the skull is located at some distance of the array.

Then, in Sec. II, the time-reversal technique improvement by the amplitude compensation and numerical backpropagation is extended to focus not only on the initial source, but also on other points surrounding the implanted source. Ultrasonic beam steering through heterogeneous nonabsorbing layers with a time reversal mirror has been recently treated by Dorme and Fink.[15] We generalize their results to absorbing media. In most ultrasonic therapeutic and imaging systems, conventional steering consists of tilting the wave front focused on the initial source in order to focus nearby. For our purposes, if the skull is close to the array, conventional steering and amplitude compensation of the time-reversed signals coming from the initial source give a good focusing quality around the new desired focal point. However, this focusing degrades as soon as the array-skull distance increases. We apply the numerical backpropagation of the wave front received from the initial point source before steering and amplitude compensation.

I. TRANSMIT FOCUSING THROUGH THE SKULL

In this section, we review the results previously obtained in Ref. 14. First, the analytical derivation is extended to medium with density, sound velocity, and absorption heterogeneities. Second, new experimental materials along with more detailed description of the effects induced by the skull are presented. Third, the method used to extend the time reversal processing into the case of absorbing layer and its limitations are reviewed.

The basic theory of the time-reversal process is based on an elementary property of wave equations in a lossless medium: The time derivative appears only at the second order in this wave equation. Indeed, in a fluid nonabsorbing medium with a sound velocity $c(r)$, a density $\rho(r)$ that varies with space, the acoustic pressure field $p(r,t)$ satisfies:

$$\rho(\mathbf{r})\mathrm{div}\left(\frac{\mathbf{grad}(p(\mathbf{r},t))}{\rho(\mathbf{r})}\right)-\frac{1}{c(\mathbf{r})^2}\frac{\partial^2 p(\mathbf{r},t)}{\partial t^2}=0. \tag{1}$$

Whatever the medium, the wave equation contains only a second-order time-derivative operator: Thus if $p(r,t)$ is a solution of Eq. (1), then $p(r,-t)$ is also another solution. If $p_{r0}(r,t)$ is the resulting pressure field that propagates from a single source located in the nonabsorbing medium, the optimal way to focus on this source consists of a time-reversal operation in the whole 3D volume, generating $p_{r0}(r,T-t)$ in the medium, where T is a time delay that ensures causality. Using Huygen's principle we reduce the time-reversal operation on the 3D volume to a time-reversal operation on a two-dimensional closed surface, and develop there the concept of closed time-reversal cavity.[16] However, a closed cavity is difficult to realize experimentally. A compromise consists in applying the time-reversal process to a portion of the cavity such as a two-dimensional piezoelectric transducers array.

However, this property becomes obsolete in an absorbing medium: Acoustic losses are taken into account in the wave equation by a first derivative in time and are not time-reversal invariant. Indeed, one way to introduce losses is to modify the equation of state to allow for a delay between the application of a sudden pressure change and the attainment of the resulting equilibrium condensation.[17] This modification yields a third term in the wave equation for an absorbing fluid medium:

$$\rho(\mathbf{r})\left(1+\tau(\mathbf{r})\frac{\partial}{\partial t}\right).\mathrm{div}\left(\frac{\mathbf{grad}(p(\mathbf{r},t))}{\rho(\mathbf{r})}\right)-\frac{1}{c(\mathbf{r})^2}\frac{\partial^2 p(\mathbf{r},t)}{\partial t^2}=0, \tag{2}$$

where τ is the relaxation time of the medium.

Time reversal of the initial wave front is no longer the optimal solution to focus on the source. However, in a particular case, when the aberrating and absorbing medium can be modeled as a thin layer placed at some distance from the transducers array, it is possible to take into account the amplitude modulation originating from acoustic losses and so, to improve the quality of TRM focusing. This assumption of very thin absorbing layer will be discussed and justified in the next part. Afterward, we assume that this condition is fulfilled by the skull.

A. Degradation of time-reversal focusing induced by the skull

In this paper, experimental data have been obtained with a real time electronic prototype previously described in Ref. 14. This electronic prototype is used with a prefocused 1-D cylindrical array of transducers. It is made of 128 piezoelectric elements working at 1.5 MHz. Each transducer element is prefocused in elevation at 100 mm and its aperture is 1 mm wide and 25 mm high. The array pitch is 1.2 mm and the elements are set out on a spherical backing of 100 mm radius of curvature. The total aperture of the array of transducers is thus equal to 140 mm. The time-reversal experiments were performed using a rectangular transducer, 1 mm wide and 10 mm high, as an active transmitting line source. This same monoelement was used in receive mode as a hydrophone to scan the pressure field. The array of transducers and the monoelement were immersed in a water tank.

The time reversal process was performed in two steps: In the first step, the monoelement worked in transmit mode. The pressure field generated by this source was recorded on the 128 transducers of the array. In a second step, the signals

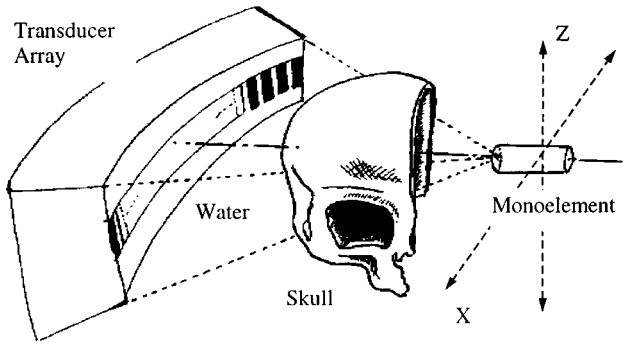


FIG. 1. Experimental setup.

recorded were time reversed and transmitted simultaneously by the 128 elements. The monoelement was then used as a hydrophone to scan the pressure field along the x axis (Fig. 1). It was moved by stepper motors to 30 mm on each side of the initial location of the source with a 0.3-mm scanning step, in order to plot the directivity pattern which represents the maximum of the pressure field for each location of the hydrophone.

In a first experiment, the optimal focusing quality of our system was measured. For this purpose a TRM experiment was conducted in homogeneous medium. In this case, time-reversal process was only used to compensate the geometrical distortion, if any, of the array of transducers. Figure 2(a) shows the waveforms received in homogeneous medium. Signal amplitude is shown linearly on a gray scale, the horizontal axis is the transducer number, and the vertical axis is the evolution in time. As the monoelement was placed at the center of curvature of the transducers array, all signals arrive at the same time on the transducers.

Figure 3 shows the directivity patterns (gray line) with the time-reversal procedure in homogeneous medium (pure water). One can notice the sharp focusing obtained with a background level at -40 dB and a beamwidth at -6 dB of 1.2 mm.

In the next experiments the source was observed through a half-skull (human). The dried skull was cut in the midsagittal plane. Afterward, it was degased and stored in a water tank. The skull was set between the source and the array of transducers, its outer side at a distance of 40 mm from the array. Conventional focusing consists in applying time shifts on each transducers of the array when assuming the medium to be homogeneous. In order to evaluate the beam degradation induced by the skull, when the conventional focusing is applied, the signals recorded in homogeneous medium were time reversed and transmitted a second time. The resulting wave front was distorted by the skull and the directivity pattern obtained is shown in Fig. 3 (dashed line). We can observe a 1-mm shift of the focus and the beam is widely spread. The -6-dB lateral resolution is about 4.8 mm and the side lobe level is very high.

Finally, in order to compare conventional and time-reversal focusing through the skull, a complete time-reversal experiment was performed: The monoelement was used as an active acoustic source. The generated wave front propagated through the skull and was distorted. The signals measured by the 128 transducers are presented in gray scale [Fig. 2(b)]. One can notice the strong phase and amplitude aberration induced by the skull. The sound velocity in the skull is greater than in water: Hence, the most attenuated part of the signal passed through a thicker part of the bone and arrived first at the transducers. This wave front was time reversed and transmitted through the skull. The directivity pattern is

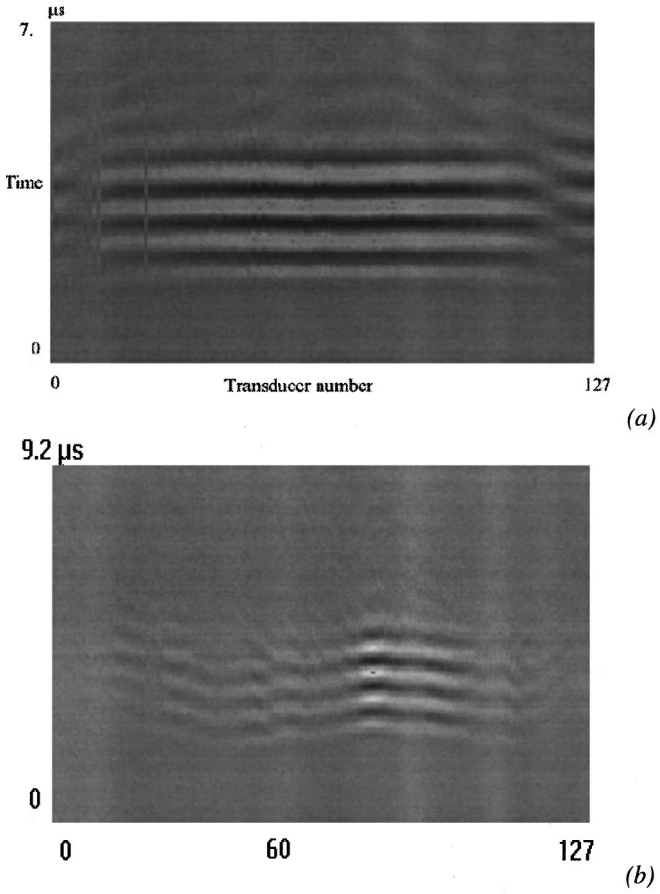


FIG. 2. Waveforms received on the transducers array: (a) in homogeneous medium; (b) through the skull. Signal amplitude is shown linearly on a gray scale, the horizontal axis is the transducer number, and the vertical axis is the evolution in time.

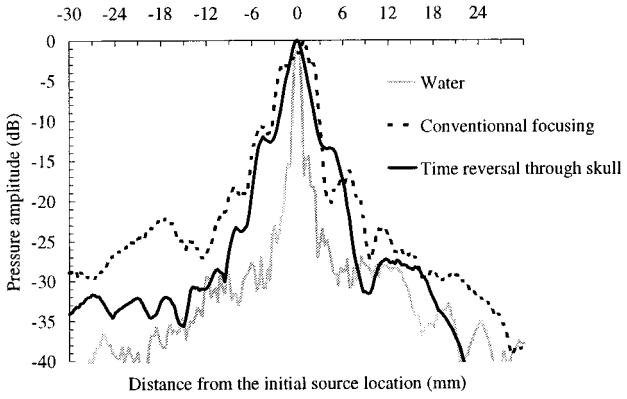


FIG. 3. Directivity patterns obtained by time reversal through pure water (gray line), time reversal through the skull of signals recorded in water (dashed line) and time reversal through the skull (solid line).

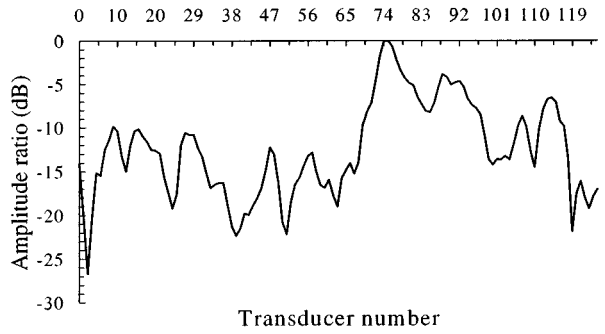


FIG. 4. Amplitude ratio between the waveforms received through the skull and the waveforms obtained in pure water.

shown Fig. 3 (solid line). This experiment shows that TRM focusing is slightly better than conventional focusing but remains far from the optimum. Indeed, the previous focus shift has been removed and the sidelobes level has been reduced, but the beam remains very large; the -6-dB beamwidth is about 3.6 mm.

We have seen that the acoustic losses induced by the skull are not time reversal invariant, and thus degrade the quality of time-reversal focusing. In practice, a transducer that received a weak signal was located in front of a strongly absorbing area of the skull. After the time-reversal operation, this weak signal backpropagated through the same area of the skull and hence, the amplitude modulation of the wave front was increased by a power of two. Thus the amplitude aberration due to the absorption in the skull was not compensated by the time-reversal process. At this point, Fig. 4 presents the amplitude ratio between the wave front propagated through the skull [Fig. 2(b)] and the one obtained through water only [Fig. 2(a)]. It shows that the losses due to absorption in the skull can vary on a scale of 25 dB.

B. Improvement of time-reversal focusing by amplitude compensation

It has been shown in a previous paper[14] that to obtain a good quality of focusing through the skull an amplitude compensation could be added to the time reversal processing. We will now review this method and the assumption it is based on. A comparison between the signals received in homogeneous medium [Fig. 2(a)] and those recorded through the skull [Fig. 2(b)], shows that the main qualitative difference is a varying time shift and attenuation. Particularly, there is no visible replica of the wave front and hence multiple scattering inside the skull is negligible. In these conditions, the simplest modeling of the skull is to assume it to a thin time shifting and absorbing layer. Thus the first method developed in Ref. 14 to compensate for this layer is to add to the time-reversal process a pretreatment taking partially into account the absorption. However, in all this paper, the absorption is assumed to be constant over the whole bandwidth of the transducers. In others words, only the spatial dependence of the absorption is taken into account. Without this strong approximation a compensation of the absorption at each frequency would be required, i.e., a deconvolution in the time domain, and the method would have to deal with the numerical instability encountered in both inverse problems: temporal and spatial.

Under these assumptions skull crossing does not modify the shape of the wave field. In other words, *if the skull is located close to the array*, the signals $s_i(t)$ received on each transducer i from an initial point source are only time delayed with a time shift τ_i, and multiplied by an absorption factor a_i, in comparison with the signals $s_i^h(t)$ that would be received in a homogeneous medium:

$$s_i(t)=(a_i\cdot\delta(t-\tau_i))\otimes s_i^h(t), \tag{3}$$

$$s_i(t)=a_i\cdot s_i^h(t-\tau_i). \tag{4}$$

This assumption of very thin layer—''phase and amplitude screen''—can now be quantitatively justified. For our purposes, Eq. (4) can be rewritten in the Fourier domain:

$$\mathbf{S}_i(f)=a_i\cdot\mathbf{S}_i^h(f)\cdot e^{j2\pi f\tau_i}. \tag{5}$$

Then, one can remark that the phase of the Fourier transform $\mathbf{S}_i(f)$ of each signal $s_i(t)$, namely, $\phi_i(f)$, has to depend linearly on the frequency f in comparison with $\phi_i^h(f)$, phase of the Fourier transform $\mathbf{S}_i^h(f)$ of the signals obtained in homogeneous medium $s_i^h(t)$:

$$\Delta\phi_i(f)=\phi_i(f)-\phi_i^h(f)=2\pi f\tau_i. \tag{6}$$

This phase difference at each frequency f was experimentally measured for each channel i and was found linearly dependent on the frequency in the limited bandwidth of the transducers. The mean standard deviation for the experimental phase difference $\Delta\phi_i(f)$ with a fitted straight line is equal to 0.118 rads.

This deviation is very small ($\sim T/50$) and we can conclude that the signals received on the array after propagation through the skull are strongly correlated with the ones obtained in the same conditions in homogeneous medium. The assumption of skull as a very thin layer is then justified.

Now, how can we correct the aberrations (a_i,τ_i) induced by the skull? The phase aberration correction is automatically accomplished by the time-reversal process and we can take into account the amplitude aberration due to attenuation by inverting the amplitude modulation of the wave field received on the array of transducers. For this purpose, the amplitude modulation a_i due to skull attenuation was estimated by comparison with the reference waveform obtained in the same conditions in homogeneous medium: Accordingly to Eq. (4), we deduce $a_i=\max_t\{|s_i(t)|\}/\max_t\{|s_i^h(t)|\}$ a ratio between the amplitudes of the wave packets (another choice could be $|\mathbf{S}_i(f=f_0)|/|\mathbf{S}_i^h(f=f_0)|$, where f_0 is the central frequency of the transducers). These estimates were then employed to invert the amplitude modulation a_i and gave the new set of corrected signals $s_i^c(t)$ to be reemitted:

$$s_i^c(t)=\overbrace{\frac{\max_t\{|s_i(t)|\}^2}{\max_t\{|s_i^h(t)|\}^2}}^{\text{Amplitude Compensation}}\cdot\underbrace{s_i(T-t)}_{\text{time reversal}}, \tag{7}$$

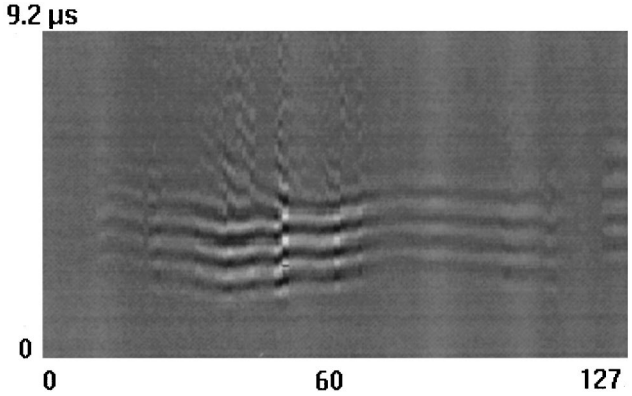


FIG. 5. Waveforms previously presented in Fig. 3 after amplitude compensation.

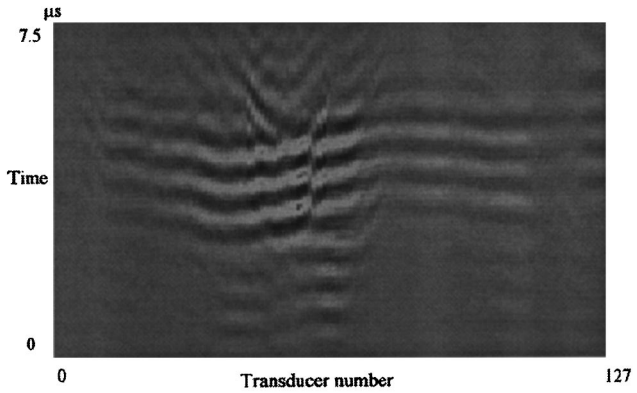


FIG. 6. Waveforms previously presented in Fig. 2(b) after numerical propagation and amplitude compensation.

$$s_i^c(t) = \frac{1}{a_i^2} \cdot s_i(T-t) = \frac{1}{a_i} \cdot s_i^h(T-t-\tau_i), \tag{8}$$

where T ensures the causality.

These new signals are reemitted by the 128 transducers of the array. As the distorting thin layer is located close to the array, the signal reemitted by each transducer i is immediately modified and became a set $s_i^{a,s}(t)$ after passing through the skull:

$$s_i^{a\,s}(t) = a_i \cdot \delta(t-\tau_i) \otimes s_i^c(t), \tag{9}$$

$$s_i^{a\,s}(t) = s_i^h(T-t). \tag{10}$$

Now we come back to a time-reversal operation in a nondissipative medium and according to the time-reversal invariance, this wave front propagates in homogeneous medium and optimally focuses on the initial source point. Figure 5 shows the wave front, recorded during the previous experiment and presented in Fig. 2(b), after this amplitude compensation.

The amplitude compensation improves the focusing and the beamwidth is perfectly corrected up to -14 dB (Fig. 7, dotted line). Nevertheless, this amplitude compensation is quite efficient only when the absorbing medium is a layer *located close to the array*. In fact, the skull-array distance was about 40 mm in this experiment and hence, additional amplitude and shape distortions may develop as the wave front propagates in the homogeneous and lossless medium after passing through the skull. These amplitude distortions are automatically compensated by the time-reversal processing and so must not be taken into account a second time in the pretreatment. The idea is to simulate a virtual array of transducers located close to the skull. For this purpose, the original diverging wavefront represented by $s_i(t)$ was time reversed and numerically backpropagated from a 100 mm to a 60 mm radius of curvature cylindrical surface respectively corresponding to the array aperture and the skull surface.[18] Indeed, since the medium between the array and the skull is lossless, the time-reversal processing achieves an inverse wave field propagator. Moreover this time-reversal processing can be easily computed as soon as the medium is assumed to be homogeneous. Amplitude distortion was then estimated on the virtual array and the estimates were used to compensate for the backpropagated field. This compensated and converging waveform can now be time reversed and numerically propagated from the approximated skull surface to the array aperture. This method, by simulating a virtual array located close to the skull, allows discrimination between these two kinds of amplitude modulations, respectively, due to the absorption in the skull and the propagation from the skull to the array. Thus we obtained a new set of 128 signals adapted to focus through the skull at the initial source location (Fig. 6).

These signals were time reversed and transmitted in the medium. A comparison of the focusing results obtained with amplitude compensation in the aperture (Fig. 7, dotted line) and with amplitude compensation after backpropagation (Fig. 7, solid line) can be made.

The focusing is now optimal up to -16 dB and the sidelobes level has been strongly decreased. This quality of focusing seems quite good for hyperthermia where the heating pattern contrast is proportional not to the pressure amplitude but rather to the ultrasound energy.

II. STEERING THROUGH THE SKULL

Although time reversal is only self-focusing onto the artificial source implanted inside the treatment volume, we

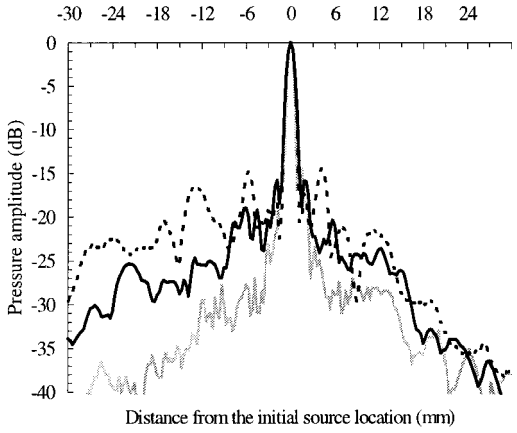


FIG. 7. Directivity patterns obtained by time reversal through pure water (gray line), time reversal with amplitude compensation (dotted line), and time reversal with backpropagation and amplitude compensation (solid line).

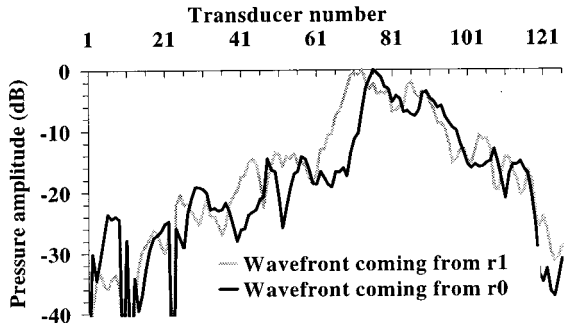


FIG. 8. Comparison between the amplitude modulation of the wave front coming from the new focal point r_1 [presented in Fig. 9(a)] and the wavefront coming from the initial source location r_0 [previously presented in Fig. 2(b)].

show how the array can be steered on points surrounding this beacon. In homogeneous medium, conventional beamsteering modifies the wave front focused on the artificial source by applying a time delay law and apodization factors, taking into account the distances between the new focus location and the array elements. Note, this process requires only to know the sound speed of the medium. However, in heterogeneous medium, this process is only valid for small angles or when the aberrations are located in a layer close to the array. Indeed, if the aberrating layer is modeled as a thin layer located close to the array, the aberrations (a_i, τ_i) induced by the skull are independent of the source location, and then the Green's functions of two different points r_0 and r_1 implanted in the medium only differ from the same time shifts and apodization factors as in homogeneous medium. On the other hand, if the thin aberrating layer is now located at some distance from the array of transducers, the two Green's functions associated to r_0 and r_1 are different: The propagation between the layer and the array modifies in a different way for these two wave fronts. The wave front coming from r_0, the artificial source location, was previously presented in Fig. 2(a) [the one coming from r_1, when located 10 mm away from r_0, is presented in Fig. 9(a)]. One can remark, as shown in Fig. 8, that the amplitude modulation of these two wave fronts are different. The two amplitude modulations, which represent the maximum of pressure measured on each transducer, are very similar, but translated in space. This translation can be easily understood by a ray theory approach.

In order to estimate the time of arrival error introduced by conventional steering, the wave front coming from r_0 was steered on r_1 assuming that the heterogeneities are located close to the array. (The difference between the time of arrival of the wave front coming from r_0 tilted to focus on r_1 and the genuine wave front coming from r_1 is presented in Fig. 11, gray line.) This time of arrival error is measured by intercorrelation between the two wave fronts, and the variance of the error is equal to 0.16 μs (to be compared with the central frequency of the signal, which is equal to 0.66 μs^{-1}). Thus as soon as the skull is not located close to the array, conventional steering of an initial wave front is not efficient if we want to construct the Green's function associated to a new point of interest.

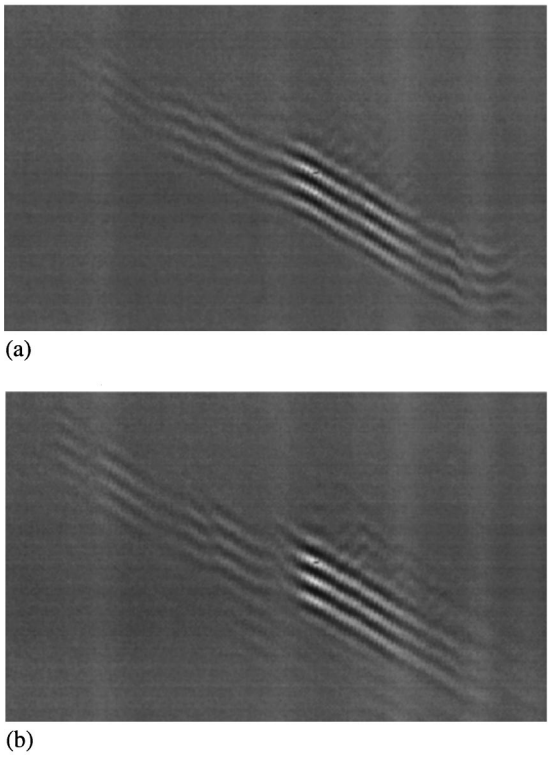


FIG. 9. Comparison between (a) the acquired Green's function of the new point of interest r_1 and (b) this Green's function calculated from the knowledge of the wavefront coming from the initial source location r_0.

Nevertheless, as well as for the amplitude compensation presented in the second part, numerical backpropagation can extend the conventional steering to aberrating layers located at any distance from the array:[15] The initial wave front coming from r_0 [Fig. 2(b)] was numerically backpropagated from the array to the surface of the skull. This wave front was then tilted in order to focus on r_1, time reversed, and numerically backpropagated from the skull to the array. Thus from the knowledge of the Green's function associated to the initial point r_0, we obtained a new calculated Green's function associated to r_1 [Fig. 9(b)].

As previously, we can compare the acquired and calculated Green's function associated to r_1. As shown in Fig. 10, their amplitude modulations are very similar. Furthermore, the variance of the time delay error is equal to 0.05 μs, and

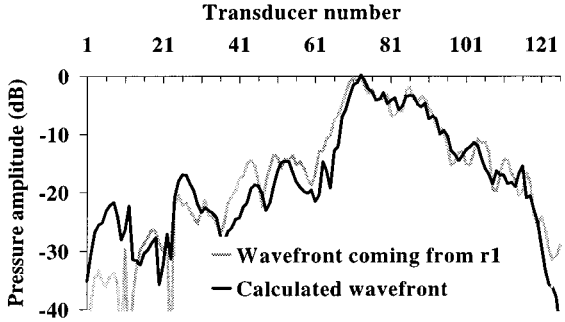


FIG. 10. Comparison between the amplitude modulation of the wave front coming from the new focal point r_1 [Fig. 9(a)] and the wave front calculated in order to focus on r_1 [Fig. 9(b)].

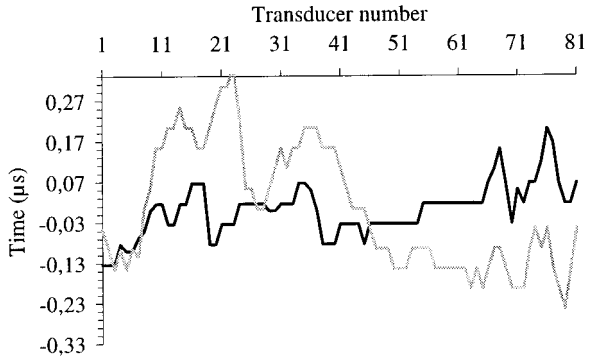


FIG. 11. Time shift error measured on each transducer of the array between the acquired wavefront coming from r_1 and this wavefront calculated without backpropagation (gray line) and with backpropagation (black line) from the knowledge of the wavefront coming from the initial source location.

thus is three times better than the variance obtained without backpropagation. One can notice that a maximum variance of $T/8$ (0.08 μs) is required to ensure a nice focusing. The time delay law error between the acquired and calculated wave fronts is presented in Fig. 11 (black line).

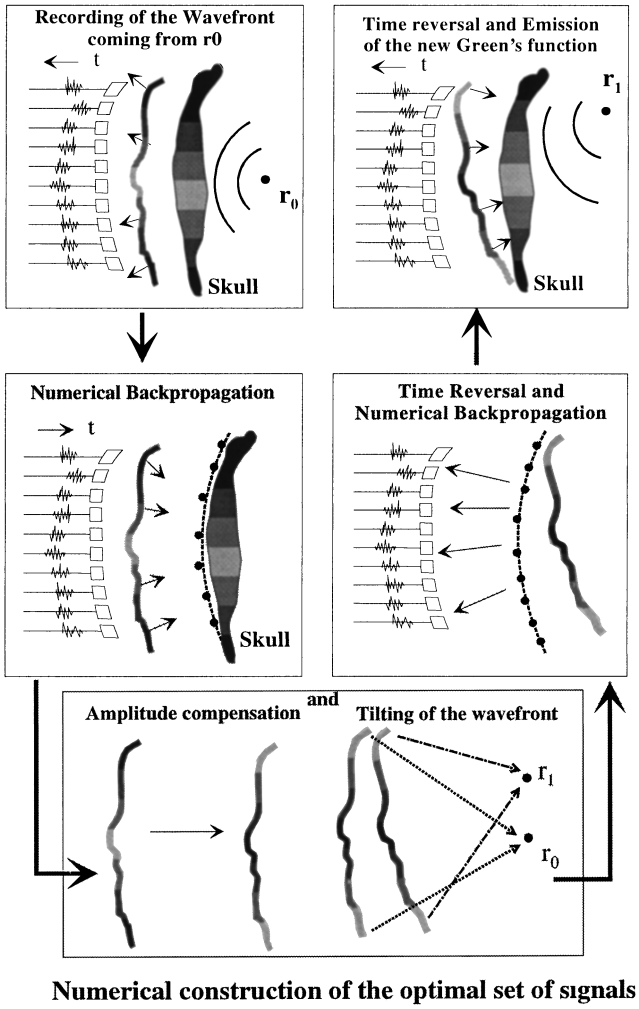
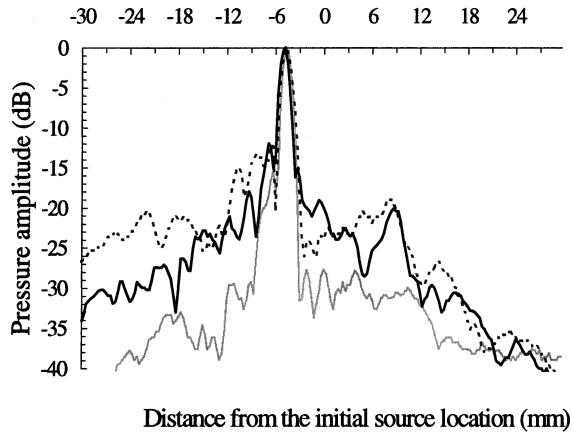
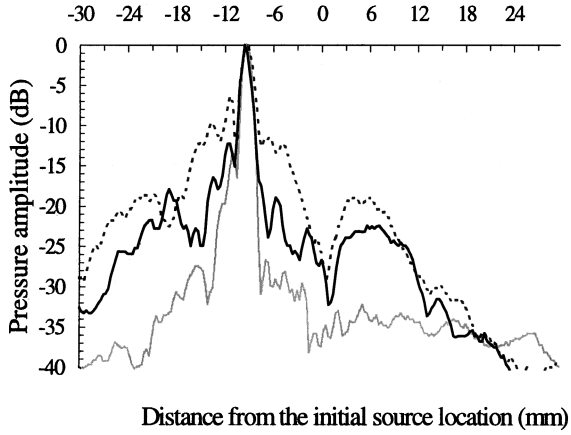


FIG. 12. Numerical construction of the Green's function associated to the point r_1 from the knowledge of the Green function associated to the initial point location r_0.





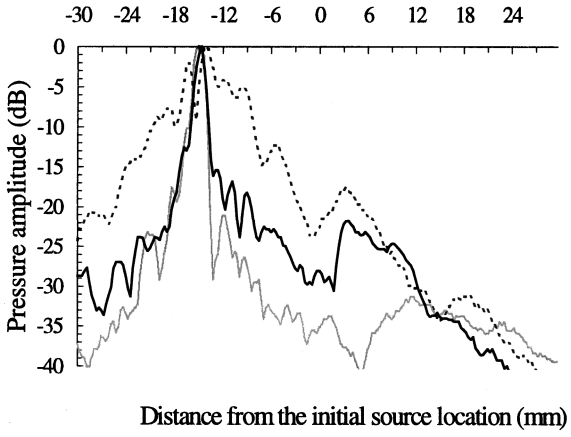


FIG. 13. Directivity patterns obtained by tilting the wavefront received from the initial point source at a distance of (a) 5 mm, (b) 10 mm, and (c) 15 mm: Tilting through pure water (gray line), time reversal of the signal received through the skull with amplitude compensation (dotted line), time reversal of the amplitude compensated signal after backpropagation (solid line).

By combining the amplitude compensation presented in the second part and the steering of the wave fronts presented in this third part, we were now able to construct the optimal set of signals in order to focus around the initial source location (Fig. 12). This process was applied to focus at 5, 10, and 15 mm away of the initial source location. We compared focusing obtained by tilting of the wave front in homogeneous medium with, on the one hand, tilting and amplitude

compensation of the wave front in the array aperture and, on the other hand, tilting and amplitude compensation of the wave front after numerical backpropagation to the skull. Directivity patterns are presented in Fig. 13. Without backpropagation, the focusing was quickly degraded as the distance from the initial source location was increased, thus sidelobes appeared, respectively, at -12 dB, -6 dB, and -2 dB (Fig. 13, dotted line). On the other hand, the focusing remained quite good when backpropagation to the skull-array distance was applied before amplitude compensation and tilting of the wave front (Fig. 13, solid line). Now, sidelobes appeared, respectively, at -12 dB, -12 dB, and -14 dB. Thus conventional beamsteering that consists of tilting the wave front without backpropagation is no longer efficient when the aberrating and absorbing layer is located some distance away of the array of transducers. Beside this limit of 15 mm, the focusing degrades more and more, even in homogeneous medium. This degradation is due to the directivity of each tranducer of the array, and we attain the maximum steering abilities of the array. Moreover in our experiment the skull is not a perfect thin layer and refraction effects are neglected which limits the maximum steer angle. Thus we are able to focus the ultrasonic beam in an area of 15 mm on each side of the initial source location in the lateral direction. Moreover, a depth of 30 mm on each side of the initial source in the axial direction can be investigated.

III. CONCLUSION

In this paper, we have shown that focusing and steering of an ultrasonic beam through the skull is possible in spite of the severe ultrasonic attenuation of the bone. We propose a method of steering combined with an amplitude compensation to improve the time-reversal focusing when an artificial acoustic source is available in the medium. Thus although time reversal is only self-focusing onto this initial source implanted inside the treatment volume, we learn how to focus the ultrasound beam onto many new others focal points in order to investigate the whole tumor. A new set of signals is calculated from the knowledge of the initial ones corresponding to the artificial source location. The proposed method is valid when the inhomogeneities can be modeled as a thin screen varying in absorption and time delay, located at any depth from the array. Experimental results show that this assumption can be done in the case of the skull.

This method could be easily improved. First, our numerical backpropagation could be improved by the acquisition of the real skull surface with a MRI or x-ray system. Second, although we used in this study only a one-dimensional array and since significant aberration exists both in elevation and azimuth, aberration correction should be performed in both dimension to be most effective. It will be done in our further work.

We want now to apply this adaptive focusing method to a new time-reversal mirror developed in our laboratory. This time-reversal mirror will be able to produce high intensity ultrasonic beams.

[1] K. Hynynen, A. Chung, T. Fjield, M. Buchanan, D. Daum, V. Colucci, P. Lopath, and F. Jolesz, ''Feasibility of using ultrasound phased arrays for MRI Monitored noninvasive surgery,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 1043 (1996).

[2] R. S. Foster, R. Bihrle, N. T. Sanghvi, F. J. Fry, and J. P. Donohue, ''High-intensity focused ultrasound in treatment of prostatic disease,'' Eur. Urol. **23**, 29–33 (1993).

[3] D. N. White, J. M. Clark, J. N. Chesebrough, M. N. White, and J. K. Campbell, ''Effect of skull in degrading the display of echoencephalographic B and C scans,'' J. Acoust. Soc. Am. **44**, 1339–1345 (1968).

[4] M. O'Donnell and S. W. Flax, ''Phase aberration measurements in medical ultrasound: Human studies,'' Ultrason. Imag. **10**, 1–11 (1988).

[5] Q. Zhu, B. D. Steinberg, and R. Arenson, ''Wavefront amplitude distortion and sidelobes levels-Part I: Theory and computer simulations,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **40**, 747–753 (1993).

[6] Q. Zhu, B. D. Steinberg, and R. Arenson, ''Wavefront amplitude distortion and sidelobes levels-Part II: *In vivo* experiments,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **40**, 754–762 (1993).

[7] F. Wu, M. Fink, R. Mallart, J.-L. Thomas, N. Chakroun, D. Cassereau, and C. Prada, ''Optimal focusing through aberrating media: A comparison between time reversal mirror and time delay correction technique,'' in Proc. IEEE 1991 Ultrason. Symp. **2**, 1195–1199 (1991).

[8] F. Wu, J.-L. Thomas, and M. Fink, ''Time reversal of ultrasonic fields-Part II: Experimental results,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 567–578 (1992).

[9] D.-L. Liu and R. C. Waag, ''Correction of ultrasonic wavefront distortion using backpropagation and a reference waveform method for time-shift compensation,'' J. Acoust. Soc. Am. **96**, 649–660 (1994).

[10] M. Fink, C. Prada, F. Wu, and D. Cassereau, ''Self-focusing with time reversal mirror in inhomogeneous media,'' Proc. IEEE 1989 Ultrason. Symp. **2**, 681–686 (1989).

[11] M. Fink, ''Time reversal of ultrasonic fields-Part I: Basic principles,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 555–566 (1992).

[12] C. Prada, F. Wu, and M. Fink, ''The iterative time reversal mirror: A solution to self-focusing in pulse-echo mode,'' J. Acoust. Soc. Am. **90**, 1119–1129 (1991).

[13] E. Ebbini, H. Wang, M. O'Donnell, and C. Cain, ''Acoustic feedback for hyperthermia phased-array applicators: Aberration correction, motion compensation and multiple focusing in the presence of tissue inhomogeneities,'' Proc. IEEE 1991 Ultrason. Symp. **2**, 1343–1346 (1991).

[14] J-L Thomas and M. Fink, ''Ultrasonic beam focusing through tissue inhomogeneities with a time reversal mirror: application to transskull therapy,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 1122–1129 (1996).

[15] C. Dorme and M. Fink, ''Ultrasonic beam steering through inhomogeneous layers with a time reversal mirror,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 167–175 (1996).

[16] D. Cassereau and M. Fink, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 579 (1992).

[17] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics* (Wiley, New York, 1982).

[18] M. Fink and C. Dorme, ''Phase aberration correction with ultrasonic time reversal mirrors,'' Proc. IEEE Ultrason. Symposium **3**, 1629–1638 (1994).

Line-focusing of ultrasonic *SV* wave by electromagnetic acoustic transducer

Hirotsugu Ogi[a] and Masahiko Hirao
Graduate School of Engineering Science, Osaka University, Machikaneyama 1-3, Toyonaka, Osaka 560-8531, Japan

Toshihiro Ohtani
Ebara Research Co., Ltd., Fujisawa, Kanagawa 251-8502, Japan

(Received 10 July 1997; accepted for publication 27 January 1998)

An electromagnetic acoustic transducer has been developed for line-focusing the shear-vertical (*SV*) wave in a metal. The EMAT consists of a permanent magnet to supply the bias magnetic field normal to the surface, and a meanderline coil to induce the dynamic field and eddy currents in the surface region of the sample. The meanderline spacing is continuously changed so that the generated *SV* waves from all segment sources become coherent on the focal line after traveling oblique paths. The printed circuit technique enables the fabrication of such a functionally spaced meanderline coil within 1-μm accuracy. The directivity is measured using a half-cylindrical sample of an aluminum alloy, which shows a much sharper radiation pattern of the focusing EMAT than the EMAT having a meanderline coil of constant spacings. This EMAT is then tested through detecting a shallow notch. The scattered signal from the notch is received by the same EMAT, which shows enough strength even for the notch of 0.2-mm depth. © *1998 Acoustical Society of America.*
[S0001-4966(98)03205-6]

PACS numbers: 43.35.Zc [HEB]

INTRODUCTION

Electromagnetic acoustic transducers (EMATs)[1–5] have definite advantages for the noncontact inspection and characterization of electrically conductive materials. They excite and detect the elastic waves through the electromagnetic transduction and they are free from any mechanical coupling, which eliminates the measurement errors associated with the contact transduction. An EMAT consists of permanent magnets (or electromagnets) to supply the bias magnetic field, and the driving coil element to excite the eddy currents and the dynamic fields. Their configuration depends on the mode of elastic wave desired.

An EMAT, however, has two basic disadvantages; the weak transduction efficiency and the broad radiation pattern of elastic waves into a solid. However, the sensitivity of the EMAT has been dramatically improved owing to the recent advances in the electronics and signal processing techniques, the permanent magnet such as the Nd–Fe–B sintered magnet, and the printed circuit techniques to fabricate the printed coils. Especially, the use of an EMAT in the ultrasonic resonant technique[6–9] overcame the low transduction efficiency by receiving highly and coherently overlapping echoes at a resonant state. On the other hand, the elastic waves generated by an EMAT propagate nearly all possible directions in the sample, which is undesirable for flaw detection purposes. For example, Fig. 1 illustrates the elastic-wave radiation by the EMAT built with a meanderline coil and a permanent magnet. The EMAT generates the logitudinal wave and the shear-vertical (*SV*) wave propagating obliquely in the sample, and the Rayleigh wave along the surface. They are generated at the same time for both sides with the broad directivity patterns and will be reflected not only by flaws but also by the sample edges. It is often difficult to distinguish the flaw signal from others. Moreover, the broad radiation decreases the ultrasonic energy to be concentrated on the target flaw.

This paper presents a technique to sharpen the directivity by line-focusing the *SV* waves. The major radiation direction by the EMAT in Fig. 1 depends on the driving frequency, the meanderline spacings, and the wave modes.[10] We choose to focus the *SV* waves at a fixed frequency by continuously changing the meanderline spacing so that the excited *SV* waves become coherent on a focal line.

Another technique is possible for focusing the shear-horizontal (*SH*) wave by the periodic-permanent-magnet EMAT (PPM-EMAT),[5] which consists of the elongated spiral coil and the periodically arrayed thin permanent magnets. The direction of the *SH* wave is governed by the spatial period of the magnet array and the frequency. The *SH* wave can be concentrated on a focal line by changing the magnet thickness. However, this method for focusing the *SH* wave is impractical for many purposes, because the megahertz-range *SH* wave needs the magnets thinner than 1 mm, which are intolerable to even a small liftoff. Besides, the present focusing technique is superior at two points. One is the high accuracy for making the meanderline coil. Unlike the permanent magnet case, the spacing of the meanderline coil can be decreased as small as the fabrication accuracy. The recent printed circuit technique makes it possible to produce the meanderline coil of theoretically generated unequal spacing within 1-μm accuracy. The second advantage is in the use of the *SV* wave itself. The *SV* wave radiated from a line source

[a]Electronic mail: ogi@me.es.osaka-u.ac.jp

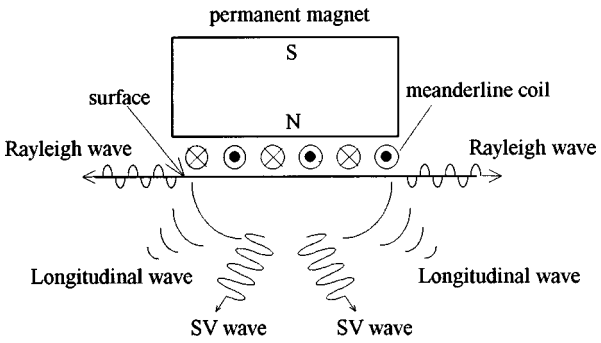


FIG. 1. Typical elastic wave radiation by an EMAT with a permanent magnet and a meanderline coil with equal spacings.

has a sharp directivity,[11] unlike the broad radiation patterns of other elastic modes.

I. DESIGN OF THE MEANDERLINE COIL

We consider a two-dimensional space, where the half-space of $z>0$ is filled with an electrically conducting material and the x-y plane coincides with the surface (Fig. 2). A meanderline coil and a permanent magnet are located near the surface. When the coil is driven by a current, the shearing forces parallel to the surface arise mainly under the straight parts of the coil due to the Lorentz force, the magnetization force, and the magnetostriction effect.[4,5] Only the Lorentz force is the possible source in a nonmagnetic metal, but all three contribute to producing the shearing force in a ferromagnetic metal.

When the meanderline coil has a constant spacing l, the angle θ for the major radiation is determined so that the elastic wave originated from each source becomes coherent:[10]

$$\sin\theta=\frac{c}{2lf}, \tag{1}$$

where c is the longitudinal or shear wave velocity and f is the frequency of the driving current. On the basis of this relation, we intend to focus the SV wave to a line in the sample using the variable spacing. Figure 3 illustrates the

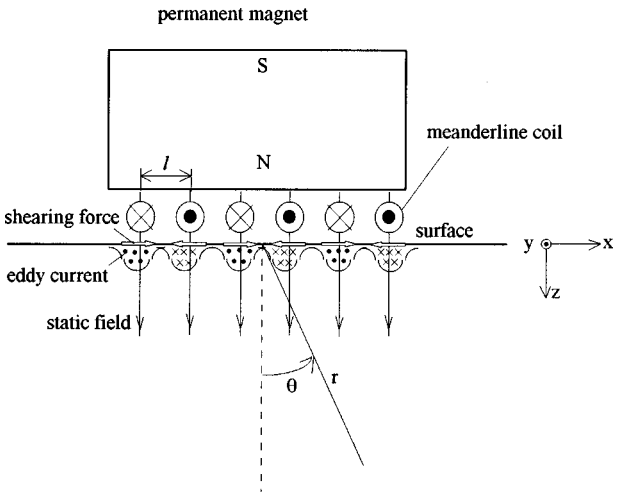


FIG. 2. Two-dimensional field for analyzing the radiation of the EMAT. Only Lorentz force is considered for simplicity.

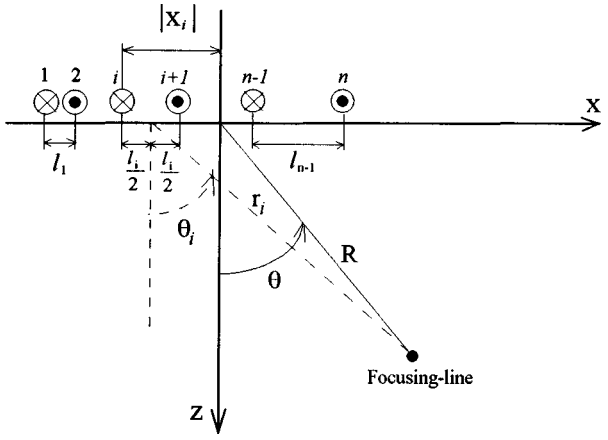


FIG. 3. Determination of the meanderline configuration.

determination of the meanderline spacing. First, we set the frequency, a focal line in the sample, and the edge of the coil at the most left side (i.e., setting f, θ, R, and x_1). The location of $(i+1)$th line (x_{i+1}) is determined so that the spacing l_i locally satisfies Eq. (1):

$$\sin\theta_i=\frac{c_S}{2l_if}=\frac{|R\sin\theta-x_i-l_i/2|}{\sqrt{(R\cos\theta)^2+(R\sin\theta-x_i-l_i/2)^2}}, \tag{2}$$

where c_S is the shear wave velocity. Equation (2) is solved for l_i by Newton's method with known x_i and then we can determine the detailed configuration of the meanderline coil. In this study, we set $f=4$ MHz, $\theta=40°$, $R=100$ mm, $x_1=-30.88$ mm, and $n=98$ with $c_S=3.2$ km/s. Figure 4 shows the meanderline spacings calculated for the above parameters.

We have fabricated the meanderline coil by printing copper on a polyimide sheet and then etching the copper following the designed pattern. This technique allows us to make a high-impedance coil of an arbitrary geometry with the accuracy of 1 μm. The liftoff of the sheet coil is only 25 μm, improving the weak coupling efficiency of the EMAT.

II. NUMERICAL SIMULATION OF RADIATION DIRECTIVITY

A numerical simulation is possible for estimating the radiation pattern of the EMAT. Considering a nonmagnetic

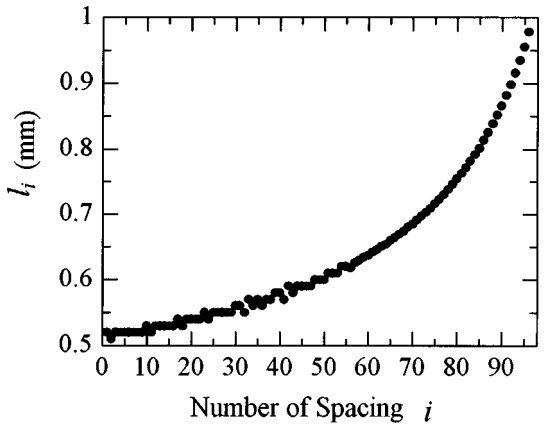


FIG. 4. Variable of the meanderline coil.

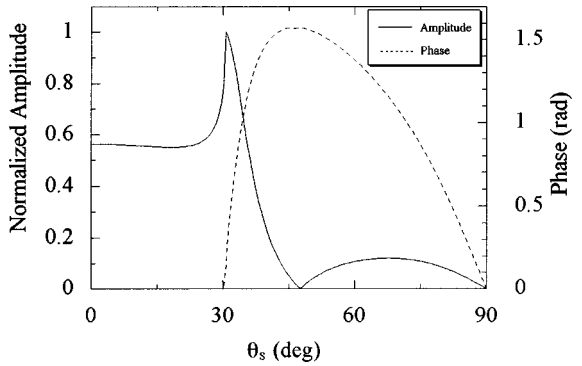


FIG. 5. Directivity of the amplitude and the phase of the *SV* wave radiated from the single strip source.

metal, the EMAT shown in Fig. 2 excites the shearing Lorentz force in the x direction in the surface region. When the sample material has a good conductivity, the shearing force takes place in the surface skin and can be considered as the traction force on the x-y plane.[12] Thompson[5] derived the Fourier transform expression for the periodic traction force and, by taking up the fundamental component, showed that it is proportional to $\cos(\pi x/l)$. We apply this formula to the small part of the coil, assuming the local uniformity of the spacings. On the surface region between the ith and $(i+1)$th lines, we approximate the traction force by the cosine curve with the characteristic length l_i (spatial frequency of $\pi x/l_i$).

We then digitize the traction-force distribution to a large number of strip sources at equal intervals to calculate the amplitude and the phase of the *SV* wave radiated from the EMAT. Each element source oscillates with the strength and phase prescribed by the cosine function. When a strip source, lying along the y axis at $x=0$ and having a very narrow width compared with the wavelength, vibrates along the x axis, the displacement u_θ perpendicular to the propagation direction of the excited *SV* wave is generated:[11]

$$u_\theta(r,\theta_S) = \frac{C}{\sqrt{r}} \frac{(\lambda+2\mu)\cos 2\theta_S \cos\theta_S}{2\mu \sin 2\theta_S \sin\theta_L \cos\theta_L + (\lambda+2\mu)\cos^2 2\theta_S} \times e^{j[(3/4)\pi - k_S r]}, \tag{3}$$

$$\cos^2\theta_L = 1 - \left(\frac{\lambda+2\mu}{\mu}\right)\sin^2\theta_S, \tag{4}$$

at a far distance r from the source. Here C is a constant, λ and μ are Lame's constants, $k_S(=\omega/c_S)$ is the wave number of the *SV* wave, and θ_S is the propagation direction of the *SV* wave. Figure 5 shows the normalized amplitude and the relative phase referring to $\theta_S=0°$ of the *SV* wave radiated from the single strip source as functions of θ_S for $c_S=3.2$ km/s, $\mu=26$ GPa, and $\lambda=58$ GPa. We see that the amplitude directivity has a sharp peak around $\theta_S=30°$, which corresponds to the critical angle of the longitudinal wave ($\theta_L=90°$), and all the exciting energy is directed toward the shear wave generation. Therefore, it is beneficial to locate the focal line so that all the sources take the angle θ_i centered around this angle, although Eq. (1) must be satisfied. The angle θ_i ranges from 23° to 51° with the meander-line coil designed in this study (Fig. 4).

By applying Eq. (3) to all the strip sources and numerically integrating their contributions,[13] we map the amplitude and the phase of the *SV* wave in the two-dimensional region. Figure 6 shows the result for the line-focusing EMAT (LF-EMAT) with the coil designing of Fig. 4 and, for comparison purposes, the constant-spaced EMAT (CS-EMAT) of $l=0.65$ mm. Attenuation is not included. The results are normalized by the maximum amplitude in the LF-EMAT case. We find that the LF-EMAT successfully focuses the *SV* wave at the position as expected ($x=64.2$ mm, $z=76.6$ mm), while the CS-EMAT produces the broad radiation pattern

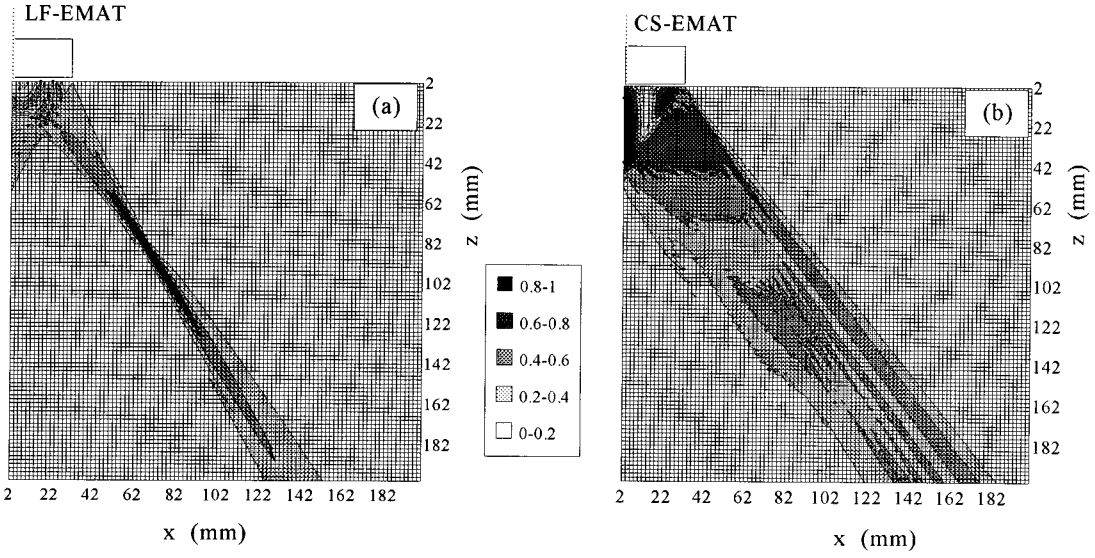


FIG. 6. Simulation of the two-dimensional distribution of the *SV*-wave amplitude radiated by (a) the LF-EMAT and (b) the CS-EMAT.

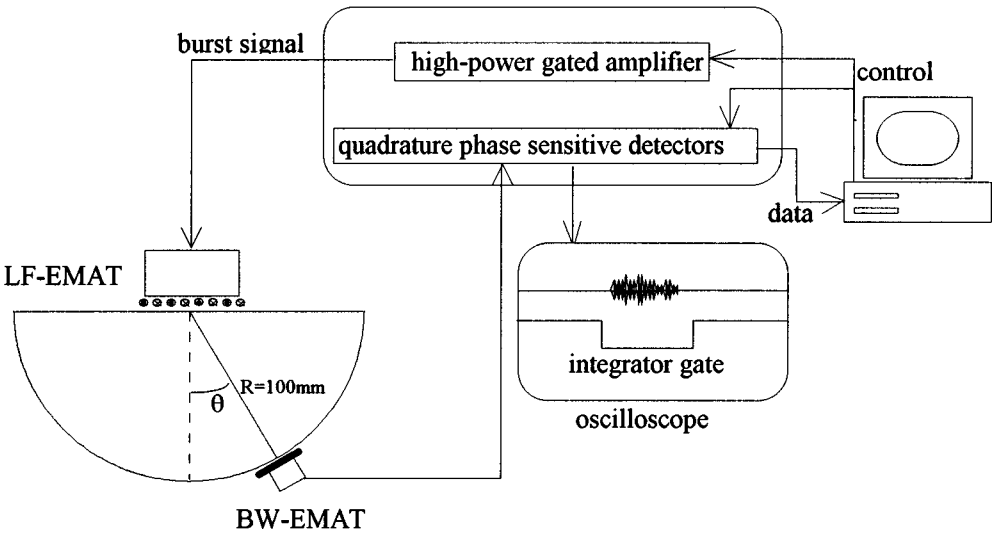


FIG. 7. Measurement setup for the directivity of the *SV*-wave amplitude.

and naturally the peak amplitude is smaller than the LF-EMAT case.

III. MEASUREMENT OF DIRECTIVITY

Figure 7 sketches the measurement setup for the radiation pattern of the EMATs. We use a half-cylindrical sample of an aluminum alloy (Al2017), whose diameter is 200 mm and axial length is 100 mm. The EMAT is located on the middle of the flat surface and is driven by high-power burst signals of 8-μs duration. The *SV* waves generated by the EMAT are received by the bulk-wave EMAT[6,8] located on the cylindrical surface. To produce a good spatial resolution for the directivity measurement, the coil face of the bulk-wave EMAT is shielded with copper sheets to make a 1-mm-wide aperture. The received signal is fed to the quadrature phase detectors to extract the in-phase and out-of-phase outputs relative to the driving signal by analog-integrator gates of 30-μs width. The amplitude is determined from the root of their squares. The details of electronics are found in the pervious work.[6]

The measurements have been performed for two EMATs, the LF-EMAT and the CS-EMAT ($l=0.65$ mm). Figure 8 shows the amplitude distribution as a function of the propagation direction. This includes the results of the numerical simulation, which shows a good agreement with the measurement, indicating the validity for estimating the radiation pattern. It is clear that the LF-EMAT is much superior to the CS-EMAT in the directivity sharpness. We also show the measured directivity pattern when the LF-EMAT is driven at 2.5 MHz. As the driving frequency decreases, the amplitude peak moves to larger angle region. This is easily understood from Eq. (1), but the directivity shows the sharpest peak at the designed angle and frequency.

We show the as-received signals at 4 MHz for three angles in Fig. 9. They also indicates the good directivity of the LF-EMAT. The signal from the LF-EMAT at the focusing angle ($\theta=40°$) is not flat, which is attributed to the transit time difference from the sources. The meanderline coil was about 60 mm long. The maximum difference occurs between the edge sources and it is estimated to be 12 μs. In the measurement, we used an 8-μs burst signal. Therefore, they are partially superimposed and the received signal takes the shape of a larger duration.

IV. DETECTION OF SHALLOW NOTCH

The LF-EMAT was then used for detecting the shallow notch on the opposite surface of an aluminum block (76-mm hight, 60-mm width, and 500-mm length) introduced by the electrodischarge technique. The notch was 20 mm long, 1.0 mm wide, and 0.2 mm deep. The notch was located at the distance of 111.5 mm from the edge of the block. The focal line of the EMAT lays on the opposite surface. By moving the EMAT on the upper surface in the length direction, we measure the amplitude of the scattering signal from the notch.

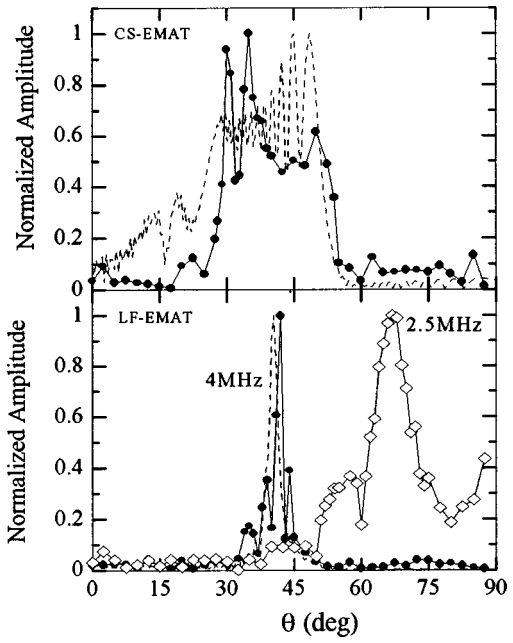


FIG. 8. Measurement and calculation of the directivity of the *SV*-wave amplitude. The solid circles are the data at 4 MHz and the open lozenges at 2.5 MHz. The broken lines show the calculation results at 4 MHz.

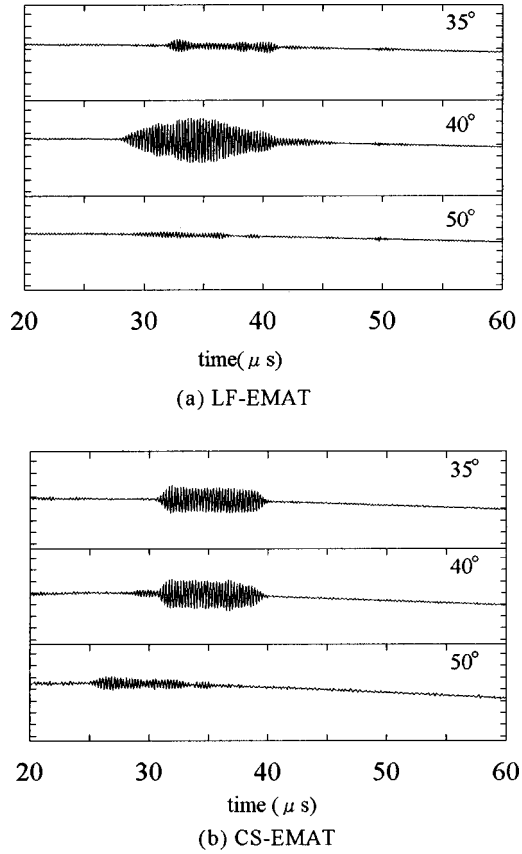


FIG. 9. Wave forms received at three angles for (a) the LF-EMAT and (b) the CS-EMAT.

The same EMAT receives the signals. The result is plotted in Fig. 10. If the focal line is away from the notch, the amplitude is as small as the noise level. Approaching the notch, the amplitude increases and, after taking the peak, it suddenly decreases back to the noise level. We show the maximum scattering signal in Fig. 11, where the signal has the flat shape and the duration is comparable to that of the input burst signals of 16 μs, unlike the previous case in Fig. 9(a). This can be explained as follows. The directivity pattern of the *SV* wave generated from the single source (Fig. 5) is also applicable to express the efficiency directivity in the receiving stage. The total transduction efficiency is governed by the square of this pattern and the peak around $\theta_S = 30°$ is much more enhanced. Therefore, the coil parts located near

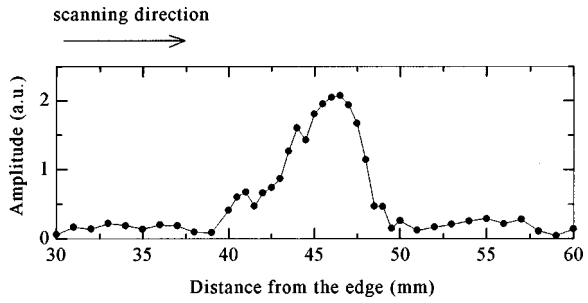


FIG. 10. Amplitude of the scattering echo from the notch detected by the LF-EMAT at 4 MHz.

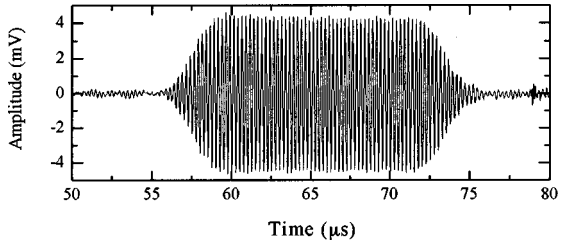


FIG. 11. Scattered echo from the notch of 0.2 mm deep.

this angle primarily contribute to the received signal. Considering only these parts, the transit time is approximately the same, which makes such a flat signal. We measured the transit time of the received signal through the phase[14] and obtained 57.8 μs. The distance between the line source located at $\theta_S = 33°$ and the notch bottom was 91.7 mm, and the transit time of the round trip is then calculated to be 57.3 μs with $c_S = 3.2$ km/s. This agreement supports the above explanation.

V. CONCLUSIONS

The line-focusing EMAT has been developed for concentrating the *SV* wave in a metal by designing the variable spacing of the meanderline coil. Owing to the high accuracy of fabricating the printing coil and the unique radiation pattern of the *SV* wave, the line-focusing EMAT establishes a sharper directivity pattern. The numerical simulation has been used for predicting the radiation pattern of the EMAT, which showed a good agreement with the measurement. The directivity measurement proved a sharper amplitude peak of the focusing EMAT than the EMAT with an equal-spacing coil. The measurement also indicated that the location of the focal line can be controlled through the frequency, though the directivity is broadened.

The line-focusing EMAT showed a pronounced sensitivity for detecting the shallow notch of 0.2-mm-depth on the opposite surface of a block. The echo from the notch was large enough and it will be measurable even for smaller notches. The line sources having the angle near 30° to the focal line give the largest contribution to the flaw detection.

[1] R. B. Thompson, IEEE Trans. Sonics Ultrason. **SU-25**, 7 (1978).
[2] W. B. Maxfield and C. M. Fortunko, Mater. Eval. **41**, 1399 (1983).
[3] K. Kawashima, IEEE Trans. Sonics Ultrason. **SU-32**, 514 (1985).
[4] A. Wilbrand, *Review of Progress in QNDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1987), Vol. 7A, p. 671.
[5] R. B. Thompson, *Physical Acoustics* (Academic, New York, 1990), Vol. 19, p. 157.
[6] M. Hirao, H. Ogi, and H. Fukuoka, Rev. Sci. Instrum. **64**, 3198 (1993).
[7] W. L. Johnson, B. A. Auld, and G. A. Alers, *Review of Progres in QNDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1994), Vol. 13, p. 1603.
[8] H. Ogi, M. Hirao, and T. Honda, J. Acoust. Soc. Am. **98**, 458 (1995).
[9] H. Ogi, M. Hirao, and K. Minoura, J. Appl. Phys. **81**, 3677 (1997).
[10] T. J. Moran and R. M. Panos, J. Appl. Phys. **47**, 2225 (1976).
[11] G. F. Miller and H. Pursey, Proc. R. Soc. London, Ser. A **223**, 524 (1954).
[12] M. R. Gaerttner, W. D. Wallace, and B. W. Maxfield, Phys. Rev. **184**, 702 (1969).
[13] H. Ogi, M. Hirao, and T. Honda, J. Acoust. Soc. Am. **98**, 1191 (1995).
[14] M. Hirao, H. Ogi, and H. Fukuoka, Res. Nondestruct. Eval. **5**, 211 (1993).

Lamb wave assessment of fatigue and thermal damage in composites

Michael D. Seale[a)]
Department of Physics, The College of William and Mary, Williamsburg, Virginia 23187

Barry T. Smith
Norfolk Academy, 1585 Wesleyan Drive, Norfolk, Virginia 23502

W. H. Prosser
NASA Langley Research Center, Mail Stop 231, Hampton, Virginia 23681

(Received 13 August 1997; accepted for publication 28 January 1998)

Among the various techniques available, ultrasonic Lamb waves offer a convenient method of evaluating composite materials. Since the Lamb wave velocity depends on the elastic properties of a structure, an effective tool exists to monitor damage in composites by measuring the velocity of these waves. Lamb wave measurements can propagate over long distances and are sensitive to the desired in-plane elastic properties of the material. This paper describes two studies which monitor fatigue damage and two studies which monitor thermal damage in composites using Lamb waves. In the fatigue studies, the Lamb wave velocity is compared to modulus measurements obtained using strain gage measurements in the first experiment and the velocity is monitored along with the crack density in the second. In the thermal damage studies, one examines samples which were exposed to varying temperatures for a three minute duration and the second includes rapid thermal damage in composites by intense laser beams. In all studies, the Lamb wave velocity is demonstrated to be an excellent method to monitor damage in composites. © *1998 Acoustical Society of America.* [S0001-4966(98)02705-2]

PACS numbers: 43.35.Zc, 43.40.Le, 43.20.Mv [HEB]

INTRODUCTION

Composite materials are being more widely used today by aerospace, automotive, sports equipment, and a number of other commercial industries because of their advantages over conventional metals. Composites have a high strength-to-weight ratio and can be constructed to meet specific design needs. In addition to satisfying the design specifications, composites offer weight savings over traditional metals without sacrificing strength. With the advantages of composite materials, their use is continually increasing and becoming more widespread to a variety of fields.[1–3]

O'Brien *et al.*[4] and Charewicz and Daniel[5] have examined the fatigue behavior of laminates and proposed lifetime prediction models based on fatigue factors such as increased matrix cracking, delamination growth, and stiffness loss. Ogin *et al.*[6,7] and Lim and Hong[8] have modeled the stiffness reduction in composites due to the effects of matrix cracking. Laws and Dvorak[9] have investigated the progression of matrix cracking as well as stiffness loss due to transverse cracking.

In addition to fatigue damage, thermal degradation of composites is of concern. Street *et al.*[10] have examined how composites respond to thermal damage of temperatures up to 350 °C for short durations, as well as exposure times as long as 30 min for temperatures of 225 °C. Herakovich and Hyer[11] and Ermer and Mall[12] have studied the effects on composites which were thermally cycled. The temperature range used was −157 °C–+121 °C in Ref. 11 and 149 °C–427 °C in Ref. 12. Stansfield and Pritchard[13] have subjected composites to rapid changes in temperature. These fluctuations, similar to the flight profile of supersonic aircraft, can lead to moisture absorption and matrix cracking.

Traditional nondestructive evaluation (NDE) techniques used on metals are often not easily adaptable to composites due to the inhomogeneous and anisotropic nature of these materials. Finding an effective means of nondestructively monitoring damage in composites is a challenging task, but it is extremely important to the safety and reliability of such structures. Several techniques are available to nondestructively evaluate damage in materials.[14–16] Salkind[14] and Stinchcomb[15] detail various methods used for damage detection in composites such as ultrasonics, x-ray radiography, thermography, and holography. In addition, Daniel *et al.*[16] examined the use of ultrasonic backscatter to monitor crack density and compared the results to those obtained using x-ray radiography.

Among the various techniques available, ultrasonic Lamb waves offer a convenient method of evaluating these composite materials. As a material is damaged, the elastic parameters of the structure change. Since the Lamb wave velocity depends on these properties, an effective tool exists to monitor damage in composites by measuring the velocity of these waves. Additionally, Lamb wave measurements are better than conventional through-the-thickness ultrasonic measurements because they can propagate over long distances and are sensitive to the desired in-plane elastic properties of the material.

[a)]Current address: NASA Langley Research Center, Mail Stop 231, Hampton, VA 23681.

Studies have been conducted which show a reduction in Lamb wave velocity due to a loss of stiffness caused by matrix cracking.[17–19] All of the authors noted that the in-plane properties of a composite are affected by cracking and that they can be monitored through the use of Lamb waves, which travel perpendicular to the crack direction. Dayal and Kinra[17] showed that for through-the-thickness measurements, the attenuation is sensitive to cracking, while the velocity is not. For Lamb wave propagation in the plane of the plate, both wave speed and attenuation were found to be sensitive to cracking. Tang and Henneke[18] noted that Lamb waves provide information about the in-plane properties of a plate, and this type of measurement is more useful due to the fact that composites are commonly designed to carry in-plane loads. Similarly, Dayal *et al.*[19] noted that the Lamb wave interaction with cracks is much stronger in the plane of the plate and, thus, provides an effective method to detect damage due to transverse matrix cracking.

Given the ability to nondestructively monitor materials using guided waves, studies have been conducted which use Lamb waves to monitor a variety of material properties in composite structures. Karim *et al.*[20] and Mal *et al.*[21] have used inversion techniques to determine the material parameters of composites from experimental Lamb wave data. Habeger *et al.*[22] studied the use of plate waves to monitor the strength of paper for the purpose of monitoring quality control on line for the paper industry. Several authors[23–25] have used Lamb waves to inspect adhesively bonded joints. Lamb wave techniques have also been used to study delaminations,[26–29] porosity,[29,30] and fiber misalignment.[30] Recently, Lamb waves have even been used by Lee *et al.*[31] to measure the temperature of silicon wafers during processing and by Pei *et al.*[32] for *in situ* measurements of thin film thickness. Both of these techniques would be of interest to the semiconductor processing industry.

While numerous studies exist which examine fatigue damage in composites using nondestructive techniques, few, if any, have been conducted which monitor thermal damage using ultrasonic nondestructive evaluation techniques. In the studies mentioned earlier,[10–13] the effects of thermal damage were measured using mainly mechanical tests and microscopic observation. This work will focused on the use of Lamb waves to nondestructively monitor fatigue as well as thermal damage in composite structures.

The following sections describe two experimental studies which monitor fatigue damage in composites using Lamb waves. In the first study, the Lamb wave velocity was compared to modulus measurements obtained using strain gage measurements. Additionally, experimental dispersion curves are compared for damaged and undamaged specimens. In the second study, the Lamb wave velocity was monitored along with the crack density. For both studies, the Lamb wave velocity measurements showed a significant decrease as the samples underwent increasing amounts of damage.

In addition to the fatigue studies, two studies were also conducted which monitor thermal damage in composites using the Lamb wave technique. The first study involves samples which were heated to varying temperatures for a three minute duration. The second study examined rapid thermal damage in composites. These samples were subjected to a very short duration (0.5 and 1.0 s) laser source. In both studies, the Lamb wave velocity was demonstrated to be an excellent method to monitor thermal damage in composite materials.

I. FATIGUE DAMAGE 1

In this experiment, studies have been conducted which monitor fatigue damage in composite samples using strain gage measurements as well as Lamb wave velocity measurements. A description of the test samples is followed by the results of two different measurements of Lamb wave velocity. The first technique is a contact measurement done at a single frequency, while the second involved an immersion study of Lamb waves in which dispersion curves were obtained. The results of the Lamb wave method for monitoring fatigue damage are compared to the damage progression measured using strain gages.

The composite samples studied were AS4/3501-6 graphite/epoxy with a stacking sequence of $[0/90_3]_S$. Two 30.5-by 38.1-cm plates were manufactured and C-scanned prior to being cut into specimens to check for any abnormalities. The scans revealed a 5%–7% level of porosity in the samples. The plates were then cut into 28-by 3.8-cm coupons. The finished coupons had an average thickness of 0.12 cm. Two 0.635-cm strain gages were attached to each sample: one axial and one transverse. From the values measured for the stress, axial strain, and transverse strain, one can obtain Young's modulus and Poisson's ratio for the sample.

Before the fatiguing process, two samples were loaded quasi-statically to failure to obtain Young's modulus, Poisson's ratio, and ultimate strength. Each of these coupons was loaded from 0 to 3250 microstrain and back twice. On the third run, the samples were loaded to failure. The modulus and Poisson's ratio were obtained by averaging the values measured for each of the three load cycles described above. The ultimate strength was defined as the load at which the samples failed.

The specimens were subjected to tension–tension fatigue in a 55-kip capacity load frame. The samples were fatigued at a frequency of 10 Hz and at an R value (minimum load/maximum load) of 0.3. The upper load was taken to be 33% of the ultimate strength, which had a value of 160 MPa. The specimens were successively raised to higher fatigue cycle values. The velocity of the lowest order symmetric Lamb mode was measured and the modulus obtained from strain gage measurements both before and after each cyclic loading. The samples were removed from the load frame at intermediate values of fatigue cycles in order to make the contact measurements.

Lamb modes can be generated by placing a longitudinal contact transducer on a plate and driving it with a sine wave tone burst. Ditri *et al.*[33] showed that periodic loading normal to the surface of a plate, as produced by a longitudinal transducer, will excite Lamb waves at distances away from the source which are large compared to the wavelength of the Lamb mode. Lamb waves are generated in most materials at a distance of roughly 1 cm or less. Since measurements are

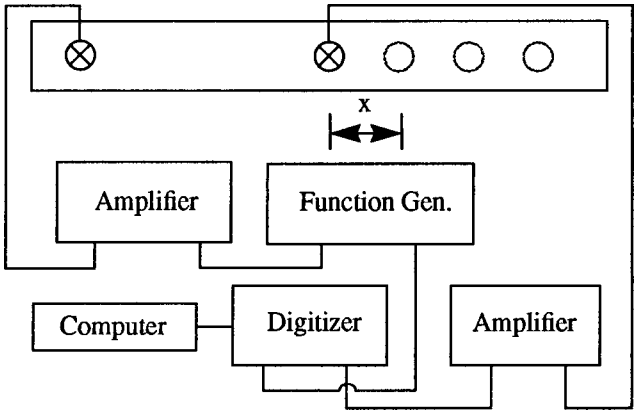


FIG. 1. Experimental arrangement for the contact measurement.

normally done anywhere from 5 to 10 cm away from the source, this does not create any difficulties in generating and measuring Lamb modes using this technique.

A Lamb wave was generated by using a 2-cycle, 0.5-MHz sine wave to excite the sending transducer. The signal was received at various positions, digitized at 10 MHz, and the waveforms saved on a computer. Both the sending and receiving transducers were broadband longitudinal wave transducers with a nominal center frequency of 1.0 MHz and a diameter of 1.27 cm. A coupling gel was used between the transducers and the plate. The experimental arrangement is shown in Fig. 1.

The time differences were measured by imposing a computational delay on the first signal to overlap the signals received at greater distances. The small thickness of the plates combined with the low frequency used to generate the Lamb wave yield a frequency–thickness product of 0.6 MHz·mm. In this region, only the S_0 and A_0 modes propagate. Therefore, the leading part of the wave was identified as the lowest order symmetric Lamb wave, which is not very dispersive. The trailing portion of the signals contained the dispersive antisymmetric wave as well as edge reflections. The distances were measured to an accuracy of 0.01 cm using a sonic ranger. A least-squares fit from a plot of time and distance was performed to obtain the velocity of the prompt S_0 mode.

As the samples underwent increasing amounts of fatigue, an increasing number of cracks developed in the 90° plies. These cracks degrade the modulus. Photomicrographs at different magnifications were taken of the edge of a fatigued composite and are shown in Fig. 2. For the lowest magnification, the fibers from the 0° laminas are the vertical bright streaks on the outer edges and the six 90° laminas are what appear as dots in between the outer layers. The ends of the fibers for these laminas is seen quite clearly at the higher magnification shown in Fig. 2(d). The size of the crack shown is approximately 1 μm.

The composite samples exhibited a decrease in the velocity as the modulus degraded due to increasing fatigue. Both the modulus and velocity squared follow the same general trend with increasing fatigue cycles. This is shown schematically in Fig. 3. In the plot, the velocity was normalized to the velocity measured before fatiguing. The normalized

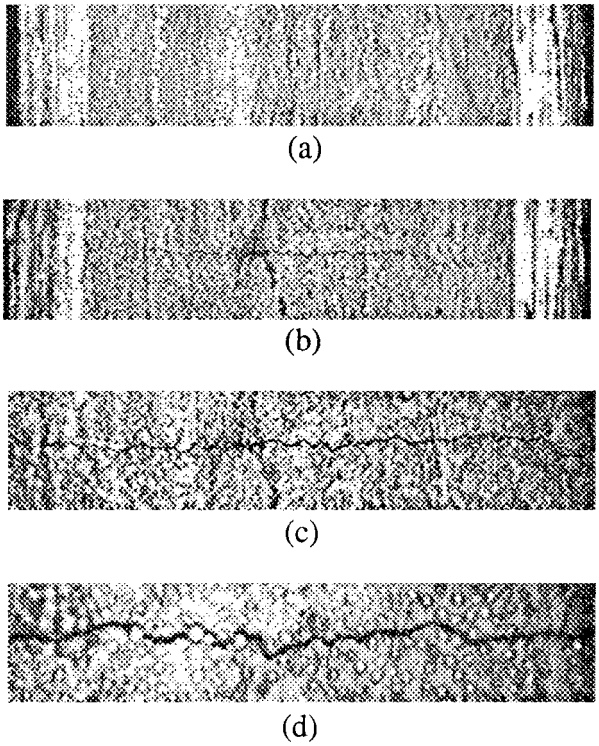


FIG. 2. Photomicrographs of the edge of a composite sample: (a) without a crack and (b)–(d) with a transverse matrix crack. The magnification for each was: (a) 5.5; (b) 5.5; (c) 11.0; and (d) 28.0. The thickness of the sample is 0.12 cm.

modulus was obtained in a similar manner. The velocity squared error bars are an estimate of the total error due to the uncertainties in the distance measurement and the digital signal overlap used to obtain the time differences. Some of values for the three measurements on the individual samples varied as much as 8%–10% from the average. Using a similar technique, Tang and Henneke[18] have reported deviations as great as 10% from the average of three measurements as well.

The velocity of the lowest order symmetric mode depends on the square root of the modulus for isotropic plates and the square root of the in-plane stiffness values for unidirectional composites.[34] Since the modulus shown in the plot was obtained from strain gage measurements of the in-

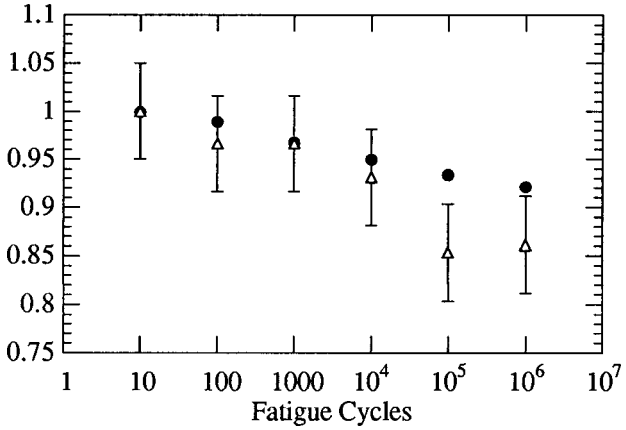


FIG. 3. Plot of normalized modulus (solid circles) and normalized velocity squared (open triangles) versus fatigue cycles for a fatigued sample.

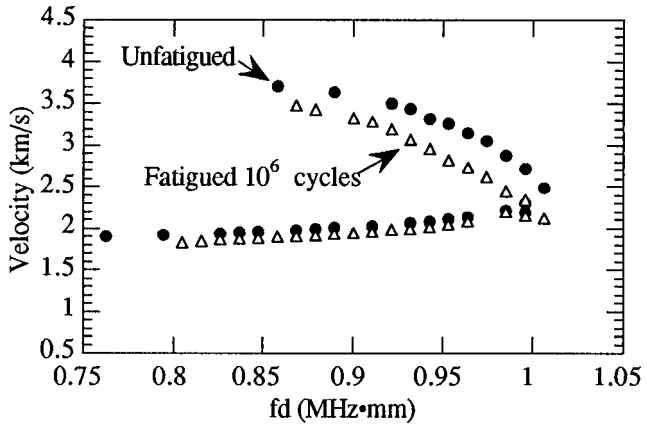


FIG. 4. Plot of the lower modes of the experimental dispersion curves for a sample which is undamaged (solid circles) and one which has been fatigued 10^6 cycles (open triangles).

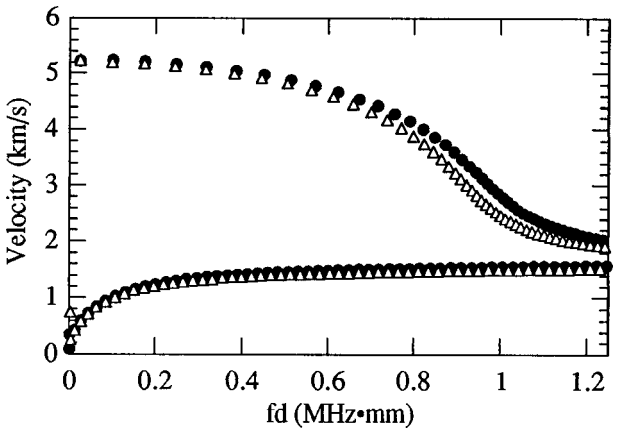


FIG. 5. Comparison of lower modes of numerical curves using full parameters (solid circles) and with c_{33} reduced by 5% (open triangles).

plane stiffness, the square of the velocity was chosen to plot against the number of fatigue cycles for convenience.

In addition to the contact measurements, which were done at a single frequency, an immersion method was used to obtain dispersion curves for two samples. The experimental arrangement for this type of measurement is detailed by Balasubramaniam *et al.*[35] In this method, the sample was immersed in water, which serves as the couplant between the transducer and the plate, and the same transducer was used to generate and receive the ultrasonic signal. An aluminum block served as a reflector for the input pulse as the specimen was rotated about the corner. When a Lamb wave is excited in the plate, part of the total energy is converted to a propagating Lamb mode. Due to this loss of energy, a minimum in the frequency spectrum of the reflected signal corresponds to the frequency at which a Lamb mode is propagating.

The reflected waveform at each angle was saved for later analysis of the frequency content. The velocity can be determined from Snell's law using the fact that only a longitudinal wave is supported in water and that the Lamb mode is propagating in the plate. After determining the frequency of the minimum for each angle, the velocity verses the frequency–thickness product can be plotted to obtain the dispersion curve. The frequency range is determined by the bandwidth and frequency of the transducer used.

In this experiment, an immersion transducer with a nominal center frequency of 1.0 MHz and diameter of 2.86 cm was used in a pulse-echo arrangement. The specimen was rotated 0.4°–60° in steps of 0.4° and the reflected waveform at each angle was saved for later analysis of the frequency content. After determining the frequency of the minimum for each angle, the velocity (from Snell's law) verses the frequency–thickness product was plotted to obtain the dispersion curves. This technique was used to obtain dispersion curves for an undamaged sample as well as one which had one million fatigue cycles. The results are shown in Fig. 4.

The dispersion curves for the fatigued sample is clearly shifted from the undamaged sample for both of the lower modes. The curves have roughly a 7% difference in velocity for the symmetric mode. This is in good agreement with the contact result for the same sample, which had a 7.2% relative decrease in velocity after one million fatigue cycles. It is also noted in the figure that the A_0 mode velocity variation was much less than that of the S_0 mode. This can be understood by examining how the velocity relates to the modulus. The velocity changes as the square root of the stiffness for the symmetric mode, but as the fourth root of the stiffness for the antisymmetric mode.[34] Therefore, the A_0 mode velocity will change less than the S_0 mode velocity for a given decrease in modulus.

For propagation across the fibers, where the 1-axis is in the fiber direction and the 3-axis is perpendicular to the fibers, a decrease in the elastic stiffness parameter c_{33} alters the dispersion curves. This is due to the fact that c_{33} is the modulus across the fibers. The transverse matrix cracks (due to fatigue damage) should, therefore, lead to a decrease in the elastic stiffness constant c_{33}. To examine the effect of the matrix cracking, a through-the-thickness finite element model, detailed in Dong and Huang[36] and Datta *et al.*,[37] is used to generate dispersion curves in laminated composites. Numerical dispersion curves were calculated in which the value of c_{33} was reduced by 5% in only these layers. A comparison was done for the numerical dispersion curves for the lower modes. The curves for the lower modes calculated using the full elastic parameters and those calculated using c_{33} reduced by 5% in the 90° layers are shown in Fig. 5. When compared in the region where the experimental data was obtained ($0.75 \leqslant fd \leqslant 1.05$ MHz·mm), the numerical results (see Fig. 5) are in good agreement with the experimental dispersion curves shown in Fig. 4. For the theoretical curves in this region, the S_0 velocity shows roughly a 7% decrease and the A_0 mode variation is much less than the S_0 velocity change. These are exactly the characteristics which were observed in the experimental curves.

It has been noted by Charewicz and Daniel[5] that strain gages only measure a local change in modulus and do not measure a global change in the specimen. Since the strain gages used in this study had a gauge length of 0.635 cm, they may not give an accurate account of damage throughout the entire plate. This would be especially true at the lower fatigue levels where very few cracks were observed in the samples. As the samples were damaged further, the cracking

became more uniform. Thus a Lamb wave technique at low frequencies used to track the modulus as a function of fatigue history may provide improved global interrogation over the propagation length of the wave.

II. FATIGUE DAMAGE 2

Additional composite samples of the same material, AS4/3501-6, and same architecture, $[0/90_3]_S$, were obtained in order to conduct a second fatigue damage study. The new composites were loaded to greater levels in order to induce damage more quickly in the samples. Two composite plates identical to those used in the previous study were obtained. The new composite plates were C-scanned prior to cutting them into coupons to check for porosity. The images were very uniform and, therefore, no large scale areas of porosity were expected. The plates were then cut into 18 3.8-by 28-cm coupons. The edges were polished and photomicrographs were taken. No appreciable amounts of porosity were found.

Similar to the first study, the specimens have been subjected to tension–tension fatigue. The samples were fatigued at a frequency of 10 Hz and at an R value of 0.35. For this study, however, the upper load was taken to be 45% of the ultimate strength, which has a value of 200 MPa. The samples were fatigued for a range of 10^n cycles ($n=3$, 4, and 5). As before, the samples were removed from the load frame at intermediate values of fatigue cycles and the velocity of the lowest order symmetric Lamb mode was measured. In addition to loading these samples at higher levels, the crack density was also monitored during the process.

A Lamb wave was generated by using a 6-cycle, 0.5-MHz sine wave to excite a longitudinal wave transducer with a nominal center frequency of 0.5 MHz and a diameter of 1.27 cm. The signal was received at various positions by an array of pinducers with active elements of 0.135 cm. The pinducers were placed in a holder at a fixed separation distance of 1.0 cm between each transducer. This receiving arrangement greatly reduces the error in the measurement of the velocity due to the uncertainty in the distance between the transducers. Since a differential measurement is being made, fixing the transducers in a holder virtually eliminates the distance error. A coupling gel was used between the transducers and the plate.

The signals were digitized at 10 MHz, averaged, and the waveforms saved to the computer. In an effort to reduce the error incurred in the time measurement by human judgement of the overlap region, the time differences between the pinducers were measured by performing a cross correlation. The leading part of both waves, which is identified as the lowest order symmetric Lamb wave, was used in the correlation. The distances between the five pinducers in the array were held fixed at 1, 2, 3, and 4 cm, respectively. The velocity of the prompt S_0 mode was obtained by a least-squares fit from a plot of time and distance.

To measure the crack density, a 2.5-cm section in the middle of the plate was marked. This was done in order to count the number of cracks occurring in the same 2.5-cm length in the sample. Keeping the same measurement area is especially important at the lower cycle levels where the cracking is not very uniform. The number of cracks was counted by examining the edges of the coupons under a microscope. The velocity measurements described above were centered over the same region where the crack density was monitored.

The influence of fatigue cycles on transverse matrix cracking and Lamb wave velocity was examined for five different samples. The velocity and crack density were obtained after each cyclic loading. Each velocity was obtained from the average of three measurements. The standard deviation for all of the samples was less than 5% from the average of the three measurements and most were within 2.5% of the average for each individual sample. Thus the reproducibility of the measurements is quite good and provides an improvement over the technique used in the first fatigue study. Also, these results are similar to those obtained by Tang and Henneke,[18] who report a variation of 1.5% using a similar technique.

As expected, the crack density for these samples increased with increasing cycles and, as with the previous study, the velocity squared decreased with fatigue cycles. Since strain gages were not used on these samples, the decrease in modulus will be estimated from the crack density. An expression which relates reduced modulus to crack density is given in Caslini *et al.*[38] For a $[0_m/90_n]_S$ laminate, the stiffness loss is given by[38]

$$\frac{E}{E_0}=\left[1+\frac{nE_2}{mE_1}\frac{\tanh(\lambda/2D)}{(\lambda/2D)}\right]^{-1}, \tag{1}$$

where

$$\lambda^2=\frac{3G_{12}E_0}{h^2E_1E_2}\frac{(n+m)}{n^2m}. \tag{2}$$

In the above equations, E and E_0 are the values of the modulus for the damaged and undamaged laminate, respectively. E_1, E_2, and G_{12} are the moduli (longitudinal, transverse, and shear, respectively) for each individual unidirectional lamina in the plate. D is the crack density, h is the lamina thickness, and m and n are the number of 0° and 90° layers, respectively, in the half-thickness of the plate.

For the case presented here, $m=1$, $n=3$, and $h=0.15$ mm. The value for the undamaged modulus, E_0, was obtained from the average of the four modulus measurements for the samples which were loaded to failure. The engineering constants for AS4/3501-6, are taken to be[1]

$$E_1=142.0,\ E_2=10.34,\ G_{12}=7.17,$$

where all values are listed in GPa.

The reduced stiffness as a function of crack density was calculated from the parameters given above using Eqs. (1) and (2). Figure 6 shows a comparison of the calculated reduced stiffness and the normalized velocity squared for all of the samples in this study. The velocity was normalized in a similar manner as described previously. Since the uncertainty in the distance and time measurements in this study is less than that of the earlier study, the velocity error bars are not as large as were previously shown. Due to the possibility of overlooking a small crack under the microscope, the error in calculation of the crack density is assumed to be a ± 1 error in the total number of cracks counted in the 2.5-cm

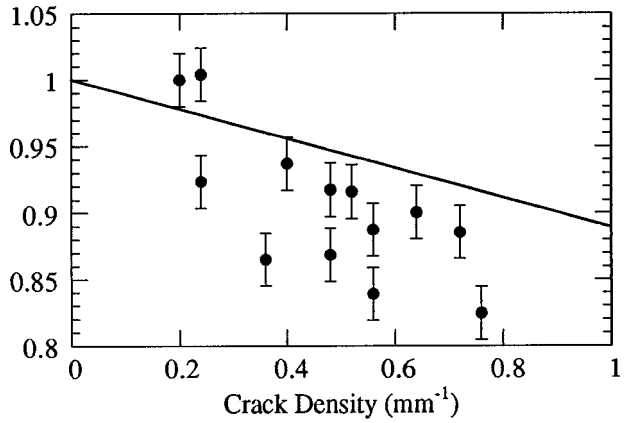


FIG. 6. Calculated normalized modulus (solid line) versus normalized velocity squared (solid circles) as a function of crack density for all samples.

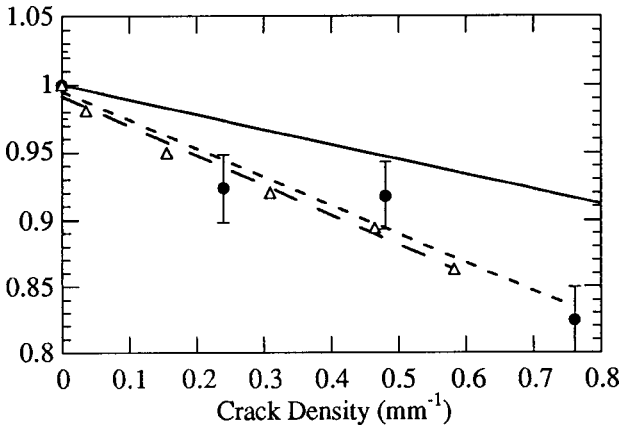


FIG. 7. A plot of the normalized velocity squared for a fatigued sample (solid circles) versus data from Dayal *et al.* (Ref. 19) (open triangles). Curve fits are shown for the data in this study (short dashes) and for the data of Dayal *et al.* (Ref. 19) (long dashes). The solid line is the theoretical modulus reduction described in the text.

length. However, for clarity, the error bars for the crack density have been omitted in the figure. Figure 6 shows a larger decrease for the normalized velocity squared than for the normalized modulus which was calculated from the crack density. Although not as pronounced as the results shown in Fig. 6, this behavior was evident in the previous study when the modulus was measured using strain gages (see Fig. 3).

These results are also in agreement with those found by Dayal *et al.*,[19] who examined the modulus degradation of an AS4/3502, $[0/90_3]_S$ composite using velocity measurement of the lowest order symmetric Lamb mode. The architecture and the fibers (AS4) of the samples used by Dayal *et al.*[19] are identical to the ones studied here. The epoxy matrix used (3502) is slightly different than the one used here, but the material properties of the two samples are very close.

Similar to the measurements done here, the tests by Dayal *et al.*[19] were also done at a frequency of 0.5 MHz. The velocity of the S_0 mode, measured using a different technique, and the crack density, measured using edge replication, were monitored as the samples underwent increasing damage. The modulus was calculated directly from the velocity measurement using the relation[19]

$$E = \rho v^2, \tag{3}$$

where E is the modulus, ρ is the density, and v is the velocity of the S_0 Lamb mode. Given this relation, the normalized modulus will also be equal to the normalized velocity squared. The results from Dayal *et al.*[19] are presented, along with the results for a representative sample from this study, in Fig. 7.

In the figure, the solid line represents the theoretical reduced modulus previously calculated using the material parameters for AS4/3501-6. However, as mention earlier, the parameters for this type of composite are extremely close to those of AS4/3502. Therefore, the theoretical curve shown in the figure is representative of the composites used in both studies. Another interesting feature in the figure involves the comparison of the two velocity measurements. Both sets of data show a larger decrease in normalized velocity squared than what is theoretically predicted for the modulus reduction. Dayal *et al.*[19] reported similar results when the modulus reduction measured using the Lamb wave technique was compared with the results of analytical modeling and static stiffness measurements conducted by other authors. Thus the Lamb wave velocity technique may provide a more sensitive method of damage detection than modulus reduction measurements.

Presented in the second fatigue study are the results of a different technique than that which was used in the previous study for determining the Lamb wave velocity in composites. Due to a reduction in the uncertainty in the distance and time measurements, the error in the velocity measurements were found to be less using this technique than those found in the first study. The crack density was also monitored as a function of fatigue cycles. A decrease in normalized velocity squared with increasing fatigue cycles was noted. This was expected and agrees with the results presented in the initial study.

The main difference between this study and the initial fatigue study was the monitoring of the crack density with fatigue cycles. The crack density was observed to increase with an increasing number of cycles, as is to be expected, and the Lamb wave velocity exhibited a decrease with increasing crack density. This drop in velocity as a function of crack density was compared with the results of Dayal *et al.*,[19] who saw a similar velocity decrease.

As was mentioned before, strain gages only measure a local change in modulus and do not measure a global change in the specimen. The strain gages used in the previous study had a gauge length of 0.635 cm and therefore, may not have given an accurate account of damage throughout the entire plate. It was shown in Caslini *et al.*[38] that the crack density can be theoretically related to stiffness reduction. Since the crack density was monitored over a 2.5-cm length, this becomes the effective gauge length for measuring the modulus reduction via the crack density.

The modulus reduction, calculated from the crack density, was compared to the results obtained using the Lamb wave velocity measurements. A more dramatic decrease was observed in the velocity measurements than in the theoretical modulus reduction. Thus a Lamb wave technique used to track damage in composites may provide improved interro-

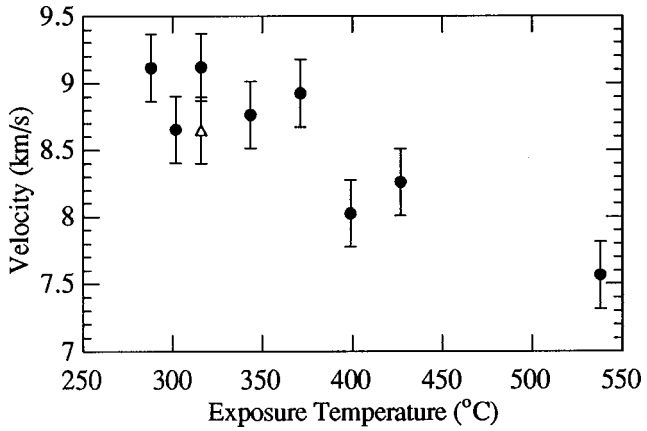


FIG. 8. Plot of S_0 velocity for a range of samples exposed to different temperatures for durations of 3 min (solid circles) and 10 min (open triangles).

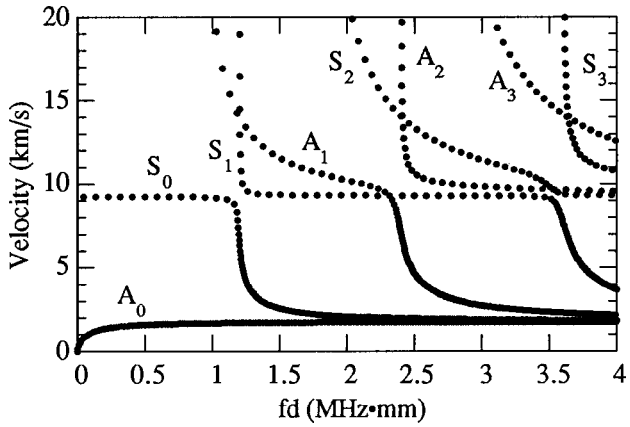


FIG. 9. Theoretical dispersion curves for an AS4/977-3 unidirectional composite.

gation of the effective stiffness of a structure over the propagation length of the wave.

III. THERMAL DAMAGE

Determining the response of composites to thermal damage is necessary for a complete understanding of the total use environment of these materials. Heat damaged samples, which simulate typical thermal damage on carbon/epoxy aircraft structures, were provided by McDonnell Douglas Aerospace. These 15.24-by 15.24-cm specimens were fabricated as 16-ply, uniaxial plates of AS4/977-3 and had an average thickness of 3.07 mm. All samples had received different levels of heat damage, had been C-scanned, and had Barcol hardness readings performed. The exposure temperatures ranged from 288 °C to 538 °C for a duration of 3 min and included a sample which was exposed to 316 °C for 10 min.

Lamb waves were generated and received with the same techniques described in the second fatigue study. The only differences were that the transmitting transducer had a nominal center frequency of 1.0 MHz and the receiving transducers had fixed separations of 4, 5, 6, and 7 cm. The signal was received at various positions and digitized at 10 MHz. The time differences were measured by imposing a computational delay on the first signal to overlap the signals received at greater distances. A least-squares fit from a plot of time and distance was performed to obtain the velocity of the prompt S_0 mode.

As in the fatigue studies, the velocity of the lowest order symmetric Lamb mode was measured. The results of the velocity measurements in the fiber direction for ten different thermally damaged samples are plotted in Fig. 8. The error bars are an estimate in the velocity error and are mainly due to uncertainty in the digital channel alignment. The decrease in velocity with increasing thermal damage is quite dramatic. The variation in velocity is on the order of 18% for the two extremes of temperature exposure. The correlation with extended temperatures is excellent and provides a method to assess thermal damage. Although the velocity showed a large change with thermal damage, the Barcol hardness numbers provided by McDonnell Douglas showed little variation over this temperature range. Therefore, measurement of the Lamb wave velocity could provide a better technique to assess thermal damage than conventional methods.

A model calculation[36,37] to determine the dispersion curves for this material system was computed using engineering constants supplied by McDonnell Douglas. The four standard engineering values (E_1, E_2, G_{12}, and ν_{12}) can be used to calculate the corresponding five elastic stiffness parameters (c_{11}, c_{13}, c_{23}, c_{33}, and c_{55}) using various techniques.[1] Since a value for G_{23} was not available, the value of c_{44} [where $c_{44}=(c_{22}-c_{23})/2$] must be estimated using common values for polymer matrix/graphite fiber composites.[39] The value of c_{44} typically ranges from 3.5 to 4.0 GPa, so a median value of 3.75 GPa was chosen for the calculation. The density of the material was taken to be 1560 kg/m^3. A five layer model calculation of the dispersion curves for propagation in the fiber direction is shown in Fig. 9.

From the dispersion curve, one can see that the S_0 mode is almost completely nondispersive for the lower frequency–thickness products. Also, the S_0 velocity in this region is around 9 km/s, which is exactly the velocity that was measured for the samples with little or no damage. Thus the model compares well with the experimentally determined velocity.

Theoretical dispersion curves were also calculated for AS4/977-3 using reduced stiffness constants. The effect of reducing different c_{IJ}'s by 20% on the low frequency, nondispersive region of the lowest order symmetric mode ($0.0 \leqslant fd \leqslant 1.0$ MHz·mm) was investigated. The only parameter which altered the velocity of the S_0 mode in this region for propagation along the fibers is c_{11}.

Figure 10 shows a plot of the experimentally determined normalized velocity versus exposure temperature, where the normalized velocity is the velocity in the damaged sample divided by the velocity in an undamaged control sample. Also plotted is the normalized velocity of the S_0 mode with reduced values of c_{11}. Here, the normalized velocity is the velocity with c_{11} reduced divided by the velocity with the full values of the elastic stiffness constants. Both velocities were taken at a constant frequency–thickness product of 0.5 MHz·mm on the dispersion curve. Theoretical calcula-

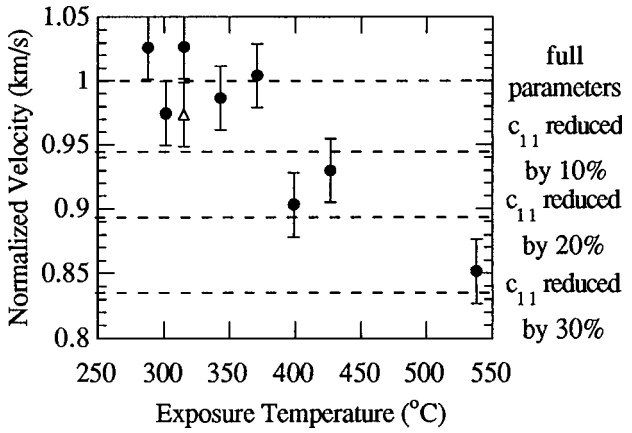


FIG. 10. Plot of the normalized S_0 velocity of samples exposed to different temperatures. The duration times are 3 min (solid circles) and 10 min (open triangle). Also shown (dashed lines) is the theoretically determined normalized S_0 velocity with reduced values of c_{11}, as described in the text.

tions with reduced values of c_{11} show that a 25%–30% reduction in this parameter lowered velocities of the S_0 mode by 15%. This value corresponds to the velocity reduction of the sample which saw an exposure temperature of 538 °C.

IV. LASER DAMAGE

The three laser damaged samples examined were fabricated as quasi-isotropic plates with a stacking sequence of $[0/+45/-45/90]_S$. The fiber/matrix used on two of the samples was IM7/SP500 and the remaining sample was IM7/PEEK. All three samples were irradiated at 100 W/cm^2. The IM7/SP500 samples were subjected to the laser source for durations of 0.5 and 1.0 s. The IM7/PEEK sample was subjected to the laser source for 0.5 s.

In this study, Lamb wave velocities were measured using the same technique described in the second fatigue study. The velocity of the lowest order symmetric Lamb mode was measured in the 0° direction in both the undamaged and damaged regions of each of the samples. The velocities as well as the damage level for each of the different samples are shown in Table I. The listed velocities are the average of three measurements performed in each region. For each of the three measurements in a particular region, the values of the velocity varied less by less than 5% from the average.

As can be seen from the data, the velocity actually increases in the area which was subjected to the laser source. This is contrary to what was seen in the thermally damaged samples discussed earlier in this paper. The main difference between the laser damage and the thermal damage previously mentioned is that the laser subjects the sample to an extremely rapid thermal cycle. This type of heating could actually cause embrittlement of the sample, which would serve to increase the velocity of the Lamb wave in that region.

TABLE I. Data for laser damaged samples.

Sample	Irradiation (W/cm^2)	Time (s)	Fluence (J/cm^2)	Undamaged Vel. (km/s)	Damaged Vel. (km/s)
IM7/SP500	100	0.5	50.0	6.01	6.10
IM7/SP500	100	1.0	100.0	5.91	6.33
IM7/PEEK	100	0.5	50.0	6.50	6.56

V. CONCLUSION

The Lamb wave velocity measurements in this study were conducted at long wavelengths. This was done for several reasons. First, if the wavelength is large compared to the diameter of the fibers, composites can be treated as homogeneous. Second, only the lowest order modes propagate in the frequency range where the contact measurements were conducted. Finally, the effective elastic properties of the material can be measured because local defects, which scatter high frequency waves, will not be observed at long wavelengths.

Lamb waves offer a useful technique for characterizing many forms of damage in composite materials. Lamb wave velocity measurements are better than conventional ultrasonic measurement schemes (i.e., through-the-thickness measurements) because they can propagate over long distances and are sensitive to the desired in-plane elastic properties of the material. The propagation of Lamb waves depends on a variety of material properties: elastic stiffness constants, density, and thickness. As a material is damaged in some fashion, one or more of these material properties are altered. Since the Lamb wave velocity is directly related to these parameters, an effective tool exists to monitor damage in composites by measuring the velocity of these waves.

The Lamb wave velocity is a quantitative measurement and it has been shown by this work to be an effective tool in monitoring both fatigue and thermal damage in composites. Thus the Lamb wave method can be used to verify the integrity of a composite structure over the service life of a component. This is an important measurement for flight qualified composite materials that may have undergone some measure of nonvisible damage, which could seriously degrade the performance of the structure. Since the Lamb wave velocity depends on a variety of material properties, an ideal technique exists to monitor composites as damage is incurred. With the continued development of damage assessment techniques such as the Lamb wave method, the safety of such structures can be assured.

ACKNOWLEDGMENTS

The authors would like to thank John Masters of Lockheed Engineering and Sciences for providing the samples used in the fatigue studies. They would also like to thank Kevin Slattery of McDonnell Douglas Aerospace for providing the thermally damaged samples and Bill Laughlin of Physical Sciences, Inc. for providing the laser damaged samples. This work was supported in part by NASA Grant No. NAG-1-1063 and NASA Task Assignment No. NAS1-19656.

[1] I. M. Daniel and O. Ishai, *Engineering Mechanics of Composite Materials* (Oxford U. P., New York, 1994), pp. 3–76.
[2] D. Stover, senior editor, ''Composites Use Increases on New Commercial Transports...,'' Advanced Composites **6**, 30–40 (1991).

[3] Composites News, ''Boron/Epoxy System to Repair Entire B-1 Fleet,'' Advanced Composites **6**, 18–20 (1991).
[4] T. K. O'Brien, M. Rigamonti, and C. Zanotti, ''Tension fatigue analysis and life prediction for composite laminates,'' Int. J. Fatigue **11**, 379–393 (1989).
[5] A. Charewicz and I. M. Daniel, ''Damage mechanisms and accumulation in graphite/epoxy laminates,'' in *Composite Materials: Fatigue and Fracture*, ASTM STP 907, edited by H. T. Hahn (American Society for Testing and Materials, Philadelphia, 1986), pp. 274–297.
[6] S. L. Ogin, P. A. Smith, and P. W. R. Beaumont, ''Matrix cracking and stiffness reduction during the fatigue of a (0/90)s GFRP laminate,'' Compos. Sci. Technol. **22**, 23–31 (1985).
[7] S. L. Ogin, P. A. Smith, and P. W. R. Beaumont, ''A stress intensity factor approach to the fatigue growth of transverse ply cracks,'' Compos. Sci. Technol. **24**, 47–59 (1985).
[8] S. G. Lim and C. S. Hong, ''Prediction of transverse cracking and stiffness reduction in cross-ply laminated composites,'' J. Compos. Mater. **23**, 695–713 (1989).
[9] N. Laws and G. J. Dvorak, ''Progressive transverse cracking in composite laminates,'' J. Compos. Mater. **22**, 900–916 (1988).
[10] K. N. Street, A. J. Russell, and F. Bonsang, ''Thermal damage effects on delamination toughness of a graphite/epoxy composite,'' Compos. Sci. Technol. **32**, 1–14 (1988).
[11] C. T. Herakovich and M. W. Hyer, ''Damage-induced property changes in composites subjected to cyclic thermal loading,'' Eng. Fract. Mech. **25**, 779–791 (1986).
[12] P. G. Ermer and S. Mall, ''Damage mechanisms in a unidirectional metal matrix composite during thermal cycling,'' in *American Society for Composites, Technical Conference*, 5th, East Lansing, MI, June 12–14, 1990, Proceedings (A91-26701 10-4) (Technomic, Lancaster, PA, 1990), pp. 528–536.
[13] K. E. Stansfield and E. Pritchard, ''Damage Generation and Healing During Composite Thermal Spikes,'' in *Proceedings of the International SAMPE Technical Conference*, 21st, Atlantic City, NJ, 25–28 Sept. 1989, pp. 120–129.
[14] M. J. Salkind, ''Early Detection of Fatigue Damage in Composite Materials,'' AIAA Paper No. 75-772, *AIAA/ASME/SAE 16th Structures, Structural Dynamics, and Materials Conference* (Denver, Colorado, 27–29 May 1975), pp. 1–8.
[15] W. W. Stinchcomb, ''Nondestructive evaluation of damage accumulation processes in composite laminates,'' Compos. Sci. Technol. **25**, 103–118 (1986).
[16] I. M. Daniel, S.-C. Wooh, and J.-W. Lee, ''Nondestructive evaluation of damage development in composite materials,'' in *Elastic Waves and Ultrasonic Nondestructive Evaluation*, edited by S. K. Datta, J. D. Achenbach, and Y. S. Rajapakse (Elsevier Science, Amsterdam, 1990), pp. 183–189.
[17] V. Dayal and V. K. Kinra, ''Ultrasonic NDE of Composites for Transverse Cracking,'' in *Optical Methods in Composites; Proceedings of the SEM Fall Conference on Experimental Mechanics*, Keystone, CO, 2–5 Nov. 1986 (A88-13876 03-24) (Society for Experimental Mechanics, Inc., Bethel, CT, 1986), pp. 17–22.
[18] B. Tang and E. G. Henneke II, ''Lamb-wave monitoring of axial stiffness reduction of laminated composite plates,'' Mater. Eval. **47**, 928–934 (1989).
[19] V. Dayal, V. Iyer, and V. K. Kinra, ''Ultrasonic evaluation of microcracks in composites,'' in *Advances in Fracture Research; Proceedings of the Seventh International Conference on Fracture* (ICF7), Houston, TX, 20–24 Mar. 1989, Vol. 5 (A90-41276 18-39) (Pergamon, Oxford, 1989), pp. 3291–3300.
[20] M. R. Karim, A. K. Mal, and Y. Bar-Cohen, ''Inversion of leaky Lamb wave data by simplex algorithm,'' J. Acoust. Soc. Am. **88**, 482–491 (1990).
[21] A. K. Mal, M. R. Gorman, and W. H. Prosser, ''Material characterization of composite laminates using low-frequency plate wave dispersion data,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1992), Vol. 11, pp. 1451–1458.
[22] C. C. Habeger, R. W. Mann, and G. A. Baum, ''Ultrasonic plate waves in paper,'' Ultrasonics **17**, 57–62 (1979).
[23] R. A. Kline and M. M. Doroudian, ''Ultrasonic evaluation of adhesively bonded composites using guided waves,'' in *Advances in Composite Materials and Structures* (A91-23701 08-24) (American Society of Mechanical Engineers, New York, 1989), pp. 63–67.
[24] P. N. Dewen and P. Cawley, ''An ultrasonic scanning technique for the quantitative determination of the cohesive properties of adhesive joints,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1992), Vol. 11, pp. 1253–1260.
[25] M. J. S. Lowe and P. Cawley, ''The applicability of plate wave techniques for the inspection of adhesive and diffusion bonded joints,'' J. Nondestruct. Eval. **13**, 185–200 (1994).
[26] Keun J. Sun, ''Application of guided acoustic waves to delamination detection,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1992), Vol. 11, pp. 1213–1219.
[27] N. Guo and P. Cawley, ''The interaction of Lamb waves with delaminations in composite laminates,'' J. Acoust. Soc. Am. **94**, 2240–2246 (1993).
[28] N. Guo and P. Cawley, ''Lamb waves for the NDE of composite laminates,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1992), Vol. 11, pp. 1443–1450.
[29] Y. Bar-Cohen and D. E. Chimenti, ''NDE of defects in composites using leaky Lamb waves,'' in *Symposium on Nondestructive Evaluation*, 15th, San Antonio, TX, 23–25 April 1985, Proceedings (A86-47129 22-38) (Nondestructive Testing Information Analysis Center, San Antonio, TX, 1986), pp. 202–208.
[30] K. Balasubramaniam and J. L. Rose, ''Guided Plate Wave Potential for Damage Analysis of Composite Materials,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1990), Vol. 9, pp. 1505–1512.
[31] Y. J. Lee, B. T. Khuri-Yakub, and K. C. Saraswat, ''Temperature measurements in rapid thermal processing using acoustic techniques,'' Rev. Sci. Instrum. **65**, 974–976 (1994).
[32] J. Pei, F. Levent Degertekin, B. T. Khuri-Yakub, and K. C. Sarawsat, ''*In situ* thin film thickness measurement with acoustic Lamb waves,'' Appl. Phys. Lett. **66**, 2177–2179 (1995).
[33] J. J. Ditri, A. Pilarski, B. Pavlakovic, and J. L. Rose, ''Generation of guided waves in a plate by axisymmetric normal surface loading,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1994), Vol. 13, pp. 133–140.
[34] W. H. Prosser, ''The propagation of characteristics of the plate modes of acoustic emission waves in thin aluminum plates and thin graphite/epoxy composite plates and tubes,'' Ph.D. thesis, The Johns Hopkins University, 1991.
[35] K. Balasubramaniam and J. L. Rose, ''Physically based dispersion curve feature analysis in the NDE of composites,'' Res. Nondestruct. Eval. **3**, 41–67 (1991).
[36] S. B. Dong and K. H. Huang, ''Edge vibrations in laminated composite plates,'' J. Appl. Mech. **52**, 433–438 (1985).
[37] S. K. Datta, A. H. Shah, and W. Karunasena, ''Wave propagation in composite media and material characterization,'' in *Elastic Waves and Ultrasonic Nondestructive Evaluation*, edited by S. K. Datta, J. D. Achenbach, and Y. S. Rajapakse (Elsevier Science, Amsterdam, 1990), pp. 159–167.
[38] M. Caslini, C. Zanotti, and T. K. O'Brien, ''Study of matrix cracking and delamination in glass/epoxy laminates,'' J. Composites Technol. Res. **9**, 121–130 (1987).
[39] W. H. Prosser, ''Ultrasonic characterization of the nonlinear elastic properties of unidirectional graphite/epoxy composites,'' NASA Contractor Report 4100, NASA Langley Research Center, October, 1987, p. 65.

A finite element analysis of an interferometric optical fiber hydrophone

Jong-in Im
Functional Ceramic Division, Research Institute of Industrial Science & Technology, Pohang 760-330, Korea

Yong-rae Roh
Department of Electronic Engineering/Department of Sensor Engineering, Kyungpook National University, 1370 Sankyung-Dong, Buk-Gu, Taegu 702-701, Korea

(Received 3 June 1997; accepted for publication 4 February 1998)

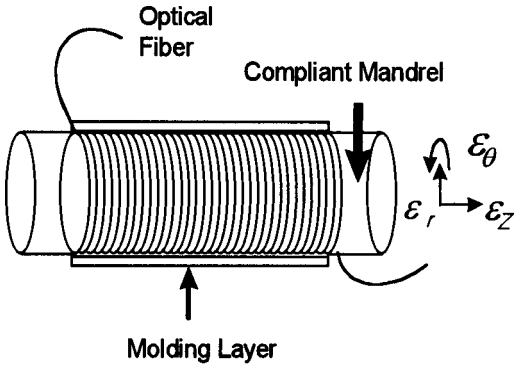
This paper investigates the influence of structural parameters on the performance of an optical fiber wound mandrel hydrophone by using the finite element method (FEM). A hydrophone should exhibit the highest possible sensitivity, the widest possible flat frequency response, and an omni-directional sensitivity pattern within the frequency range. The parameters studied included the mandrel geometry, the thickness of the molding coated over the fiber, and the material properties of the constituent parts of the hydrophone. The analysis results showed that the pressure sensitivity of a hydrophone increases in relation to the length of the mandrel and the thickness of the molding. A higher pressure-sensitivity also requires a mandrel or molding material with a relatively low Young's modulus and Poisson's ratio. On the other hand, the flat frequency response of a hydrophone increases when either the mandrel length is shortened or the mandrel material is hardened. The omni-directional characteristic is also improved with a shorter mandrel. Therefore, a hydrophone with the best performance must balance the tradeoff between the geometrical and material parameters. The analysis discussion is focused on a representative specification of a frequency range of up to 5 kHz. © *1998 Acoustical Society of America.* [S0001-4966(98)03905-8]

PACS numbers: 43.38.Ar, 43.30.Yj, 43.58.Fm [SLE]

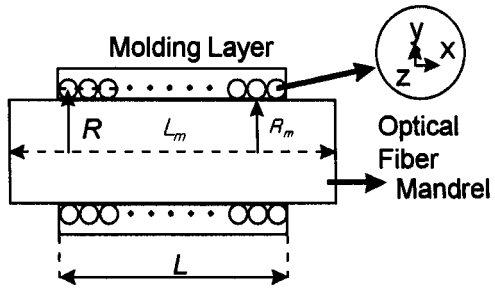
INTRODUCTION

Interferometric optical fiber hydrophones exploit the changes in an optical path length induced by sound in the optical fibers. One of the major areas of application for hydrophones is in underwater acoustic sensing. There are numerous reasons for this interest which range from cost and performance to the geometric versatility of the sensing head. There are two main approaches to coupling an acoustically induced phase shift in an optical fiber to the detection process: a mandrel sensor or a coated fiber sensor.[1–3] Both of these approaches are viable up to ~50 kHz. In coated fiber sensors the optical fiber is wound in a coil as a sensing element and the size is comparatively large. Since the acousto-optic coupling of the fiber is weak, a long fiber is generally used to increase the induced phase change. Pressure sensitivity is a complex function of a Young's modulus, a bulk modulus, and the cross-sectional area of a secondary coating.[2] In the case of a mandrel sensor, a thin jacket fiber is typically wrapped around a compliant mandrel. The optical fiber then measures the pressure-induced strain in the mandrel. It is important to maximize the scale factor or the acoustic phase response of the sensor in order to maintain a high sensitivity and a flat acoustic response for its frequency range. Mandrel sensors appear to be ideal for point sensors even though they have somewhat limited geometrical versatility.

This paper investigates the influence of structural parameters on the performance of a mandrel hydrophone by using the finite element method (FEM). Acoustic response characteristics such as pressure sensitivity, bandwidth, and directivity depend upon geometry and the material properties of the sensor element. Previous research on mandrel sensors have analytically analyzed and concentrated on compliant plastic mandrels, typical examples being nylon and Teflon.[1,4] Africk analytically determined sensor response by calculating a pressure response term which neglected the cylinder-end response by separately determining the term associated with the radial strain of the mandrel.[4] His results demonstrate a very good agreement between theory and experiment for the case of a relatively hard mandrel like Teflon. However, the calculated results are significantly higher than the experimental results. This is attributed to the neglect of the circumferential deformation of the mandrel. In this paper, the full three-dimensional behavior of a mandrel hydrophone is analyzed by using the FEM without any prior assumptions or simplifications of the structure. A hydrophone should exhibit the highest possible sensitivity, the widest possible flat frequency response, and an omni-directional sensitivity pattern within the frequency range. However, to facilitate a more practical discussion of the analysis results, the frequency range was restricted to 5 kHz as a representative case. The parameters studied included the mandrel geometry, the molding thickness, and the material properties of the constituent parts of the hydrophone. The pressure sensitivity, the frequency range, and the directivity pattern of the hydrophone were analyzed as a function of the structural parameters, and the results discussed.



(a) Genernal view of the hydrophone



(b) Cross section of the hydrophone

FIG. 1. Schematic structure of the optical fiber wound mandrel hydrophone.

I. PRESSURE SENSITIVITY OF A MANDREL HYDROPHONE

In a fiber wound mandrel sensor, a thin jacket fiber is typically wrapped around a compliant mandrel and a thin polymer, less than 1 mm thick, is coated over the fiber as protection during operation. The optical fiber measures the pressure-induced strain in the mandrel and protecting layer. Mandrel sensors are important because they are easy to produce and they exhibit a high sensitivity and an amenability to spatial shading.

Consider the mandrel sensor shown in Fig. 1(a). A homogeneous cylindrical mandrel of length L_m and radius R_m is radially wrapped with a single-mode fiber of length l over a length L. The molding layer is assumed to be only coated on the fiber layer and has a thickness of less than 1 mm. Assuming that the fiber sensor response is driven solely by the mandrel response, it is clear that a circumferential change in the mandrel couples directly to a length change in the fiber. The phase of light propagating through an optical fiber is defined as

$$\phi=\beta\star l=n\star k_0\star l, \tag{1}$$

where β is the propagation constant in the fiber length direction, l is the total fiber length wrapped around the mandrel equal to $2\pi R^*N$, N is the number of fiber windings round the mandrel, R is the distance from the center of the mandrel to that of the fiber as denoted in Fig. 1(b), n is the refractive index of the fiber core, and k_0 is the free space propagation constant. The change in the optical phase of the fiber which is subjected to a unit of pressure may be written as

$$\Delta\phi=\beta\star\Delta l+\frac{d\beta}{dn}\star\Delta n\star l+\frac{d\beta}{dD}\star\Delta D\star l, \tag{2}$$

where D is the fiber diameter.[5,6] The change Δl in the fiber length using the cylindrical coordinate system is written as

$$\Delta l=l\star(\varepsilon_r+\varepsilon_\theta), \tag{3}$$

where ε_r is the radial strain, and ε_θ is the axial strain in the fiber with the axis convention with respect to the mandrel in Fig. 1. Since the refractive indices of the core and cladding differ only slightly, one can use the approximation of

$$\frac{d\beta}{dn}\approx k_0. \tag{4}$$

The change in the refractive index Δn can be written in terms of the change in the optical indicatrix given by the well-known expression[7,8]

$$\Delta\left(\frac{1}{n^2}\right)_i=\sum_{j=1}^{3}P_{ij}\varepsilon_j, \quad i,j=1,2,3, \tag{5}$$

where p_{ij} are elements of the strain optic tensor and ε_j are principal strains. Since fluid pressure is transmitted to the hydrophone only at right angles to its boundary at every point, the components containing a shear strain have been assumed to be zero in Eq. (5). A typical optical fiber consists of a core and a cladding fabricated from glasses which have similar properties. Hence, the fiber may be considered homogeneous and isotropic, leading to

$$p_{ij}=\begin{bmatrix} p_{11} & p_{12} & p_{12} \\ p_{12} & p_{11} & p_{12} \\ p_{12} & p_{12} & p_{11} \end{bmatrix}. \tag{6}$$

Thus, the Δn in Eq. (2) for the light propagating in the z direction[6] is

$$\begin{aligned}\Delta n_z&=-\frac{n^3}{2}\star\Delta\left(\frac{1}{n^2}\right)_{x,y}\\&=-\frac{n^3}{2}\star[(p_{11}+p_{12})\star\varepsilon_r+p_{12}\star\varepsilon_\theta].\end{aligned} \tag{7}$$

The third term in Eq. (2) can be shown to be negligible compared with the first two.[9]

Therefore, the acoustic sensitivity of the mandrel sensor per unit of sound pressure becomes

$$\begin{aligned}\frac{\Delta\phi}{\phi}&=\sum\varepsilon_r+\sum\varepsilon_\theta\\&\quad-\frac{n^2}{2}\left[p_{11}\star\sum\varepsilon_r+p_{12}\star\left(\sum\varepsilon_\theta+\sum\varepsilon_r\right)\right].\end{aligned} \tag{8}$$

Equation (8) means that the phase change of the mandrel sensor can be found once we determine the appropriate strain distribution in relation to the unit of applied pressure, which then leads to the analysis of the hydrophone performance. In this paper, the strain distribution is calculated by using the finite element method. Equation (8) is written in the form of a summation of the strains distributed over all the discrete elements in the finite element model. The following values

TABLE I. Material properties as well as the geometry of the fiber and its equivalent tube.

	Material and geometry	E (10^9 Pa)	Density (kg/m^3)	Poisson's ratio
Core	$99SiO_2-1GeO_2$ Diam.$=4\times10^{-6}$ m	72.45	2202	0.17
Cladding	$95SiO_2-5B_2O_3$ Diam.$=125\times10^{-6}$ m	65.14	2187	0.149
Inner jacket	Si-rubber Diam.$=187\times10^{-6}$ m	0.0035	1050	0.4995
Outer jacket	PMMA Diam.$=245\times10^{-6}$ m	6.24	1190	0.34
Equivalent tube	Thick.$=245\times10^{-6}$ m	19.56	1404.4	0.3417

for Pockel's coefficients of the unclad fiber are assumed in the calculations of the acoustic sensitivity with the finite element method:[2,8]

$$p_{11}=0.121, \quad p_{12}=0.27.$$

II. COMPUTATIONAL CONSIDERATIONS AND STUDIED PARAMETERS

A typical optical fiber consists of a core and a cladding fabricated from glasses which have similar properties. The single-mode optical fiber under investigation in this paper is composed of a fused silica core with traces of GeO_2 and a cladding of 5% B_2O_3- and 95% of SiO_2. This glass waveguide is coated first with an inner soft elastomer and then with an outer hard polymer jacket to preserve some level of fiber strength.

The core diameter is about 4 μm, while the clad diameter about 125 μm.[5,6,10,11] The fiber inner and outer jackets consist of 60-μm layers of silicon and PMMA polymer, respectively. The composition and the geometry of the fiber are kept constant throughout this paper. If the optical fiber is precisely described in the finite element modeling of the mandrel sensor and includes a circular cross section of the fiber, there are too many elements in the model and it is quite time consuming to get any results. Hence, the fiber layer has been modeled as an isotropic tube with a flat cross section possessing the same material properties as a fiber composite which adheres to the mixture rule of a volumetric fraction for each contributor. Table I lists the material properties as well as the geometry of the fiber and its equivalent tube.

In this study, the mandrel sensor is assumed to have a frequency range of up to 5 kHz. The optical fiber has a constant length of 10 m. The parameters studied in order to satisfy the frequency range requirement and to realize the highest sensitivity are the geometry, the material properties of the mandrel, and the molding. Before rigorous investigation, some preliminary analysis is performed so that the range of the main calculation is established even though a rough estimate. Above all, the fundamental resonance frequency of the mandrel should be at least higher than 10 kHz considering the fact that the operating frequency range of a nonresonant type sensor is usually less than half of the resonance frequency. Hence, the maximum length of the mandrel is determined where its fundamental mode frequency reaches 10 kHz. Typically, a mandrel is more compliant than the optical fiber, and, when wrapped by the fiber, the resonance frequency of the whole sensor will increase. Thus this is a rather safe and conservative way of preserving the resonance frequency above 10 kHz. Figure 2(a) shows the variation of the fundamental mode frequency of a mandrel in relation to its length for a representative radius and material. For the material, we selected one of the most compliant materials reported as appropriate for fabricating a mandrel, Teflon.[12] The fundamental mode frequency of a Teflon mandrel of 2-cm radius and 5.1-cm height is about 10 kHz. Therefore, the maximum length of the cylindrical mandrel is 5.1 cm. The minimum length $L_{\min}$ is determined to be equal to the length of the optical fiber winding over the mandrel as shown in Eq. (9) where the fiber is supposed to be wound as a single layer:

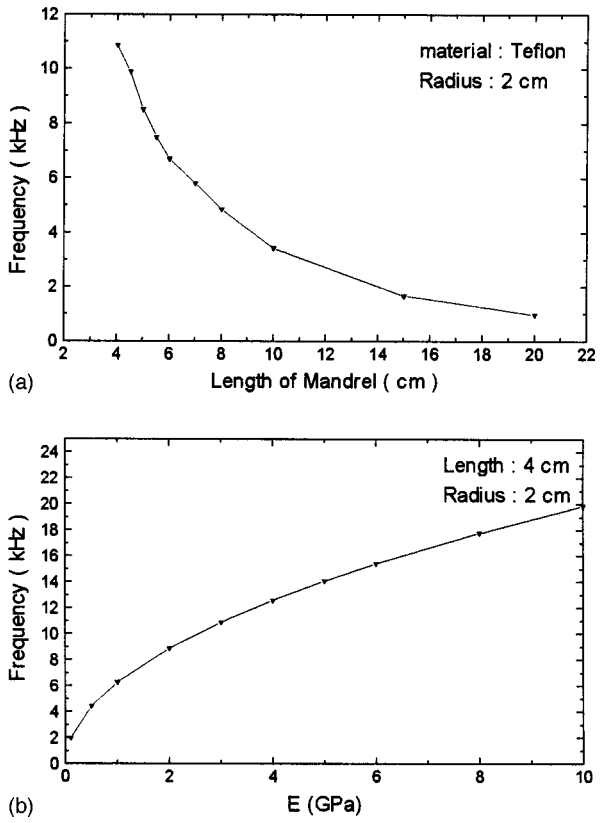


FIG. 2. Variation of the fundamental resonance frequency of a mandrel in relation to its length and Young's modulus. (a) Fundamental resonant frequency versus mandrel length. (b) Fundamental frequency versus Young's modulus of the mandrel.

TABLE II. Studied parameters of a mandrel hydrophone.

Name	Constitutional parts	Parameters
Geometry	Mandrel	Length=1.56–5.1 cm Radius=0.75–2.5 cm
	Molding	Thickness=0–1 mm
Material properties	Mandrel	Young's modulus=1–5 GPa Bulk modulus=0.5–8 GPa Poisson's ratio=0.2–0.45
	Molding	Young's modulus=0.1–5 GPa Poisson's ratio=0.2–0.45

$$L_{\min}=\frac{l\star D}{2\star\pi\star(R_m+D/2)}. \tag{9}$$

For example, if the mandrel radius is 2.5 cm, its minimum length must be 1.6 cm to achieve a constant fiber length of 10 m. On the other hand, the maximum diameter of a mandrel must be equal to its maximum length to achieve a resonance frequency of not lower than 10 kHz. The minimum diameter corresponds to the case in Eq. (9) where the mandrel length is at its maximum because the fiber length is constant.

The material parameters of the acoustic sensing part are a Young's modulus, a bulk modulus, and a Poisson's ratio. However, the independent variables of the material are either a Young's modulus and a bulk modulus, or a Young's modulus and a Poisson's ratio due to the cross-relationships of these parameters. Typical mandrel materials are Teflon, Nylon, Si, and rubber, while the molding materials are polyethylene, polypropylene, PVC, and rubber.[1,2,5,12] The Young's modulus of these materials lies between 0.1 and 6 GPa, and the bulk modulus between 1 and 8 GPa. The density of these materials has a nearly constant value of about 1500 kg/m^3. Figure 2(b) shows a variation of the fundamental mode of a mandrel that has a length of 4 cm and a radius of 2 cm as the representative geometry. The mandrel with a Young's modulus of 1 GPa has a fundamental mode around 10 kHz. Therefore, the variation range of a Young's modulus is selected from 1 to 6 GPa. All the literature reports Poisson's ratio as ranging from 0.2 to 0.45. Therefore, the bulk modulus of a mandrel automatically ranges from 0.5 to 8 GPa due to the elastic inter-relationships between the parameters.

For the molding parts working as a protection layer for the fiber, the range of Young's modulus is set from 0.1 to 5 GPa as they are generally more flexible than the mandrel. We assume that the thickness of the molding is less than or equal to 1 mm. Table II summarizes all of the above concerning the materials, the geometry parameters, and their variation range for designing a fiber wound mandrel sensor.

III. FINITE ELEMENT MODELING AND ANALYSIS

To analyze the response of a mandrel hydrophone in an underwater environment, two separate finite element models were constructed. The first model, shown in Fig. 3, analyzes the transient and frequency response of the hydrophone. The second model, shown in Fig. 4, analyzes the directional pattern of the mandrel sensor. Due to the symmetrical characteristics of the sensor, only a quarter of the whole sensor is

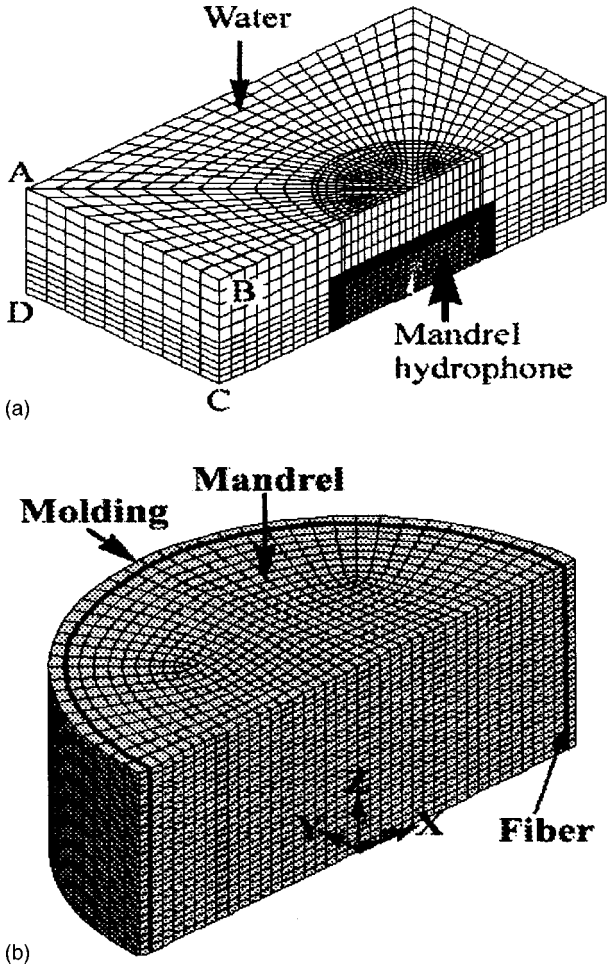
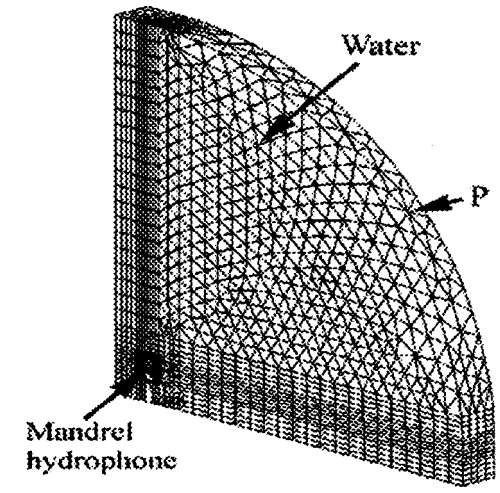
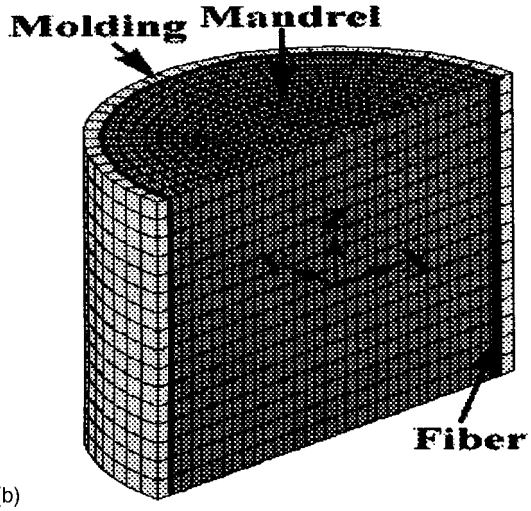


FIG. 3. Finite element model of a mandrel hydrophone for a pressure sensitivity and a frequency response spectrum analysis. (a) Course model. (b) Sub-model.

modeled. The modeling and analysis is carried out using the commercial finite element analysis software, ANSYS, which employs a particular modeling technique of substructuring in order to achieve an acceptable accuracy with the minimum computer processing time. First, a coarse three-dimensional model is constructed that includes the acoustic fluid and describes the hydrophone as a simple cylinder. Then the mandrel sensor is fine meshed again as a submodel and its response is analyzed in detail based on the calculation results of the coarse model. In Fig. 3, the wall ABCD works as a plane wave source and applies a 1-kHz acoustic pressure to the sensor. Figure 3(b) is an enlarged view of the hydrophone model showing the mandrel/fiber/molding area as described in detail in Fig. 1(b). In Fig. 4, each point P along the fluid boundary works as a point source and applies 5 kHz of acoustic pressure. To overcome the limitation given by the finite dimension of the fluid medium, a pressure relief condition is applied to all the fluid boundaries. In this analysis, no artificial constraint or simplification has been made on the deformation of the hydrophone. All the possible strain components, including the shear parts of the hydrophone, are calculated first and then their principal values are determined by means of an internal calculation function of the ANSYS. A sensitivity computation is carried out with the principal strains of the optical fiber using Eq. (8).



(a)



(b)

FIG. 4. Finite element model of a mandrel hydrophone for a directivity pattern analysis. (a) Coarse model. (b) Sub-model.

IV. RESULTS AND DISCUSSION

A. Influence of the mandrel on the hydrophone sensitivity

For the mandrel, the material parameters that have an effect on the pressure sensitivity are the bulk modulus, Young's modulus, and Poisson's ratio. Figures 5 and 6 show the variations of pressure sensitivity for material properties with fixed mandrel geometries of a 5-cm length and a 3-cm diameter. Teflon is molded over the fiber with a thickness of 1 mm. In Fig. 5, with a Young's modulus E, the pressure sensitivity decreases until a certain point, after which it starts to increase until it eventually reaches a saturation point. However, a mandrel with a low bulk modulus shows the opposite trend. With a bulk modulus, the sensitivity decreases in the high-E range, while in the low-E range, especially between 1 and 2 GPa, no general trend is observed. On the other hand, Fig. 6 shows a stable trend, i.e., the pressure sensitivity decreases while either Poisson's ratio or Young's modulus increases. When considering the cross-relationships of the elastic parameters, the complicated behavior in the low bulk modulus region of Fig. 5 could be explained as an

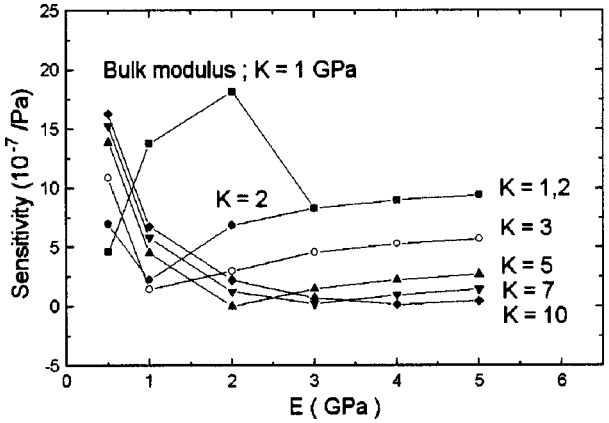


FIG. 5. Variation of pressure sensitivity of a hydrophone in relation to the Young's modulus and the bulk modulus of the mandrel at 1 kHz (length=5 cm, diameter=3 cm).

inconsistent change of Poisson's ratio. From this point of view, the bulk modulus is thus an inefficient parameter for controlling the performance of the hydrophone. It is, therefore, concluded that a hydrophone with the highest sensitivity requires a mandrel that possesses both the lowest Poisson's ratio and the lowest Young's modulus.

The effect of mandrel geometry on pressure sensitivity is shown in Figs. 7 and 8. The results reflect a mandrel of constant properties, a Young's modulus of 2 GPa, and a Poisson's ratio of 0.2. Figure 7 shows the variation of sensitivity in relation to a change of mandrel diameter. The change in diameter does not seem to have much influence on the pressure sensitivity, but it does not show any consistent relationship. Thus the diameter of a mandrel is also not an efficient parameter which improves the performance of a hydrophone. Figure 8 shows the pressure sensitivity changes in relation to mandrel length. The sensitivity steadily increases with the mandrel length to a certain level and then converges to a constant value. Therefore, the length of a mandrel is more efficient as a parameter that can improve sensitivity, and it would appear that the mandrel should be as long as possible unless it violates the flat frequency response requirement of up to 5 kHz.

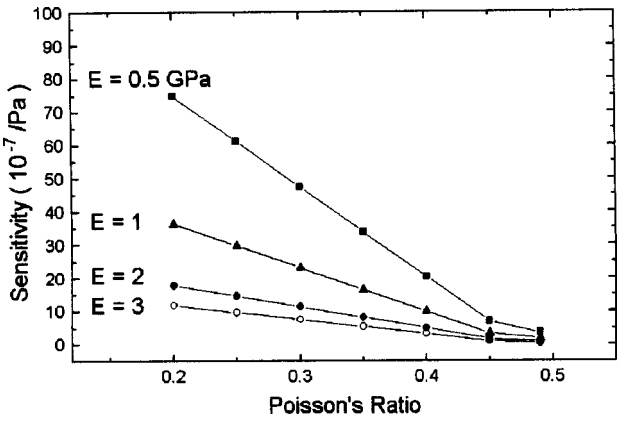


FIG. 6. Variation of pressure sensitivity of a hydrophone in relation to the Poisson's ratio and the Young's modulus of the mandrel at 1 kHz (length=5 cm, diameter=2.5 cm).

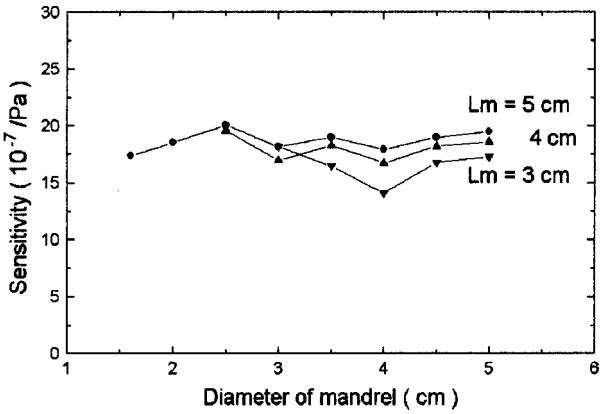


FIG. 7. Variation of pressure sensitivity of a hydrophone in relation to the diameter of the mandrel (Lm=mandrel length).

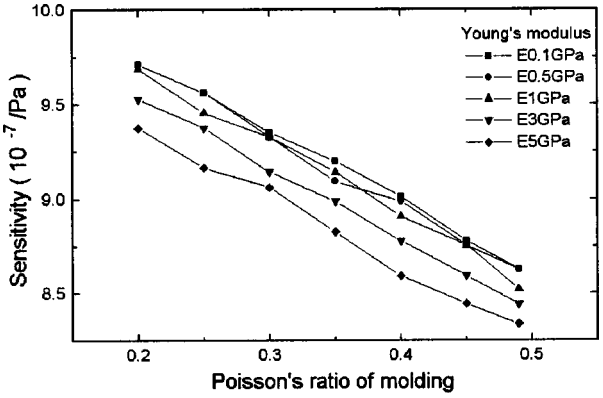


FIG. 9. Variation of the pressure sensitivity of a hydrophone in relation to the Poisson's ratio and Young's modulus of the molding.

B. Influence of the molding layer on the hydrophone sensitivity

The influence of the molding layer on the sensitivity of a hydrophone is represented in Figs. 9 and 10. Figure 9 is the variation of the sensitivity in relation to Poisson's ratio, and Fig. 10 in relation to the thickness of the molding layer. Figure 10 has been drawn for two representative values of E. The calculations have been performed fixing all the parameters of the mandrel at exemplary values, that is, a length of 3.15 cm, a diameter of 2.5 cm, a Young's modulus of 2 GPa, and a Poisson's ratio of 0.3. As shown in Fig. 9, the hydrophone has higher sensitivity with the decrease of both Young's modulus and Poisson's ratio. The pressure sensitivity also increases with the thickness in Fig. 10, although not to any great amount. These results indicate that the more flexibility in the molding layer, the more deformation induced in the optical fiber. Therefore, the molding layer with the lowest Young's modulus and Poisson's ratio should be coated as thickly as possible in order to improve the pressure sensitivity of a hydrophone.

C. Frequency response characteristics and directivity patterns

Using the above results, that is, the geometry and material properties of the constituent parts, we then analyzed the frequency response and the directivity pattern of a hydrophone. The frequency response analysis employs the model in Fig. 3 and investigates the sensor response from 500 Hz to 15 kHz. As a representative case the mandrel geometry is selected at a 4-cm length and a 2.5-cm diameter, and its material properties as a Young's modulus of 2 GPa and a Poisson's ratio of 0.3. A molding layer was selected with a 1-mm thickness, a Young's modulus of 1 GPa, and a Poisson's ratio of 0.3. Figure 11 is the result. The first resonance frequency of the hydrophone is 11.5 kHz, and it satisfies the requirement of a flat response of up to 5 kHz. The pressure sensitivity is about 27×10^{-7} per pascal up to 7 kHz, and then shows a rather unstable characteristic above that frequency. In the case when a wider frequency range is desired, the resonance frequency should then be increased. A higher resonance frequency requires a shorter mandrel and a higher Young's modulus as shown in Fig. 2. However, a higher Young's modulus has the undesirable effect of decreasing the sensitivity of the sensor, and therefore the selection should be made carefully.

An analysis of the sensitivity of the directivity pattern of a hydrophone used the finite element model in Fig. 4. The mandrel geometry and material properties are the same as those used in the frequency response analysis. Figure 12 is the directivity pattern of a hydrophone at 5 kHz. The hydrophone has the sensitivity of $(30\pm3)\times10^{-7}$ per pascal in all the directions and shows a nearly omni-directional pattern at

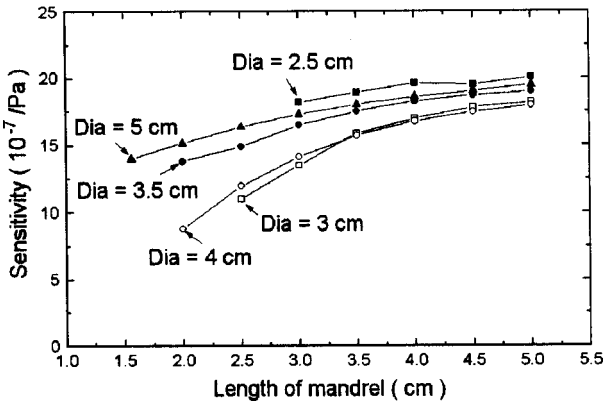


FIG. 8. Variation of pressure sensitivity of a hydrophone in relation to the length of the mandrel.

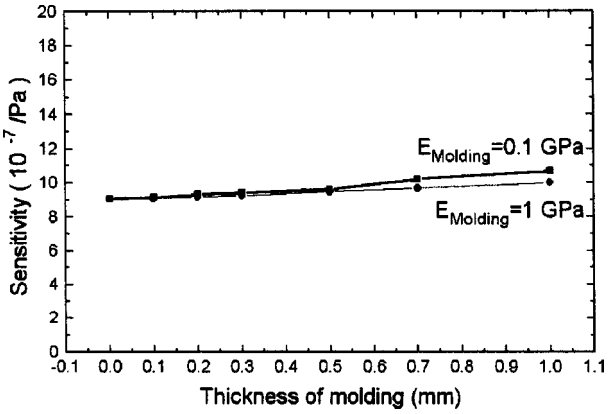


FIG. 10. Variation of the pressure sensitivity of a hydrophone in relation to the thickness of the molding.

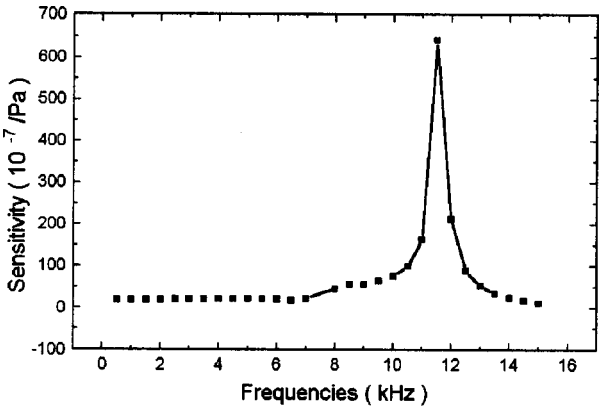


FIG. 11. Frequency response characteristics of a hydrophone.

5 kHz. Within the geometrical variation range in Table II, all the models have similar patterns with a directional fluctuation of less than 10%. If a more nearly uniform omnidirectionality is desired, a smaller geometry should be employed. However, a smaller geometry will reduce the sensitivity of the sensor as shown in Fig. 8.

V. CONCLUSION

This paper has investigated the influence of the structural parameters of an optical fiber wound mandrel hydrophone on its sensitivity by using the finite element method. A hydrophone should have as high a sensitivity as possible, a flat frequency response of up to 5 kHz, and an omnidirectional sensitivity pattern within the frequency range. The parameters studied include the mandrel geometry, the molding thickness, and the material properties of the constitutional parts of a hydrophone. The analysis results show that the pressure sensitivity of a hydrophone increases as both the mandrel length and the molding thickness increase. Also a higher pressure sensitivity requires a mandrel or a molding material with both a low Young's modulus and a low Poisson's ratio. On the contrary, a flat frequency response increases as either the mandrel length is shortened or the mandrel becomes harder. A shorter mandrel also provides a better omni-directional characteristic, that is, when a wider frequency range or a more nearly uniform omnidirectionality is desired, the structural parameters must be selected at the expense of sensitivity. These results indicate that the best sensitivity for a hydrophone is accomplished when the mandrel length and the molding thickness are as large as possible and when its constituent materials are as compliant as possible unless it violates the flat response requirement up to a given frequency range. The results of the present work are thought to provide valuable tools for making the optimum selection of structural parameters for a mandrel-type fiber optic hydrophone.

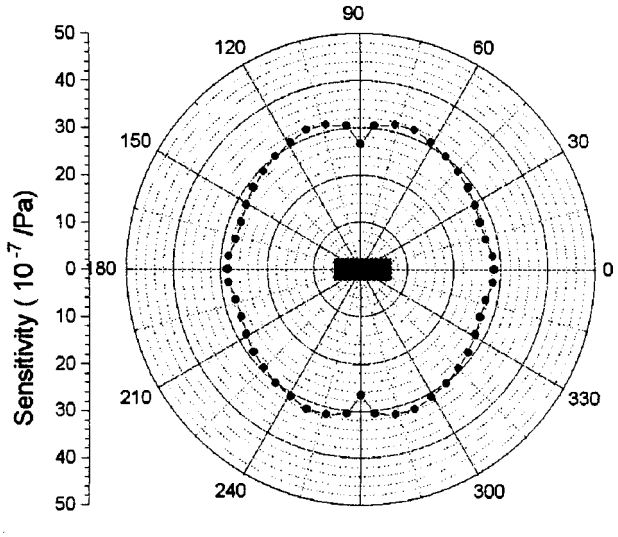


FIG. 12. Directivity patterns of a hydrophone sensitivity at 5 kHz.

[1] J. A. Bucaro, N. Lagakos, J. H. Cole, and T. G. Giallorenzi, ''Fiber optic acoustic transduction,'' in *Physical Acoustics*, Vol. XVI, edited by R. H. Thurston (Academic, New York, 1982), p. 385–457.
[2] N. Lagakos, E. U. Schnaus, J. H. Cole, J. Jarzynski, and J. A. Bucaro, ''Optimizing fiber coatings for interferometric acoustic sensors,'' IEEE J. Quantum Electron. **18**(4), 683–689 (1982).
[3] J. Jarzynski, R. Hughes, T. R. Hickman, and J. A. Bucaro, ''Frequency response of interferometric fiberoptic coil hydrophone,'' J. Acoust. Soc. Am. **69**, 1799–1808 (1981).
[4] S. Africk, T. Burton, P. Jameson, and A. Ordubadi, ''Design studies for fiber optic hydrophones,'' Report No. 4658, Bolt, Beranek & Newman, Cambridge, MA, 1981.
[5] R. Hughes and J. Jarzynski, ''Static pressure sensitivity amplification in interferometric fiberoptic hydrophones,'' Appl. Opt. **19**(1), 98–107 (1980).
[6] V. S. Sudarshanam and K. Srinivasan, ''Static phase change in a fiber optic coil hydrophone,'' Appl. Opt. **29**(6), 855–863 (1990).
[7] J. F. Nye, *Physical Properties of Crystals* (Clarendon, Oxford, 1976).
[8] G. B. Hocker, ''Fiber-optic sensing of pressure and temperature,'' Appl. Opt. **18**(9), 1445–1448 (1979).
[9] H. L. Price, ''On the mechanism of transduction in optical fiber hydrophone,'' J. Acoust. Soc. Am. **66**(4), 976–979 (1979).
[10] J. A. Bucaro and T. R. Hickman, ''Measurement of sensitivity of optical fibers for acoustic detection,'' Appl. Opt. **18**(6), 938–940 (1979).
[11] N. Lagakos, J. A. Bucaro, and R. Hughes, ''Acoustic sensitivity predictions of single mode optical fibers using Brillouin scattering,'' Appl. Opt. **19**(21), 3666–3670 (1980).
[12] D. W. Van Krevelen, *Properties of Polymers* (Elsevier, New York, 1976).

Ultrasonic three-dimensional probe modeling in anisotropic solids

A. Jonas Niklasson
Division of Mechanics, Chalmers University of Technology, S-412 96 Göteborg, Sweden

(Received 12 March 1996; revised 11 December 1997; accepted 30 January 1998)

Three-dimensional wave propagation and probe modeling in a homogeneous, linear elastic, anisotropic half-space is discussed. The probe is modeled by modifying the traction vector at the interface between an isotropic half-space and an anisotropic half-space. The traction vector is obtained by considering a plane wave in the isotropic solid incident upon the interface. The probe can be of any type, angle, frequency, and size. The probe is attached to an anisotropic half-space and the wave propagation problem is solved by means of Fourier transform techniques. Numerical examples illustrate some of the properties of the probe model and wave propagation in anisotropic solids. © *1998 Acoustical Society of America.* [S0001-4966(98)02905-1]

PACS numbers: 43.38.Ar, 43.35.Zc [SLE]

INTRODUCTION

Nondestructive testing (NDT) is a vital part of the security and maintenance of, for example, nuclear power plants. A widely used testing method is ultrasonic NDT. The method has proven to be very successful when dealing with isotropic solids where a lot of intuitive knowledge can be used. In the last decades, the use of anisotropic materials has increased rapidly. When these materials are introduced a lot of the intuitive knowledge is lost, one reason being that the phase and group velocities no longer coincide. This makes it difficult to perform and evaluate ultrasonic NDT on components made of anisotropic materials. Therefore, there are good reasons to model test situations for anisotropic materials.

Wave propagation in isotropic solids has been studied for a long period of time and there are numerous results (both theoretical and numerical) in the literature. The number of work dealing with wave propagation in anisotropic materials is significantly lower. Some basic theorems and results are given by Wheeler.[1] Auld[2] reviews the theory of wave propagation in anisotropic solids and discusses different types of materials. Van der Hijden[3] discusses propagation of transient elastic waves in stratified anisotropic media using the method of Cagniard–de Hoop. Lord *et al.*[4] use FEM to solve mainly two-dimensional (2-D) wave propagation problems in various media. Green's functions in anisotropic solids are derived by Wang and Achenbach.[5]

In order to simulate complete test situations, models of probes and various defects are needed. In the isotropic case, complete models are given by, for example, Boström and Wirdelius[6] and Chapman.[7] In the anisotropic case little has been done in this area. A 2-D *SH* model, including a probe and a striplike interface flaw, is given by Mattsson and Niklasson.[8] Rajapakse and Gross[9] study the transient response of an orthotropic medium with a cavity. The model is 2D and the waves are generated by a point force. Fellinger *et al.*[10] use the elastodynamic finite integration technique (EFIT) to solve mainly 2-D wave propagation in various materials containing defects. Wave propagation through a bimetallic weld is studied by Minachi *et al.*[11] and Nouailhas *et al.*[12] Minachi *et al.*[11] study beam propagation and compare the Gauss–Hermite beam model and a finite element method. Nouailhas *et al.*[12] compare a ray tracing code and a finite element code and use various experiments to validate the codes. In both of these last two papers, the computations are performed in 2D.

Probe modeling in isotropic solids has been done by, for example, Boström and Wirdelius,[6] Kažys and Mažeika,[13] and Schmerr and Sedov.[14] The probe by Boström and Wirdelius[6] prescribes the traction on an isotropic half-space. Schmerr and Sedov[14] use a similar approach in their modeling, but the transducer is restricted to have a circular shape. Kažys and Mažeika[13] use spatial impulse responses and FEM to model the field inside the probe. In the anisotropic case, a 2-D *SH* probe is modeled by Niklasson.[15] This probe is modeled in the same way as the probe by Boström and Wirdelius.[6]

The plan of the present paper is as follows. In Sec. I, some basic results in wave propagation in anisotropic solids are reviewed. The probe is modeled by the traction it is causing on the surface of an anisotropic half-space in Sec. II. The approach here is similar to the approaches by Boström and Wirdelius[6] and Niklasson.[15] The traction is obtained by considering a plane wave in an isotropic half-space incident upon an anisotropic half-space. The probe is modeled by a modification of the traction at the interface. The probe is attached to an anisotropic half-space in Sec. III and the corresponding wave propagation problem is solved by means of Fourier transform techniques. Various numerical examples are presented in Sec. IV which illustrate some of the probe characteristics and properties of wave propagation in anisotropic solids.

I. PRELIMINARIES

Consider a homogeneous, lossless, anisotropic solid. The existence of a strain energy function is assumed. The equations of motion in the absence of body forces are

$$\nabla \cdot \sigma = \rho_A \ddot{\mathbf{u}}, \tag{1}$$

where $\mathbf{u}$ is the displacement vector, σ is the (symmetric) stress tensor, and ρ_A is the (constant) density and dot denotes

partial differentiation with respect to time t. Since a strain energy function is assumed to exist, Hooke's law in abbreviated indices is

$$\begin{pmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{xz} \\ \sigma_{xy} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{pmatrix} \begin{pmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{zz} \\ 2\epsilon_{yz} \\ 2\epsilon_{xz} \\ 2\epsilon_{xy} \end{pmatrix}, \tag{2}$$

where $C=(C_{ij})$ is the positive definite stiffness matrix and

$$\epsilon_{ij}=(u_{i,j}+u_{j,i})/2, \quad i,j=x,y,z, \tag{3}$$

is the strain tensor. If Eqs. (1)–(3) are combined, the wave equations expressed in the components of $\mathbf{u}$ are obtained.

A plane wave is defined by $\mathbf{u}=A\mathbf{d}e^{i(qx+py+hz-\omega t)}$, where A is the amplitude, $\mathbf{d}$ is the polarization vector, q,p,h are the wave numbers in the x,y,z directions, respectively, and ω is the circular frequency. Direct insertion into the equations of motion leads to the homogeneous linear system of equations called the Christoffel equation[2]

$$D\mathbf{d}=0 \quad \text{where} \quad D=LCL^T-\rho_A\omega^2 I,$$

$$L=\begin{pmatrix} q & 0 & 0 & 0 & h & p \\ 0 & p & 0 & h & 0 & q \\ 0 & 0 & h & p & q & 0 \end{pmatrix}, \tag{4}$$

where I is the 3-by-3 identity matrix and the superscript T denotes transpose. If q, p, and ω are known, Eq. (4) splits up into

$$\begin{aligned} &\text{I. Find } h \text{ such that } \det D=0. \\ &\text{II. Given } h \text{ from I, find } \mathbf{d} \text{ such that } D\mathbf{d} \\ &\quad =0 \text{ and } |\mathbf{d}|=1. \end{aligned} \tag{5}$$

If wave propagation in, for example, a half-space is considered, three values of h and $\mathbf{d}$ are admissible. These values corresponds to different types of waves in the half-space. One talks about quasi P (qP) and quasi S (qS) waves since these waves in general are mixtures of P and S waves. The relation $\det D=0$ is the characteristic equation for the anisotropic solid. From this relation, the group velocity $\mathbf{v}$, which is equal to the energy velocity in a homogeneous lossless medium, is easily obtained[2]

$$\mathbf{v}=-\frac{\nabla_k(\det D)}{\partial(\det D)/\partial\omega}, \quad \nabla_k F=\left(\frac{\partial F}{\partial q},\frac{\partial F}{\partial p},\frac{\partial F}{\partial h}\right). \tag{6}$$

The group velocity plays an important role when wave propagation in infinite bodies is considered. Then, the admissible values of h are selected from radiation conditions on the energy (see below).

II. THE PROBE MODEL

Consider an anisotropic, homogeneous half-space. On the surface of the half-space, an ultrasonic probe is attached. The origin of the coordinate system (x,y,z) is located on the surface (see Fig. 1). In this section the aim is to model a

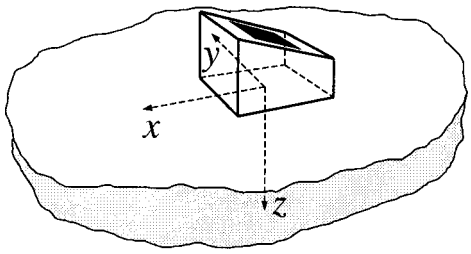


FIG. 1. A probe on a half-space.

piezoelectric contact probe. The probe model will prescribe the traction beneath the probe and not the displacement. This approach has the advantage that the boundary condition on the surface of the half-space will not be mixed. This simplifies the calculations a great deal when Fourier methods are used. The approach in this paper is similar to the approaches by Boström and Wirdelius[6] and Niklasson.[15]

Consider the case when the probe is embedded in an isotropic half-space made of the same material as the wedge of the probe (see Fig. 2). λ and μ are the Lamé constants and ρ_1 the density of the wedge and the isotropic half-space. The surface S is the part of the xy plane which is affected by the field in the wedge. In order to model the probe, a plane wave in the isotropic half-space incident upon the interface is considered. The main purpose of considering a plane wave is to obtain the phases and relative amplitudes of the different waves which are transmitted into the anisotropic half-space (one qP wave and two qS waves). The direction of propagation of the plane wave is normal to the piezoelectric crystal. If the wave is propagating in the xz plane then

$$\begin{aligned} &\mathbf{u}^{(i)}=\mathbf{u}^{(0)}=A_0(d_x^{(0)},d_y^{(0)},d_z^{(0)})e^{i\eta_0}, \\ &\eta_0=k_0(\sin\,\theta_0 x+\cos\,\theta_0 z)-\omega t. \end{aligned} \tag{7}$$

Here, three types of incident waves are considered

$$\begin{aligned} &P \text{ wave } (P \text{ probe}): \quad k_0=\omega\sqrt{\rho_I/(\lambda+2\mu)}, \\ &\qquad\qquad\qquad\quad \mathbf{d}^{(0)}=(\sin\,\theta_0,0,\cos\,\theta_0). \end{aligned} \tag{8a}$$

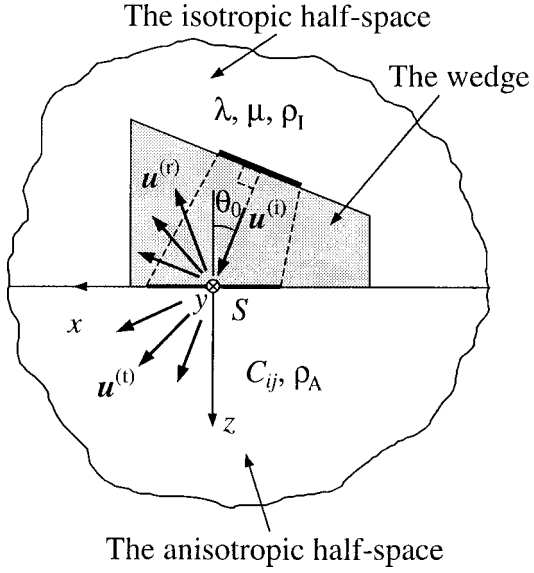


FIG. 2. The embedded probe.

$$SV \text{ wave } (SV \text{ probe}): \quad k_0=\omega\sqrt{\rho_1/\mu}, \quad \mathbf{d}^{(0)}=(\cos\theta_0,0,-\sin\theta_0). \tag{8b}$$

$$SH \text{ wave } (SH \text{ probe}): \quad k_0=\omega\sqrt{\rho_I/\mu}, \quad \mathbf{d}^{(0)}=(0,1,0). \tag{8c}$$

By choosing an incident P, SV, or SH wave, it is possible to model a P, SV, or SH probe, respectively. The following ansatz is made for the reflected and transmitted fields

$$\mathbf{u}^{(r)}=\mathbf{u}^{(1)}+\mathbf{u}^{(2)}+\mathbf{u}^{(3)}, \quad z<0, \tag{9a}$$

$$\mathbf{u}^{(t)}=\mathbf{u}^{(4)}+\mathbf{u}^{(5)}+\mathbf{u}^{(6)}, \quad z>0, \tag{9b}$$

where

$$\mathbf{u}^{(1)}=A_1(\sin\theta_1,0,-\cos\theta_1)e^{i\eta_1}, \quad \eta_1=k_1(\sin\theta_1 x-\cos\theta_1 z)-\omega t, \tag{10a}$$

$$\mathbf{u}^{(2)}=A_2(\cos\theta_2,0,\sin\theta_2)e^{i\eta_2}, \quad \eta_2=k_2(\sin\theta_2 x-\cos\theta_2 z)-\omega t, \tag{10b}$$

$$\mathbf{u}^{(3)}=A_3(0,1,0)e^{i\eta_3}, \quad \eta_3=k_3(\sin\theta_3 x-\cos\theta_3 z)-\omega t, \tag{10c}$$

$$\mathbf{u}^{(4)}=A_4(d_x^{(4)},d_y^{(4)},d_z^{(4)})e^{i\eta_4}, \quad \eta_4=q^{(4)}x+h^{(4)}z-\omega t, \tag{10d}$$

$$\mathbf{u}^{(5)}=A_5(d_x^{(5)},d_y^{(5)},d_z^{(5)})e^{i\eta_5}, \quad \eta_5=q^{(5)}x+h^{(5)}z-\omega t, \tag{10e}$$

$$\mathbf{u}^{(6)}=A_6(d_x^{(6)},d_y^{(6)},d_z^{(6)})e^{i\eta_6}, \quad \eta_6=q^{(6)}x+h^{(6)}z-\omega t. \tag{10f}$$

Since $\mathbf{u}^{(1)}$, $\mathbf{u}^{(2)}$, and $\mathbf{u}^{(3)}$ correspond to a reflected P, SV, and SH wave, respectively, the following hold

$$k_1=\omega\sqrt{\rho_I/(\lambda+2\mu)}, \quad k_2=k_3=\omega\sqrt{\rho_I/\mu}. \tag{11}$$

The transmitted fields cannot be interpreted in the same simple fashion.

If perfect (welded) contact between the half-spaces is assumed, the following hold at the interface

$$\mathbf{u}^{(i)}+\mathbf{u}^{(r)}=\mathbf{u}^{(t)}, \quad z=0, \tag{12a}$$

$$(\sigma^{(i)}+\sigma^{(r)})\cdot\mathbf{e}_z=\sigma^{(t)}\cdot\mathbf{e}_z, \quad z=0, \tag{12b}$$

where $\mathbf{e}_z$ is the unit vector in the z direction. If the fields (7) and (9) are inserted into (12) it is concluded that

$$\sin\theta_1=(k_0/k_1)\sin\theta_0, \quad \sin\theta_2=(k_0/k_2)\sin\theta_0, \quad \theta_2=\theta_3, \quad q^{(4)}=q^{(5)}=q^{(6)}=k_0\sin\theta_0. \tag{13}$$

The values of $h^{(j)}$ and $\mathbf{d}^{(j)}$ $(j=4,5,6)$ are obtained as the solutions of (5) combined with radiation conditions on the energy. The radiation conditions are

$$\mathbf{v}\cdot\mathbf{e}_z>0, \quad \text{Im } h=0, \tag{14a}$$

$$\text{Im } h>0, \quad \text{Im } h\neq 0. \tag{14b}$$

where Im means the imaginary part as usual. In addition, insertion into (12) yields a system of equations for the amplitudes A_j $(j=1,...,6)$ expressed in the amplitude of the incident wave A_0. When this system of equations is solved the reflected and transmitted fields are determined.

At this point, the double spatial Fourier transform and the temporal Fourier transform are introduced as

$$\hat{f}(q,p)=\int\int_{\mathbf{R}^2} f(x,y)e^{-i(qx+py)}\,dx\,dy, \quad \tilde{f}(\omega)=\int_{\mathbf{R}} f(t)e^{i\omega t}\,dt, \tag{15a}$$

and their inverses as

$$f(x,y)=\frac{1}{4\pi^2}\int\int_{\mathbf{R}^2}\hat{f}(q,p)e^{i(qx+py)}\,dq\,dp, \quad f(t)=\frac{1}{2\pi}\int_{\mathbf{R}}\tilde{f}(\omega)e^{-i\omega t}\,d\omega, \tag{15b}$$

where $\mathbf{R}$ denotes the set of real numbers. As mentioned above, the main purpose of considering the plane wave incident upon the interface was to obtain the phases and relative amplitudes of the waves transmitted into the anisotropic half-space. In order to model the probe, the traction caused by the plane waves above at the interface $z=0$ is modified. The expression for the modified traction is

$$\begin{aligned}(\tilde{\sigma}\cdot\mathbf{e}_z)_{z=0}=i\{&\mu(k_0A_0(d_x^{(0)}\cos\theta_0+d_z^{(0)}\sin\theta_0)\\&-k_1A_1\sin 2\theta_1-k_2A_2\cos 2\theta_2)\delta_x\mathbf{e}_x\\&+\mu(k_0A_0d_y^{(0)}\cos\theta_0-k_3A_3\cos\theta_3)\delta_y\mathbf{e}_y\\&+[k_0A_0(\lambda(d_x^{(0)}\sin\theta_0+d_z^{(0)}\cos\theta_0)\\&+2\mu d_z^{(0)}\cos\theta_0)+k_1A_1(\lambda+2\mu\cos^2\theta_1)\\&-\mu k_2A_2\sin 2\theta_2]\delta_z\mathbf{e}_z\}e^{ik_0\sin\theta_0 x}g(x,y)\tilde{f}(\omega).\end{aligned} \tag{16}$$

The probe model (16) is obtained from the traction in (12b), in this case the traction expressed in the fields in the isotropic half-space, multiplied with $\tilde{f}$ and g. The factor $e^{-i\omega t}$ is omitted and the traction is taken as the temporally Fourier transformed traction. The function $g(x,y)$ in Eq. (16) models the spatial distribution of the traction on the surface and its support is the surface S. $\tilde{f}(\omega)$ models the frequency spectrum of the probe. The constants δ_x, δ_y, and δ_z allow weighing of the components of the traction. The contact is perfect if $\delta_x=\delta_y=\delta_z=1$. Another important special case is when $\delta_x=\delta_y=0$ and $\delta_z=1$. This corresponds to (inviscid) fluid contact between the probe and the half-space. In this case, the traction is

$$\begin{aligned}(\tilde{\sigma}\cdot\mathbf{e}_z)_{z=0}=i\{&k_0A_0(\lambda(d_x^{(0)}\sin\theta_0+d_z^{(0)}\cos\theta_0)\\&+2\mu d_z^{(0)}\cos\theta_0)+k_1A_1(\lambda+2\mu\cos^2\theta_1)\\&-\mu k_2A_2\sin 2\theta_2\}e^{ik_0\sin\theta_0 x}g(x,y)\tilde{f}(\omega)\mathbf{e}_z.\end{aligned} \tag{17}$$

The constants A_1 and A_2, in this case, are determined by the boundary conditions governing a smooth contact between the half-spaces

$$(\mathbf{u}^{(i)}+\mathbf{u}^{(r)})\cdot\mathbf{e}_z=\mathbf{u}^{(t)}\cdot\mathbf{e}_z, \quad z=0, \tag{18a}$$

$$((\sigma^{(i)}+\sigma^{(r)})\cdot\mathbf{e}_z)\cdot\mathbf{e}_z=(\sigma^{(t)}\cdot\mathbf{e}_z)\cdot\mathbf{e}_z, \quad z=0, \tag{18b}$$

$$((\sigma^{(i)}+\sigma^{(r)})\cdot\mathbf{e}_z)\cdot\mathbf{e}_x=(\sigma^{(t)}\cdot\mathbf{e}_z)\cdot\mathbf{e}_x=0, \quad z=0, \tag{18c}$$

$$((\sigma^{(i)}+\sigma^{(r)})\cdot\mathbf{e}_z)\cdot\mathbf{e}_y=(\sigma^{(t)}\cdot\mathbf{e}_z)\cdot\mathbf{e}_y=0, \quad z=0, \tag{18d}$$

which replace the boundary conditions (12).

As mentioned earlier, the approach above is similar to the approach by Boström and Wirdelius.[6] In fact, the probe model reduces to the probe model by Boström and Wirdelius[6] if the anisotropic half-space is taken as isotropic and the material of the wedge is identical to the material of the anisotropic half-space. As seen above, the labeling of the probe types is only related to the type of incident plane wave in the wedge. This does not mean, for example, that a *P* probe always generates *P* waves in an isotropic half-space since the angle θ_0 may exceed the critical angle for the transmitted *P* wave.

III. FORMAL SOLUTION

In this section, the wave propagation problem in an anisotropic half-space, subjected to the probe model discussed in the previous section, is solved. The solution is obtained by means of Fourier transform techniques.

Applying the Fourier transforms (15a) to Eqs. (1)–(3) and solving for $\hat{\tilde{\mathbf{u}}}$ yields the general solution as

$$\hat{\tilde{\mathbf{u}}}=B_1\mathbf{d}_1e^{ih_1z}+B_2\mathbf{d}_2e^{ih_2z}+B_3\mathbf{d}_3e^{ih_3z}, \tag{19}$$

where $\mathbf{d}_j$ and h_j are obtained from (5) subjected to radiation conditions on the energy (14). Applying the spatial Fourier transform to Eq. (16) yields

$$\begin{aligned}(\hat{\tilde{\sigma}}\cdot\mathbf{e}_z)_{z=0}=i\{&\mu(k_0A_0(d_x^{(0)}\cos\theta_0+d_z^{(0)}\sin\theta_0)-k_1A_1\sin 2\theta_1-k_2A_2\cos 2\theta_2)\delta_x\mathbf{e}_x\\&+\mu(k_0A_0d_y^{(0)}\cos\theta_0-A_3k_3\cos\theta_3)\delta_y\mathbf{e}_y+(k_0A_0(\lambda(d_x^{(0)}\sin\theta_0+d_z^{(0)}\cos\theta_0)+2\mu d_z^{(0)}\cos\theta_0)\\&+k_1A_1(\lambda+2\mu\cos^2\theta_1)-\mu k_2A_2\sin 2\theta_2)\delta_z\mathbf{e}_z\}\hat{g}(q-k_0\sin\theta_0,p)\tilde{f}(\omega).\end{aligned} \tag{20}$$

If the traction is calculated from (19) for $z=0$ and combined with (20), a linear system of equations for the B_j's is obtained. When this is solved, the transformed solution is obtained.

When the Fourier transformations are to be inverted some problems arise. In the qp plane, there may exist a large number of branch points and in some cases poles. Here, the integration is carried out using polar coordinates in the qp plane, i.e., $q=k_pr\cos\phi$ and $p=k_pr\sin\phi$ with k_p defined below. The integration path for r is deformed into the complex r-plane (in order to avoid the possible Rayleigh-like pole) as $r=s(1-i\alpha e^{-\beta s})$ where α and β are positive parameters and $s\in[0,\infty)$. The minimum of the imaginary part is attained at $s=1/\beta$ with $\mathrm{Im}\,r=-(\alpha/\beta)e^{-1}$. Here, the choices $\beta=1/r_b$, where r_b is the (real) radius at the outmost branch point, and $\alpha=\epsilon\beta$ (with ϵ between 0.01 and 0.1) are made. When q and/or p are complex, the radiation condition differs from the usual one.[3] Therefore, along the complex contour the admissible wave numbers h_j $(j=1,2,3)$ are selected by means of continuity provided that the integration starts on a physical Riemann sheet. This is easily accomplished for large values on s.

IV. NUMERICAL EXAMPLES

The complicated nature of the problem makes it in general impossible to perform analytical calculations. Therefore, it is important to have good numerical algorithms for root finding, solution of linear systems of equations and quadrature. The root finding is carried out using Laguerre's method and the linear systems of equations are solved using Gaussian elimination. The integration in the qp plane is performed using Gaussian quadrature and polar coordinates. The integration in r is carried out as described above with approximately 500 quadrature points in s and the integration in ϕ is carried out with approximately 300 quadrature points. The ω integration is performed with 100 quadrature points and the trapezoidal rule.

In the numerical examples below, a transversely isotropic graphite–epoxy composite is considered. The stiffness matrix, expressed in the crystal axes system with the z axis as the symmetry axis, is given by[16]

$$C=\begin{pmatrix}13.92 & 6.92 & 6.44 & 0 & 0 & 0\\ 6.92 & 13.92 & 6.44 & 0 & 0 & 0\\ 6.44 & 6.44 & 160.73 & 0 & 0 & 0\\ 0 & 0 & 0 & 7.07 & 0 & 0\\ 0 & 0 & 0 & 0 & 7.07 & 0\\ 0 & 0 & 0 & 0 & 0 & 3.50\end{pmatrix}\ \text{GPa}, \tag{21}$$

and the density is $\rho_A=1578\ \mathrm{kg/m^3}$. The wedge of the probe is made of polymethyl methacrylate (PMMA) which is assumed to be isotropic and lossless. The material properties of the wedge are $\lambda=2.6$ GPa, $\mu=2.1$ GPa, and $\rho_I=1180\ \mathrm{kg/m^3}$. The normalized slowness surfaces for the graphite–epoxy are shown in Fig. 3(a) in the plane $p=0$ and the corresponding normalized wave surfaces are shown in Fig. 3(b) in the plane $v_y=0$. In the figures, $c_p=\sqrt{C_{11}^c/\rho_A}$, $k_p=\omega/c_p$ and the material is considered in the crystal axes system with h as the wave number along the symmetry axis (the z axis). C_{11}^c is element (1,1) of the stiffness matrix expressed in the crystal axes system, i.e., $C_{11}^c=13.92$ GPa in

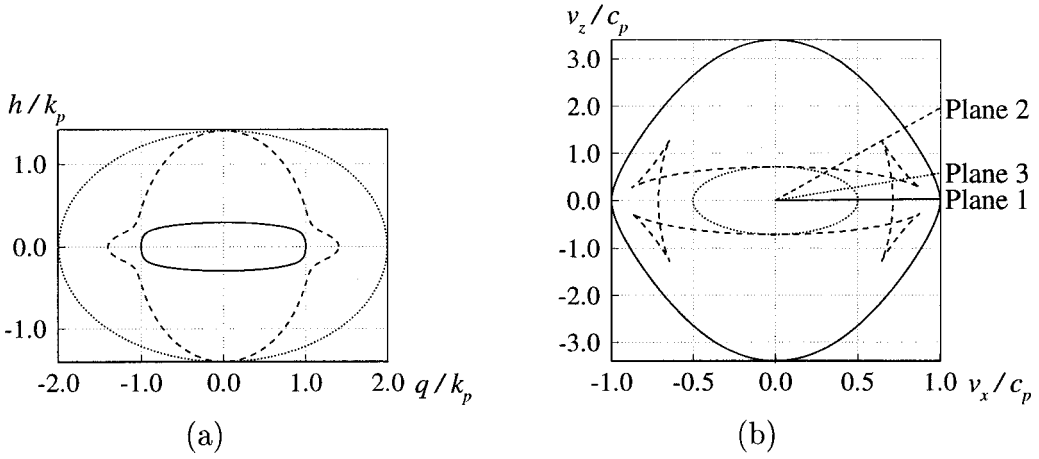


FIG. 3. The slowness and wave surfaces of the graphite-epoxy. (a) The slowness surfaces. (b) The wave surfaces.

this case. The slowness surfaces are obtained from the characteristic equation (5) and the three surfaces represent different waves. The three wave surfaces are obtained from the group velocity expression Eq. (6). In the figures, the solid curve is the quasi P (qP) wave, the dashed is the quasi SV (qSV) wave and the dotted is the SH wave. Pure SH waves exist due to the material symmetry considered, i.e., transverse isotropy. The complete slowness and wave surfaces are obtained as the surfaces of revolution around the h and v_z axes, respectively, due to this material symmetry. In Fig. 3(b), three straight lines are included. These lines are only included in the figure as aid in the identification of the waves (see Fig. 10) and they are not part of the wave surfaces. If the coordinate system (x,y,z) does not coincide with the crystal axes system, the elastic constants must be transformed. This is done with Bond matrices[2] and Euler angles. The Euler angles φ_1, φ_2 and φ_3 are defined according to Fig. 4. In Fig. 4, (X,Y,Z) represents the crystal axes system and (x,y,z) the probe system as before.

Two simple and convenient choices of the effective probe surface S are

$$S_1=\{(x,y)|(x,y)\in[-a,a]\times[-b,b]\},$$
$$S_2=\{(x,y)|(x/a)^2+(y/b)^2\leqslant 1\}, \tag{22}$$

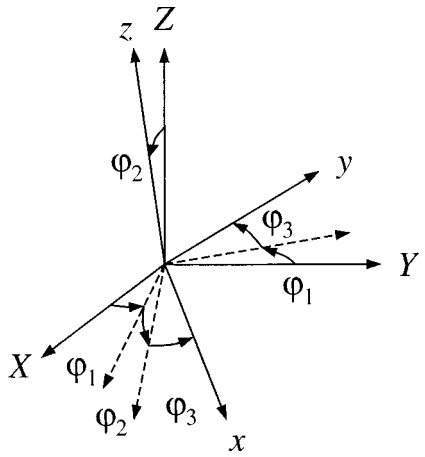


FIG. 4. The Euler angles.

i.e., a rectangular and an elliptical surface. If the characteristic function of a set A is introduced as

$$\chi_A=\begin{cases}1, & x\in A,\\ 0, & x\notin A,\end{cases} \tag{23}$$

some suitable choices of the spatial shape function $g(x,y)$ are

$$g_1(x,y)=\chi_{S_1},\quad g_2(x,y)=\sqrt{1-(x/a)^2)(1-(y/b)^2}\,\chi_{S_1},$$
$$g_3(x,y)=\chi_{S_2},\quad g_4(x,y)=\sqrt{1-(x/a)^2-(y/b)^2}\,\chi_{S_2}. \tag{24}$$

These choices of $g(x,y)$ are such that they can be Fourier transformed analytically and seem realistic. In the numerical examples, however, only $g_1(x,y)$ (the piston model with rectangular support) is used. This shape function is the simplest and the most widely used. In addition, the support is restricted to a square with side $2a$. This means that the probe is situated at $(x,y)\in[-a,a]\times[-a,a]$. The spatially Fourier transformed g_1 is (when $b=a$)

$$\hat{g}_1(q,p)=4\sin(qa)\sin(pa)/qp. \tag{25}$$

The frequency spectrum $\tilde{f}(\omega)$ is chosen as

$$\tilde{f}(\omega)=\sin^2\left(\frac{|\omega|-\omega_1}{\omega_2-\omega_1}\,\pi\right)\chi_\Omega,$$
$$\Omega=\{\omega|\omega\in[-\omega_2,-\omega_1]\cup[\omega_1,\omega_2]\}. \tag{26}$$

The reason for this choice of $\tilde{f}(\omega)$ is that it has a compact support and can approximate a pulse rather well.

In the figures, the displacement is represented in either the frequency domain as $|\mathrm{Re}\,\tilde{\mathbf{u}}|$ or in the time domain as $|\mathbf{u}|$. The fields are presented in contour plots consisting of 15 levels with black representing the largest value in each figure. The nondimensional frequency $\bar{\omega}$ and the nondimensional time τ are introduced as

$$\bar{\omega}=k_pa=\omega a\sqrt{\rho_A/C_{11}^c},\quad \tau=c_pt/a=(t/a)\sqrt{C_{11}^c/\rho_A}, \tag{27}$$

with a as above and all lengths are normalized with a. This means that the probe is about 7.1 S (SV or SH) wavelengths and 3.9P wavelengths wide at $\bar{\omega}=10$. Note that these are

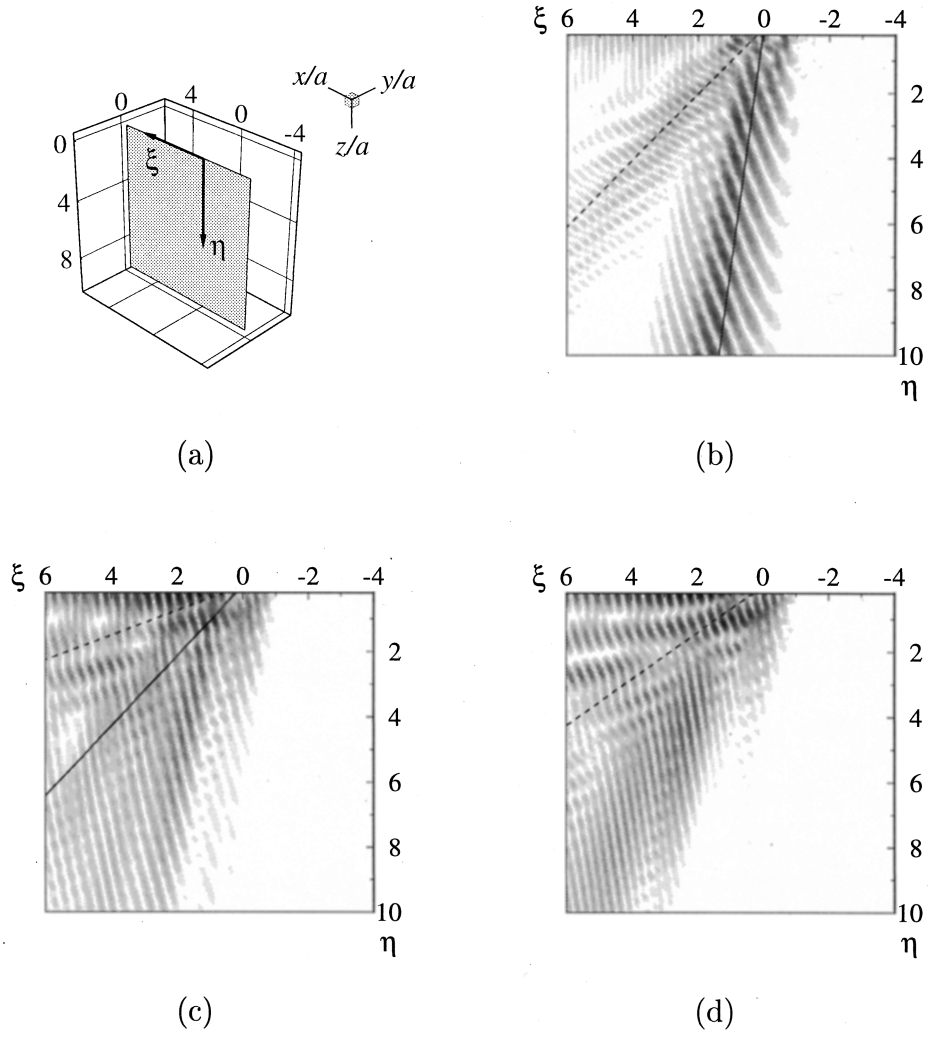


FIG. 5. $|\mathrm{Re}\,\tilde{\mathbf{u}}|$ for a P probe when $\theta_0=30°$, 50°, 60°, and $\bar{\omega}=10$. (a) The plane. (b) $\theta_0=30°$. (c) $\theta_0=50°$. (d) $\theta_0=60°$.

wavelengths in the isotropic wedge. The plot planes are chosen as the planes spanned by the group velocities belonging to the probe model (the transmitted plane waves) and the z axis. Here it is worth noting that there may be less than three group velocities. This is the case when the angle of incidence is larger than one or more of the critical angles. In the planes spanned by the group velocities, the coordinates ξ and η are used. They are defined by

$$\xi=\sqrt{x^2+y^2}/a, \quad \mathbf{e}_\xi=\mathbf{v}^{xy}/|\mathbf{v}^{xy}|,$$
$$\mathbf{v}^{xy}=(\mathbf{v}\cdot\mathbf{e}_x)\mathbf{e}_x+(\mathbf{v}\cdot\mathbf{e}_y)\mathbf{e}_y, \qquad \eta=z/a, \quad \mathbf{e}_\eta=\mathbf{e}_z. \tag{28}$$

Note that the probe will not be situated at $\xi\in[-1,1]$ unless the ξ axis coincides with the x or the y axis. In each of the different cases (but not in all of the figures), a plot showing the orientation of the planes is included. In addition, lines (solid, dashed or dotted) are included to indicate the group velocity directions of the main waves generated by the probe as rays originating from the origin.

In Fig. 5, the displacement in the frequency domain for a P probe is shown for three different values of the probe angle θ_0. The contact is perfect, i.e., $\delta_x=\delta_y=\delta_z=1$, the frequency is $\bar{\omega}=10$ and (x,y,z) coincides with the crystal axes system. The symmetry axis of the material is parallel to the z axis which means that the stiffness matrix is given by (21). The result is that the group velocities of the plane waves associated with the probe all lie in the xz plane and the $\xi\eta$ axes and the xz axes coincide. In the figures the solid lines indicate the group velocity direction of the qP wave and the dashed ones the direction of the qSV wave. Since the xz plane is a plane of elastic symmetry, no SH waves exist in this plane for a P probe. In the case when $\theta_0=30°$, which is far away from the critical angle, the direction of the qP wave is close to the axis of the greatest stiffness, i.e., the z axis. Here, the waves coincide very well with the group velocity directions in the figure and the difference in the direction of the phase and group velocity is very well illustrated. As θ_0 is increased, the directions of the waves are getting closer to the probe surface and it gets difficult to identify distinct waves. When θ_0 is 50° and 60°, large parts of the field are located close to the probe surface. Finally, when $\theta_0=60°$, $q/k_p=(k_0/k_p)\sin\theta_0=1.07$, and the critical angle for the qP wave is exceeded [see Fig. 3(a)]. Therefore only one ray (dashed) representing the qSV wave is included in the last plot.

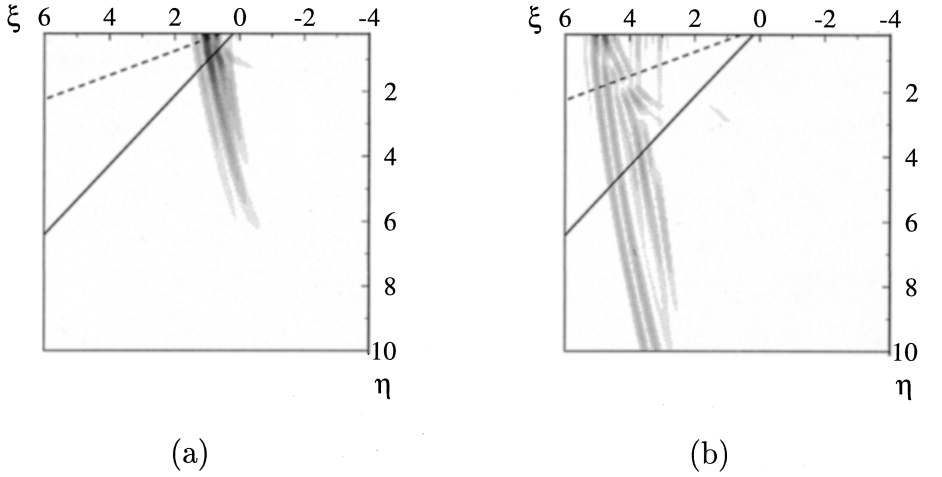


FIG. 6. Snapshots of $|\mathbf{u}|$ in the same situation as in Fig. 5(c). (a) $\tau=1$. (b) $\tau=5$.

The same situation as in Fig. 5(c) is considered in Fig. 6. Here, two snapshots of $|\mathbf{u}|$ are shown when $\tau=1$ and $\tau=5$. The lower and upper frequencies are $\overline{\omega_1}=1$ and $\overline{\omega_2}=19$, respectively. In Fig. 6(b), two waves can be identified, a qP wave with 80° between the phase velocity and the η axis (with group velocity along the solid line) and a qSV wave with 57° between the phase velocities and the η axis (with group velocity along the dashed line). Note that the angles of the phase velocity of the qP and the qSV waves are measured along the solid and dashed lines, respectively. It is evident from the figure that the high degree of anisotropy ''stretches'' the wave front a great deal.

In Fig. 7, fields from the two remaining probe types, the SV and SH probes, are shown in the frequency domain. These figures should be compared to the field in Fig. 5(b), which is the same situation for a P probe. Again, the symmetry of the material result in no SH waves for the SV probe and no qP and qSV waves for the SH probe in the xz plane. In the case of an SV probe, $q/k_p=(k_0/k_p)\sin\theta_0=1.11$ and the critical angle for the qP wave is exceeded [see Fig. 3(a)]. Therefore only one ray (solid) representing the qSV wave is included in the field from the SV probe. Note that as in the case with a P probe a large part of the field is located near the surface. For the SH probe one very distinct SH wave is identified (traveling along the solid line). Since the SH wave decouples from the two other waves in this case, this is expected. In Fig. 7(b) the near and far field are very nicely illustrated. When the distance to the probe is greater than approximately $3a$ the field seems to be radiated from a point source located at the origin.

In Fig. 8, wave propagation for a P probe is shown in the frequency domain for $\theta_0=30°$, $\delta_x=\delta_y=\delta_z=1$, and $\overline{\omega}=10$. Here, the crystal axes system does not coincide with (x,y,z). The transformation from the crystal axes system to the (x,y,z) system is accomplished with the Euler angles $\varphi_1=0°$, $\varphi_2=60°$, and $\varphi_3=45°$ (see Fig. 4). The transformed stiffness matrix is given by

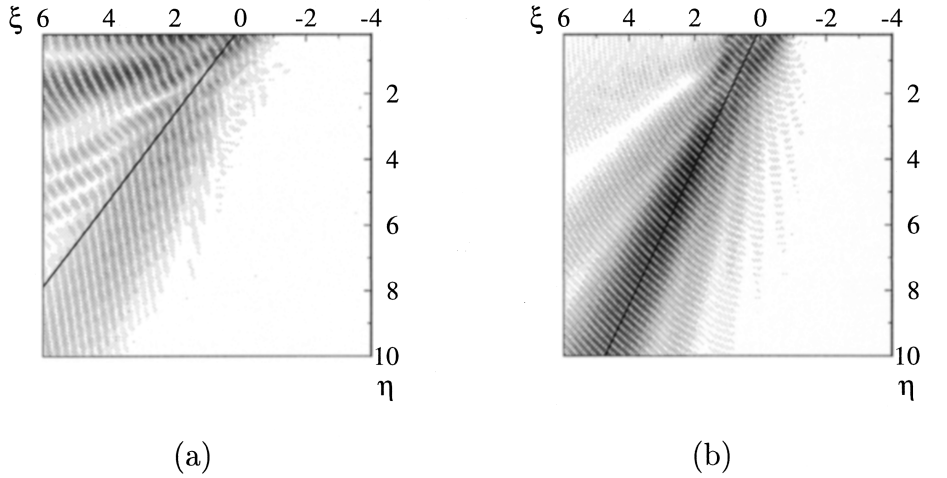


FIG. 7. $|\mathrm{Re}\,\tilde{\mathbf{u}}|$ for an SV and SH probe with $\theta_0=30°$ and $\overline{\omega}=10$. (a) SV probe. (b) SH probe.

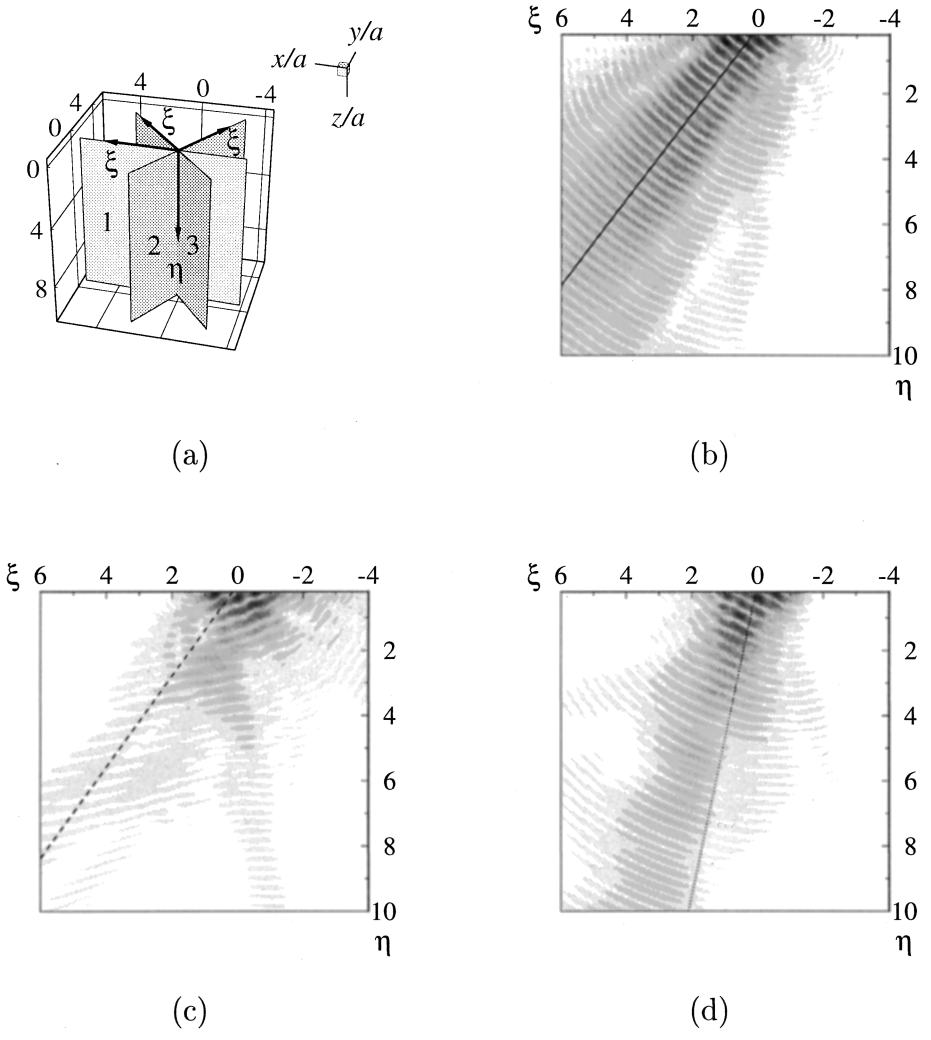


FIG. 8. $|\mathrm{Re}\,\tilde{\mathbf{u}}|$ for a P probe when $\theta_0=30°$, $\bar{\omega}=10$, $\varphi_1=0°$, $\varphi_2=60°$, and $\varphi_3=45°$. (a) The planes. (b) Plane 1. (c) Plane 2. (d) Plane 3.

$$C=\begin{pmatrix} 37.69 & 25.33 & 19.13 & 15.18 & -17.37 & -21.27 \\ 25.33 & 37.69 & 19.13 & 17.37 & -15.18 & -21.27 \\ 19.13 & 19.13 & 25.59 & 12.26 & -12.26 & -12.33 \\ 15.18 & 17.37 & 12.26 & 18.25 & -13.85 & -16.42 \\ -17.37 & -15.18 & -12.26 & -13.85 & 18.25 & 16.42 \\ -21.27 & -21.27 & -12.33 & -16.42 & 16.42 & 24.95 \end{pmatrix} \text{ GPa.} \tag{29}$$

Since (x,y,z) does not coincide with the crystal axes system, the three group velocities do not lie in one plane and the figure consists of the field plotted in three different planes. Plane 1 is spanned by the group velocity of the qP wave and the z axis, plane 2 by the group velocity of the qSV wave, and the z axis and plane 3 by the group velocity of the SH wave and the z-axis. In addition, the wave types do not decouple as they do in Figs. 5–7 and all three types will in general be present at the same time. In all of the figures, the solid line is the group velocity direction of the qP wave, the dashed line the direction of the qSV wave and the dotted line the direction of the SH wave. In all of the figures, a dominant part is seen close to the surface. Since this contribution comes from the qP wave, the field is largest in plane 1 where the qP wave is strongest. In plane 2, a weak qSV wave seems to be traveling along the dashed line. The third wave type, the SH wave, is much more difficult to identify.

In Fig. 9, the same situation as in Fig. 8 is considered except that the contact is inviscid, i.e., $\delta_x=\delta_y=0$ and $\delta_z=1$. This means that the directions of the main waves are the same here as in Fig. 8. Since Figs. 8 and 9 look very much alike, the normal part of the traction should be the dominat-

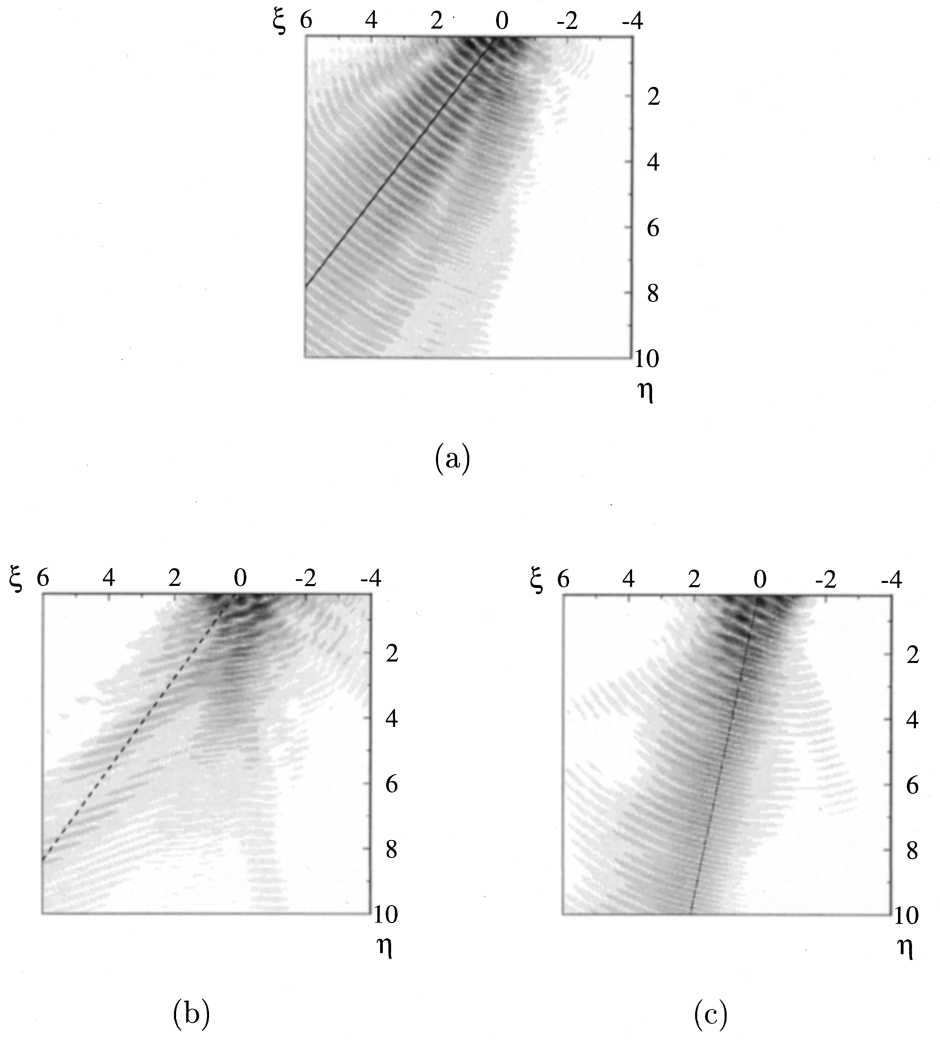


FIG. 9. $|\mathrm{Re}\,\tilde{\mathbf{u}}|$ for the same situation as in Fig. 8 except for $\delta_x=\delta_y=0$ and $\delta_z=1$. (a) Plane 1. (b) Plane 2. (c) Plane 3.

ing generator of the field in these case. Contrary to the case in Fig. 8, it is possible to identify the *SH* wave propagating along the dotted line in plane 3.

The same situation as in Fig. 8 is shown in Fig. 10 but in the time domain instead of the frequency domain. The field is shown at $\tau=3$ and $\tau=6$. The lower and upper frequencies are $\overline{\omega_1}=1$ and $\overline{\omega_2}=19$, respectively. The planes used and the directions of propagation are the same as in Fig. 8. Here, it is seen that the field is strongest in plane 1 and very weak in plane 2, as would be expected from Fig. 8. In order to identify the main waves, the absolute values and the directions of the group velocities are used. The directions of the main waves shown in the figures are also shown as straight lines in Fig. 3(b) (note that the directions are transformed to the crystal axes system). If the values of $|\mathbf{v}|$ where the straight lines cross the wave surfaces are multiplied with $\tau=6$, the result is the following:

Plane 1: *qP* at $(\xi,\eta)\approx(3.6,4.8)$;

qSV at $(\xi,\eta)\approx(2.6,3.4)$;

SH at $(\xi,\eta)\approx(1.8,2.4)$.

Plane 2: *qP* at $(\xi,\eta)\approx(6.1,8.4)$;

qSV at $(\xi,\eta)\approx(5.0,6.9)$;

SH at $(\xi,\eta)\approx(2.3,3.1)$.

Plane 3: *qP* at $(\xi,\eta)\approx(1.4,6.6)$;

qSV at $(\xi,\eta)\approx(1.2,5.6)$,

$(\xi,\eta)\approx(1.1,5.2)$,

$(\xi,\eta)\approx(1.1,4.9)$;

SH at $(\xi,\eta)\approx(0.7,3.2)$.

With this information at hand it is possible to identify the main waves from the figures. Note that, for example, only the main part of the *qP* wave will be located at the point calculated in plane 1. The other two waves will possibly prefer other directions of propagation but there will still be contributions at the calculated points which enables identification of the waves. Also note that due to the cusps in Fig. 3(b), three *qSV* waves are obtained along the dotted line. Moreover, contributions to the field apart from the main

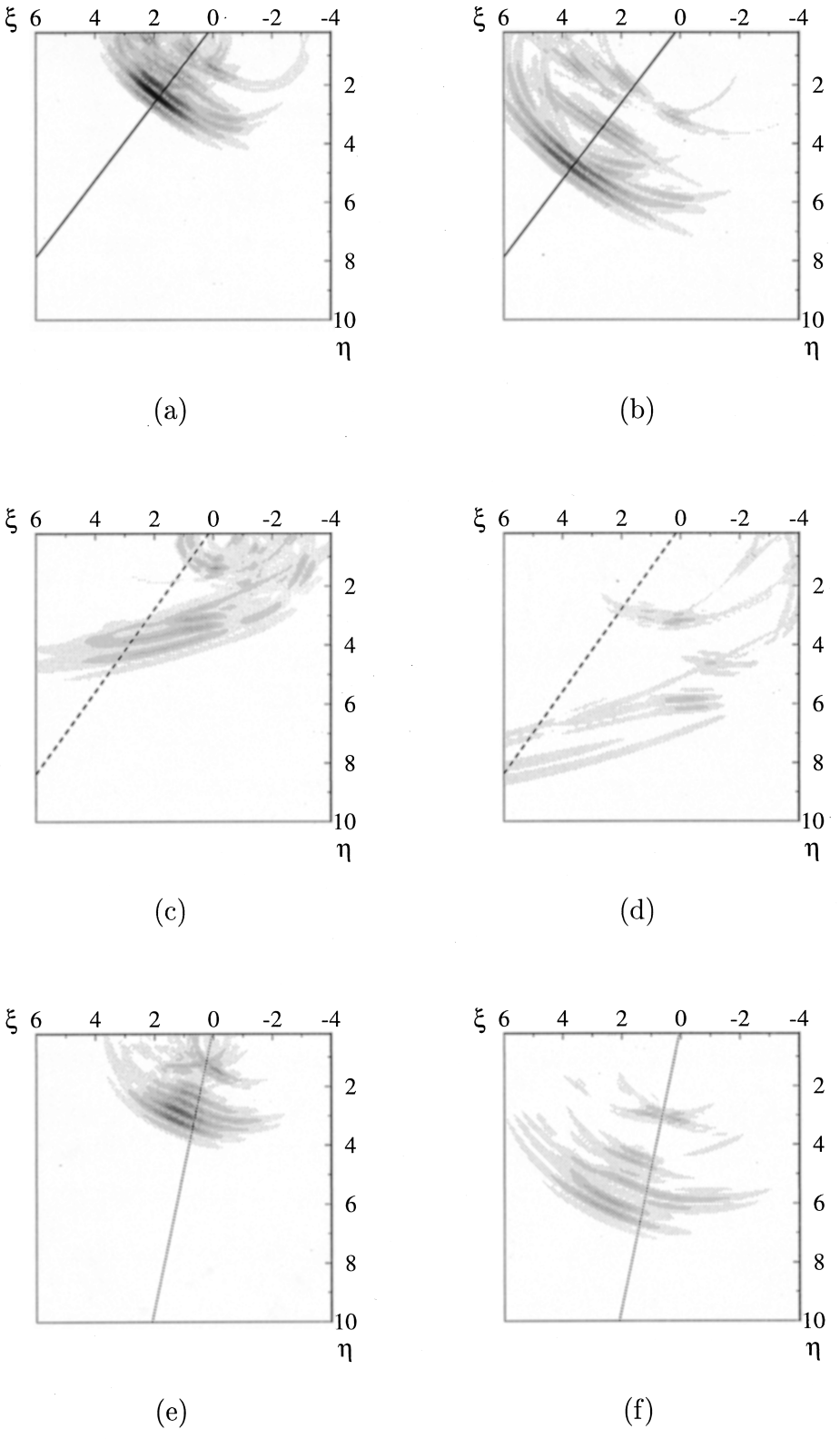


FIG. 10. Snapshots of $|\mathbf{u}|$ for the same situation as in Fig. 8. (a) $\tau=3$ in plane 1. (b) $\tau=6$ in plane 1. (c) $\tau=3$ in plane 2. (d) $\tau=6$ in plane 2. (e) $\tau=3$ in plane 3. (f) $\tau=6$ in plane 3.

waves are seen. From the figures, it is seen that these contributions come from the edges of the probe.[6,14] This is not surprising since the piston model implies a discontinuity in the traction here. This effect is somewhat reduced if the shape function is chosen as $g=g_2$ since this function has a tapering off at the edges.

V. CONCLUDING REMARKS

In this paper, a three-dimensional (3D) piezoelectric contact probe of *P*, *SV*, or *SH* type for anisotropic media is modeled. The model prescribes the traction beneath the probe and one advantage of the model is its simplicity. The

approach is especially suitable when Fourier methods are used. The model includes three parameters, δ_x, δ_y, and δ_z, and two functions, $g(x,y)$ and $\tilde{f}(\omega)$.

The probe is attached to an anisotropic half-space and the wave propagation problem is solved by means of Fourier transform techniques. The nature of the problem does not allow many analytical calculations, i.e., almost everything has to be calculated by means of numerical algorithms. Another property of wave propagation in anisotropic solids, which further complicates the situation, is the fact that the group velocities of the three wave types do not in general span one plane. This results in visualization problems and the fields are represented in 1–3 planes in this paper. It also results in interpretation problems since the different wave types are strongly mixed.

In the numerical examples, a transversely isotropic material and the piston model are used. There are examples of *P*, *SV*, and *SH*, probes, but mainly *P* probes are used. Radiated fields are shown in both the frequency and time domain. In the figures it is evident that the wave separation problem is large. This problem is somewhat reduced if the fields are examined in the time domain, but even here it can be hard to identify different waves. The mixing should make evaluation of ultrasonic test results much more difficult than in the isotropic case.

ACKNOWLEDGMENTS

This work has been financially supported by the Swedish Nuclear Power Inspectorate (SKI) and this is gratefully acknowledged. The author also wishes to thank his advisor Prof. Anders Boström for valuable discussions and comments on this paper.

[1] L. T. Wheeler, ''Some results in the linear dynamical theory of anisotropic elastic solids,'' J. Appl. Mech. **28**, 91–101 (1970).
[2] B. A. Auld, *Acoustic Fields and Waves in Solids* (Krieger, Malabar, FL, 1990), Vol. 1.
[3] J. H. M. T. Van der Hijden, *Propagation of Transient Elastic Waves in Stratified Anisotropic Media* (North-Holland, Amsterdam, 1987).
[4] W. Lord, R. Ludwig, and Z. You, ''Developments in ultrasonic modeling with finite element analysis,'' J. Nondestruct. Eval. **9**, 129–143 (1990).
[5] C.-Y. Wang and J. D. Achenbach, ''Three-dimensional time-harmonic elastodynamic Green's functions for anisotropic solids,'' Proc. R. Soc. London, Ser. A **449**, 441–458 (1995).
[6] A. Boström and H. Wirdelius, ''Ultrasonic probe modeling and nondestructive crack detection,'' J. Acoust. Soc. Am. **97**, 2836–2848 (1995).
[7] R. K. Chapman, ''A system model for the ultrasonic inspection of smooth planar cracks,'' J. Nondestruct. Eval. **9**, 197–211 (1990).
[8] J. Mattsson and A. J. Niklasson, ''Ultrasonic 2-D *SH* crack detection in anisotropic solids,'' J. Nondestruct. Eval. **16**, 31–41 (1997).
[9] R. K. N. D. Rajapakse and D. Gross, ''Transient response of an orthotropic elastic medium with a cavity,'' Wave Motion **21**, 231–252 (1995).
[10] P. Fellinger, R. Marklein, K. J. Langenberg, and S. Klaholz, ''Numerical modeling of elastic wave propagation and scattering with EFIT—Elastodynamic finite integration technique,'' Wave Motion **21**, 47–66 (1995).
[11] A. Minachi, J. Mould, and R. B. Thompson, ''Ultrasonic beam propagation through a bimetallic weld—A comparison of predictions of the Gauss–Hermite beam model and finite element method,'' J. Nondestruct. Eval. **12**, 151–158 (1993).
[12] B. Nouailhas, G. V. C. Nguyen, F. Pons, and S. Vermersch, ''Ultrasonic modeling and experiments: An industrial case: Bimetallic weld in nuclear power plant,'' J. Nondestruct. Eval. **9**, 145–153 (1990).
[13] R. J. Kažys and L. Mažeika, ''Transients in piezoelectric transducers of finite dimensions,'' J. Acoust. Soc. Am. **88**, 12–18 (1990).
[14] L. W. Schmerr and A. Sedov, ''An elastodynamic model for compressional and shear wave transducers,'' J. Acoust. Soc. Am. **86**, 1988–1999 (1989).
[15] A. J. Niklasson, ''Ultrasonic 2-D *SH* probe modeling in anisotropic materials,'' J. Acoust. Soc. Am. **100**, 2132–2138 (1996).
[16] A. K. Mal, C.-C Yin, and Y. Bar-Cohen, ''Analysis of acoustic pulses from fiber-reinforced composite laminates,'' J. Appl. Mech. **59**, S136–S144 (1992).

Modal excitation and control of membrane waves on cylindrical shells

Joseph E. Bondaryk
Department of Ocean Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 24 June 1997; revised 10 November 1997; accepted 19 January 1998)

A set of in-air experiments were made on two, thin, cylindrical shells to understand and control the propagation of membrane waves, which are primary determinants of target strength in the mid-frequency, high-aspect angle regimes. A circumferential source array was used to drive single modes in all wave types, compressional, shear, and flexural. Resulting waves were collected using an axial line array of laser Doppler velocimeter measurements and the wave types separated by array processing. Comparison between a clean plastic shell and a steel shell with a keel-like feature suggest on the order of 10-dB reduction of compressional wave response via coupling into higher order flexural waves by the nonaxisymmetric discontinuity. The conversion coupled with subsequent damping of the resulting flexural waves makes this a potentially useful mechanism for control of membrane waves. © *1998 Acoustical Society of America.* [S0001-4966(98)01705-6]

PACS numbers: 43.40.At, 43.40.Ey [CBB]

INTRODUCTION

Recent structural acoustics research on finite-length submerged shells has quantified the importance of quasi-shear and quasi-compressional membrane waves in radiation and scattering. For example, Fig. 1 shows the monostatic target strength in dB *re:* 1 m versus nondimensional frequency and aspect angle of a thin, finite cylindrical shell, collected for MIT by the Acoustic Scattering Group at the Naval Research Laboratory in 1991. It has been shown[1,2] that the strong features in the frequency range, $2<ka<12$, where k is the water wave number and a is shell radius, and in the aspect-angle range $60°>\phi>85°$ are due to reradiation of trace-matched membrane waves. On this empty shell, flexural waves are poorly coupled to the acoustic medium, due to the absence of periodic frames.

Only major changes to the shell can alter the membrane wave behavior. Their long wavelengths tend to ''ride over'' small and local treatments. For example, small periodic frames induce flexural Floquet waves,[3] which contribute to scattering at higher ka, but leave the membrane wave region relatively unaffected. An example of the effect of a major modification is shown in Fig. 2, the monostatic target strength in dB *re:* 1 m for the same shell with four, unequally spaced, axisymmetric stiffeners. These are major discontinuities, with the mass of the stiffeners equal to mass of shell. Note that the membrane wave region has been extended by at least 10°.

The ultimate objective of this research is to explore various modifications to the basic shell structure that might provide a designer with options in optimizing a hull for combined hydrostatic strength, shock resistance, and acoustic performance. Specifically, those modification schemes appropriate to the control of membrane waves in the above frequency and aspect-angle range are being studied experimentally. These include heavy bulkheads, neckdown sections, and keel-like structures.

Section I of this paper describes the experimental setup and methodology used to understand membrane wave propagation on shells. A novel source array configuration allows single mode excitation of shells. The resulting data are collected at an array of shell locations, which allows the separation of wave types via array processing. With this approach, single mode propagation as well as wave type and mode coupling can be studied in detail. The experiments described here were conducted in air for simplicity. Complexities such as water-mass loading and acoustic-resistive coupling can be appended to the results analytically via reasonably well known physics. Section II describes data derived from experimentation including observation of the source-selected modes propagating on a plastic shell. Section III describes several strategies for the control of membrane waves and some preliminary data on a steel shell with a keel-like feature. Data suggests significant reduction of the compressional response, on the order of 10 dB, via coupling of energy into higher order flexural modes.

I. EXPERIMENTAL SETUP AND PROCESSING

The overall experimental setup is comprised of three major systems: (1) the cylindrical shell itself; (2) the source array; and (3) the receiver array. The particular shell under test will be discussed in Secs. II and III, the latter two systems are described here.

The sources are 6.35×10^{-3} m (1/4 in.) o.d. piezoceramic rings made of lead zirconate titanate 5500 by Channel Industries. P holes were drilled into the shell at $360°/P$ intervals around the shell circumference and the rings bonded in with epoxy, as shown in Fig. 3 for the example of $P=12$. When a voltage is applied to the inner and outer surface of the ring, it expands radially and tends to excite compressional waves in the shell. A workstation is used to drive each source signal independently via 12-bit D/A cards. The elements were driven in a swept-sine mode, 256 frequencies over the range 0–25 kHz for ≈100-Hz resolution, coherently with amplitude shading over element number to

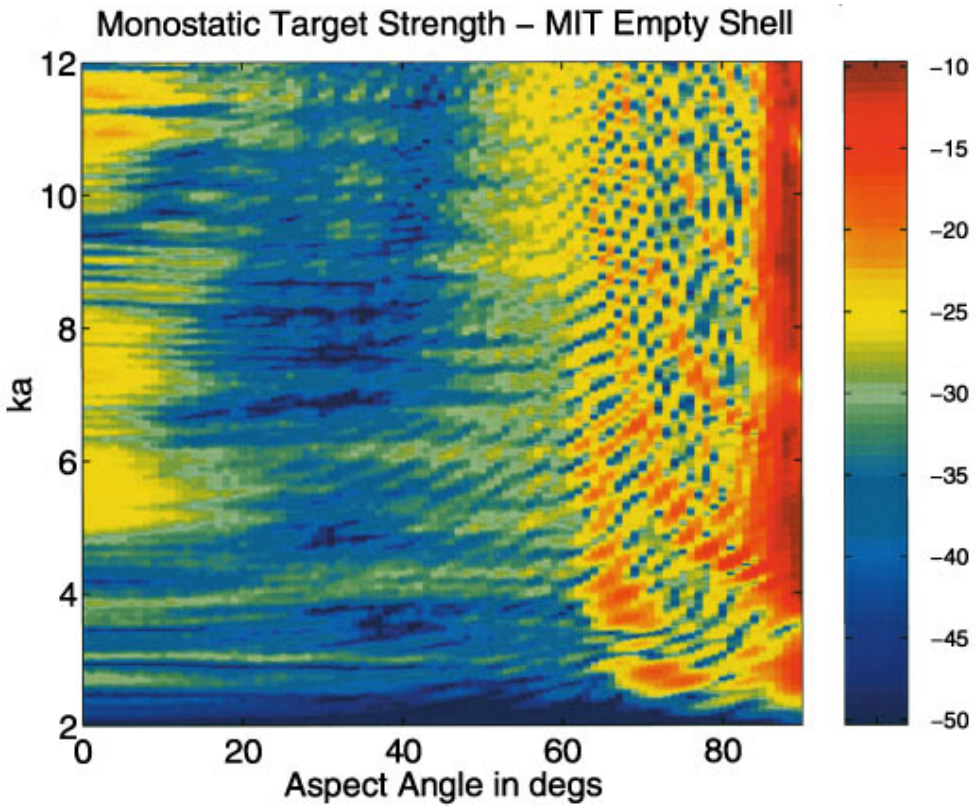


FIG. 1. Monostatic target strength for MIT empty shell in dB *re:* 1 m. Thickness to radius = 1%.

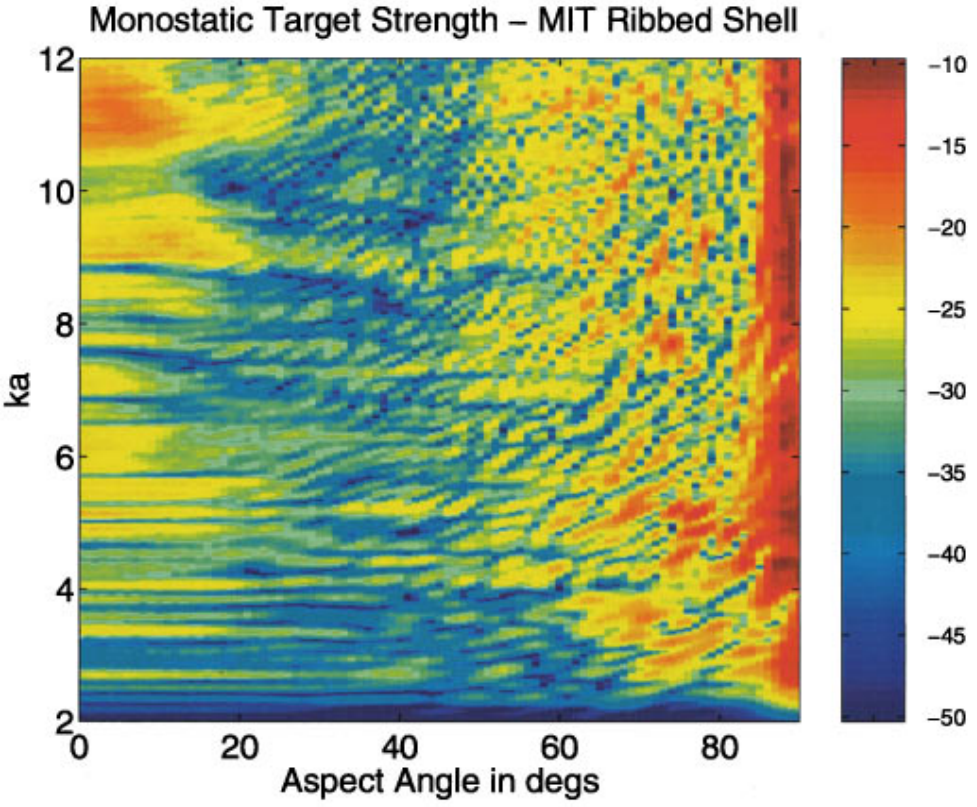


FIG. 2. Monostatic target strength for MIT ribbed shell in dB *re:* 1 m. Four unequally spaced ring stiffeners, total mass equal to mass of shell.

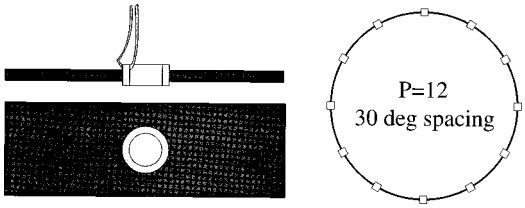


FIG. 3. Piezoceramic ring source embedded in shell and source array configuration, e.g., $P=12$ sources.

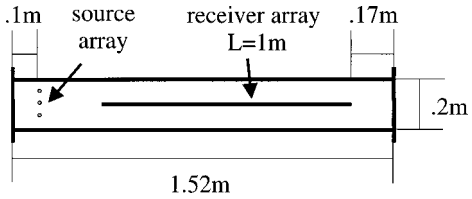


FIG. 4. PVC test shell with array locations.

achieve modal excitation. The pth amplitude shading is given by $a(p)=\cos(2\pi Mp/P)$, where P is the total number of source elements and M is the desired mode number. The voltages of these signals are then amplified by $10\times$ via step-up transformers, since the piezoceramic rings are capacitive in nature, and sent on to the ring sources.

With this driving system, it is possible to drive single modes in all wave types on the shell. By the Nyquist sampling theorem, P sources around the circumference can drive single unaliased modes up to $M=P/2$. As we shall see, this discrete mode drive must excite modes at higher harmonics. Specifically, for a P element source, exciting mode M, where $M<P/2$, the excited modes are modes M, $(l\times P)+M$ and $(l\times P)-M$ for $l=1,2,3,....$ A mode will only exist if the frequency is above the cutoff frequency for that mode. This is seldom the case for the higher order membrane waves in the frequency range of interest but is usually the case for the flexural waves, due to their dense frequency spacing. In theory, all wave types with the excited mode number should be generated by this circumferential source array. However, the sources elements themselves are compressional in nature and thus the array tends to excites compressional waves primarily.

Output velocity is measured at points on the shell by a single-axis Polytec PI laser Doppler velocimeter (LDV),

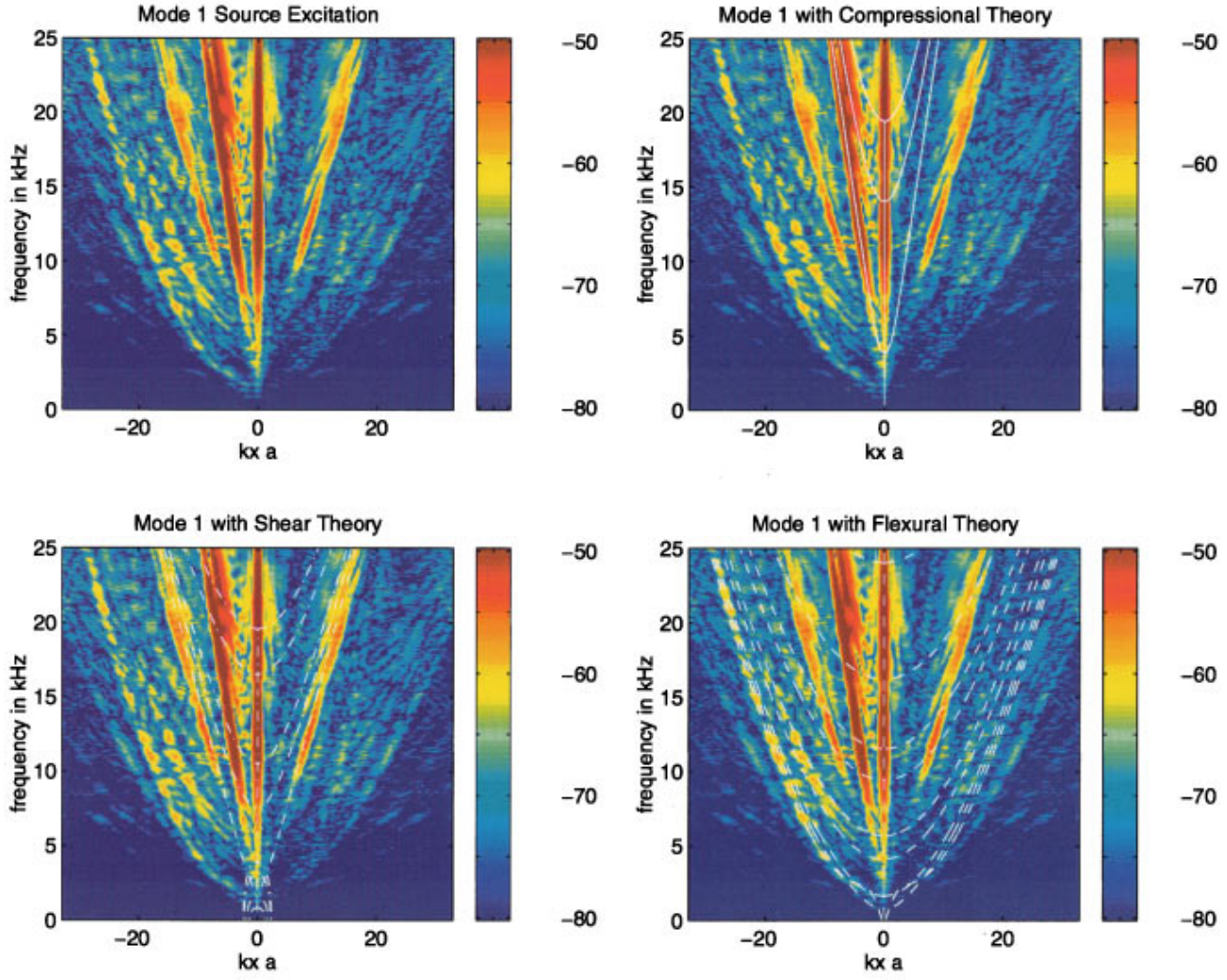


FIG. 5. Frequency wave number spectrum in dB *re:* 1 m/s^2/V_{in} for PVC shell. Mode $M=1$ source excitation with overlays of thick shell theory for dispersion relations of three wave types—compressional, shear, and flexural.

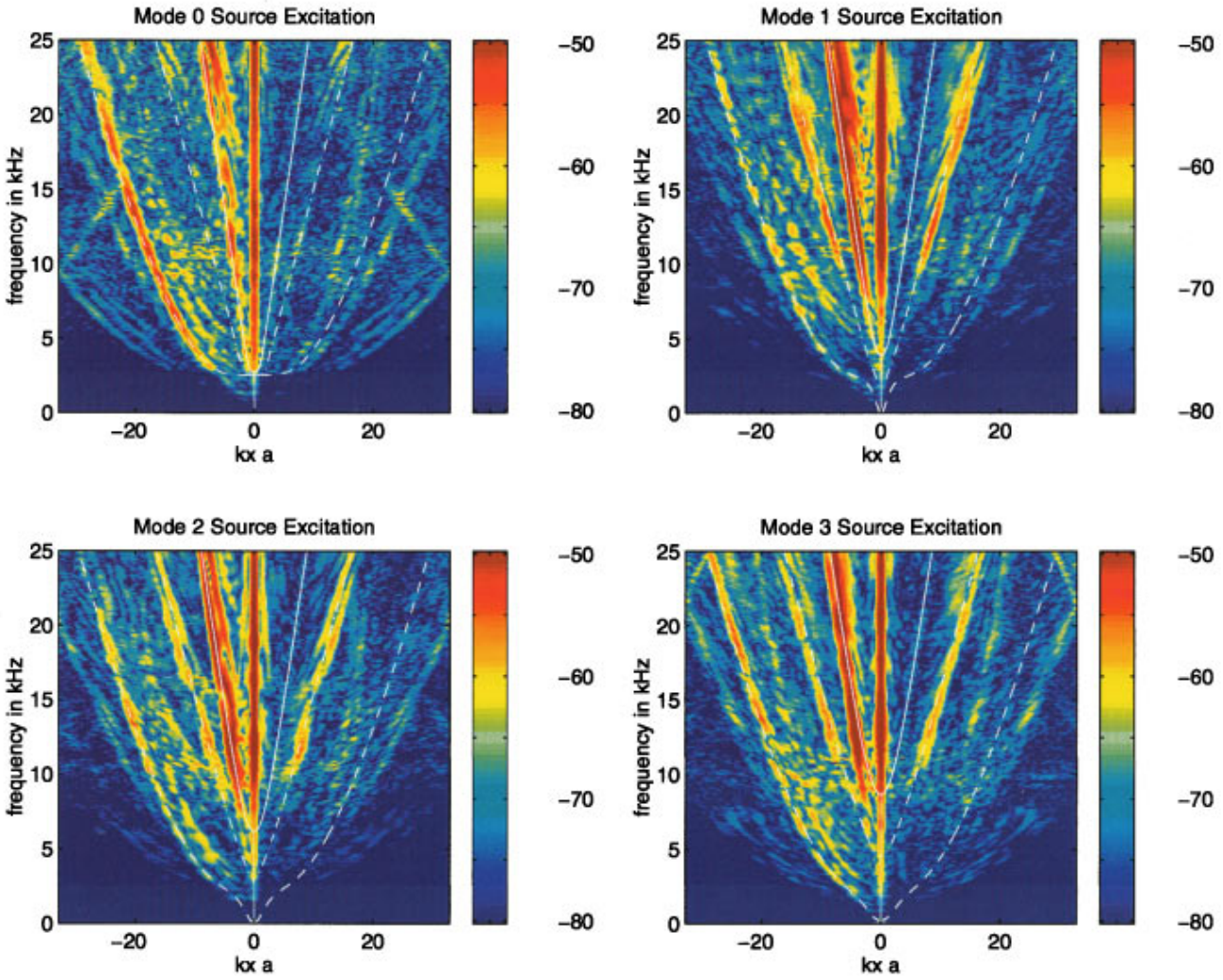


FIG. 6. Frequency wave number spectrum in dB *re:* $1\ \mathrm{m/s^2}/V_{\mathrm{in}}$ for PVC shell. Mode $M=0-4$ source excitation with overlays of thick shell theory for dispersion relations of three wave types—compressional (solid), shear (dot-dashed), and flexural (dashed).

OFV353 laser head and OFV3001 controller. The system is capable of measuring velocity along the beam over a velocity spectral range of $0.5\ \mu\mathrm{m/s}\sqrt{\mathrm{Hz}}-10\ \mathrm{m/s}\sqrt{\mathrm{Hz}}$ and a frequency range of 0.5 Hz–1.5 MHz. Radial velocity measurements are made by aiming the beam at 90° to the shell, axial velocity measurements are made by aiming the beam at 45° to the shell axis, and transverse velocity measurements are made by aiming the beam at 45° to the shell transverse axis. The latter two measurements contain both the desired and radial components. These components are not separated by subtracting laser data directly. Rather, subsequent array processing is used to separate wave types in frequency wave number space. Thus the measurements represent scaled wave amplitudes, rather than particle velocity in a particular direction.

The workstation is used to acquire the LDV output and the shell input signal via a 12-bit A/D board. Data are antialias filtered, sampled at 50 kHz and cross spectra between the LDV output and shell input signal are acquired over 100 ensembles.

A Parker linear positioning table under control of the workstation is used to move the LDV beam along a line down the shell axis by an increment of dx. This increment controls spatial aliasing, which occurs at an axial wave number of $k_x=\pi/dx$. The above measurements are made at each of the N array positions. For a given dx to control aliasing, N determines the array length. The spatial frequency resolution is inversely proportional to array length, L.

For each frequency, the N complex spectral values are collected into vector $\mathbf{x}(f)=[x_1(f)w_1,\ldots,x_N(f)w_1]$, where w_n is a standard Hanning taper function and $x_n(f)$ is the complex spectral value at frequency f from position z_n. The spectral matrix is formed via $S_x(f)=\mathbf{x}(f)\mathbf{x}^{\dagger}(f)$, where ''†'' denotes complex conjugate transpose. The array is then steered over axial wave number, k_x in the unaliased axial wave number range $-\pi/dx<k_x<\pi/dx$. The steering vector $\mathbf{e}(k_x)[e^{j2\pi k_x\cdot z_1},\ldots,e^{j2\pi k_x\cdot z_N}]$ is applied to the data for all k_x to form the output frequency wave number surface, $P(f,k_x)=\mathbf{e}^{\dagger}(k_x)S_x(f)\mathbf{e}(k_x)$. This results in a $\geqslant$30 dB SNR over the frequency range $100\ \mathrm{Hz}<f<25\ \mathrm{kHz}$. As will be shown in the next section, the three wave types overlap completely in frequency, but are well separated in axial wave number, k_x.

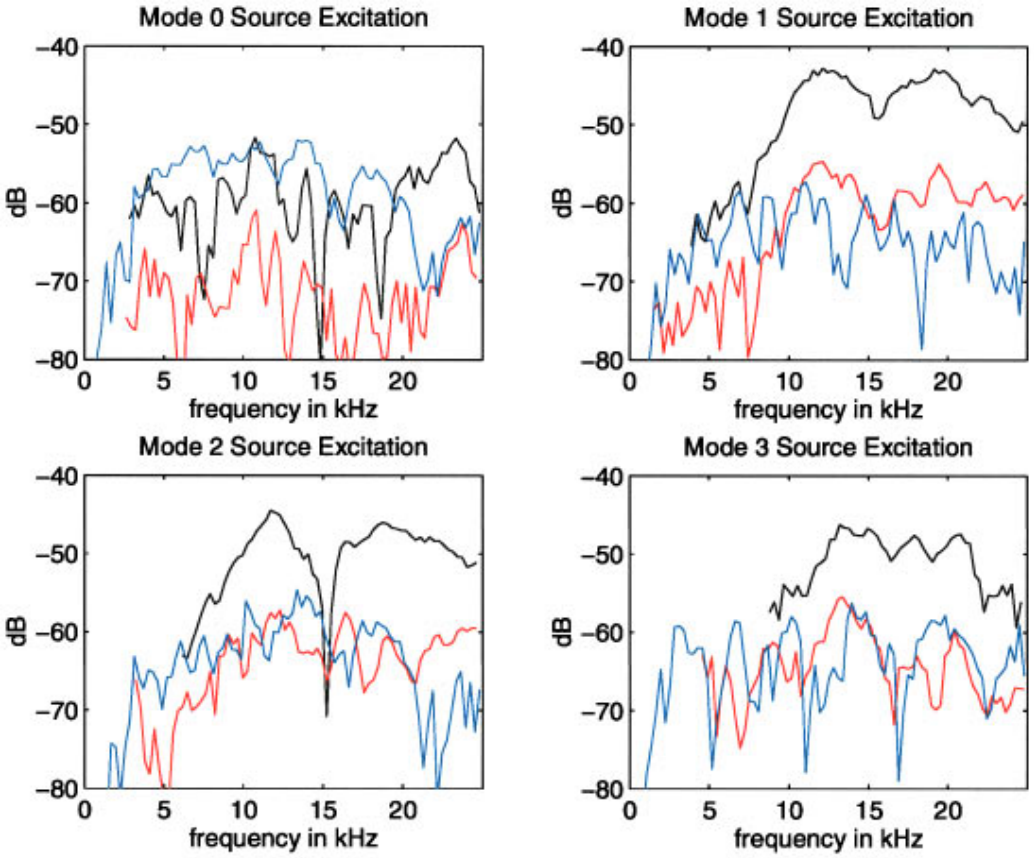


FIG. 7. Spectral amplitude of right-going waves on PVC shell by wave type in dB *re:* $1\ \mathrm{m/s^2}/V_{\mathrm{in}}$ for modes $M=0-3$. Black line is compressional, red is shear, and blue is flexural level. 300-Hz band average.

II. RESULTS—CLEAN SHELL

For this section, testing was conducted on a 1.525 m (5 ft) section of radius 0.1 m (4 in.), 4.76×10^{-3} m (3/16 in.) thick, type I, PVC pipe. The shell has flat-plate endcaps of 6.35×10^{-3} m (1/4 in.) thick, type I, PVC attached with epoxy. The measured compressional wave speed was 1660 m/s and Poisson's ratio was assumed to be $\nu=0.4$. The shell was suspended from a wooden frame by bungie straps, which provide vibrational isolation above approximately 10 Hz. The source and receiver array locations are shown in Fig. 4. The source array has $P=6$ elements at 60° spacing and the receiver array has $N=101$ elements at $dx=1$ cm spacing.

Figure 5 shows the resulting data for a mode $M=1$ source excitation, measured by an axial/radial measurement. The upper left plot shows the acceleration level on the shell surface versus frequency in kHz on the y axis and normalized axial wave number, k_xa on the x axis, where a is the shell radius. Negative values of k_xa indicate propagation from left to right w.r.t. Fig. 4, positive values of k_xa indicate propagation from right to left. Color levels are in dB *re:* $1\ \mathrm{m/s^2}/V_{\mathrm{in}}$, where V_{in} is the input voltage to the piezo sources. The other three plots of Fig. 5 show the same data overlayed with theoretical dispersion curves for the expected fundamental and higher order aliased modes separated by wave type. Hermann–Mirsky thick shell theory[4] with the above parameters was used to generate these curves. The compressional overlay shows modes 1, 5, and 7. Higher modes cutoff above this frequency range. The shear overlay shows these modes plus modes 11 and 13. The flexural overlay shows these modes plus modes 17, 19, 23, and 26. The strong line at $k_xa=0$ is due to electronic leakage of the extremely high voltages used to drive the sources into the acquisition system and should be ignored.

For the compressional wave, data matches theory for the entire mode 1. Mode 5 is clearly present above 20 kHz. Mode 7 is not seen. Note, that propagation for this wavetype is limited to the negative k_xa range. The plastic shell has high intrinsic damping. Auxiliary measurements over the 25-kHz range indicate $\eta=0.2$ for compressional waves. This is the direct wave from the source, the wave reflected from the far end of the shell has decayed too much to be observed in the data.

The shear wave mode 1 data matches theory above 10 kHz. There is a hint of modes 5 and 7, but not higher modes. This wave type seems to propagate both forward and backward on the shell. This seems unlikely due to the high damping of the shell. However, since the data must be correlated with the source to appear in the plot at all and since anomalies caused by array processing problems would apply equally to all wave types, this must be true phenomenon. The damping for shear waves in this plastic shell is under investigation. The flexural wave shows strong negative k_xa propagation in modes 1, 5, and 7.

Figure 6 shows the data for this shell for source excitation of modes $M=0-3$. The theoretical overlays are shown only for the fundamental excited mode for the three wave

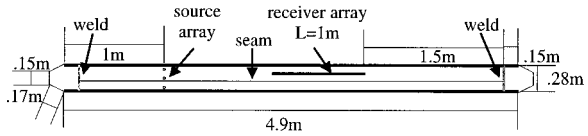


FIG. 8. Steel test shell with conical endcaps and keel-like feature with array locations.

types. Note that there is good agreement between the data and the theoretical dispersion curves. This plot shows that single mode excitation of the cylindrical shell has been achieved. As expected for mode $M=0$ excitation, there is a sharp reduction in energy below about 2.7 kHz, which is the ring frequency for this shell. The compressional wave is the strongest wave present in all modes, which is consistent with the compressional nature of the source elements. Note that the shear wave is not present in $M=0$ as expected, but is present for higher order M. This supports the above statement that the data shows a true shear wave and not a processing or measurement anomaly. Most easily seen in mode 0 is a wave which exists outside of the flexural ''U'' and aliases at about 9 kHz. Since the flex wave is the slowest structural wave which exists on this shell, this must be some other phenomenon. Analysis indicates that this very slow wave is the air-borne duct mode inside the closed pipe.

For a more quantitative view the data, consider Fig. 7. It shows the spectral acceleration level in dB *re:* 1 m/s^2/V_{in} for each of the wave types on the PVC shell. These curves were made from Fig. 6 by extracting the values of the data under the theoretical curve for the fundamental mode in the right-going direction only. The black curve represents the compressional level, red is shear and blue is flexural level. Notice that the compressional wave is most directly excited in modes $M=1$, 2, and 3 with 10 dB of separation from the other curves. This will be in contrast to the behavior of the shell with a keel in the next section.

III. MEMBRANE WAVE CONTROL—SHELL WITH KEEL

Membrane wave control could be achieved by taking advantage of one or several basic wave phenomenon: (1) damping; (2) dispersion; (3) path blocking; and (4) wave conversion. Damping is difficult to achieve against membrane waves, since their particle velocities are mostly unaffected by traditional constrained layer or bulk fill damping

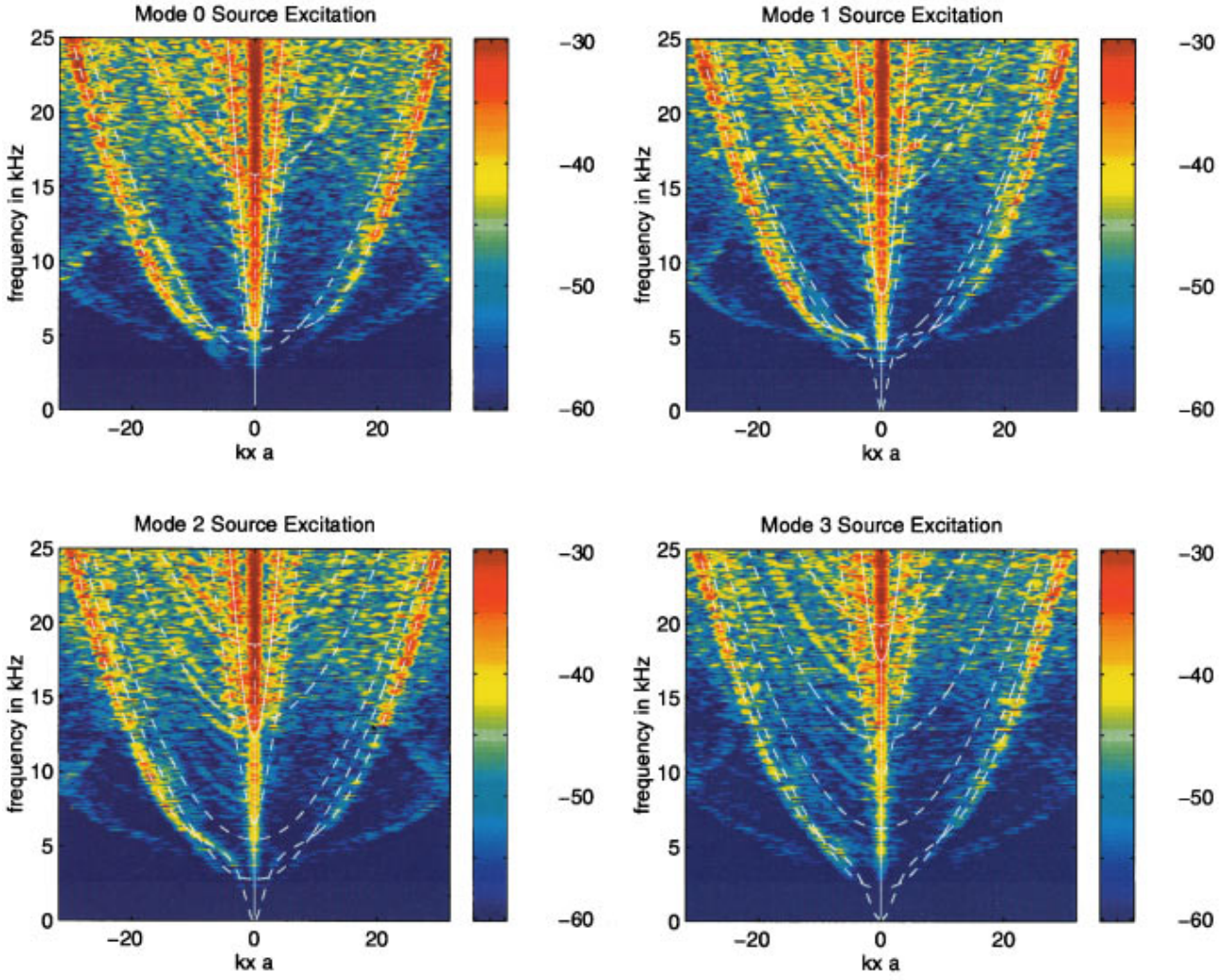


FIG. 9. Frequency wave number spectrum in dB *re:* 1 m/s^2/V_{in} for steel shell. Modes $M=0-4$ source excitation with overlays of thick shell theory for dispersion relations of three wave types—compressional (solid), shear (dot-dashed), and flexural (dashed).

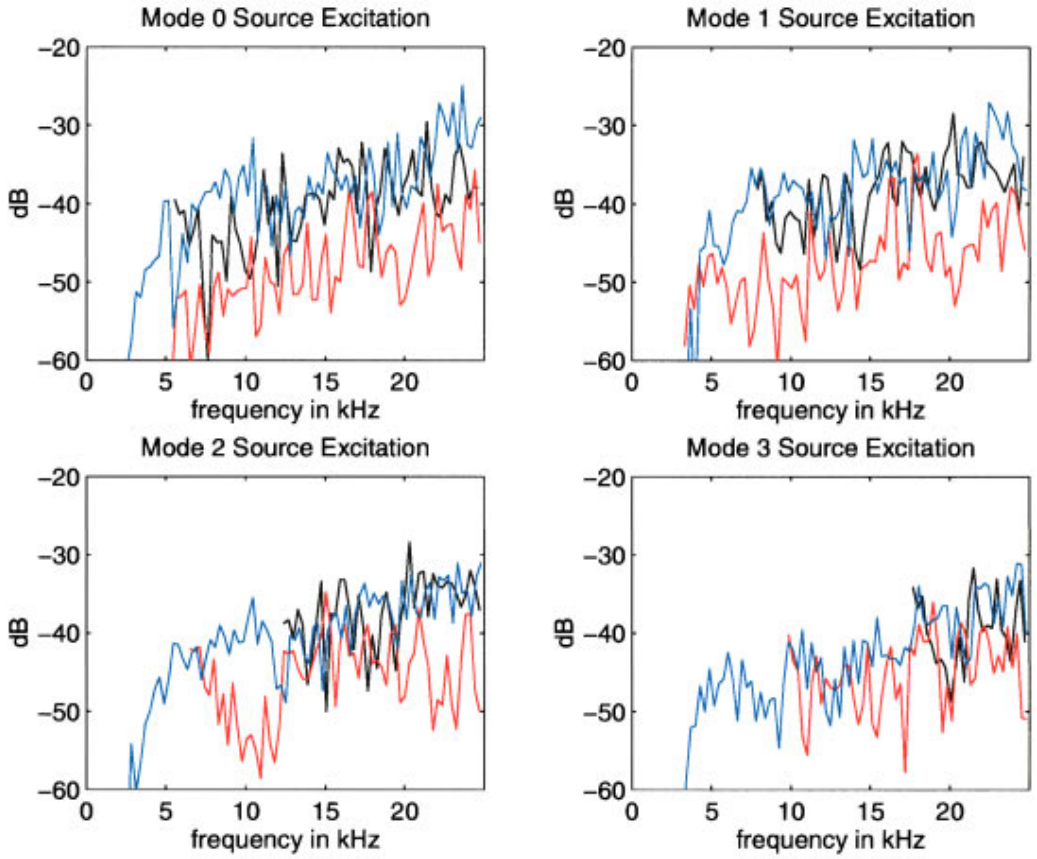


FIG. 10. Spectral amplitude of right-going waves on steel shell by wave type in dB *re:* 1 m/s^2/V_{in} for modes $M=0-3$. Black line is compressional, red is shear, and blue is flexural level. 300-Hz band average.

approaches. These waves are also nondispersive. This leaves the latter two approaches. An example of a major modification which should cause such behavior is a nonaxisymmetric axial discontinuity or ''keel'' in the shell.

The testing described in this section was conducted on the steel pipe depicted in Fig. 8. The shell radius is 0.152 m (6 in.) and is 2.54×10^{-3} m (0.1 in.) thick with a measured compressional wave speed of 5048 m/s and assumed Poisson's ratio of $\nu=0.35$. This cylindrical shell has a keel-like feature in the form of a cusped weld seam along its axis, about twice the thickness of the shell itself. The shell has open conical endcaps. The conical sections have a material and slope discontinuity where they meet the straight cylinder section, but are axisymmetric. The shell was suspended from a wooden frame by bungie straps which provide vibrational isolation above 10 Hz. The source array has $P=12$ elements at 30° spacing and the receiver array has $N=67$ elements at $dx=1.5$ cm.

Figure 9 shows the frequency wave number response of the steel shell for a source excitation of modes $M=0-3$. Since there are $P=12$ sources in this shell, the spacing between the expected aliased modes is larger. Only the fundamental compressional and shear waves are below their cutoff in this frequency range. The theoretical dispersion curves for the fundamental compressional and shear modes are overlayed on the plots as well as the those for *all* expected flexural modes in the frequency range. The thick shell theory is approximate here since the shell is no longer axisymmetric, but is reasonable as the keel's effect on wave speed is probably small. Again the line at $k_xa=0$ is due to electronic noise and should be ignored. Due to the low intrinsic damping of steel, acceleration levels per input volt on this shell are 20 dB higher than those found on the PVC shell. This also accounts for the fact that waves propagate in both directions as evidence by high levels seen in both the $+k_xa$ and $-k_xa$ sides of the plots. Excited waves reflect many times at both endcaps before decaying into the ambient noise level.

Coupling between wave types at the endcaps is certainly occurring as it did in the plastic shell. However, since all endcaps are axisymmetric, this coupling should be restricted to the same circumferential order, M. Indeed, energy in all wave types is 5–10 dB higher above the fundamental compressional mode cutoff frequency than below, seen most easily in the plots for excited modes $M=2$ and 3. The sources primarily excites the compressional mode, M. This mode could couple into the same mode, M, in shear and flexure at the endcaps. However, there are many higher order flexural waves excited, besides the expected directly excited and aliased modes. These are seen as the series of ''U''-shaped curves stacked in frequency, which are not overlayed by theoretical lines. These modes are coupled by the keel either directly from the fundamental compressional mode generated at the source or indirectly from the secondary shear and flexural modes generated at the endcaps. Either way, the energy equilibrium that exists with these higher order flexural

modes, must reduce the energy of the primary compressional wave.

Further evidence of coupling can be found in Fig. 10. This figure shows the spectral acceleration level in dB *re:* $1\ \mathrm{m/s^2}/V_{\mathrm{in}}$ for each of the wave types on the steel shell. These curves were made from Fig. 9 by extracting the values of the data under the theoretical curve for the fundamental mode in the right-going direction only. The black curve is for the compressional wave, red is for shear and blue is for flexural level. The compressional and flexural levels are fairly close across frequency and mode, there are no longer the 10-dB differences seen in the clean shell. This is consistent with the theory that compressional energy is being wicked away into the higher order flexural modes. This preliminary evidence suggests that the keel may be an effective coupling mechanism between these wave types. More direct comparisons, i.e., the same shell with and without a variety of keel configurations, are underway to investigate coupling mechanisms and their controlling parameters. If proven, this mechanism for the coupling of membrane energy into flexural waves is potentially beneficial, as this wave type can be controlled effectively by classical damping treatments.

IV. CONCLUSIONS

This paper has shown that the experimental procedure of synthetic LDV array measurements and subsequent frequency wave number processing are sufficient to measure and separate membrane and flexural wave types on cylindrical shells. It is possible to excite a select few modes on a cylindrical shell through the use of a properly shaded discrete circumferential array. The characteristics of the these experimentally measured modes agree well with theoretical thick-shell dispersion curves.

Experiments were conducted on a clean shell and a shell with a keel-like feature. It was suggested that there is on the order of 10-dB reduction of the excited compressional mode by the keel through coupling into higher order flexural modes. These initial results are encouraging, since flexural modes can be effectively damped by conventional methods, constrained layer for example. This could provide a useful mechanism for the control of membrane wave propagation on cylindrical shells. The coupling mechanisms and important parameters of this coupling, for example whether the mass or bending stiffness of the keel of is of primary importance, are topics of continuing research.

ACKNOWLEDGMENTS

The author acknowledges the Office of Naval Research for financial support of this work. Substantial thanks go to Ira Dyer for technical and personal support during this project. The particular numerical implementation of the thick-shell dispersion relations was made by Mark Hayner at MIT. The most excellent data of Figs. 1 and 2 were collected for MIT by the Acoustic Scattering Group at the Naval Research Laboratory under the direction of Brian Houston.

[1] M. L. Rumerman, ''Contribution of membrane wave reradiation to scattering from finite cylindrical steel shells in water,'' J. Acoust. Soc. Am. **93**, 55–65 (1993).
[2] C. N. Corrado, ''Mid-Frequency Acoustic Backscattering From Finite Cylindrical Shells and the Influence of Helical Membrane Waves,'' Ph.D. thesis, Massachusetts Institute of Technology, January 1993.
[3] B. H. Houston, J. A. Bucaro, and D. M. Photiadis, ''Scattering from flexural waves on a ribbed cylindrical shell,'' J. Acoust. Soc. Am. **98**, 2851–2853 (1995).
[4] G. Borgiotti and E. Rosen, ''The state vector approach to the wave and power flow analysis of the forced vibrations of a cylindrical shell. Part I. Infinite cylinders in vacuum,'' J. Acoust. Soc. Am. **91**, 911–925 (1992).

Sound radiation by a simply supported unbaffled plate

B. Laulagnet
Laboratoire Vibrations Acoustique, Institut National des Sciences Appliquées de Lyon, 20, avenue Albert Einstein, 69621 Villeurbanne Cedex, France

(Received 14 October 1996; revised 23 September 1997; accepted 4 February 1998)

Sound radiation by a simply supported unbaffled plate is presented using a double layer integral representation of the acoustic pressure. Both the pressure jump and the plate displacement are developed in series of the simply supported plate modes. This leads to modal coupling coefficients, which are numerically calculated for this specific unbaffled problem. The unbaffled problem is then transformed in the presentation of the baffled one which allows us to define original modal radiation impedance coefficients to the unbaffled plate. Baffled and unbaffled radiation impedances are then compared versus nondimensional frequency, in real and imaginary parts for a set of representative modes. An approximate expression for the direct unbaffled radiation impedance is presented, which is useful for light fluid loading. Finally, the influence of the baffle on sound radiation both in light and in heavy fluid, on the acoustic radiated power, the plate velocity, and the radiation factor is investigated. © *1998 Acoustical Society of America.* [S0001-4966(98)04105-8]

PACS numbers: 43.40.Rj [CBB]

INTRODUCTION

Plate sound radiation has been a problem of practical interest for many years. As a matter of fact, a lot of industrial structures can be modeled, to a first approximation, as a finite flat plate.

Usually, one of the simplest models available is the one of the baffled, simply supported plate. The simply supported plate is chosen because of the simplicity of the eigenmodes, which can be analytically determined. The rigid baffle allows us to transform the double layer integral representation into a simple one and it renders the vibroacoustical problem solvable. Nevertheless, in most of the practical cases, a rigid infinite baffle does not exist.

Early in the 1960s and 1970s, some authors studied the radiated sound from a baffled, simply supported, plate. First, Maidanik[1] investigated a single-mode radiation resistance of a finite, simply supported baffled plate, and Wallace[2] obtained an expression for the real part of the radiation coefficients versus frequency, for only the diagonal terms. Maidanik[3] gave an asymptotic formula for these coefficients which is still useful and widely used. Davies in Refs. 4 and 5 gave an approximate formula for the imaginary part of the radiation coefficient to allow for the calculation of the fluid added modal mass effect. This formula is widely used, when for example, a heavy fluid is considered and when one wants to calculate a first approximation of the resonance frequencies of a baffled plate in water. More recently, Graham[6] gave an approximate formula for the cross-modal radiation impedance terms in order to calculate sound radiation by a simply supported baffled plate excited by a turbulent boundary layer.

These radiation coefficients have been calculated for simply supported baffled cylindrical shells by Harari and Sandman,[7] Stepanishen,[8] and Laulagnet and Guyader[9] for heavy fluid radiation. In Ref. 8, the cross-modal radiation terms are derived. In Ref. 9, a comprehensive study of the radiation phenomenon in air and in water was investigated. Finally Sandman, in Ref. 10, investigated the influence of the baffle on the radiation coefficients for a cylindrical shell. His conclusion is that the baffle is of little influence on the sound radiation, and hence on the radiation coefficients. Although this conclusion is drawn from radiation impedances, rather than in impedance levels, the sound radiated by a cylindrical shell is not representative of the one for a flat plane.

Recently, Attala *et al.*[11] investigated the sound radiation by an unbaffled plate with general boundary conditions, using a light fluid assumption, i.e., neglecting the pressure jump when calculating the plate velocity. The pressure jump and the plate displacement are expanded in polynomial series. This allows the authors to describe analytically any boundary condition on the plate's edge, which renders the method attractive. However, the numerical implementation of the method suffers from instabilities, which leads to an overestimate of the unbaffled radiation efficiency of the plate with increasing frequency. However, in the low-frequency range, theory and experiment are presented which show a good agreement in radiation efficiency for an unbaffled free plate.

The objective of this paper is to address the sound radiation problem for an unbaffled plate with simply supported end conditions when submerged in light or heavy fluid without neglecting the pressure jump in the plate velocity, and to resolve the question of the influence of the baffle on the acoustic radiation, in a wide frequency range. The original set of cross-modal radiation impedance of the simply supported unbaffled plate are plotted versus a wide range of nondimensional frequencies.

In the first part of the paper, the equations of the vibroacoustic problem are presented and are solved using a modal decomposition for both the pressure jump and the plate displacement. This leads to the definition of the acoustical modal coupling coefficient of the unbaffled plate problem. In the second part, the equations are transformed in order to compare those coefficients with those of the baffled problem.

Then, the set of cross-modal radiation impedance coefficients are deduced for the unbaffled plate and compared to a baffled one. Numerical comparisons are presented which exhibit the differences in behavior between the two plates. Approximate formulas for the unbaffled coefficients are presented, which can be useful, in practical cases of light fluid loading.

In the last part of this paper, the influence of the rigid baffle is evaluated for a light fluid like air and in a heavy fluid like water. The plate is a 1-m^2 steel plate and represents a normalized case of the ''commission de validation des logiciels,'' a French organization supported by the French Acoustical Society and the French Mechanical Engineering Society[12] charged with validation of vibroacoustical software of simple basic cases. For this normalized case, the influence of the baffle is evaluated in terms of the sound radiated power, the plate velocity, and the radiation efficiency.

I. GOVERNING EQUATIONS

Consider a thin plate, excited by a harmonic driving force of angular frequency ω, located in the plane $z=0$. The plate is unbaffled, simply supported, with a surface S_p. The acoustic medium is infinite. The force is normal to the plate producing flexural waves governed by the classical equation:

$$D\nabla^4 w(x,y) - m\omega^2 w(x,y) = F(x,y) + \overline{p}(x,y), \tag{1}$$

where D is the rigidity modulus, m the mass per unit area, $w(x,y)$ the transverse displacement, $F(x,y)$ the driving force, and $\overline{p}(x,y)$ the difference of the acoustic surface pressure. The expression of the pressure jump due to the transverse motion of the plate is given by

$$\overline{p}(x,y) = p^-(x,y) - p^+(x,y), \tag{2}$$

where $p^-(x,y)$ and $p^+(x,y)$ are, respectively, the surface pressures on the side z negative and z positive of the plate.

The acoustic pressure is governed by the classical homogeneous Helmholtz equation:

$$\nabla^2 p(x,y,z) + k^2 p(x,y,z) = 0, \tag{3}$$

where $p(x,y,z)$ is the acoustic pressure anywhere in the acoustic fluid, k the acoustic wave number, ω/c, c the sound velocity, and ∇^2 the Laplacian operator.

The continuity between the normal acceleration of the plate and the fluid particle is given by Euler's equation:

$$\left.\frac{\partial p(x,y,z)}{\partial z}\right|_{z=0} = \rho_0 \omega^2 w(x,y), \tag{4}$$

where ρ_0 is the fluid density.

II. INTEGRAL REPRESENTATION FOR THE ACOUSTIC PRESSURE

Applying the Green's formula to a volume entirely closing the plate, one can write for the pressure anywhere in the volume: (cf. Fig. 1)

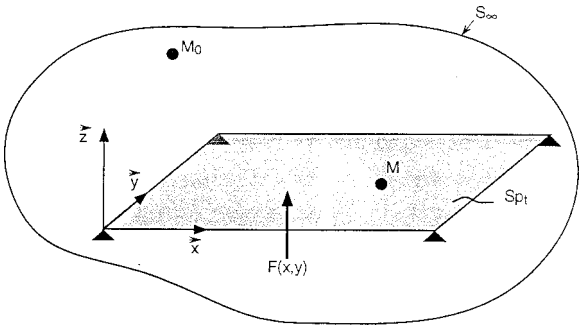


FIG. 1. Schematic view of the supported plate in the infinite medium.

$$p(M_0) = \int_{S_T} \left(\frac{\partial G}{\partial n_M}(M,M_0) p(M) - \frac{\partial p}{\partial n_M}(M) G(M,M_0) \right) dS_T, \tag{5}$$

where M_0 belongs to V, $\mathbf{n}_M$ is the normal to the total surface taken at M composed by S_{p_t} and S_∞ pointing outward, and $G(M,M_0)$ is the Green's function satisfying:

$$\nabla^2 G(M,M_0) + k^2 G(M,M_0) = \delta(x-x_0)\delta(y-y_0) \times \delta(z-z_0), \tag{6}$$

where δ is the Dirac delta distribution. The expression for the Green's function can be found using the bidimensional spatial Fourier transform of Eq. (6):

$$G(x,x_0,y,y_0,z,z_0) = \frac{j}{8\pi^2} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{e^{jk_x(x-x_0)} e^{jk_y(y-y_0)} e^{jk_z|z-z_0|}}{k_z} \times dk_x\, dk_y, \tag{7}$$

where k_z is

$$k_z = \sqrt{k^2 - k_x^2 - k_y^2}, \tag{8}$$

where k_x, k_y are the real Fourier's variables in the bidimensional wave-number domain. To ensure that G, when $|z-z_0|$ tends to infinity, is finite, the definition of k_z is taken as

$$\begin{aligned} k_z &= \sqrt{k^2 - k_x^2 - k_y^2} \quad \text{if } k^2 \geqslant k_x^2 + k_y^2, \\ k_z &= j\sqrt{k_x^2 + k_y^2 - k^2} \quad \text{if } k^2 \leqslant k_x^2 + k_y^2. \end{aligned} \tag{9}$$

In the following, the expression of the pressure in Eq. (5) will be simplified if we consider that there is no reflected wave from the surface S_∞. Since the acoustical sources are located on the plate, the surface integral will be reduced to S_{p_t}, where S_{p_t} is the whole surface S_{p+} plus S_{p-} (cf. Fig. 2).

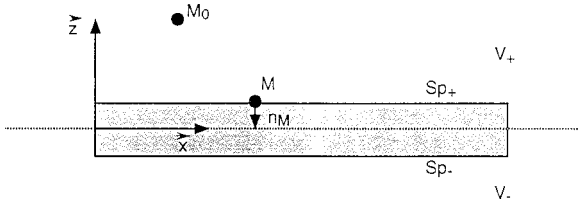


FIG. 2. Detail of the contour of integration on the plate.

In order to calculate the pressure at M_0, we need to know the relative location of M_0 vs M which belongs to the plate surface. Two cases appear which determine the sign of the absolute value in the G function (7):

$$\begin{aligned} &e^{-jk_z(z-z_0)} \quad \text{if } z_0 \geqslant z=0, \\ &e^{jk_z(z-z_0)} \quad \text{if } z_0 \leqslant z=0. \end{aligned} \tag{10}$$

In the following it is sufficient to estimate the pressure at a point M_0 always located above the plate. Consequently we will calculate the Green's function with the first inequality defined in Eq. (10). (The problem is an antisymmetric problem and to know the pressure in the half-space $z_0 \leqslant 0$ one has only to change the sign of the pressure calculated in the half-space $z_0 \geqslant 0$.)

Identifying the normal $\mathbf{n}_M$ to $\mathbf{z}$ or $-\mathbf{z}$ depending on which side of the plate is being considered, Eq. (5) can be decomposed as

$$\begin{aligned} p(M_0) = &\int_{S_{p^+}} -\frac{\partial G}{\partial z_M}(M,M_0)p^+(M)dS_{p^+} \\ &+ \int_{S_{p^-}} \frac{\partial G}{\partial z_M}(M,M_0)p^-(M)dS_{p^-} \\ &+ \int_{S_{p^+}} \frac{\partial p^+}{\partial z}(M)G(M,M_0)dS_{p^+} \\ &- \int_{S_{p^-}} \frac{\partial p^-}{\partial z}(M)G(M,M_0)dS_{p^-}, \\ &M_0 \in V^+, \end{aligned} \tag{11}$$

where V^+ is the half-space located ''above'' the plate.

Since both $G(M,M_0)$ and the gradiant of the acoustic pressure are continuous functions on the plate midsurface S_p, the simple layer contribution vanishes when S_{p+} and S_{p-} tend to S_p:

$$\int_{S_p} \left(\frac{\partial p^+(M)}{\partial z} - \frac{\partial p^-(M)}{\partial z} \right) G(M,M_0)dS_p = 0, \quad M \in S_p. \tag{12}$$

Substituting (1) into (11), one can write the following double layer integral of the pressure:

$$\begin{aligned} p(M_0) = &\int_{S_p} [D\nabla^4 w(M) - m\omega^2 w(M) - F(M)] \\ &\times \frac{\partial G}{\partial z_M}(M,M_0)dS_p, \quad M \in S_p, \quad M_0 \in V^+. \end{aligned} \tag{13}$$

The integral representation (13) depends on two unknown fields, the acoustic field $p(M_0)$ and the vibratory field of the plate $w(M)$. These fields can be connected using Euler's equation in the plane of the plate at $z=0$:

$$\begin{aligned} \left. \frac{\partial p}{\partial z}(M_0) \right|_{z=0} &= \rho_0 \omega^2 w(M_0) \\ &= \int_{S_p} [D\nabla^4 w(M) - m\omega^2 w(M) - F(M)] \\ &\quad \times \frac{\partial^2 G_{(M,M_0)}}{\partial z_M \, \partial z_{M_0}} dS_p, \\ &\quad M_0 \in S_p, \quad M \in S_p, \end{aligned} \tag{14}$$

where $\partial/\partial z_M$ and $\partial/\partial z_{M_0}$ denote, respectively, the partial derivatives versus $\mathbf{z}$ at M and M_0. Equation (14) allows us to find a solution for the displacement field, by eliminating the pressure.

III. SOLUTION FOR THE INTEGRAL REPRESENTATION USING THE MODAL BASIS OF THE SIMPLY SUPPORTED PLATE

Let us expand the shell displacement in a series of the plate modes:

$$w(x,y) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{nm} \phi_{nm}(x,y), \tag{15}$$

where a_{nm} is the displacement modal amplitude and $\phi_{nm}(x,y)$ is the modal shape given by

$$\phi_{nm}(x,y) = \sin \frac{n\pi x}{a} \sin \frac{m\pi y}{b}. \tag{16}$$

The same kind of expansion can be written for the driving mechanical force:

$$F(x,y) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} F_{nm} \phi_{nm}(x,y), \tag{17}$$

where F_{nm} is the generalized force density.

Substituting the series (15) and (16) into (14) and using the properties of the function ϕ_{nm} versus the bi-Laplacian operator (cf. Appendix A), one can transform (14) as follows:

$$\begin{aligned} &\rho_0 \omega^2 \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{nm} \phi_{nm}(x_0,y_0) \\ &= \int_{S_p} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} [m(\omega_{nm}^2 - \omega^2) a_{nm} - F_{nm}] \phi_{nm}(x,y) \\ &\quad \cdot \frac{\partial^2 G}{\partial z \, \partial z_0}(x,x_0,y,y_0,z=z_0=0)dx \, dy, \end{aligned} \tag{18}$$

where ω_{nm} is the resonance frequency of the mode n,m. The density of the generalized force is fairly simple to calculate, in the case of a driving point force applied at M_e of coordinates (x_e,y_e):

$$\begin{aligned} F_{nm} &= \frac{4}{S_p} \int F_0 \delta(x-x_e)\delta(y-y_e)\phi_{nm} dx \, dy \\ &= \frac{4F_0}{S_p} \phi_{nm}(x_e,y_e), \end{aligned} \tag{19}$$

where F_0 is the amplitude of the point force.

To find the set of unknowns a_{nm}, it is necessary to project Eq. (18) on a particular modal shape ϕ_{pq}. After the use of the orthogonality relationships one obtains:

$$\rho_0\omega^2 \frac{S_p}{4} a_{pq} = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} (m(\omega_{nm}^2-\omega^2)a_{nm} - F_{nm})\gamma_{pq\,nm}, \tag{20}$$

where $\gamma_{pq\,nm}$ are coupling terms defined as

$$\gamma_{pq\,nm} = \int_{S_p}\int_{S_p} \phi_{pq}(x_0,y_0) \frac{\partial^2 G}{\partial z\ \partial z_0}(x,x_0,y,y_0,z=z_0=0) \times \phi_{nm}(x,y)dx\ dx_0\ dy\ dy_0. \tag{21}$$

Introducing the G function in (21) leads to the following formula for the $\gamma_{pq\,nm}$ terms once the two integrals over S_p are done:

$$\gamma_{pq\,nm} = \frac{j}{8\pi^2} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} k_z \tilde{\phi}_{pq}^*(k_x,k_y)\tilde{\phi}_{nm}(k_x,k_y)dk_x\ dk_y, \tag{22}$$

where $\tilde{\phi}_{nm}(k_x,y_y)$ denotes the Fourier transform of the modal shape $\phi_{nm}(x,y)$ and $\tilde{\phi}_{pq}^*$ is the conjugate transform defined by

$$\tilde{\phi}_{nm}(k_x,k_y) = \int_0^a\int_0^b \sin\frac{n\pi x}{a}\sin\frac{m\pi y}{b} e^{-jk_x x}e^{-jk_y y}dx\ dy \tag{23}$$

whose calculation can be made analytically (cf. Appendix B). One has to note that when we deal with the analog problem of the sound radiation by a baffled plate, the modal radiation impedance terms (cf. Refs. 1–6) are derived by the same integral formula as (22) but where the k_z term appears in the denominator. This is due to the fact that the integral representation is a ''simple layer one.'' Consequently the Green's function is not differentiated at all. Here the Green's function is differentiated twice.

The $\gamma_{pq\,nm}$ terms are the acoustical cross-modal coupling terms. Similar to the cross-modal radiation impedances terms of a baffled plate, these are symmetric and they have the same properties versus the mode's parity. They vanish when the modes are of different parities. They are only different from zero in the following parity cases: odd odd odd odd, even even even even, even odd even odd, and odd even odd even.

Consequently, similar to the baffled plates, the acoustic radiation leaves the modes of different parities independent, and renders those of same parity dependent.

The calculation of the expression (22) can be completed and one finds for modes of same parity:

$$\gamma_{nm\,pq} = \gamma_{pq\,nm} = \frac{2jS_p^2}{pq\ nm\ \pi^6} \int_0^\infty\int_0^\infty \frac{\sqrt{k^2-k_x^2-k_y^2}(1-(-1)^p\cos k_x a)(1-(-1)^q\cos k_y b)}{(a^2k_x^2/n^2\pi^2-1)(a^2k_x^2/p^2\pi^2-1)(b^2k_y^2/q^2\pi^2-1)(b^2k_y^2/m^2\pi^2-1)} \cdot dk_y\ dk_x. \tag{24}$$

To obtain expression (24) one notes that the integrand is even both in k_x and k_y, reducing the domain of integration only to the first quadrant of the k_x,k_y plane.

We present in Appendix B the numerical method for calculating these cross-coupling terms, namely the numerical criteria which allows us to overcome these rather difficult oscillating infinite integrals in the k_x,k_y domain.

Once the $\gamma_{pq\,nm}$ terms are numerically evaluated, one can find a solution for the modal amplitudes, solving the following linear system:

$$[\gamma_{pq\,nm}]\{m(\omega_{nm}^2-\omega^2)a_{nm}\} - \rho_0\omega^2\frac{S_p}{4}[I]\{a_{nm}\} = [\gamma_{pq\,nm}]\{F_{nm}\}, \tag{25}$$

where $[I]$ is the identity matrix of size N by N, and $[\gamma_{pq\,nm}]$ is the matrix of the cross-coupling terms.

The size N of this linear system depends, as in modal methods, on the truncation order of the series, which must be high enough to ensure the convergence of the pressure and the vibratory fields of the plate. In a light fluid, it is well known that this size N can be determined by the criteria of the resonant mode at the current frequency of excitation ω. That is to say, all modes whose eigenfrequencies are smaller or equal to ω must be taken into account in the calculation. This gives the size N of modes to be included. In the FORTRAN algorithm, it is convenient to increase the current frequency to the range $\omega+\Delta\omega$, where $\Delta\omega$ is an adjustable frequency band which allows us to improve the accuracy of the criteria, taking into account the small contribution of stiffness modes.

When the fluid is heavy, like water, the resonant mode criteria is no more sufficient, since it is based on the plate in vacuum resonance frequency which does not take into account the added modal masses.

Nevertheless, in the algorithm, it is easy to increase the adjustable frequency band $\Delta\omega$, to take into account heavy fluid loading. Thus the size N will be larger in the case of the water loading than air loading for a given plate. Numerical examples of water loading and air loading will be presented later.

Using the parity case property, one can solve independently four linear systems of smaller size, one for each type of parity. These linear system are solved using a classical

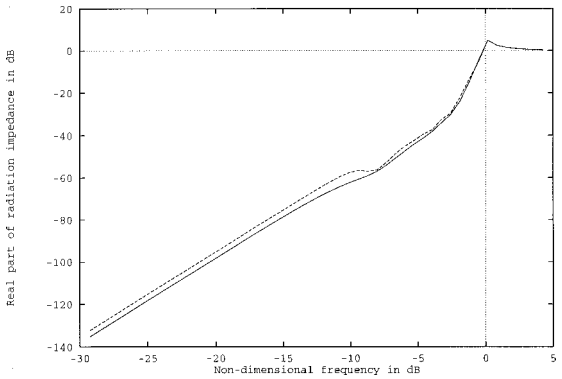


FIG. 3. Comparison between the exact and approximated calculation in the real part of the unbaffled radiation impedance; ——— approximated, -------- exact.

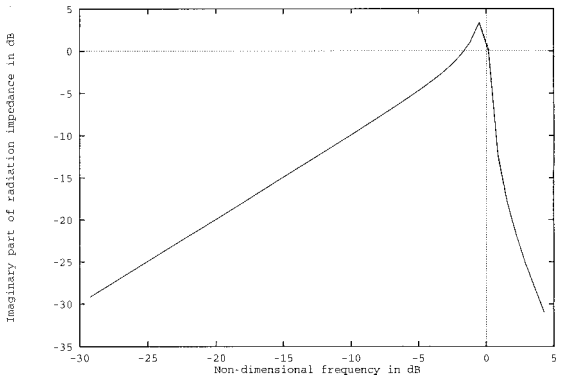


FIG. 4. Comparison between the exact and approximated calculation on the imaginary part of the unbaffled radiation impedance; ——— approximated, -------- exact.

Gauss algorithm program. Numerical examples will be presented in the following.

IV. DEFINITION OF THE CROSS-MODAL RADIATION IMPEDANCE TERMS FOR AN UNBAFFLED SIMPLY SUPPORTED PLATE AND COMPARISON WITH THE BAFFLED ONES

The goal of this section is to define a new set of radiation impedances for the unbaffled, simply supported plate and to compare it with the well-known baffled plate. Although it is not necessary to define these unbaffled radiation impedances to solve the linear system (25), it is interesting to calculate them in order to exhibit the main difference that occurs versus frequency.

Inverting the linear system (25), one can write:

$$[I]m(\omega_{nm}^2-\omega^2)\{a_{nm}\}-\rho_0\omega^2(S_p/4)[\gamma_{pq\,nm}]^{-1}\{a_{nm}\} =\{F_{nm}\}. \tag{26}$$

Multiplying by the constant $S_p/4$, it is possible to identify the equivalent baffled modal equations, cf. Refs. 1–6:

$$-\rho_0\omega^2\left(\frac{S_p}{4}\right)^2[\gamma_{pq\,nm}]^{-1}\{a_{nm}\}=[Z^0_{pq\,nm}]\{j\omega a_{nm}\}. \tag{27}$$

The linear system (27) allows us to define the cross-modal radiation impedance terms for the unbaffled plate from the $\gamma_{pq\,nm}$ coefficients as

$$[Z^0_{pq\,nm}]=j\omega\rho_0\left(\frac{S_p}{4}\right)^2[\gamma_{pq\,nm}]^{-1}. \tag{28}$$

Finally it is now possible to define the matrix normalized to $\rho_0 c\ S_p/4$ and to define the normalized cross-modal radiation impedance terms for a simply supported unbaffled plate:

$$[z^0_{pq\,nm}]=jk\,\frac{S_p}{4}\,[\gamma_{pq\,nm}]^{-1}. \tag{29}$$

As in the baffled case, the $z^0_{pq\,nm}$ terms are complex. The real part is associated with acoustic radiation and the imaginary part is the added mass to the plate, which is responsible for the decrease of the resonance frequencies of the plate. The coupling between modes is introduced by the off-diagonal terms of (29), but is limited to modes of the same parity.

Contrary to the baffled case where the radiation impedance formula appears directly in the equations to be solved, the $z^0_{pq\,nm}$ terms do not appear directly. It is necessary, first to calculate the $\gamma_{pq\,nm}$ terms and second, to invert this matrix.

Consequently, the quality of the $z^0_{pq\,nm}$ terms presented below depends on the accuracy of the calculation of the $\gamma_{pq\,nm}$ terms themselves but also on the size of the matrix of $\gamma_{pq\,nm}$ to be inverted. Tests of convergence have been done which show that the diagonal terms $z^0_{pq\,pq}$ are not sensitive to the size of the matrix $[\gamma_{pq\,nm}]$ to be inverted. The off-diagonal terms are sensitive and their calculation must be performed with particular care.

These last remarks allow us to think, since the diagonal terms can be calculated to a first approximation without the long computation needed for the off-diagonal terms $\gamma_{pq\,nm}$, they can be deduced from the simple approximate formula:

$$z^0_{pq\,pq}\approx jk\,\frac{S_p}{4}\,\frac{1}{\gamma_{pq\,pq}}. \tag{30}$$

This formula renders the calculation of the diagonal terms very easy, as easy as for the baffled case, which do not necessitate a matrix inversion. The formula (30) can be used to approximate the radiation of an unbaffled plate, neglecting, the cross-modal coupling writing independently for each mode (p,q):

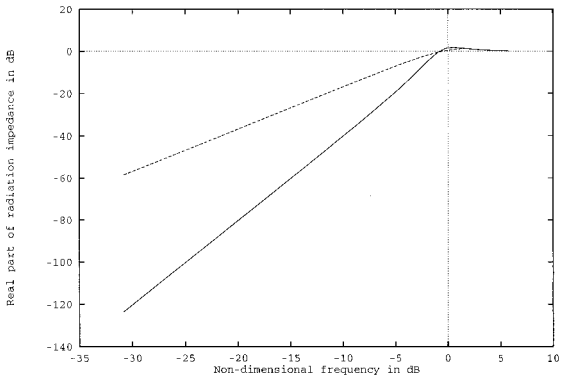


FIG. 5. Radiation impedance in the real part of the mode (1,1); ——— unbaffled, -------- baffled.

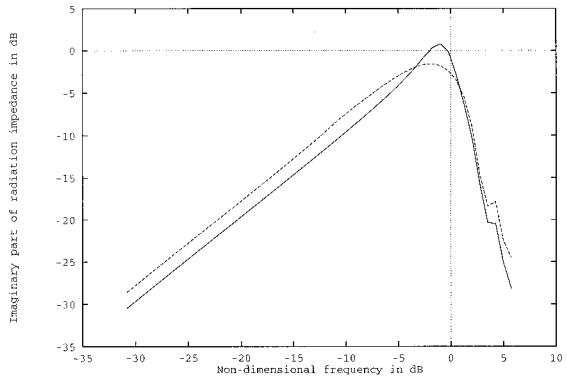


FIG. 6. Radiation impedance in the imaginary part of the mode (1,1); ——— unbaffled, -------- baffled.

$$m_{pq}(\omega_{pq}^2-\omega^2)a_{pq}-\rho_0\omega^2\left(\frac{S_p}{4}\right)^2\frac{1}{\gamma_{pq\,pq}}a_{pq}=\mathscr{F}_{pq}, \tag{31}$$

where m_{pq} and $\mathscr{F}_{pq}$ denote, respectively, the generalized mass and the generalized force of the mode (p,q) (cf. Appendix A).

It will be shown in the next section that the solution (31) when compared to the exact solution (25) is a very good approximation in air and a not so bad approximation in water.

Moreover, what is interesting in the approximate solution (31) is it allows us to substitute in a program the baffled impedance routine library and the unbaffled one. This program is a powerful tool, in industrial applications, for one of the more commonly encountered cases; in which the plate cannot be considered as baffled.

Finally to render the problems of the baffled and unbaffled plates equivalent, we have either to multiply by two the baffled radiation impedances, since they take into account only the fluid loading on one side of the plate, or to divide by two the $z^0_{pq\,nm}$ terms since the unbaffled plate is loaded on its two sides. The second option has been chosen.

Figures 3 and 4 deal with the comparison between an exact calculation using (29) and the approximate one using (30), respectively, for the real and imaginary part of the radiation impedance of the mode (7,7) for an unbaffled plate. In these figures and all the subsequent figures, the normalized radiation impedances are in dB and plotted versus the

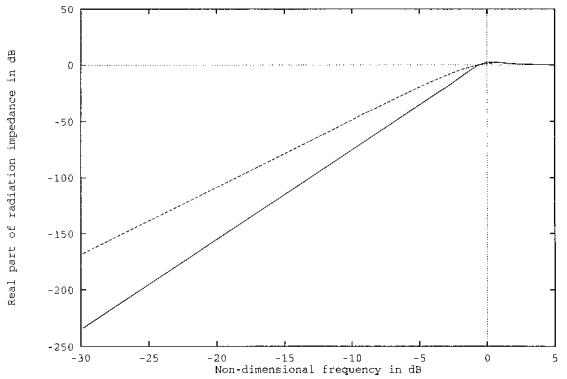


FIG. 7. Radiation impedance in the real part of the mode (2,2); ——— unbaffled, -------- baffled.

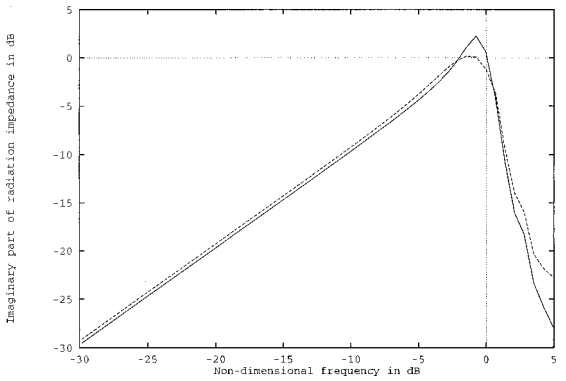


FIG. 8. Radiation impedance in the imaginary part of the mode (2,2); ——— unbaffled, -------- baffled.

nondimensional wave number k/k_{nm} in dB, where k is the acoustical wave number and k_{nm} is the mechanical modal wave number defined by

$$k_{nm}=\left[\left(\frac{n\pi}{a}\right)^2+\left(\frac{n\pi}{b}\right)^2\right]^{1/2}, \tag{32}$$

where a and b denote the dimensions of the plate. It can be seen that differences occur in the real part to within 3 dB. However, the difference between the exact and ''diagonal approximation'' is not sensitive in the imaginary part over the whole frequency range considered.

As was mentioned above, the diagonal terms can be easily calculated to a first approximation using the approximate formula (30), which neglects the cross-modal terms of the matrix $\gamma_{pq\,nm}$.

Figures 5 and 6 show the comparison of the radiation impedance when the plate is baffled and not baffled, for the direct diagonal term of the mode (1,1). The calculation of the unbaffled terms, in these comparisons, is not approximate. It can be seen that, at low frequencies, large differences do occur (more than 60 dB in the real part at a nondimensional frequency of -30 dB). Consequently, as is well-known experimentally, the baffle raises the radiation loading at low frequencies. This is due to the transformation of the equivalent monopole source of the mode (1,1) into a pure dipole source. Removing the baffle renders the mode a particularly ineffective radiator, increasing considerably the short-circuit effect, from one side to the other side of the plate.

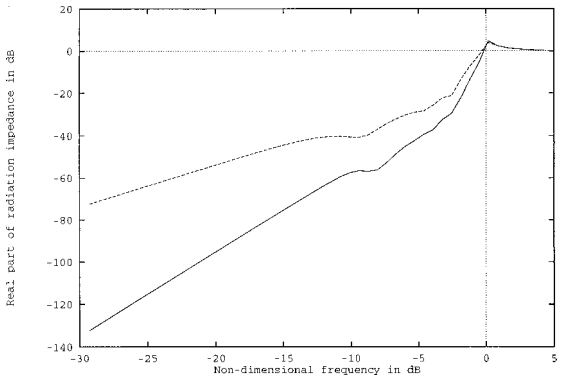


FIG. 9. Radiation impedance in the real part of the mode (7,7); ——— unbaffled, -------- baffled.

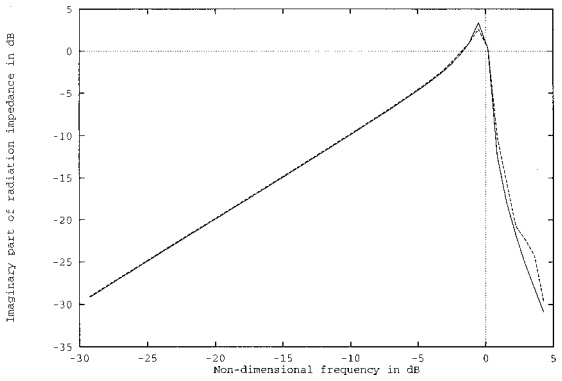


FIG. 10. Radiation impedance in the imaginary part of the mode (7,7); ——— unbaffled, -------- baffled.

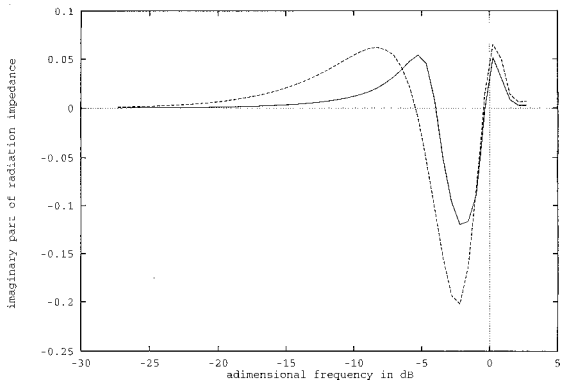


FIG. 12. Cross-modal radiation impedance in the imaginary part of the modes 11 13; ——— unbaffled, -------- baffled.

Conversely, the differences vanish at high frequencies and for the baffled plate the real part of the unbaffled radiation term tends to unity. The differences in the imaginary parts are not so strong since only 4 dB can be seen around the maximum value. Since the imaginary part is related to an added mass phenomenon, i.e., nonpropagating acoustical waves in the near-field of the plate, the presence or absence of the baffle strongly affects the propagating waves at low frequencies but only slightly the nonpropagating ones. Nevertheless, one can say that removing the baffle decreases a little the added mass phenomenon. This will be obvious in the computed velocity results of a plate immersed in water.

Figures 7 and 8 deal with the same kind of impedances but for the mode (2,2) to check the influence of the parity of the mode on the conclusions deduced from the first odd-odd mode. The behavior is the same: a large difference at low frequencies in the real part and little difference for the imaginary part. As for the baffled plate, the even-even modes radiate less than the odd-odd modes of the unbaffled plate.

Figures 9 and 10 show the results for the odd-odd mode of order (7,7). The same conclusions can be reached for both the real and the imaginary parts. Nevertheless, the differences on the imaginary part are now negligible over the whole frequency range $[-30,5]$ dB. Consequently, the differences in the imaginary part vanish when one increases the order of the mode.

Finally the study would not be complete without examining the cross-modal coefficients of the real part in Fig. 11 and the imaginary part in Fig. 12. These terms can be positive or negative. So, they are presented in linear versus the nondimensional modal wave number. Since the cross-modal coefficients involve two modes [here the modes (1,1) and (1,3)], the mechanical wave number chosen to render the wave number axis nondimensional is the larger. [In this example it is the mode (1,3)].

Although the behavior is the same, it can be seen that differences do occur between the baffled and the unbaffled plates, both in the real and the imaginary parts. What the figures do not exhibit, because of the linear presentation, is that, at low frequencies, the differences are of most importance, but the coefficients themselves tend to zero rapidly.

V. INFLUENCE OF THE BAFFLE ON THE SOUND RADIATED BY A SIMPLY SUPPORTED PLATE IN AIR AND IN WATER

A. Radiation simulation in heavy fluid

The interest of this simulation is that it allows us to check the whole set of modal coupling coefficient $\gamma_{pq\,nm}$, since all the coefficients (diagonal and off-diagonal, in real or imaginary parts) need to be calculated to get the correct simulation. The plate under consideration is the one defined in the organization (commission de validation des logiciels en acoustique) (cf. Ref. 12), i.e., a square 1-cm-thick steel plate, of $1\,\mathrm{m}^2$, excited at (0,7,0,2) by a 1 N driving point force. The damping coefficient η of the plate is fixed to be 10^{-2}. The water characteristics are a fluid density of

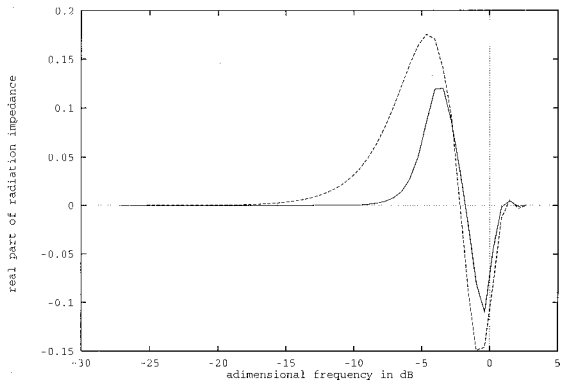


FIG. 11. Cross-modal radiation impedance in the real part of the modes 11 13; ——— unbaffled, -------- baffled.

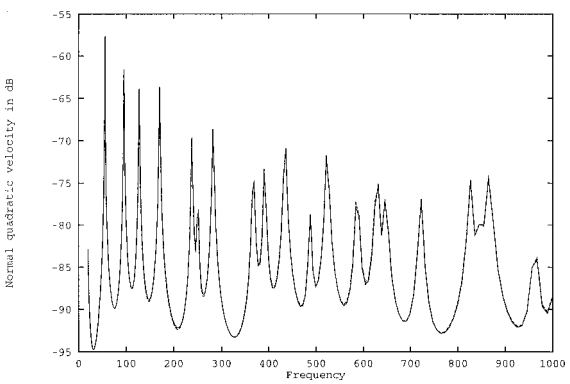


FIG. 13. Normal quadratic velocity for the unbaffled plate in water; ——— approximated, -------- exact.

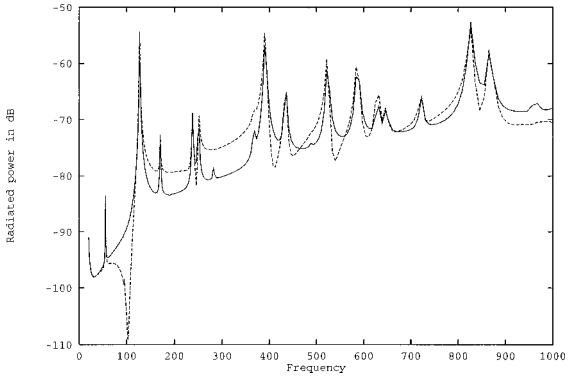


FIG. 14. Radiated power for the unbaffled plate in water; ——— approximated, -------- exact.

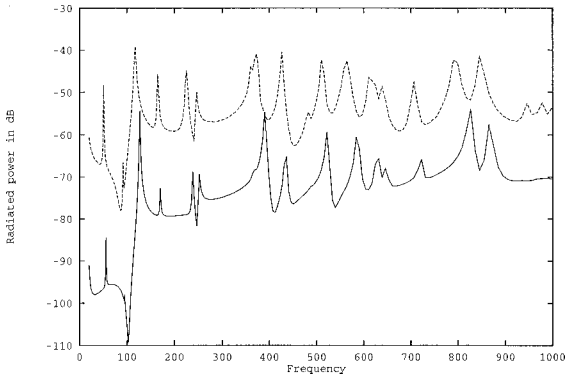


FIG. 16. Radiated power of the plate in water; ——— unbaffled, -------- baffled.

1000 kg/m^3 and a sound velocity of 1500 m/s. The solution is derived from the linear system (25) for the exact solution and from (31) for the approximate, diagonal-only solution. The structural damping is introduced in both systems by making complex the eigenfrequencies as follows:

$$\omega_{pq}^{2*} = \omega_{pq}^2 (1 + j\,\eta). \tag{33}$$

To reach the 1000-Hz frequency range and to take into account the fluid added mass, about 90 modes of the *in-vacuo* simply supported plate are needed to converge reasonably. Definitions of the acoustical radiated power, normal quadratic velocity, and radiation factor are given in Appendix C.

The two solutions are compared in Fig. 13 for the velocity and in Fig. 14 for the radiated power, the diagonal solution is computed by not calculating the off-diagonal terms $\gamma_{pq\,nm}$ and making them zero. As for the baffled problem, exact and diagonal solutions give similar results in velocity. However, there is a great sensitivity in the radiated power with and without the off-diagonal terms. Nevertheless it seems not to affect the frequency location of the peaks, but only the levels (both in negative and positive offsets).

This numerical simulation gives an estimation of the error introduced by neglecting the off-diagonal terms for an unbaffled plate in water, leading to 7 or 8 dB error in some frequency range only in the radiated power, while the velocity levels are not quite as affected. These remarks were already true for the baffled plate (cf. Ref. 12).

Figure 15 for the velocity, Fig. 16 for the radiated power, and Fig. 17 for the radiation factor show the comparison of the baffled and unbaffled plate, using the exact solution in both cases. One has to say that, to render this comparison meaningful, the baffled plate has to be loaded on its two sides. This is done in the FORTRAN program by multiplying by two the radiation impedance of the baffled plate (cf. Ref. 12). It can be seen that the velocity of the plate is quite similar. One can see that the resonance frequencies of the baffled plate are a little lower than the one of the unbaffled plate. This is due to the small differences between the imaginary parts of the radiation impedances in Figs. 6 and 8. Consequently, the fluid loading is a little smaller when unbaffled. On the contrary, the radiated power is much lower when unbaffled.

This difference is very important because the relative position of the excitation frequency to the critical frequency of the plate is given by

$$f_c = \frac{c^2}{2\pi h}\sqrt{\frac{12\rho(1-\nu^2)}{E}}, \tag{34}$$

where ρ, ν, E, and h are, respectively, the density, the Poisson ratio, the Young's modulus, and the thickness of the plate. The plate's critical frequency in water is about 23 kHz. Consequently, one can consider calculating the ratio of the current frequency at 1000 Hz to f_c (equal to about 4%), we are in a very low-frequency range in terms of sound radiation. The big decrease in radiated power (more than 20 dB in

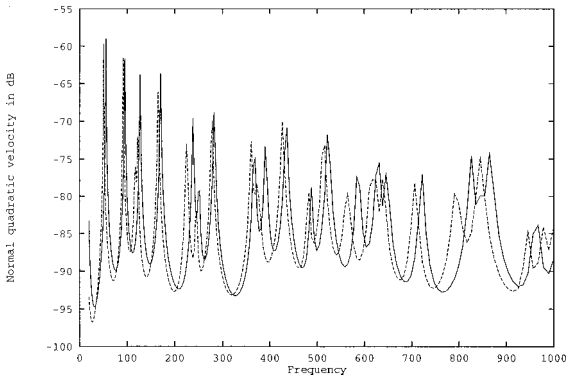


FIG. 15. Normal quadratic velocity for the plate in water; ——— unbaffled, -------- baffled.

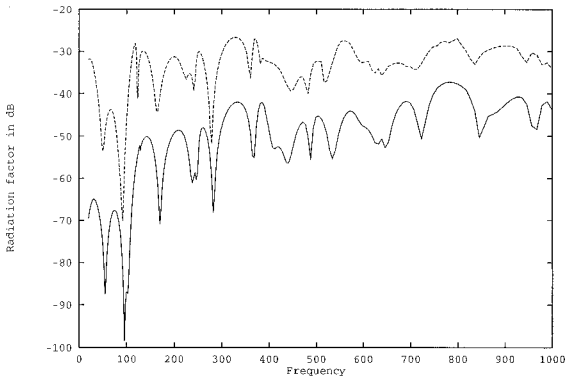


FIG. 17. Radiation coefficient of the plate in water; ——— unbaffled, -------- baffled.

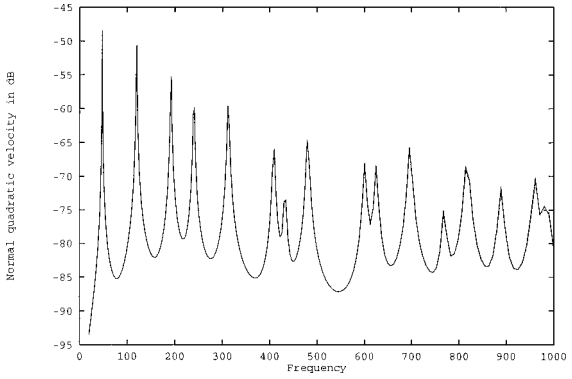


FIG. 18. Normal quadratic velocity of the plate in air; ——— unbaffled, -------- baffled.

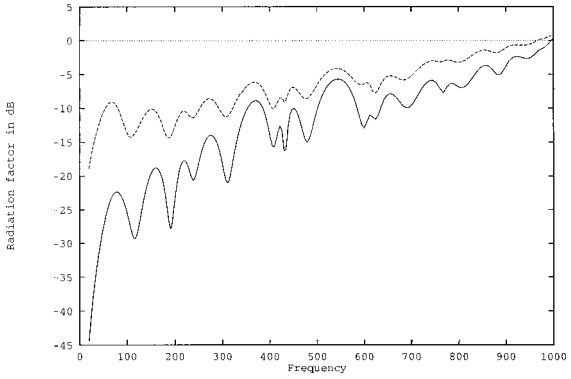


FIG. 20. Radiation factor of the plate in air; ——— unbaffled, -------- baffled.

average) is explained by the same proportional decrease in the radiation factor.

B. Radiation simulation in light fluid

The case under study is a plate immersed in air of fluid density 129 kg/m^3 and of sound velocity 340 m/s. For the purpose of light fluid simulation, only the diagonal approximation (31) is used. The exact solution (25) gives exactly the same numerical results for this 1-cm steel plate under consideration. The diagonal solution is, what is called, a ''good approximation.'' Obviously, in air, it can be seen that the velocity is not affected by the baffle, Fig. 18.

The plate critical frequency in air is about 1200 Hz. The ratio of the current frequency to the critical frequency is much higher in air than in water, belonging to [0,01,0,8]. Consequently, the power decrease in air when unbaffled is smaller than in water, only 5 dB at 1000 Hz, Fig. 19. The decrease of the radiation efficiency of the plate is in the same proportion, Fig. 20.

In conclusion, the acoustic power decrease is sensitive below the plate critical frequency, both in air and in water which is intuitively well corroborated.

VI. CONCLUSION

A model of acoustic radiation from an unbaffled simply supported plate has been presented and compared to the well-known baffled problem both in light and heavy fluid. Removing the baffle decreases the radiation efficiency of the plate and its radiated power, but is of little effect on its velocity, even in heavy fluid. The radiation impedances of the unbaffled plate have been presented and compared to the well-known baffled ones. Large differences appear in the real part, and smaller differences in the imaginary part.

The influence of the modal coupling is exhibited both in air and in water. For the baffled problem, the modal coupling in air is of small effect. Conversely, the modal coupling in water is very important, including small errors in velocity but large errors in the acoustic radiated power.

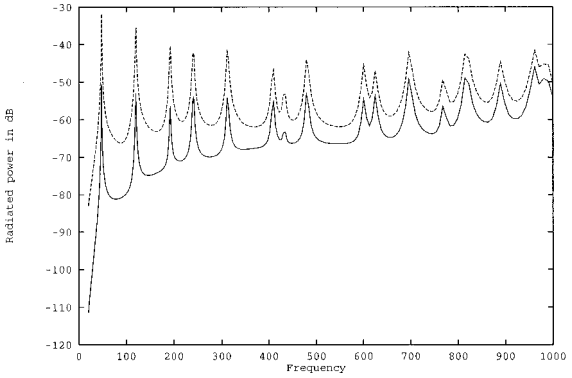


FIG. 19. Radiated power of the plate in air; ——— unbaffled, -------- baffled.

APPENDIX A

The modes $\phi_{nm}(x,y)$ are the modes of the unloaded simply supported plate. They satisfy:

$$D\nabla^4\phi_{nm}(x,y)=m\omega_{nm}^2\phi_{nm}(x,y), \tag{A1}$$

where ∇^4 is defined by

$$\nabla^4=\frac{\partial^4}{\partial x^4}+\frac{2\partial^4}{\partial x^2\,\partial y^2}+\frac{\partial^4}{\partial y^4} \tag{A2}$$

and where ω_{nm} is the *in-vacuo* eigenfrequency of a simply supported mode given by

$$\omega_{nm}=\left(\frac{D}{m}\right)^{1/2}k_{nm}^2, \tag{A3}$$

where k_{nm} denotes the modal mechanical wave number of the mode (n,m):

$$k_{nm}=\left[\left(\frac{n\pi}{a}\right)^2+\left(\frac{m\pi}{b}\right)^2\right]^{1/2}, \tag{A4}$$

(a,b) being the dimensions of the rectangular plate.

The orthogonality relationship is the following:

$$\int_{S_p}D\nabla^4\phi_{nm}\cdot\phi_{pq}dS_p=\int_{S_p}m\omega_{nm}^2\phi_{nm}\cdot\phi_{pq}dS_p$$

$$=m\,\frac{S_p}{4}\,\omega_{pq}^2\delta_{np}\delta_{mq}, \tag{A5}$$

where δ denotes here the Kronecker's delta.

The product $m(S_p/4)$ denotes also the generalized mass of the mode, which for a simply supported plate is independent of the order of the mode:

$$m_{pq}=m\cdot\frac{S_p}{4},\quad \forall_p\forall_q. \tag{A6}$$

The generalized force of the mode (p,q) can be calculated using:

$$\mathscr{F}_{pq}=\int_{S_p}F(x,y)\,\phi_{pq}(x,y)dx\;dy. \tag{A7}$$

This leads to the following relationship when a driving mechanical point force F_0 is applied at M_e at coordinates (x_e,y_e):

$$\mathscr{F}_{pq}=F_0\phi_{pq}(x_e,y_e)=\frac{S_p}{4}F_{pq}. \tag{A8}$$

Force and density of force (here a pressure) are related by the constant $S_p/4$.

APPENDIX B

The calculation of the Fourier transforms of the modal shapes $\phi_{pq}(x,y)$:

$$\tilde{\phi}_{pq}(k_x,k_y)=\int_0^a\int_0^b\sin\frac{n\pi}{a}x\,\sin\frac{m\pi}{b}y e^{-jk_xx}e^{-jk_yy}dx\;dy \tag{B1}$$

are evaluated analytically:

$$\tilde{\phi}_{pq}(k_x,k_y)=ab\pi^2pq\,\frac{[(-1)^pe^{-jk_xa}-1]}{a^2k_x^2-p^2\pi^2}\,\frac{[(-1)^qe^{-jk_yb}-1]}{b^2k_y^2-q^2\pi^2} \tag{B2}$$

and for the conjugate form of order (n,m):

$$\tilde{\phi}^*_{nm}(k_x,k_y)=ab\pi^2nm\,\frac{[(-1)^ne^{jk_xa}-1]}{a^2k_x^2-n^2\pi^2}\,\frac{[(-1)^me^{jk_yb}-1]}{b^2k_y^2-m^2\pi^2}. \tag{B3}$$

From the definition of $\gamma_{pq\,nm}$:

$$\gamma_{pq\,nm}=\frac{j}{8\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}k_z\tilde{\phi}_{pq}(k_x,k_y)\tilde{\phi}^*_{nm}(k_x,k_y)dk_x\;dk_y \tag{B4}$$

it can be seen that the integrand is odd with respect to k_x if the orders (n,p) versus the x direction are of different parity and that the integrand is odd with respect to ky for the same reasons versus the mode orders (q,m). It can be seen the integrand is even in the remaining cases of parity. This leads to the following properties:

$$\gamma_{pq\,nm}=\frac{j}{2\pi^2}\int_0^\infty\int_0^\infty k_z\tilde{\phi}_{pq}(k_x,k_y)\tilde{\phi}^*_{nm}(k_x,k_y)dk_x\;dk_y,$$

when p and n of same parity, and q and m of same parity. (B5)

$$\gamma_{pq\,nm}=0 \text{ when } p \text{ and } n \text{ of different parity, or } q \text{ and } m \text{ of different parity.} \tag{B6}$$

Finally, when the parity is the same one can write for the $\gamma_{pq\,nm}$ terms:

$$\gamma_{pq\,nm}=\frac{2ja^2b^2}{pq\;nm\;\pi^6}\int_0^\infty\int_0^\infty\sqrt{k^2-k_x^2-k_y^2}\times\frac{(1-(-1)^p\cos k_xa)}{(a^2k_x^2/n^2\pi^2-1)(a^2k_x^2/p^2\pi^2-1)}\times\frac{(1-(-1)^q\cos k_yb)}{(b^2k_y^2/q^2\pi^2-1)(b^2k_y^2/m^2\pi^2-1)}\,dk_x\;dk_y. \tag{B7}$$

Substituting the square root dependence given in Eq. (9), one can write the imaginary part as

$$\mathrm{Im}(\gamma_{pq\,nm})=\frac{2a^2b^2}{pq\;nm\;\pi^6}\int_0^k dk_x\int_0^{\sqrt{k^2-k_x^2}}\sqrt{k^2-k_x^2-k_y^2}\times\frac{(1-(-1)^p\cos k_xa)}{(a^2k_x^2/n^2\pi^2-1)(a^2k_x^2/p^2\pi^2-1)}\times\frac{(1-(-1)^q\cos k_yb)}{(b^2k_y^2/q^2\pi^2-1)(b^2k_y^2/m^2\pi^2-1)}\,dk_y \tag{B8}$$

and the real part as

$$\mathrm{Re}\{\gamma_{pq\,nm}\}=\frac{-2a^2b^2}{pq\;nm\;\pi^6}\int_0^\infty dk_x\int_{\sqrt{k^2-k_x^2}\geqslant 0}^\infty\sqrt{k_x^2+k_y^2-k^2}\times\frac{(1-(-1)^p\cos k_xa)}{(a^2k_x^2/n^2\pi^2-1)(a^2k_x^2/p^2\pi^2-1)}\times\frac{(1-(-1)^q\cos k_yb)}{(b^2k_y^2/q^2\pi^2-1)(b^2k_y^2/m^2\pi^2-1)}\,dk_y. \tag{B9}$$

These two integrals are evaluated numerically using the criteria which will be described below.

The numerical integration over the singularities of the denominators in (B8) and (B9) is done using the following asymptotic formulas:

if $n=p$

$$\begin{cases}\dfrac{1-(-1)^p\cos k_xa}{[a^2k_x^2/n^2\pi^2-1]^2}\to\left(\dfrac{p\pi}{2}\right)^2\\ k_x\to\dfrac{n\pi}{a},\end{cases} \tag{B10}$$

if $q=m$

$$\begin{cases} \dfrac{1-(-1)^q \cos k_y b}{[b^2k_y^2/m^2\pi^2-1]^2} \to \left(\dfrac{m\pi}{2}\right)^2 \\ k_y \to \dfrac{m\pi}{b}, \end{cases} \tag{B11}$$

if $n \neq p$

$$\begin{cases} \dfrac{1-(-1)^p \cos k_x a}{[a^2k_x^2/n^2\pi^2-1][[a^2k_x^2/p^2\pi^2-1]]} \to 0 \\ \text{if } k_x \to \dfrac{n\pi}{a} \text{ or } k_x \to \dfrac{p\pi}{a}, \end{cases} \tag{B12}$$

if $m \neq q$

$$\begin{cases} \dfrac{1-(-1)^q \cos k_y b}{[b^2k_y^2/q^2\pi^2-1][[b^2k_y^2/m^2\pi^2-1]]} \to 0 \\ \text{if } k_y \to \dfrac{q\pi}{b} \text{ or } k_y \to \dfrac{m\pi}{b}. \end{cases} \tag{B13}$$

To ensure the numerical convergence of these oscillating integrals both the k_x and the k_y domains have to be discretized based on wavelength criteria. In this paper, the criteria are fixed to six points per wavelengths in k_x and k_y. These lead to the following mesh:

$$\Delta k_x = \frac{2\pi}{a} \cdot \frac{1}{6} = \frac{\pi}{3a}, \quad \Delta k_y = \frac{2\pi}{b} \cdot \frac{1}{6} = \frac{\pi}{3b}. \tag{B14}$$

Criteria (B14) allows us to ensure good accuracy in the results and a good ratio accuracy to time consuming.

An additional criterion must be stated to limit the domain of integration which must be truncated to a finite value, both in k_x and k_y. From (B9) it can be seen that the integrand tends to zero when k_x tends to infinity or k_y tends to infinity, respectively (following an $O(k_x^{-3})$ or an $O(k_y^{-3})$). Thus the domain is limited versus k_x and k_y at values such as

$$k_x^{\text{stop}} = \frac{n\pi}{a} + K\,\frac{n\pi}{a}, \quad k_y^{\text{stop}} = \frac{m\pi}{b} + K\,\frac{m\pi}{b}, \tag{B15}$$

where K is an arbitrary integer and n or m is the majorant of the couples of integer (n,p) or (q,m). In the numerical results presented here K is fixed to 4 which is sufficient to give a good accuracy. A similar process of integration has already been used for the calculation of the radiation impedances of a simply supported baffled plate which involves very similar integrands (the k_z term is at the denominator and not at the numerator).

APPENDIX C

The radiated power is defined by

$$W = \frac{1}{2} \int_{S_{Pt}} R_e\{p \cdot \mathbf{v}^*\}\mathbf{n}\, dS_{p_t}, \tag{C1}$$

where S_{p_t} denotes S_{p+} plus S_{p-}. Splitting the integral (C1) this gives:

$$W = \frac{1}{2} \int_{S_{p+}} R_e\{p^+ \cdot \mathbf{v}^*\}\mathbf{n}\, dS_{p+} + \frac{1}{2} \int_{S_{p-}} R_e\{p^- \cdot \mathbf{v}^*\}\mathbf{n}\, dS_{p-}. \tag{C2}$$

Since $\mathbf{n}$ denotes the normal to the surfaces pointing inward to V, we can associate $\mathbf{n}$ to $\mathbf{z}$ or $-\mathbf{z}$ depending on the side of the plate under consideration giving:

$$W = \frac{1}{2} \int_{S_p} R_e\{\{p^- - p^+\}v_z^*\}dS_p, \tag{C3}$$

where v_z denotes the velocity component along $\mathbf{z}$ of a point on the plate.

Using a Fourier decomposition in a series of the modal shapes of the plate, the pressure jump can be written as

$$\overline{p}(x,y) = \sum_{l=1}^{\infty} \sum_{s=1}^{\infty} p_{ls}\phi_{ls}(x,y). \tag{C4}$$

The plate velocity can be written as

$$v_z^* = \frac{\partial w^*}{\partial t} = -j\omega \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{nm}^* \phi_{nm}(x,y). \tag{C5}$$

This gives the following formula, when the integral on S_p is performed:

$$W = \frac{\omega}{2}\frac{S_p}{4} R_e\left\{ j \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} p_{nm} a_{nm}^* \right\}. \tag{C6}$$

We can now write the p_{nm} coefficients function of the displacement amplitude a_{nm} using the plate equilibrium equation:

$$D\nabla^4 \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{nm}\phi_{nm}(x,y) - m\omega^2 \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{nm}\phi_{nm}(x,y) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} F_{nm}\phi_{nm}(x,y) + \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} p_{nm}\phi_{nm}(x,y). \tag{C7}$$

Using the orthogonality relationships between the modal shapes, this gives:

$$m(\omega_{pq}^2 - \omega^2)a_{pq} = F_{pq} + p_{pq}, \tag{C8}$$

and finally introducing (C8) into (C6):

$$W = \frac{\omega S_p}{8} R_e\left\{ j \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} [m(\omega_{nm}^2 - \omega^2)a_{nm} - F_{nm}]a_{nm}^* \right\}. \tag{C9}$$

The structural losses can be introduced by making complex the eigenfrequencies of the modes, by replacing in (C9) ω_{nm}^2 by ω_{nm}^{2*} defined by:

$$\omega_{nm}^{2*} = \omega_{nm}^2(1 + j\eta), \tag{C10}$$

where η denotes the structural damping coefficient.

The normal quadratic volocity is given by

$$\langle v^2 \rangle = \frac{1}{2S_p} \int |v(x,y)|^2 dx\, dy, \tag{C11}$$

where $v(x,y)$ is the local velocity of a point on the plate of the surface S_p. Using an expansion in the series of plate modes for the velocity:

$$v(x,y) = j\omega \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{nm}\phi_{nm}(x,y), \tag{C12}$$

this gives

$$\langle v^2\rangle = \frac{\omega^2}{8}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}|a_{nm}|^2 \tag{C13}$$

and for the radiation factor

$$\sigma = \frac{W}{\rho_0 c S_p \langle v^2\rangle}, \tag{C14}$$

where W is the radiated power given by (C9).

The velocity, radiated power and radiation factor are presented in dB as follows:

$$\begin{aligned} L_w &= 10\log W, \\ L_v &= 10\log\langle v^2\rangle, \\ L_\sigma &= 10\log\sigma. \end{aligned} \tag{C15}$$

[1] G. Maidanik, ''Response of ribbed panels,'' J. Acoust. Soc. Am. **34**, 809–835 (1962).

[2] G. Maidanik, ''Vibrational and radiative classifications of modes of a baffled finite panel,'' J. Sound Vib. **34**, 447–455 (1974).

[3] C. E. Wallace, ''Radiation resistance of rectangular panel,'' J. Acoust. Soc. Am. **53**, 946–952 (1972).

[4] H. G. Davies, ''Acoustic radiation from fluid loaded rectangular plates,'' MIT Report No. 71476-1.

[5] H. G. Davies, ''Sound from turbulent boundary layer excited panels,'' J. Acoust. Soc. Am. **49**, 878–889 (1971).

[6] W. R. Graham, ''High frequency vibration and acoustic radiation of fluid loaded plates,'' Philos. Trans. R. Soc. London, Ser. A **352**, 1–43 (1995).

[7] A. Harari and B. E. Sandman, ''Vibratory response of laminated cylindrical shells embedded in an acoustic fluid,'' J. Acoust. Soc. Am. **60**, 117–128 (1976).

[8] P. R. Stepanishen, ''Radiated power and radiation loading of cylindrical surfaces with non uniform velocity distributions,'' J. Acoust. Soc. Am. **63**, 328–338 (1978).

[9] B. Laulagnet and J. L. Guyader, ''Modal analysis of a shell's acoustic radiation in light and heavy fluid,'' J. Sound Vib. **131**, 397–415 (1989).

[10] B. E. Sandman, ''Fluid loading influence coefficients for a finite cylindrical shell,'' J. Acoust. Soc. Am. **60**, 1256–1264 (1976).

[11] N. Atalla, J. Nicolas, and C. Gauthier, ''Acoustic radiation of an unbaffled vibrating plate with general elastic boundary conditions,'' J. Acoust. Soc. Am. **99**, 1484–1494 (1996).

[12] B. Laulagnet and J. L. Guyader, Commission for the prediction of radiated noise of the French Mechanical Engineering Society (S.F.M.) and the French Acoustical Society (S.F.A.). Reference results in plate acoustic radiation, 4 numerical simulations, 1 experiment, 1996.

Measurement of the reflection matrix of a discontinuity in a duct

Mohamed Akoum and Jean-Michel Ville
Laboratoire LG2mS, UPRES A CNRS 6066 Département de Génie Mécanique, Université de Technologie de Compiègne, BP 20 529, 60205 Compiègne Cedex, France

(Received 25 September 1996; accepted for publication 30 January 1998)

An original technique of measuring reflection (or transmission) matrices of duct discontinuities is described. The technique is based on measuring the acoustic pressure by microphone pairs in two closed cross sections of a constant area circular duct, which separates incident and reflected modes after a modal decomposition by a Fourier–Lommel transform. Theory, experimental setup, and data treatment are presented. Then experimental reflection and conversion coefficients at an open duct end for plane wave and higher order modes are presented as an application and compared with analytical results. © *1998 Acoustical Society of America.* [S0001-4966(98)02805-7]

PACS numbers: 43.58.Vb, 43.20.Mv [SLE]

LIST OF SYMBOLS

a	duct radius
d	distance between measurement cross sections
r,θ,z	cylindrical coordinates
t	time
$c=340\ \mathrm{m\cdot s^{-1}}$	speed of sound
f	frequency
f^c_{mn}	cut-on frequency of the mode (m,n)
m	azimuthal wave number
n	radial number
$\omega=2\pi f$	pulsation
$k=\omega/c$	total wave number
k_{mn}	axial wave number
λ_{mn}	axial wavelength
p	acoustic pressure
J_m	Bessel function of first kind and order m
χ_{mn}	nth zero of J'_m, first derivative of J_m
A^i_{mn}	amplitude of incident mode (m,n)
A^r_{mn}	amplitude of reflected mode (m,n)
$R_{mn,pq}$	reflection coefficient of incident mode (p,q) over reflected mode (m,n)
$[\ ]$	matrix $N\times N$
$\{\}$	vector

INTRODUCTION

In the studies of mufflers and similar devices, the reflection and transmission properties are of interest. In the general case, with N propagating modes in a duct, these properties have to be described by $[N\times N]$ matrices. Indeed, in a reflection matrix, diagonal terms are the reflection coefficients and other terms are conversion coefficients. Until now, most of the experimental methods provide only the reflection coefficients, because plane wave propagation is assumed.[1] Also, when a spinning mode synthesizer generates one mode at a time,[2] the reflection coefficients of the selected higher order modes are determined.[3] In 1989, Abom, working[4] on a modal decomposition technique based on transfer function measurements between microphone pairs, proposed a method to measure the reflection matrix. However, because of a major problem in the generation of the required number of independent incident fields, experimental results were limited to plane wave propagation.

The method which we are presenting here is similar to that described by Abom.[4] The pressure distribution is measured in two closed cross sections of a constant circular area duct before the discontinuity by a pair of microphones. A specified modal decomposition and an original process of independent incident fields generation are used.

I. THEORY

A. Description of the sound field

Consider acoustic propagation in a semi-infinite hard wall cylindrical duct (Fig. 1) without mean flow. The fluid is assumed to be ideal and linear acoustic theory is valid. Then it is well known that, at time t, the acoustic pressure distribution in the duct can be written in cylindrical coordinates (r,θ,z) as

$$p(r,\theta,z,t)=\sum_{m=-\alpha}^{+\alpha}\sum_{n=0}^{\alpha} C_{mn}(z)p_{mn}(r,\theta)e^{-i\omega t}, \tag{1}$$

with the eigenfunctions

$$p_{mn}(r,\theta)=J_m\left(\frac{\chi_{mn}}{a}\right)e^{im\theta} \tag{2}$$

and the total modal coefficients in the cross section z

$$C_{mn}(z)=[A^i_{mn}e^{ik_{mn}z}+A^r_{mn}e^{-ik_{mn}\cdot z}], \tag{3}$$

where A^i_{mm} and A^r_{mn} are, respectively, the incident and reflected modal coefficients, (m,n) are integer numbers representing azimuthal wave number and radial number, (0,0) corresponds to the plane wave mode, and $k^2_{mn}=k^2-(\chi_{mn}/a)^2$ is the square of the axial wave number.

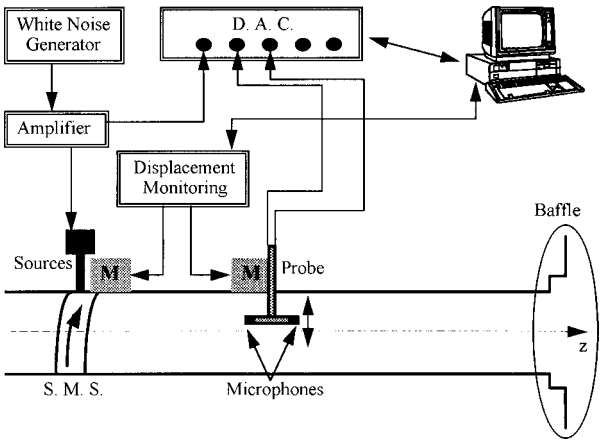


FIG. 1. Duct configuration.

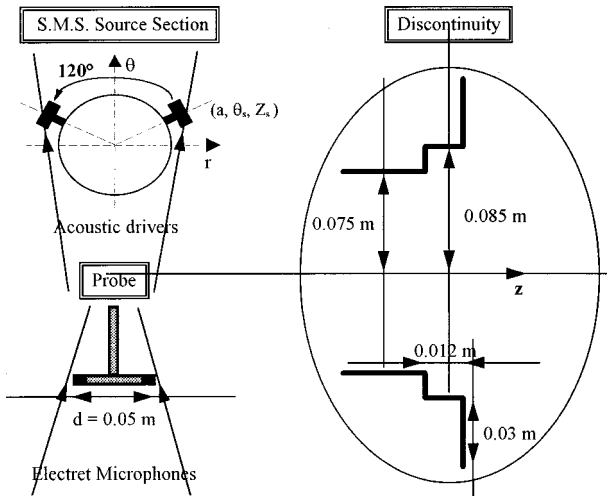


FIG. 2. Test configuration.

B. Determination of incident and reflected modal coefficients A^i_{mn} and A^r_{mn}

In the constant circular area duct, assuming that the sound pressure fields p_1 and p_2 are known in two cross sections located at z_1 and z_2 (Fig. 1), then:

$$p_1(r,\theta,t)=\sum_{m=-u}^{\mu}\sum_{n=0}^{\nu} C^1_{mn}p_{mn}(r,\theta)e^{-i\omega t},$$

$$C^1_{mn}=C_{mn}(z_1), \tag{4}$$

$$p_2(r,\theta,t)=\sum_{m-u}^{\mu}\sum_{n-0}^{\nu} C^2_{mn}p_{mn}(r,\theta)e^{-i\omega t},$$

$$C^2_{mn}=C_{mn}(z_2). \tag{5}$$

In Eqs. (4) and (5), all (u,v) modes, above and below the cutoff, which have significant levels at the z_1 and z_2 measurement positions should be included in the distribution.

A Fourier–Lommel transform[5] is applied to both of the total sound pressure distributions p_1 and p_2 in order to determine the total modal coefficients, $C^1_{mn}(f)$ and $C^2_{mn}(f)$. Incident and reflected modal coefficients, $A^i_{mm}(f)$ and $A^r_{mn}(f)$ are solutions of the following system of equations:

$$\begin{aligned} C^1_{mn}(f)&=A^i_{mn}(f)e^{ik_{mn}z_1}+A^r_{mn}(f)e^{-ik_{mn}z_1},\\ C^2_{mn}(f)&=A^i_{mn}(f)e^{ik_{mn}z_2}+A^r_{mn}(f)e^{-ik_{mn}z_2}. \end{aligned} \tag{6}$$

Equations (6) can also be written in a matrix form as

$$\begin{Bmatrix} A^i_{mn}\\ A^r_{mn}\end{Bmatrix}=M^{-1}_{mn}\begin{Bmatrix} C^1_{mn}\\ C^2_{mn}\end{Bmatrix}, \tag{7}$$

where

$$M_{mn}=\begin{bmatrix} e^{ik_{mn}z_1} & e^{-ik_{mn}z_1}\\ e^{ik_{mn}z_2} & e^{-ik_{mn}z_2}\end{bmatrix}. \tag{8}$$

Equation (7) has solutions if $\det[M_{mn}]\neq 0{:}d\neq q(\lambda_{mn}/2)$, $q=1,2,...$ [$d=z_2-z_1$ is the distance between the cross sections; $\lambda_{mn}=2\pi/k_{mn}$ is the axial wave length of the mode (m,n)]. The influence of errors on the two microphone transfer function technique has already been discussed.[6] The generalization of this technique to a modal propagation results in the following condition on d:

$$\lambda_{mn}/20<d<(2/5)\lambda_{mn}. \tag{9}$$

As the plane axial wavelength is the smallest among all modes, the upper limit in the previous inequality specifies an upper limit $f_{\max}$ to the frequency domain of measurement:

$$f<f_{\max}=\frac{2}{5}\frac{c}{d}. \tag{10}$$

Near mode cut-off frequencies, a large dispersion appears because d is very small compared to an axial wavelength. A frequency domain where error becomes very important is then delimited for each mode, the upper limit of that is deduced from the lower limit of the inequality [Eq. (9)]:

$$\forall(m,n),\ \ f^c_{mn}<f<f^c_{mn}\sqrt{1+\left(\frac{1}{20}\frac{c}{df^c_{mn}}\right)^2}, \tag{11}$$

if $m=0$ and $n=0$, $0<f<c/(20d)$ where f^c_{mn} is the cut–on frequency of the mode (m,n). (In Table II, the associated ka values are calculated with the experimental distance d.)

If the acoustic pressure is distributed over N modes, the overall solution can be written as two vectors A^i and A^r with N components (order is given by cut-off frequency order):

$$A^i=\{A^i_{mn}\} \tag{12}$$

and

$$A^r=\{A^r_{mn}\}. \tag{13}$$

C. Determination of reflection and conversion coefficients

1. Reflection and conversion matrix expression

If the duct can be regarded as a linear, time invariant and passive system in the positive z direction, then there exist a linear relationship between A^i and A^r:

$$\{A^r\}=[R]\cdot\{A^i\}, \tag{14}$$

where $[R]$ is the $[N\times N]$ reflection matrix.

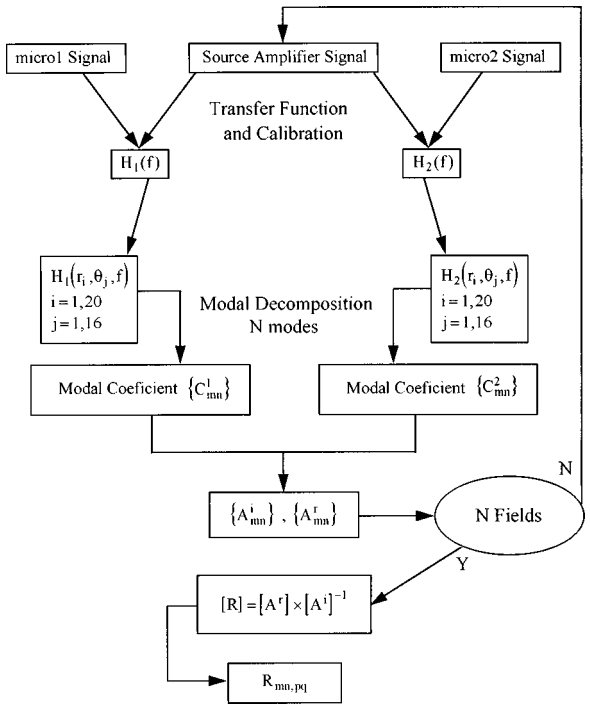


FIG. 3. The data treatment diagram.

Coefficients $R_{pq,mn}$ of matrix $[R]$ are the pressure reflection coefficients: reflection of the incident mode (m,n) on reflected mode (p,q). When $m=p$ and $n=q$ they are called the reflection coefficients of mode (m,n) and when $m\neq p$ or $n\neq q$ they are called the conversion coefficients: the reflection of the incident mode (m,n) to the reflected mode (p,q). As shown by Eq. (15) below, the reflection of the incident mode (m,n) and the conversion of all other incident modes on (m,n) contribute to the amplitude of the reflected mode (m,n):

$$A^r_{mn}=R_{mn,mn}A^i_{mn}+\sum_{mn\neq pq}R_{mn,pq}A^i_{pq}. \tag{15}$$

N independent incident acoustic pressure fields have to be generated in order to calculate N^2 coefficients $R_{pq,mn}$ from the system:

$$[R]=[A^r][A^i]^{-1}. \tag{16}$$

A column j of $[A^i]$ is filled with $A^{i,j}=\{A^{i,j}_{mn}\}$, a vector shown by Eq. (12) but for the jth incident field configuration. The $N\times N$ matrix $[A^i]$ is given by

$$[A^i]=[A^{i,1}_{mn}\quad\cdots\quad A^{i,j}_{mn}\quad\cdots]. \tag{17}$$

The matrix of the reflected pressure fields, $[A^r]$, is also determined and written in the same manner as $[A^i]$.

2. Generation of N independent incident fields

Let $r_s=a$, θ_s, and z_s be cylindrical coordinates of the sth source (Fig. 2) and A_s and φ_s be the amplitude and phase. The modal coefficients A^i_{mn} of the incident acoustic pressure field produced by these N_s point sources in an infinite cylindrical hard wall duct can be written[2] as the product of a coefficient D_{mn} ($D_{mn}=1/2ik_{mn}N_{mn}$, where N_{mn} is the normalization factor) by a function of the location and physical source characteristics:

$$A^i_{mn}=D_{mn}\left(\sum_{s=0}^{N_s-1}A_s e^{i(m\theta_s+k_{mn}z_s+\varphi_s}\right). \tag{18}$$

When N modes propagate, the total incident field is written as a vector:

$$\{A^i\}=\{A^i_{mn}\}=\mathrm{diag}(D_{mn})\left\{\left(\sum_{s=0}^{N_s-1}A_s e^{i(m\theta_s+k_{mn}z_s+\varphi_s)}\right)\right\}. \tag{19}$$

Modification of the source configuration is performed by rotating all sources by an azimuthal angle $j\Delta\theta$ (j is an integer). Then two acoustic incident pressure fields $j=j_1$ and $j=j_2$ are linearly independent if $j_1\neq j_2$ as shown by the following relationship:

$$\{A^{i,j}_{mn}\}=\mathrm{diag}(D_{mn})\left\{\left[\sum_{s=0}^{N_s-1}A_s e^{i(m\theta_s+k_{mn}z_s)}e^{imj\Delta\theta}\right]\right\}. \tag{20}$$

For the case of a pressure field that consists of 4 modes (m_1, 0); (m_2, 0); (m_3, 0); (m_1; 1). Then the rotation is repeated 3 times ($j=1$ to $N-1=3$) and the incident pressure field matrix $[A^i]$ which is based on the 4 configurations ($j=0$ is an initial configuration) is then given by

$$[A^i]=\mathrm{diag}\left(D_{mn}\sum_{s=0}^{N_s-1}A_s e^{i(m\theta_s+k_{mn}z_s+\varphi_s)}\right)\times\begin{bmatrix}1 & e^{im_1\Delta\theta} & e^{im_1 2\Delta\theta} & e^{im_1 3\Delta\theta}\\ 1 & e^{im_2\Delta\theta} & e^{im_2 2\Delta\theta} & e^{im_2 3\Delta\theta}\\ 1 & e^{im_3\Delta\theta} & e^{im_3 2\Delta\theta} & e^{im_3 3\Delta\theta}\\ 1 & e^{im_1\Delta\theta} & e^{im_1 2\Delta\theta} & e^{im_1 3\Delta\theta}\end{bmatrix}. \tag{21}$$

This example points out that, when two modes have the same m but different n, two lines are identical and the matrix $[A^i]$ is not reversible. To avoid this problem, in the fourth field, a source ($\Delta\theta=0$) is translated by Δz from its axial position, as shown below,

$$[A^i]=\mathrm{diag}\left(D_{nm}\sum_{s=0}^{N_s-1}A_s e^{i(m\theta_s+k_{mn}z_s+\varphi_s)}\right)\times\begin{bmatrix}1 & e^{im_1\Delta\theta} & e^{im_1 2\Delta\theta} & e^{ik_{m_1 0}\Delta z}\\ 1 & e^{im_2\Delta\theta} & e^{im_2 2\Delta\theta} & e^{ik_{m_2 0}\Delta z}\\ 1 & e^{im_3\Delta\theta} & e^{im_3 2\Delta\theta} & e^{ik_{m_3 0}\Delta z}\\ 1 & e^{im_1\Delta\theta} & e^{im_1 2\Delta\theta} & e^{ik_{m_1 1}\Delta z}\end{bmatrix}. \tag{22}$$

The N^2 coefficients $R_{pq,mn}$ are determined by solving Eq. (16).

II. THE REFLECTION MATRIX MEASUREMENT TECHNIQUE TEST

A. Experimental setup and procedure

The purpose of the experiment is the measurement of the reflection matrix of an open duct end discontinuity shown on

TABLE I. Source configurations.

Source configuration No.	Source 1 $\theta°$	Source 1 z(m)	Source 2 $\theta°$	Source 2 z(m)	Note
1	0	0	120	0	Experiment
2	15	0	135	0	Simulation
3	30	0	150	0	Simulation
4	40	0	160	0	Simulation
5	50	0	170	0	Simulation
6	0	0	120	0.15	Experiment

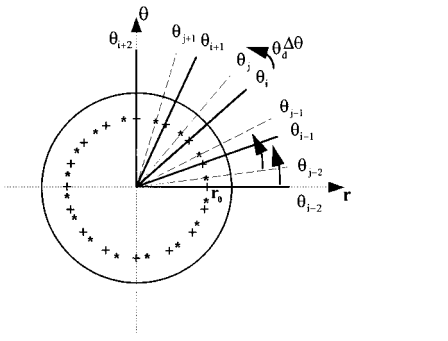
Fig. 2. The duct diameter is 0.15 m. Two acoustic drivers mounted in a duct section which can rotate over 360° are located at the periphery of the wall in the same cross section or at two different axial positions. During an experiment, they are driven in phase with a white noise signal that is band limited to 0–3000 Hz. For this duct radius and frequency range the parameter ka (a is the duct radius) is lower than 4, therefore only 6 modes (0,0); ($\pm$1,0); ($\pm$2,0) and (0,1) are cut-on at 3000 Hz.

Two 0.6-mm-diam electret microphones are mounted back to back on a support which traverses the duct diameter. The axial separation between microphones was set at 5 cm [according to Eq. (10) $f_{\max}=2760$ Hz].

A 320-points discretization of the acoustic pressure field measurement in cross sections z_1 and z_2 (Refs. 7, 8) is achieved by rotating the source section at 16 positions equally space by 22.5° and displacing the probe to 20 radial positions. This allows modal decomposition from $m=-7$ to $+7$ and $n=0$ to 9.

As shown in Fig. 3, transfer function measurements between signals from each microphone and source amplifier are achieved simultaneously in both duct cross sections for each of the 16×20 microphone positions. Then, these transfer functions are Fourier–Lommel transformed[5] for each frequency, yielding the total modal coefficients $C^1_{mn}(f)$ and $C^2_{mn}(f)$ with $-7<m<7$; $n<9$. The incident and reflected modal coefficients vectors $\{^{A^i_{mn}}_{A^r_{mn}}\}$ are then calculated using Eq. (7).

As the measurement section is far from the discontinuity, only the $N=6$ cut-on modes at 3000 Hz will be taken into account to determine the reflection matrix coefficients. Therefore, six source configurations (Table I) which satisfy the conditions explained in Sec. II C2 should be set up.



From 16 Radius (320 mesureament positions) ➡ 16 Interpolated Radius

$\theta_{i+1}-\theta_i=360°/16$ (i = 1, 16) is the experimental azimuthal incrementation

$\theta_{j+1}-\theta_j=360°/16$ (j = 1, 16) is the interpolated azimuthal incrementation

$\Delta\theta=\theta_j-\theta_i$

FIG. 4. Azimuthal interpolation of pressure field.

TABLE II. Cut-on frequency parameter ka.

(m,n)	(0,0)	($\pm$1,0)	($\pm$2,0)	(0,1)	($\pm$3,0)
nondimensional cut-on frequency ka	0	1.84	3.05	3.83	4.2
nondimensional upper frequency limit ka in Eq. (11)	0.47	1.90	3.08	3.86	4.23

To reduce the time for the experimental procedure, the acoustic pressure fields generated by rotation (No. 2, 3, 4, 5) are deduced from the measured field. As the configuration is axisymetric, there is reciprocity between rotating the sources and rotating the probe. In order to simulate the rotation of the sources, a computation of $\Delta\theta$ rotation of the probe is achieved by a B_3 uniform splines function[9] interpolation (Fig. 4) from a measured pressure field.

B. Results

Experimental results are compared with an analytical model based on modal theory of wave propagation in a semi-infinite duct without flow[10] ended by an infinite flange. The radiation impedance condition at the end of the duct are calculated as given by Zorumski.[11] As in the experiment, the duct is ended by a short flange (Figs. 1 and 2), therefore, the theoretical results are expected to slightly overestimate the experimental results.[12]

The modulus and phase of the coefficients are plotted versus ka, for their cut-on frequency to $ka=4$. Results are easier to analyze if the cut-on frequency parameters are followed as indicated in Table II. All results are diverging

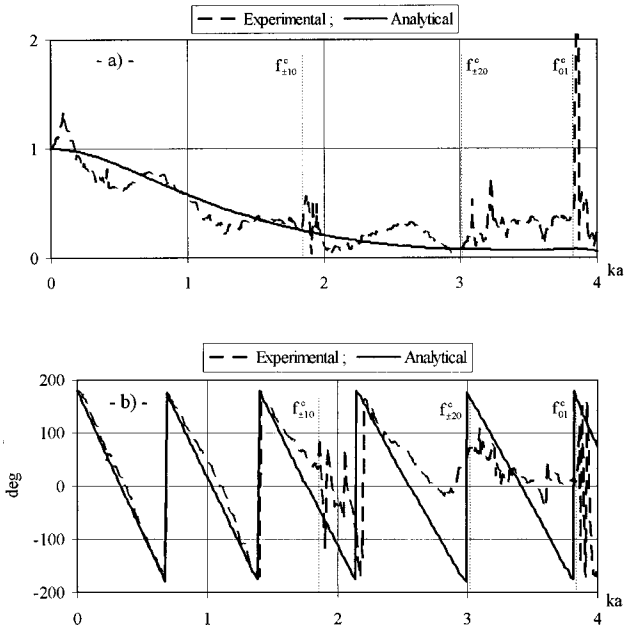


FIG. 5. Reflection coefficient spectrum $R_{00,00}$: (a) modulus; (b) phase.

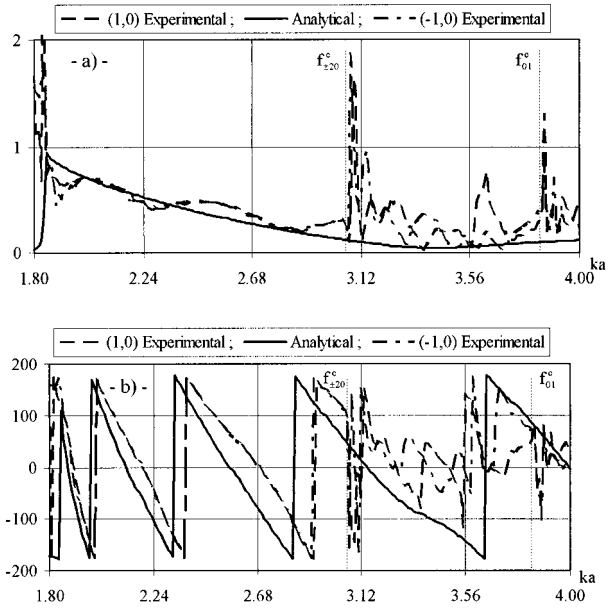


FIG. 6. Reflection coefficient spectrum $R_{\pm 10,\pm 10}$: (a) modulus; (b) phase.

around the cut-off frequencies as pointed out by remarks in Sec. I B.

1. Plane wave reflection coefficient: $R_{00,00}$

The complex coefficient $R_{00,00}$ is plotted (Fig. 5) from $ka=0$ to 4 and compared with theory. Agreement is quite good, but results are diverging around the cut-on frequencies of modes $(\pm 1,0)$, $(\pm 2,0)$, and particularly for (0,1) because its cut-on frequency is higher than $f_{\max}$ [Eq. (10)]. An oscillation is evident which is the characteristic of the diffraction effect[13] by the probe. This oscillation is of the order of the difference between cases with and without the flange. So, one can conclude that our comparison may not give an absolute estimation due to the difference in the cases concerned, i.e., numerical and experimental cases.

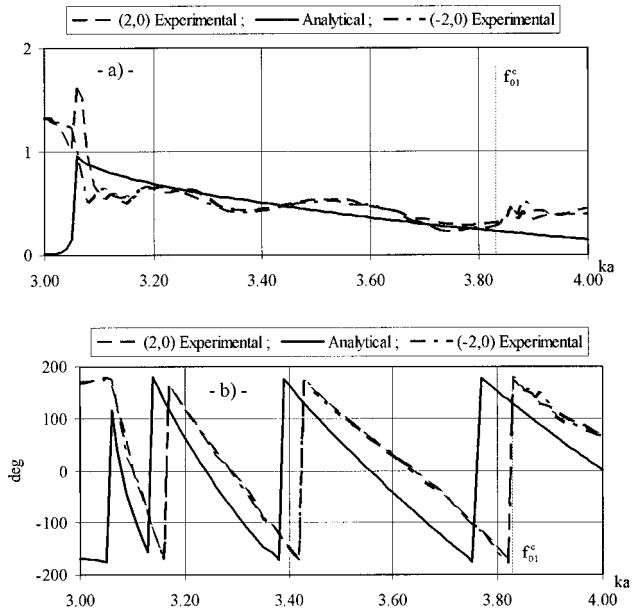


FIG. 7. Reflection coefficient spectrum $R_{\pm 20,\pm 20}$: (a) modulus; (b) phase.

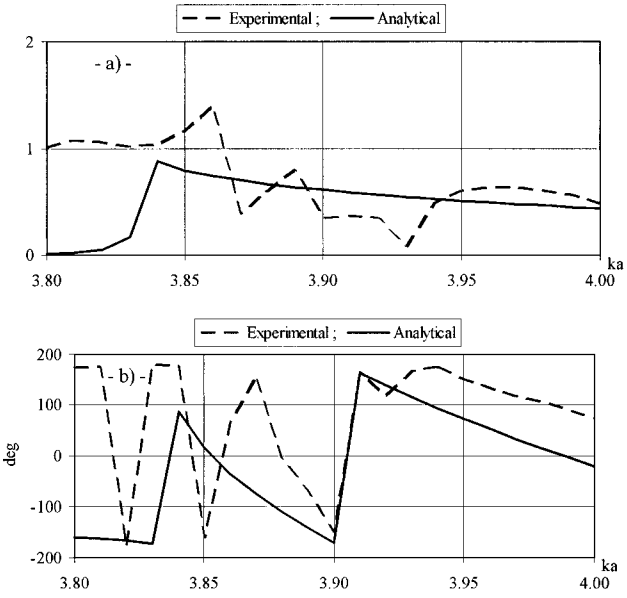


FIG. 8. Reflection coefficient spectrum $R_{01,01}$: (a) modulus; (b) phase.

2. Modes (±1,0) reflection coefficients: $R_{10,10}$ and $R_{10,-10}$

Complex coefficients $R_{10,10}$ and $R_{10,-10}$ are plotted (Fig. 6) from $ka=1.8$ to 4 and compared with theory. The agreement is remarkable except around the cuton frequencies of modes $(\pm 2,0)$, and (0,1).

3. Modes (±2,0) reflection coefficients: $R_{20,20}$ and $R_{20,-20}$

Complex coefficients $R_{20,20}$ and $R_{20,-20}$ are plotted (Fig. 7) from $ka=3$ to 4 and compared with theory. Good agreement is achieved away from the cuton frequency of mode (0,1).

4. Modes (0,1) reflection coefficients: $R_{01,01}$

The complex coefficient $R_{01,01}$ is plotted (Fig. 8) from $ka=3.8$ to 4 and compared with theory. This result is shown over a very short frequency domain around its cuton frequency. Agreement between experiment and theory is not so good as it is for modes with an axial symmetry described previously, because the frequency is higher than $f_{\max}$ [Eq. (10)].

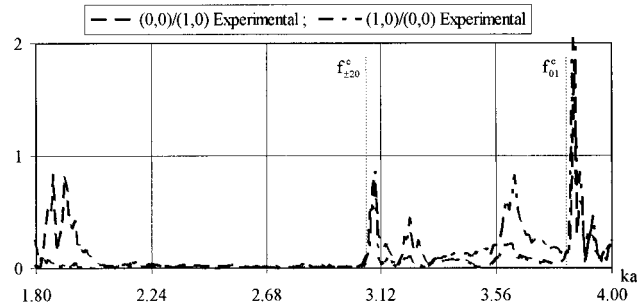


FIG. 9. Conversion coefficient spectra $R_{10,00}$ and $R_{00,10}$: modulus.

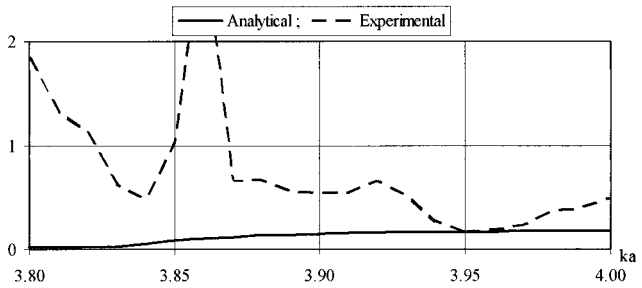


FIG. 10. Conversion coefficient spectra $R_{01,00}$: modulus.

5. Conversion coefficients

As the duct is axisymetric, only conversion between radial modes occurs. In the frequency domain studied, only $R_{00,10}$ and $R_{10,00}$ should then be different from 0. This assumption is experimentally verified on Fig. 9 where $R_{00,10}$ and $R_{10,00}$ are found to be close to zero except around the cut-on frequencies of modes $(\pm 2,0)$ and (0,1). The conversion coefficient $R_{01,00}$ is plotted (Fig. 10) and compared with theory. The agreement is not good with the experimental conversion coefficient being higher than predicted. The lack of agreement can be expected because the analysis is not very reliable since, due to the processing of the experimental data, the frequency range is higher than $f_{\max}$ and around the cut-on frequency of the mode.

III. CONCLUSIONS

A method of measuring pressure reflection coefficient matrices of duct discontinuities has been developed. It can be regarded as an extension of the Chung-Blaser[1] method from plane wave propagation to high order mode propagation and an improvement of the technique suggested by Abom.[4] Indeed, to perform the reflection matrix determination, transfer functions between microphone pairs traversing two duct cross sections are processed in an original way: in addition to a different other modal decomposition technique, a new method of creating linearly incident pressure fields is proposed in order to solve the system.

An experimental test facility has been set up. The experimental design allows the acoustic pressure reflection matrix measurement of an unflanged duct opening to be conducted up to a frequency parameter ka of 4. Experimental data of $R_{pq,mn}$ are presented and compared with theory. A good agreement is found between experiment and theory for reflection coefficients away from cut-on frequencies at which the method is diverging. Also, the conversion coefficients between azimuthal modes with the same radial number are found to be close to zero as expected because of the duct axisymmetry. Results obtained on the radial mode (0,1) are not very reliable because the study is conducted in its cut-on frequency range where the method diverges.

This method will now be used to study other geometrical discontinuities and lined ducts. We also intend to extend it to air flow environments. Improvement on the probe has to be done to avoid the experimental diffraction effects observed and to perform measurements with flow.

[1] J. Y. Chung and D. A. Blaser, ''Transfer function method for measuring in-duct acoustic properties. I. Theory, II. Experiment,'' J. Acoust. Soc. Am. **68**, 907–921 (1980).
[2] J. M. Ville, ''Definition and realization of a spinning mode simulator,'' ESA-TT-514 (1978).
[3] J. M. Ville and R. J. Silcox, ''Experimental investigation of radiation of sound from an unflanged duct and a bellmouth, including flow effect,'' NASA TP 1697 (1980).
[4] M. Abom, ''Modal decomposition in ducts based on transfer function measurements between microphone pairs,'' J. Sound Vib. **135**, 95–114 (1989).
[5] J. M. Auger and J. M. Ville, ''Measurement of liner impedance based on the determination of duct eigenvalues by Fourier–Lommel's transform,'' J. Acoust. Soc. Am. **88**, 19–22 (1990).
[6] M. Abom and H. Boden, ''Error analysis of two microphone measurements in ducts with flow,'' J. Acoust. Soc. Am. **83**, 2429–2438 (1988).
[7] J. M. Auger and J. M. Ville, ''Flow effects on the measurement of the modal decomposition of acoustic field in a hard wall cylindrical duct,'' in *Aero and Hydroacoustics IUTAM Symposium Lyon* (1985) (Springer-Verlag, Berlin, 1986), pp. 437–443.
[8] Z. Lazreq, M. Ben Tahar, J. M. Ville, and J. M. Auger, ''A mixed experimental and theoretical method for noise propagation and radiation prediction of a cowled axial fan,'' J. Acoust. Soc. Am. **100**, 166–177 (1996).
[9] C. de Boor, ''A practical guide to splines,'' in *Applied Mathematical Sciences* (Springer-Verlag, New York, 1987), Vol. 27.
[10] A. Roure, ''Propagation guidée. Etude des discontinuités,'' thesis, Université Aix Marseille II (1976).
[11] W. E. Zorumski, ''Generalized radiation impedances and reflection coefficients of circular and annular ducts,'' J. Acoust. Soc. Am. **54**, 1667–1673 (1973).
[12] Y. C. Cho, ''Sound radiation from a hyperboloïdal inlet duct,'' paper 79-0677, AIAA 5th Aeroacoustics Conference, Seattle, WA (1979).
[13] F. J. Fahy, *Sound Intensity* (Elsevier Science, New York, 1990).

Known source detection predictions for higher order correlators

Lisa A. Pflug
Naval Research Laboratory, Code 7173, Stennis Space Center, Mississippi 39529-5004

George E. Ioup and Juliette W. Ioup
Department of Physics, University of New Orleans, New Orleans, Louisiana 70148

(Received 15 January 1996; accepted for publication 21 January 1998)

The problem addressed in this paper is whether higher order correlation detectors can perform better in white noise than the cross correlation detector for the detection of a known transient source signal, if additional receiver information is included in the higher order correlations. While the cross correlation is the optimal linear detector for white noise, additional receiver information in the higher order correlations makes them nonlinear. In this paper, formulas that predict the performance of higher order correlation detectors of energy signals are derived for a known source signal. Given the first through fourth order signal moments and the noise variance, the formulas predict the SNR for which the detectors achieve a probability of detection of 0.5 for any level of false alarm, when noise at each receiver is independent and identically distributed. Results show that the performance of the cross correlation, bicorrelation, and tricorrelation detectors are proportional to the second, fourth, and sixth roots of the sampling interval, respectively, but do not depend on the observation time. Also, the SNR gains of the higher order correlation detectors relative to the cross correlation detector improve with decreasing probability of false alarm. The source signal may be repeated in higher order correlations, and gain formulas are derived for these cases as well. Computer simulations with several test signals are compared to the performance predictions of the formulas. The breakdown of the assumptions for signals with too few sample points is discussed, as are limitations on the design of signals for improved higher order gain. Results indicate that in white noise it is difficult for the higher order correlation detectors in a straightforward application to achieve better performance than the cross correlation. [S0001-4966(98)01805-0]

PACS numbers: 43.60.Gk, 43.60.Cg [JLK]

INTRODUCTION

The potential advantages over conventional detection methods that may be obtained using higher order moment and related spectral techniques have received much attention in recent years. Higher order techniques show promise in applications for stationary signals and also for short-time transients where only a single occurrence of a signal may be available for detection (Dwyer, 1984; Hinich, 1990; Hinich and Wilson, 1990; Kletter and Messer, 1990; Sangfelt and Persson, 1993; Delaney, 1994; Tague *et al.*, 1994; Baugh and Hardwicke, 1994; Nuttall, 1994). The latter case, for correlation detectors, has been investigated in previous papers by the authors using both computer simulations (Pflug *et al.*, 1992b, 1994b) and more recently for unknown source detection, using theoretical performance predictions for the case of uncorrelated noise (Pflug *et al.*, 1995b). The theoretical performance predictions are extended in this paper to include predictions for known source detection, of which active detection is the most common application. The formulas can be used to determine under what conditions higher order correlations perform better than the cross correlation detector in uncorrelated noise, if the higher order correlations include more than one hydrophone, or channel, of data. Although the cross correlation is the optimal linear detector in white noise, the inclusion of additional channels of data in the higher order correlations makes them nonlinear detectors, which makes possible improvement over the cross correlation. Preliminary results using these formulas have been included in two abstracts (Pflug *et al.*, 1994a; Ioup *et al.*, 1995) and a proceedings article (Pflug *et al.*, 1995a).

After a background discussion given in Sec. I, the formulas for known source detection are derived in Sec. II. Comparison of the formula predictions with simulations using a set of various test signals is given in Sec. III. Section IV presents a discussion of signal design limitations for higher order gain and the applicability of the prediction formulas. Repeating the source signal in higher order correlation detectors is addressed in Sec. V. Finally, a summary of the findings appears in Sec. VI.

I. DEFINITIONS AND ASSUMPTIONS

The detection criteria used in this paper are based on the second through fourth order moments of a deterministic transient signal $s(t)$, defined by

$$m_p^s = \Delta t \sum_{k=0}^{N_s-1} s^p(t), \tag{1}$$

in which p is the moment order, $t = k\Delta t$, and N_s is the number of points in the source signal. The second through fourth order moments correspond to central ordinate, or zero time lag, values for the cross correlation, bicorrelation, and tricorrelation, respectively. Although the definition of m_p^s includes the factor Δt, changing the sampling rate of a signal results

in a corresponding change in the summation of Eq. (1), leaving the value of m_p^s the same, if the signal is adequately sampled.

The underlying noise process is assumed to be zero-mean stationary, independent, and identically distributed (i.i.d.), and therefore higher order white and uncorrelated. Less restrictive assumptions concerning the noise can be found in Appendices A and B of Pflug *et al.* (1995b). These assumptions can be modified in a straightforward manner to substitute the known source signal for one of the noise sequences. No averaging is used in the detection process, since the transients are nonstationary over their domains. Thus the noise is treated as an energy signal in the process of detecting an energy transient and has moments defined as in Eq. (1).

It is assumed for all but Sec. V that the source signal is simultaneously recorded on each of p-1 spatially separated sensors for the pth order moment. When the source signal is present, the second, third, and fourth order correlation detectors, also called the cross correlation (CC), bicorrelation (BC), and tricorrelation (TC) detectors, have detection statistics

$$\mathrm{CC}=\sum_{k=0}^{N_s-1} s(t)[s(t)+n_1(t)]\Delta t, \tag{2a}$$

$$\mathrm{BC}=\sum_{k=0}^{N_s-1} s(t)[s(t)+n_1(t)][s(t)+n_2(t)]\Delta t, \tag{2b}$$

and

$$\mathrm{TC}=\sum_{k=0}^{N_s-1} s(t)[s(t)+n_1(t)][s(t)+n_2(t)] \times[s(t)+n_3(t)]\Delta t. \tag{2c}$$

In these equations $n_i(t)$ represents the noise at one of the p-1 spatially separated sensors. The first terms in the extended sums of Eqs. (2a), (2b), and (2c) are simply the moments of the signal, as given in Eq. (1). When the source signal is absent, the correlation detectors are given by

$$\mathrm{CC}=\sum_{k=0}^{N_s-1} s(t)n_1(t)\Delta t, \tag{3a}$$

$$\mathrm{BC}=\sum_{k=0}^{N_s-1} s(t)n_1(t)n_2(t)\Delta t, \tag{3b}$$

and

$$\mathrm{TC}=\sum_{k=0}^{N_s-1} s(t)n_1(t)n_2(t)n_3(t)\Delta t. \tag{3c}$$

The amplitude signal-to-noise ratio (SNR) for energy signals is defined by

$$\mathrm{SNR}=\frac{\sigma_s}{\sigma_n}. \tag{4}$$

Conversion of the amplitude SNR in Eq. (4) to power SNR in decibels is accomplished using $20\log_{10}(\sigma_s/\sigma_n)$. The variance of a deterministic signal $s(t)$ is

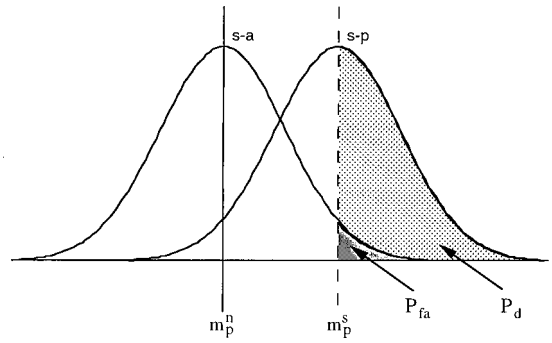


FIG. 1. The signal-absent (s-a) and signal-present (s-p) PDFs that determine the $P_d=0.5$ point of a ROC curve.

$$\sigma_s^2=\frac{\Delta t}{T}\sum_{k=0}^{N_s-1}[s(t)-\bar{s}]^2=\frac{1}{T}\left[m_2^s-\frac{(m_1^s)^2}{T}\right], \tag{5}$$

where the mean is $\bar{s}=m_1^s/T$. For known source detection, the processing window duration, $T=\Delta t N_s$, is by definition equal to the transient signal duration, denoted by T_s. The noise variance, σ_n^2, is defined in the same way as σ_s^2 because the correlations are defined for energy transients. The noise ensemble is ergodic and differences between the sample and population noise means and variances are assumed to be small for comparison of theoretical and simulated results.

II. KNOWN SOURCE DETECTION PREDICTION FORMULAS

Derivations of formulas that predict the SNR at which a passive correlation detector achieves the minimum detectable level (MDL) are given by Pflug *et al.* (1995b). The MDL is the SNR at which a detector achieves a probability of detection (P_d) equal to 0.5 for a selected probability of false alarm (P_{fa}). The derivations are based on the areas beneath the signal-absent (s-a) and signal-present (s-p) probability density functions (PDFs) of the zero lag correlation values that define a receiver operating characteristic (ROC) curve (see Fig. 1). For the pth order correlation, the mean of the s-p PDF is at m_p^s (the pth order signal correlation), and the mean of the s-a PDF is at m_p^n (the pth order noise correlation), which is zero under the assumptions in this paper. The s-a PDF moments are consistent with Gaussian moments, as shown in the Appendix.

Since the s-p PDF is symmetric, the mean is equal to the median of the PDF and the $P_d=0.5$ threshold occurs at the mean, m_p^s (see Fig. 1). For a fixed SNR, this threshold also defines the P_{fa}. The P_{fa}, or area to the right of the threshold, is related to a standardized score, called z_n, by

$$z_n=\frac{m_p^s-m_p^n}{\sqrt{\alpha_p^n}}, \tag{6}$$

where α_p^n represents the variance of the s-a PDF. The standardized score is the abscissa value which defines the tail area for a normalized Gaussian distributed PDF having zero-mean and unit variance. This makes it possible to use one formula or table of areas to find the P_{fa} for any Gaussian distribution, regardless of its defining parameters. A table of these areas can be found in most basic statistics books or

computed numerically. Here, z_n and P_{fa} are inversely related. When α_p^n in Eq. (6) and the corresponding SNR are smaller, z_n is larger, and the two PDFs in Fig. 1 are narrower and have less overlap, resulting in a smaller P_{fa}. For more details on the relationship between z scores and ROC curves, see Egan (1975).

For $P_d=0.5$, Eq. (6) is related to the remaining ROC curve parameters through α_p^n, SNR, and P_{fa}. Since m_p^s and m_p^n are known, if α_p^n can be expressed theoretically as a function of signal and noise parameters, then Eq. (6) can ultimately be used as a prediction formula for detection performance. The derivations of α_p^n for $p=2$, 3, and 4 follow.

The PDF variance, α_2^n, of cross correlation realizations of a finite-length transient signal, $s(t)$, and members of an infinite ensemble of finite-length noise sequences, $n_i(t)$, is given by

$$\alpha_2^n=E\left\{\left[\sum_{k=0}^{N_s-1} s(t)n_{i_1}(t)\Delta t\right]^2\right\} - E^2\left\{\sum_{k=0}^{N_s-1} s(t)n_{i_1}(t)\Delta t\right\}. \tag{7}$$

The expectation operator represents averaging the large number of realizations generated in performing Monte Carlo simulations over the noise ensemble to evaluate detector performance, and is not related to any averaging in the detector itself. Since the signal and noise are approximately uncorrelated across the finite-time interval of interest and the noise is assumed to be zero mean, the second term can be neglected. Then,

$$\alpha_2^n=E\left\{(\Delta t)^2\sum_{k_1=0}^{N_s-1}\sum_{k_2=0}^{N_s-1} s(t_1)s(t_2)n_{i_1}(t_1)n_{i_1}(t_2)\right\} \tag{8a}$$

$$=E\left\{(\Delta t)^2\sum_{k_1=0}^{N_s-1} s^2(t_1)n_{i_1}^2(t_1)\right\} \tag{8b}$$

$$=(\Delta t)m_2^sE\{n_{i_1}^2(t_1)\} \tag{8c}$$

$$=\Delta t m_2^s\sigma_n^2. \tag{8d}$$

The summation over k_2 in the derivation of Eq. (8b) from Eq. (8a) is a result of the assumption that the noise is i.i.d. $E\{n^2(t)\}$ represents the infinite ensemble variance of the noise, which is equivalent to σ_n^2 in Eq. (8d) for an infinite time sum due to the ergodicity of the noise. This derivation also follows from the result given in the Appendix and is based on the assumption that finite-time averages are equal to infinite-time averages, which is approximately true for large enough N_s (see the discussion in Sec. IV for practical restrictions on the signal). For the bicorrelation, noise received from two sensors is used in the s-a PDF variance.

$$\alpha_3^n=E\left\{\left[\sum_{k=0}^{N_s-1} s(t)n_{i_1}(t)n_{i_2}(t)\Delta t\right]^2\right\} - E^2\left\{\sum_{k=0}^{N_s-1} s(t)n_{i_1}(t)n_{i_2}(t)\Delta t\right\} \tag{9a}$$

$$=E\left\{(\Delta t)^2\sum_{k_1=0}^{N_s-1}\sum_{k_2=0}^{N_s-1} s(t_1)s(t_2)n_{i_1}(t_1) \times n_{i_1}(t_2)n_{i_2}(t_1)n_{i_2}(t_2)\right\} \tag{9b}$$

$$=\Delta t m_2^s\sigma_n^4. \tag{9c}$$

Again, N_s must be large. For the tricorrelation, noise received from three sensors is used in the s-a PDF variance.

$$\alpha_4^n=E\left\{\left[\sum_{t=0}^{N_s-1} s(t)n_{i_1}(t)n_{i_2}(t)n_{i_3}(t)\Delta t\right]^2\right\} - E^2\left\{\sum_{t=0}^{N_s-1} s(t)n_{i_1}(t)n_{i_2}(t)n_{i_3}(t)\Delta t\right\} \tag{10a}$$

$$=\Delta t m_2^s\sigma_n^6. \tag{10b}$$

These s-a PDF variances can be substituted in Eq. (6). As mentioned previously, the cross correlation s-p PDF has mean equal to m_2^s, and the expression for z_n becomes

$$z_n=\frac{m_2^s-m_2^n}{\sqrt{\alpha_2^n}}=\frac{\sqrt{m_2^s}}{\sigma_n\sqrt{\Delta t}}. \tag{11}$$

Using Eq. (4) gives

$$\mathrm{SNR}_{\mathrm{CC}}=\frac{\sigma_s z_n\sqrt{\Delta t}}{\sqrt{m_2^s}}. \tag{12}$$

Here, $\mathrm{SNR}_{\mathrm{CC}}$, the MDL for the cross correlation detector, is interpreted as the SNR required to achieve the detection level of $P_d=0.5$ at the P_{fa} corresponding to z_n, the standardized threshold value. Since σ_s is proportional to $\sqrt{m_2^s}$, m_2^s dependence is canceled in the derivation, and $\mathrm{SNR}_{\mathrm{CC}}$ is constant for all signals. This is not to be interpreted as saying that all signals of equal duration and sampling are equally detectable using the cross correlation. The correct interpretation is that the cross correlation detector achieves $P_d=0.5$ for these signals at the same SNR, which requires varying levels of noise for different signals.

The bicorrelation s-p PDF has mean equal to m_3^s, which for simplicity is assumed to be positive. Evaluating the z_n gives

$$z_n=\frac{m_3^s-m_3^n}{\sqrt{\alpha_2^n}}=\frac{m_3^s}{\sigma_n^2\sqrt{\Delta t m_2^s}}. \tag{13}$$

Substituting to obtain the SNR yields the formula

$$\mathrm{SNR}_{\mathrm{BC}}=\sigma_s\sqrt{\frac{z_n^{\mathrm{BC}}\sqrt{\Delta t m_2^s}}{m_3^s}}. \tag{14}$$

A similar analysis for the tricorrelation yields

$$z_n = \frac{m_4^s - m_4^n}{\sqrt{\alpha_2^n}} = \frac{m_4^s}{\sigma_n^3 \sqrt{\Delta t m_2^s}}, \tag{15}$$

and

$$\mathrm{SNR}_{\mathrm{TC}} = \sigma_s \sqrt[3]{\frac{z_n^{\mathrm{TC}} \sqrt{\Delta t m_2^s}}{m_4^s}}. \tag{16}$$

For zero-mean energy signals, the SNR prediction formulas reduce to

$$\mathrm{SNR}_{\mathrm{CC}} = \frac{z_n \sqrt{\Delta t}}{\sqrt{T_s}}, \tag{17}$$

$$\mathrm{SNR}_{\mathrm{BC}} = \sqrt{\frac{(m_2^s)^{3/2} z_n \sqrt{\Delta t}}{m_3^s T_s}}, \tag{18}$$

and

$$\mathrm{SNR}_{\mathrm{TC}} = \sqrt[3]{\frac{(m_2^s)^2 z_n \sqrt{\Delta t}}{m_4^s (T_s)^{3/2}}}. \tag{19}$$

For zero-mean signals, $\mathrm{SNR}_{\mathrm{BC}}$ and $\mathrm{SNR}_{\mathrm{TC}}$ can also be written as functions of signal skewness and kurtosis (Pflug *et al.*, 1995a).

To compare the three detection methods, $\mathrm{SNR}_{\mathrm{CC}}$, $\mathrm{SNR}_{\mathrm{BC}}$, and $\mathrm{SNR}_{\mathrm{TC}}$ should be evaluated at the same P_{fa}. The difference in dB levels is referred to as the SNR gain of the bicorrelation or tricorrelation detector over the cross correlation detector. Using one source signal and p-1 distinct channels of received data, the bicorrelation SNR gain in dB is

$$\mathrm{BCG} = 20 \log_{10} \left[\frac{(m_3^s)^2 z_n^2 \Delta t}{(m_2^s)^3} \right]^{1/4}, \tag{20}$$

and the tricorrelation SNR gain in dB is

$$\mathrm{TCG} = 20 \log_{10} \left[\frac{(m_4^s)^2 z_n^4 (\Delta t)^2}{(m_2^s)^4} \right]^{1/6}. \tag{21}$$

The SNR gain formulas for zero-mean and nonzero-mean signals are the same, since the only dependence on the signal mean in the nonzero mean formulas is through the standard deviation, and the dependence of $\mathrm{SNR}_{\mathrm{CC}}$, $\mathrm{SNR}_{\mathrm{BC}}$, and $\mathrm{SNR}_{\mathrm{TC}}$ on the standard deviation cancels in the gain ratios.

The gain formulas show no functional dependence on T_s. Known source detection, in contrast to unknown source detection (see Pflug *et al.*, 1995a, b) is independent of observation time, T, provided the entire signal is in the processing window. This is because the multiplication of the noise by the signal sets the noise to zero outside the signal window. Known source detection performance does, however, depend on Δt. In particular, for known source detection, $\mathrm{SNR}_{\mathrm{CC}}$, $\mathrm{SNR}_{\mathrm{BC}}$, and $\mathrm{SNR}_{\mathrm{TC}}$ are proportional to the second, fourth, and sixth root of the sampling interval, respectively. Performance improves for all the correlation detectors with decreasing sampling interval. However, as sampling intervals increase, the bicorrelation and tricorrelation detectors degrade less quickly than the cross correlation detector. For a given signal, noise level, and P_{fa}, each of the three detectors is associated with a range of sampling intervals for which it will perform best. For example, if a zero-mean signal has moments $m_2^s = 0.0025$ s, $m_3^s = 0.0022$ s, and $m_4^s = 0.0015$ s, then the bicorrelation and tricorrelation SNR gains versus sampling interval will be as shown in Fig. 2 for $P_{\mathrm{fa}} = 0.001$ ($z_n = 3.09$). For a positive bicorrelation gain, the sampling interval must be 0.000 34 s or larger, and for a positive tricorrelation gain, the sampling interval must be 0.000 44 s or larger.

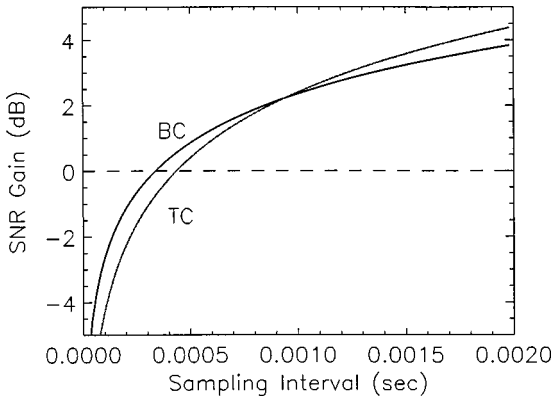


FIG. 2. Bicorrelation and tricorrelation SNR gain versus sampling interval for a zero-mean signal with moments $m_2^s = 0.0025$, $m_3^s = 0.0022$, $m_4^s = 0.0015$, and $P_{\mathrm{fa}} = 0.001$.

Detection performance improves (lower SNR) for the three detectors as z_n decreases, or the tolerance for false alarm increases. However, if the tolerance for false alarm is small, as it often is in practice, the bicorrelation and tricorrelation SNR formulas, which give increasing gain with decreasing P_{fa}, indicate that higher order detectors could outperform the cross correlation.

TABLE I. Moments m_p^s in amplitude units to the pth power times seconds for the eight test signals.

Signal	T_s (s)	m_1^s	m_2^s	m_3^s	m_4^s
Whale transient	1.00	-9.13×10^{-5}	2.78×10^{-1}	1.07×10^{-3}	1.75×10^{-1}
Low frequency whale	1.00	-1.42×10^{-2}	2.80×10^{-1}	2.31×10^{-2}	1.78×10^{-1}
5 Hz sinusoid	0.75	7.51×10^{-3}	7.31×10^{-2}	2.93×10^{-2}	4.05×10^{-2}
Narrow pulse	0.03	3.12×10^{-3}	2.68×10^{-3}	2.20×10^{-3}	1.93×10^{-3}
50 Hz sinusoid	1.72	-5.78×10^{-6}	1.77×10^{-1}	7.25×10^{-8}	9.38×10^{-2}
49–51 Hz sinusoid	1.21	3.25×10^{-3}	4.19×10^{-2}	2.25×10^{-3}	1.83×10^{-2}
FM linear sweep	1.72	7.42×10^{-6}	1.77×10^{-1}	-3.05×10^{-8}	9.38×10^{-2}
Nonlinear FM sweep	1.72	1.11×10^{-4}	1.77×10^{-1}	2.67×10^{-8}	9.39×10^{-2}

TABLE II. Known source detection SNR_{CC} in dB for $P_{fa}=0.001$ calculated from computer simulations and the prediction formulas, and the absolute dB difference between the two methods.

Signal	Simulated SNR_{CC}	Predicted SNR_{CC}	Absolute difference
Whale transient	−19.49 dB	−20.30 dB	0.81 dB
Low frequency whale	−19.47	−20.30	0.83
5 Hz sinusoid	−16.77	−17.29	0.52
Narrow pulse	−5.64	−5.67	0.03
50 Hz sinusoid	−17.96	−18.64	0.68
49–51 Hz sinusoid	−17.89	−17.99	0.10
FM linear sweep	−18.39	−18.64	0.25
Nonlinear FM sweep	−18.27	−18.63	0.36

TABLE IV. Known source detection SNR_{TC} in dB for $P_{fa}=0.001$ calculated from computer simulations and the prediction formulas, and the absolute dB difference between the two methods.

Signal	Simulated SNR_{TC}	Predicted SNR_{TC}	Absolute difference
Whale transient	−8.64 dB	−9.14 dB	0.50 dB
Low frequency whale	−8.81	−9.14	0.33
5 Hz sinusoid	−9.36	−9.62	0.26
Narrow pulse	−6.20	−8.31	2.11
50 Hz sinusoid	−9.98	−10.36	0.38
49–51 Hz sinusoid	−12.80	−13.31	0.51
FM linear sweep	−10.10	−10.36	0.26
Nonlinear FM sweep	−10.01	−10.35	0.34

III. COMPARISONS BETWEEN PREDICTIONS AND COMPUTER SIMULATIONS

Pflug *et al.* (1994b) present detection results using hypothesis testing and Monte Carlo simulations with 10 000 Gaussian noise realizations for eight energy signals. Note that the formulas do not require the noise to be Gaussian distributed, but they require the noise ensemble to be i.i.d. The signals are varied in nature, generally of low frequency, and they are described in detail in the above reference. Known source detection simulations are performed with the resulting P_d vs SNR curves interpolated at $P_d=0.5$. These simulations involve only one replica of the source signal in the bicorrelation and tricorrelation. While the focus of the paper by Pflug *et al.* (1994b) is on the advantage of prefiltering in higher order correlations, results for detection simulations with no prefiltering are also given. These results are used to corroborate the known source prediction formulas derived in this paper, which assume there is no signal-specific or situation-specific passband filtering.

Table I lists the signal duration and first through fourth order signal moments used in detection performance prediction for the eight test signals. The test signals include a model 20 Hz finback whale transient, the whale transient shifted to 12 Hz and called the low frequency whale, an amplitude modulated 5 Hz sinusoid, a narrow-time pulse which has a flat magnitude spectrum to approximately 80 Hz and a smooth rolloff to 256 Hz, an amplitude modulated 50 Hz sinusoid, a 49–51 Hz beating sinusoid, an FM linear sweep from 49 to 51 Hz, and a nonlinear FM sweep from 120 to 0 Hz. For the computer simulations, the first four signals are sampled with interval (1/1024) s, the last four with (1/500) s. The low frequency whale transient, 41–59 Hz FM linear sweep, and 120–0 Hz nonlinear FM sweep do not have bicorrelation maximum magnitudes at zero time lag. Also, the whale transient and 50 Hz sinusoid have approximately zero bispectra since the lowest frequency present in the signal is essentially half the highest frequency. This makes the bicorrelation value at zero time-lag inappropriate as a detection criterion (Ioup *et al.*, 1989; Pflug, 1990; Pflug *et al.*, 1992a).

Tables II–IV contain the results of the computer simulations and the theoretical predictions for known source detection at $P_{fa}=0.001$. Excepting the narrow pulse signal, the predictions and simulations agree reasonably well, with the absolute differences between the computer simulations and formula predictions of SNR ranging from 0.10 dB to 0.83 dB for the three detectors. For these seven test signals, the cross correlation detector performs better than the higher order correlation detectors, as shown by the predicted BCG and TCG values in Table V. For the narrow pulse, both the bicorrelation and tricorrelation detectors perform better than the cross correlation. However, the gain achieved in the simulations is significantly lower than predicted with the formulas. The reasons are discussed in detail in the following section.

The difference between the computer simulations and the predictions, which generally fluctuates between 0 and 0.5 dB for the test signals except the narrow pulse, is not removed by finer sampling, which implies that for those signals there is a sufficient number of signal samples for the finite summation to approximate the infinite summation in the formula derivations. For example, in Fig. 3, the computer

TABLE III. Known source detection SNR_{BC} in dB for $P_{fa}=0.001$ calculated from computer simulations and the prediction formulas, and the absolute dB difference between the two methods.

Signal	Simulated SNR_{BC}	Predicted SNR_{BC}	Absolute difference
Whale transient	···	···	···
Low frequency whale	···	···	···
5 Hz sinusoid	−8.41 dB	−8.84 dB	0.43 dB
Narrow pulse	−6.08	−7.51	1.43
50 Hz sinusoid	···	···	···
49–51 Hz sinusoid	−3.34	−3.59	0.25
FM linear sweep	···	···	···
Nonlinear FM sweep	···	···	···

TABLE V. Predicted bicorrelation and tricorrelation SNR gains over the cross correlation detector.

Signal	BCG (dB)	TCG (dB)
Whale transient	···	−11.16
Low frequency whale	···	−11.16
5 Hz sinusoid	−8.45	−7.67
Narrow pulse	1.84	2.64
50 Hz sinusoid	···	−8.28
49–51 Hz sinusoid	−14.40	−4.68
FM linear sweep	···	−8.28
Nonlinear FM sweep	···	−8.28

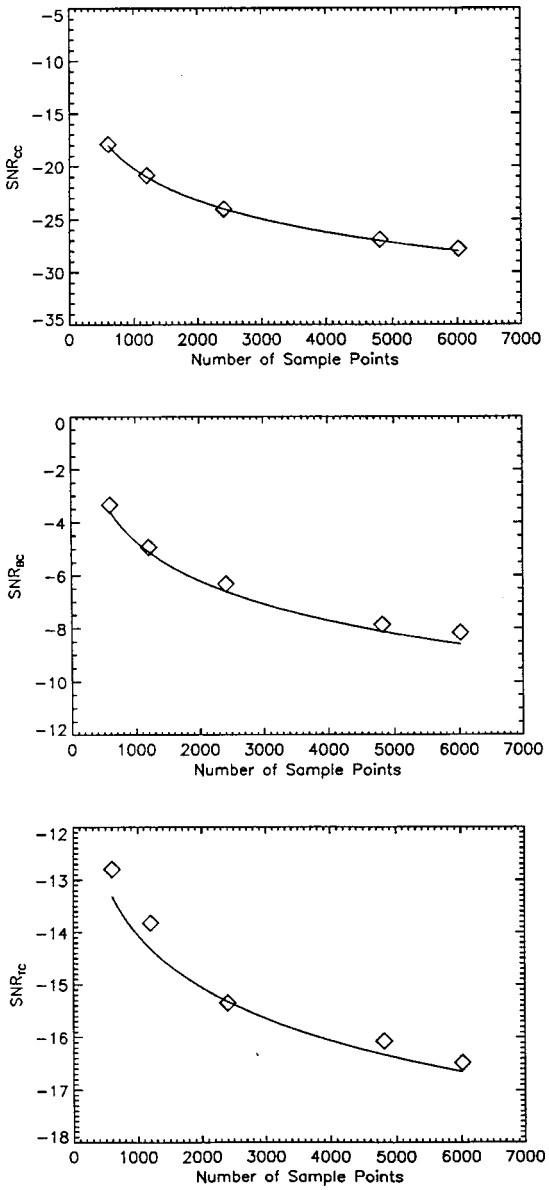


FIG. 3. The minimum detectable level as a function of the number of sample points for the 49–51 Hz sinusoid. The solid curves give the formula predictions and the symbols give the computer calculations.

simulated and predicted SNRs are shown versus number of sample points for the 49–51 Hz sinusoid. The differences fluctuate randomly with increased sampling, but remain less than 0.5 dB for each detector.

Another possible source of error in the computer simulations is that Gaussian noise generators may not perform well when very low levels of noise, corresponding to the tails of the noise distribution, are needed. A potential limitation at low P_{fa} occurs due to the finite number of realizations used in a computer simulation. If 10 000 realizations are used, as in this paper, the minimum nonzero P_{fa} value possible is 1×10^{-4}, or one false alarm out of 10 000 on average. Thus interpolation of the P_d vs P_{fa} curve for a fixed SNR at P_{fa} values lower than 1×10^{-3} is generally not statistically reliable. The number of realizations required to obtain a discrete point on the ROC curve at the desired P_{fa} increases with decreasing P_{fa}. For example, to obtain a P_d vs P_{fa} point at a P_{fa} value of 1×10^{-6}, at least one million realizations would be required.

IV. LIMITATIONS ON THE DESIGN OF SIGNALS FOR POSITIVE HIGHER ORDER GAIN

The physically motivated definition of the signal moments, which includes the time dimension, Δt, in the summation given in Eq. (1), has the intrinsic advantage that the signal moments do not change if the signal is interpolated or decimated in time. Of course, this will only be true if there is sufficient sampling (Pflug *et al.*, 1993), and the interpolation is bandlimited (Bracewell, 1986). The above holds whether or not the moment is normalized by dividing by T_s, and also for the variance, skewness, and kurtosis.

Since the moments do not change with interpolation or decimation, subject to the given restrictions, the entire effect of either of these two operations on the higher order detection gain is contained in the Δt factor of Eqs. (20) and (21). Larger Δt increases the higher order advantage, while smaller Δt does the opposite, favoring the cross correlation detector (matched filter). In fact, in the limit of the continuous-time signal (Δt goes to zero), the gain tends to negative infinity, and the cross correlation is best for any signal if the signal and noise satisfy the necessary assumptions. For a given signal, it is straightforward to determine whether there is any Δt large enough, within the sampling limitations, to give positive higher order gain. It should be noted that experimental conditions may also put limitations on Δt, and these are not considered here.

Another design question which may be asked of the gain formulas is how varying the width of a signal affects the gain. That is, if the signal samples are moved closer together or further apart (Δt is decreased or increased) without changing the signal ordinate values (but changing the moments), how does the gain change? The functions are related by similarity (Bracewell, 1986). As Δt is varied, an upper limit on Δt is determined by the sampling requirement (Pflug *et al.*, 1993). The Similarity Theorem (Bracewell, 1986) describes the effect of varying Δt on the Fourier transform. To analyze the effect of a width change, consider the tricorrelation gain formula

$$\mathrm{TCG}=20\log_{10}\left[\frac{(m_4^s)^2 z_n^4(\Delta t)^2}{(m_2^s)^4}\right]^{1/6}$$
$$=20\log_{10}\left[\frac{z_n^{2/3}(\Sigma_{k=0}^{N-1}s^4(t))^{1/3}}{(\Sigma_{k=0}^{N-1}s^2(t))^{2/3}}\right]. \qquad (22)$$

All the Δt dependence cancels, leading to the conclusion, which agrees with intuition, that as long as the ordinate values of $s(t)$ are unchanged, the gain does not change with a similarity transformation. The same Δt cancellation occurs for the bicorrelation gain.

For many signals, there is a restriction on the use of these formulas for large Δt. First, $N_S=T_S/\Delta t$ must be large enough to ensure that the summation over the product of signal and noise is uncorrelated and that the finite sums approximate the infinite sums. Second, and more specifically, there must be a sufficient number of significant nonzero sig-

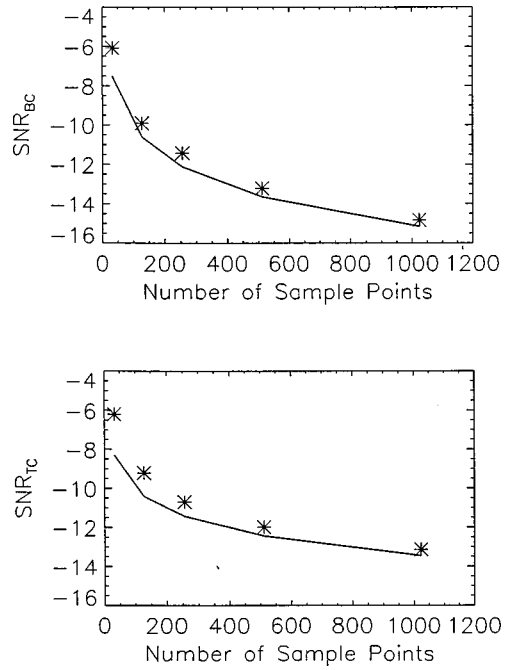


FIG. 4. Formula predictions (continuous curves) and computer simulations (asterisks) of SNR_{BC} and SNR_{TC} for the narrow pulse versus number of sample points within the original signal duration.

nal values, since these determine how many terms contribute significantly to the sums of Eqs. (7), (9a), and (10a). The gain prediction formulas may therefore fail to apply even for a Δt which is small enough to satisfy the sampling requirements.

Consider the case of the narrow pulse. The differences between the predicted SNR and the computer simulated SNR for SNR_{CC}, SNR_{BC}, and SNR_{TC} are 0.03, 1.43, and 2.11 dB, respectively. Differences between the formula predictions and computer simulations grow larger with increasing detector order, and the differences for SNR_{BC} and SNR_{TC} are larger than the differences for the other seven signals in Tables II–IV. Investigation into the source of the difference for the narrow pulse reveals that the bicorrelation and tricorrelation s-a PDFs are noticeably nonGaussian, which violates the assumptions required for the prediction formula derivation. At its original sampling, there are only 32 nonzero signal points in the narrow pulse, causing the breakdown of the assumption that finite-time averages should approximate infinite-time averages, and resulting in a non-Gaussian s-a PDF. Additional computer simulations support this hypothesis. The original 32-point signal is interpolated to contain 128, 256, 512, and 1024 points within in the original 0.03-s signal duration, and the computer simulated values for SNR_{BC} and SNR_{TC} are calculated. The predicted and simulated values become closer as the number of signal samples increases (Fig. 4) and the s-a PDFs become increasingly Gaussian. With 1024 points, the difference between the predicted and simulated SNR values decreases to 0.32 dB for both the bicorrelation and tricorrelation detectors.

For a realistic P_{fa}, the narrow pulse signal does not show a positive SNR gain for the bicorrelation or tricorrelation detectors over the ordinary correlation detector, if Δt is chosen small enough for the prediction formulas to apply. With 128 points within the signal duration, fewer than necessary for the prediction formulas to apply, both the predictions and simulations show that the cross correlation detector performs best. While it is possible to select a large enough Δt to get a positive gain for this signal in simulations, it is simply not possible to predict that gain well with the gain formulas.

V. REPEATED SOURCE REPLICAS IN HIGHER ORDER CORRELATION DETECTION

Repeating the source signal when using bicorrelation and tricorrelation detectors can have beneficial or detrimental effects, depending on the source signal (Pflug *et al.*, 1992b). Derivation of bicorrelation and tricorrelation detector prediction formulas for repeated sources is analogous to the derivations for a single source. If the source is repeated once in the tricorrelation so that signals from only two sensors are used, then the s-p and s-a tricorrelations are

$$\text{s-p:}\quad \text{TC}=\sum_{k=0}^{N_s-1} s^2(t)[s(t)+n_1(t)][s(t)+n_2(t)]\Delta t, \tag{23a}$$

$$\text{s-a:}\quad \text{TC}=\sum_{k=0}^{N_s-1} s^2(t)n_1(t)n_2(t)\Delta t, \tag{23b}$$

and α_4^n, as given by Eq. (A5) in the Appendix, is $\alpha_4^n = \Delta t m_4^s \sigma_n^4$, and the predicted SNR is

$$\text{SNR}_{\text{TC}}^{(2,2)} = \sigma_s \sqrt{\frac{z_n^{\text{TC}}\sqrt{\Delta t}}{\sqrt{m_4^s}}}, \tag{24}$$

where (2,2) denotes the power on the known source in the correlation and the number of channels used in the correlation [see Eq. (23)]. With this notation, all previous formulas would have superscripts of $(1,p-1)$ for the moment of order p. The corresponding SNR gain over the cross correlation detector is

$$\text{TCG}^{(2,2)} = 20\log_{10}\left[\frac{z_n^2 m_4^s \Delta t}{(m_2^s)^2}\right]^{1/4}. \tag{25}$$

This formula shows that for fixed Δt and P_{fa}, the gain depends explicitly on only the signal moments and that the simple calculation of these moments predicts whether repeating the source is advantageous or not.

Improved detection may result from repeating the source signal in the bicorrelation detector. However, at zero lag this higher order repeated source case, with noise from only one sensor, is equivalent to a correlation of a signal filter with one received signal. When only one received signal is used in the detection process, the cross correlation or matched filter is expected to be optimal for detection in white noise (Gardner, 1986). For the bicorrelation with a repeated source, the s-p and s-a bicorrelations are

$$\text{s-p:}\quad \text{BC}=\sum_{k=0}^{N_s-1} s^2(t)[s(t)+n_1(t)]\Delta t, \tag{26a}$$

$$\text{s-a:}\quad \text{BC}=\sum_{k=0}^{N_s-1} s^2(t)n_1(t)\Delta t, \tag{26b}$$

and $\alpha_3^n=\Delta t m_4^s \sigma_n^2$. This leads to

$$\text{SNR}_{\text{BC}}^{(2,1)}=\frac{z_n^{\text{BC}}\sigma_s\sqrt{m_4^s\Delta t}}{m_3^s}, \tag{27}$$

and the corresponding gain over the cross correlation detector which is

$$\text{BCG}^{(2,1)}=20\log_{10}\left[\frac{(m_3^s)^2}{m_2^s m_4^s}\right]^{1/2}. \tag{28}$$

If the source signal is repeated twice in the tricorrelation, then the s-a and s-p tricorrelations are

$$\text{s-p:}\quad \text{TC}=\sum_{k=0}^{N_s-1} s^3(t)[s(t)+n_1(t)]\Delta t, \tag{29a}$$

$$\text{s-a:}\quad \text{TC}=\sum_{k=0}^{N_s-1} s^3(t)n_1(t)\Delta t, \tag{29b}$$

and $\alpha_4^n=\Delta t m_6^s \sigma_n^2$. The predicted SNR is

$$\text{SNR}_{\text{TC}}^{(3,1)}=\frac{z_n^{\text{TC}}\sigma_s\sqrt{m_6^s\Delta t}}{m_4^s}, \tag{30}$$

and the SNR gain for this case is

$$\text{TCG}^{(3,1)}=20\log_{10}\left[\frac{(m_4^s)^2}{m_2^s m_6^s}\right]^{1/2}. \tag{31}$$

The SNR gains for the $\text{SNR}_{\text{BC}}^{(2,1)}$ and $\text{SNR}_{\text{TC}}^{(3,1)}$ are independent of P_{fa} and sampling interval, while the SNR gain for $\text{SNR}_{\text{TC}}^{(2,2)}$ is not. This means that $\text{SNR}_{\text{BC}}^{(2,1)}$ and $\text{SNR}_{\text{TC}}^{(3,1)}$ have the same dependence on the sampling interval and P_{fa} as SNR_{CC}, which is consistent with the fact that these higher order detectors are ordinary filters.

For the seven test signals other than the narrow pulse, the tricorrelation detector with the source repeated once shows a higher SNR gain than the tricorrelation detector without a repeated source for $P_{\text{fa}}=0.001$. Repeating the source twice improves the SNR gain even more. Repeating the source in the bicorrelation detector for these signals results in improvement for some signals, but not others, but the cross correlation still performs best. The predicted bicorrelation gain for the narrow pulse is 1.83 dB without a repeated source, but drops to -0.29 dB with a repeated source. The tricorrelation gain for the narrow pulse is 2.64 dB without a repeated source, drops to 1.98 dB with the source is repeated once, and is -0.60 dB when the source is repeated twice.

VI. CONCLUSIONS

Formulas that predict cross correlation, bicorrelation, and tricorrelation known source detector performance at the minimum detectable level for any probability of false alarm are derived and corroborated with computer simulations. The prediction formulas indicate that there are circumstances under which the higher order correlation detectors could outperform the cross correlation detector for the known source case, and give the relationships to the variables upon which the detection performance depends. They show that the cross correlation, bicorrelation, and tricorrelation detectors are proportional to the second, fourth, and sixth roots of the sampling interval. The SNR gain formulas indicate that the higher order detectors are more likely to exhibit better detection performance than the cross correlation detector when the tolerance for false alarm is low. A simulation result is given in which there is positive SNR gain for the higher order correlations, but the prediction formula assumptions are not satisfied for this case. While it is theoretically possible for the higher order correlation detectors to perform better than the cross correlation detector, only a limited class of signals exists for which they do.

ACKNOWLEDGMENTS

The authors wish to acknowledge funding from NRL Grant No. N00014-89-J-6002 and ONR Grant No. N00014-95-1-0648, and thank Dr. Ananthram Swami of the U.S. Army Research Laboratory for helpful discussions.

APPENDIX: GAUSSIAN CHARACTER OF CORRELATION CENTRAL ORDINATE PDFS

Following Isserlis (1918) and Gardner (1986), the qth ensemble moments of the pth order correlation central ordinate PDFs are shown to be consistent with a zero-mean Gaussian density where the correlation consists of $(p\text{-}r)$ zero-mean noise sequences correlated with a known signal raised to the rth power. That is, all odd order ensemble moments are zero, and all even order ensemble moments greater than two are appropriately proportional to powers of the second moment, if the assumptions discussed in Sec. I are satisfied.

The qth order ensemble moment of the source signal and $p-r$ sequences drawn from an infinite noise ensemble is

$$\begin{aligned}
M_p^q &= E\left\{\left[\sum_{k=0}^{N_s-1} s^r(t)n_{i_1}(t)n_{i_2}(t)\cdots n_{i_{p-r}}(t)\Delta t\right]^q\right\} \\
&= (\Delta t)^q E\left\{\left[\sum_{k_1=0}^{N_s-1} s^r(t_1)n_{i_1}(t_1)n_{i_2}(t_1)\cdots n_{i_{p-r}}(t_1)\right]\right. \\
&\quad\times\left[\sum_{k_2=0}^{N_s-1} s^r(t_2)n_{i_1}(t_2)n_{i_2}(t_2)\cdots n_{i_{p-r}}(t_2)\right]\cdots \\
&\quad\left.\times\left[\sum_{k_q=0}^{N_s-1} s^r(t_q)n_{i_1}(t_q)n_{i_2}(t_q)\cdots n_{i_{p-r}}(t_q)\right]\right\} \qquad \text{(A1)} \\
&= (\Delta t)^q E\left\{\sum_{k_1=0}^{N_s-1}\sum_{k_2=0}^{N_s-1}\cdots\sum_{k_q=0}^{N_s-1} s^r(t_1)s^r(t_2)\cdots s^r(t_q)\right. \\
&\quad\times n_{i_1}(t_1)n_{i_1}(t_2)\cdots n_{i_1}(t_q)\cdot n_{i_2}(t_1) \\
&\quad\left.\times n_{i_2}(t_2)\cdots n_{i_2}(t_q)\cdots n_{i_{p-r}}(t_1)n_{i_{p-r}}(t_2)\cdots n_{i_{p-r}}(t_q)\right\}, \qquad \text{(A2)}
\end{aligned}$$

with the number of signal points N_s and $t_j=k_j\Delta t$. This expression is zero when q is odd, and whenever all t_j are dis-

tinct. It is only nonzero when q is even and times are equal in pairs. Using delta function notation, the moments are nonzero when

$$M_p^q=(\Delta t)^q E\Bigg\{\sum_{k_1=0}^{N_s-1}\sum_{k_2=0}^{N_s-1}\cdots\sum_{k_q=0}^{N_s-1} s^r(t_1)s^r(t_2)\cdots s^r(t_q)$$
$$\times n_{i_1}(t_1)n_{i_1}(t_2)\cdots n_{i_1}(t_q)n_{i_2}(t_1)n_{i_2}(t_2)\cdots n_{i_2}(t_q)\cdots$$
$$\times n_{i_{p-r}}(t_1)n_{i_{p-r}}(t_2)\cdots n_{i_{p-r}}(t_q)$$
$$\times\Bigg[\sum\ \delta(t_{j_1}-t_{j_2})\cdots\delta(t_{j_{q-1}}-t_{j_q})\Bigg]\Bigg\} \tag{A3}$$

is nonzero, where the summation over the product of delta functions is taken over all possible ways of dividing q integers into $q/2$ combinations of pairs. There are $(1)(3)(5)\cdots(q-3)(q-1)$ terms in the summation. Applying the delta summation, exchanging summations and expectations where appropriate, and using the uncorrelatedness assumptions results in

$$M_p^q=[(1)(3)(5)\cdots(q-3)(q-1)]$$
$$\times(\Delta t)^{q/2}(m_{2r}^s)^{q/2}E\{n_{i_1}^2(t_1)n_{i_1}^2(t_2)\cdots n_{i_1}^2(t_{q/2})\}$$
$$\times E\{n_{i_2}^2(t_1)n_{i_2}^2(t_2)\cdots n_{i_2}^2(t_{q/2})\}\cdots$$
$$\times E\{n_{i_{p-r}}^2(t_1)n_{i_{p-r}}^2(t_2)\cdots n_{i_{p-r}}^2(t_{q/2})\}. \tag{A4}$$

Since the square of the noise is uncorrelated in time, this is equal to

$$M_p^q=[(1)(3)(5)\cdots(q-3)(q-1)]$$
$$\times(\Delta t)^{q/2}(m_{2r}^s)^{q/2}E\{n_{i_1}^2(t_1)\}E\{n_{i_1}^2(t_2)\}\cdots E\{n_{i_1}^2(t_{q/2})\}$$
$$\times E\{n_{i_2}^2(t_1)\}E\{n_{i_2}^2(t_2)\}\cdots E\{n_{i_2}^2(t_{q/2})\}E\{n_{i_{p-r}}^2(t_1)\}$$
$$\times E\{n_{i_{p-r}}^2(t_2)\}\cdots E\{n_{i_{p-r}}^2(t_{q/2})\}$$
$$=[(1)(3)(5)\cdots(q-3)(q-1)]$$
$$\times(\Delta t)^{q/2}(m_{2r}^s)^{q/2}[E\{n^2(t)\}]^{(p-r)q/2}, \tag{A5}$$

which is the even qth order ensemble moment of the correlation of p-r zero-mean noise sequences and the signal raised to the rth power. These moments are consistent with the moment relationships for a Gaussian distributed density.

Baugh, K. W., and Hardwicke, K. R. (**1994**). ''On the detection of transient signals using spectral correlation,'' Circuits Syst. Signal Process. **13**, 467–479.

Bracewell, R. N. (**1986**). *The Fourier Transform and Its Applications* (McGraw-Hill, New York), pp. 189–198.

Delaney, P. A. (**1994**). ''Signal detection using third order moments,'' Circuits Syst. Signal Process. **13**, 481–496.

Dwyer, R. F. (**1984**). ''Use of the kurtosis statistic in the frequency domain as an aid in detecting random signals,'' IEEE J. Ocean. Eng. **9**, 85–92.

Egan, J. P. (**1975**). *Signal Detection Theory and ROC Analysis* (Academic, New York), pp. 61–63.

Gardner, W. A. (**1986**). *Introduction to Random Processes With Applications to Signals and Systems* (Macmillan, New York), pp. 269–272.

Hinich, M. J. (**1990**). ''Detecting a transient signal by bispectral analysis,'' IEEE Trans. Acoust. Speech Signal Process. **38**, 1277–1283.

Hinich, M. J., and Wilson, G. R. (**1990**). ''Detection of non-Gaussian signals in non-Gaussian noise using the bispectrum,'' IEEE Trans. Acoust. Speech Signal Process. **38**, 1126–1131.

Ioup, G. E., Ioup, J. W., Barnes, K. H., Field, R. L., Leclere, J. H., and Rayborn, G. H. (**1989**). ''Evaluation of bicorrelations for transient detection,'' Proceedings of the Workshop on Higher-Order Spectral Analysis, Vail, CO, 46–51.

Ioup, G. E., Pflug, L. A., and Ioup, J. W. (**1995**). ''Effect of ergodicity failure on detection performance prediction,'' J. Acoust. Soc. Am. **97**, 3311(A).

Isserlis, L. (**1918**). ''On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables,'' Biometrika **12**, 134–139.

Kletter, D., and Messer, H. (**1990**). ''Suboptimal detection of non-Gaussian signals by third-order spectral analysis,'' IEEE Trans. Acoust. Speech Signal Process. **38**, 901–909.

Nuttall, A. H. (**1994**). ''Detection performance of power-law processors for random signals of unknown location, structure, extent, and strength,'' NUWC-NPT Technical Report 10,751.

Pflug, L. A. (**1990**). ''Higher order correlations and spectra for the detection of deterministic transients,'' M.S. thesis, University of New Orleans.

Pflug, L. A., Ioup, G. E., Ioup, J. W., and Field, R. L. (**1992a**). ''Properties of higher order correlations and spectra for bandlimited deterministic transients,'' J. Acoust. Soc. Am. **91**, 975–988.

Pflug, L. A., Ioup, G. E., Ioup, J. W., Barnes, K. H., Field, R. L., and Rayborn, G. H. (**1992b**). ''Detection of oscillatory and impulsive transients using higher-order correlations and spectra,'' J. Acoust. Soc. Am. **91**, 2763–2776.

Pflug, L. A., Ioup, G. E., and Ioup, J. W. (**1993**). ''Sampling requirements and aliasing for higher-order correlations,'' J. Acoust. Soc. Am. **94**, 2159–2172.

Pflug, L. A., Ioup, G. E., and Ioup, J. W. (**1994a**). ''Active detection prediction using higher-order moments,'' J. Acoust. Soc. Am. **95**, 2982.

Pflug, L. A., Ioup, G. E., Ioup, J. W., and Field, R. L. (**1994b**). ''Prefiltering for improved correlation detection of bandlimited transient signals,'' J. Acoust. Soc. Am. **95**, 1459–1473.

Pflug, L. A., Ioup, G. E., and Ioup, J. W. (**1995a**). ''Performance prediction formulas for higher order correlation detection of energy signals,'' Proceedings of the IEEE Signal Processing/ATHOS Workshop on Higher-Order Statistics, Begur, Spain, 152–156.

Pflug, L. A., Ioup, G. E., Ioup, J. W., and Field, R. L. (**1995b**). ''Prediction of SNR gain for passive higher order correlation detection of energy transients,'' J. Acoust. Soc. Am. **98**, 248–260.

Sangfelt, E., and Persson, L. (**1993**). ''Experimental performance of some higher-order cumulant detectors for hydroacoustic transients,'' Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics, South Lake Tahoe, CA, 182–186.

Tague, J. A., Pike, C. M., and Sullivan, E. J. (**1994**). ''Active sonar detection in multipath: A new bispectral analysis approach,'' Circuits Syst. Signal Process. **18**, 455–466.

Reference-beam detection for scanning laser acoustic microscopy

Moisés Cywiak, Cristina Solano,[a] and Glen Wade
Centro de Investigaciones en Optica, A.C., Apdo. Postal 1-948, León, Gto, México 37000

S. Isakson
Department of Electrical and Computer Engineering, University of California, Santa Barbara, California 93106

(Received 30 November 1996; accepted for publication 1 December 1997)

Reference-beam detection is inherently superior to the knife-edge detection currently in use in scanning laser acoustic microscopy. This new detector makes use of a reference beam, retarded 90 degrees, which is mixed coherently in a photodiode with the acoustically obtained image-modulated beam. The new detector has an isotropic transfer function which is circularly symmetrical around its highest value, namely the zero-frequency point in the spatial spectrum. This property makes it possible to detect spatial frequencies in all directions simultaneously and with equal sensitivity and simplifies the associated electronics. It also makes possible the employment of acoustic evanescent-wave detection so that ultrasound of low temporal frequency can be used and at the same time high spatial frequencies can be detected for obtaining high resolution. Oblique insonification, required for best operation in the knife-edge detector, is thus not preferred in the reference-beam detector and the resultant Doppler shift in the detected frequency of the transmitted zero-order acoustic waves is avoided. © *1998 Acoustical Society of America.* [S0001-4966(98)01905-5]

PACS numbers: 43.60.Pt, 43.35.Sx [JLK]

INTRODUCTION

Scanning laser acoustic microscopy (SLAM) has become a reliable option for obtaining high-quality images on a microscopic scale with good resolution and contrast. It is the only acoustic microscope that produces its images in real time. The conventional technique for image data readout in SLAM is to deflection modulate a laser beam by scanning it over a solid surface containing the acoustic field scattered from the object and then to demodulate the beam by means of a knife-edge detector (KED). The probing beam is reflected from the surface of a mirrored coverslip onto which scattered acoustic waves, transmitted through the object, impinge (see Fig. 1). The image information, is encoded as deflection modulation on the reflected beam and is detected by the KED. As shown, the system consists basically of a knife edge placed in the focal field of a lens with a photodiode positioned behind the knife edge.

The need for higher resolution and more sensitive methods of acoustic data appropriation and image reconstruction of microscopic objects has grown. In the literature a number of systems for detecting scanned laser acoustic images are described.[1–4] Slam and other acoustic microscopes with scanning laser data acquisition are widely used in many applications[5] and are the subject of a strong international research effort. Thus novel techniques for laser-scanned data procurement continue to be researched and developed.[6–10]

This paper gives a brief description of the operation of the KED in SLAM in a more intuitive way than appears in the literature and by means of which a major weakness of the detector is made apparent. We explain precisely how and why the KED is sensitive not only to the spatial frequencies of the sinusoidal components in the object but also to the direction in which the sinusoidal variations take place. We show that the KED transfer function is antisymmetric with zero response at zero spatial frequency and negative response at negative frequencies. Because of the nature of the transfer function, flexibility in choosing the optimum laser spot size is limited. In addition, an obliquely incident insonification at an optimum angle of incidence is ordinarily employed in SLAM, with a resultant Doppler shift in the detected frequency of the transmitted zero-order acoustic beam, along with single side band detection. In spite of the use of these measures, the bandwidth of the spatial frequency spectrum used in reconstructing the image must be restricted so that it does not extend beyond the positive part of the transfer function and avoids and null response at zero frequency. Therefore, the resolution of the reconstructed image is correspondingly limited. Thus while the KED provides acceptable results for imaging when the spatial variations in the object are in the preferred direction and at the preferred frequencies, it is quite unsatisfactory for variations at other spatial frequencies and in other directions and has a severely limited resolution.

Following that we describe and analyze the new reference-beam detector (RBD) and show that its transfer function is symmetric with no negative response anywhere in the spatial frequency spectrum. This function is highest at zero frequency. Thus both side bands can be demodulated so that no signal power is lost. Neither null responses nor negative frequencies have to be avoided and therefore the detected wave field can have a greater bandwidth of the spatial-frequency spectrum than with the KED.

We also show that because of the above characteristics of the transfer function, evanescent-wave detection is possible with the RBD. The maximum spatial frequency for the

[a]Electronic mail: csolano@foton.cio.mx

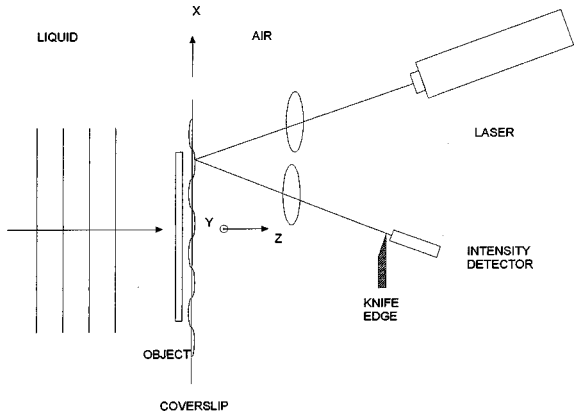


FIG. 1. Basic set up for the knife-edge detector.

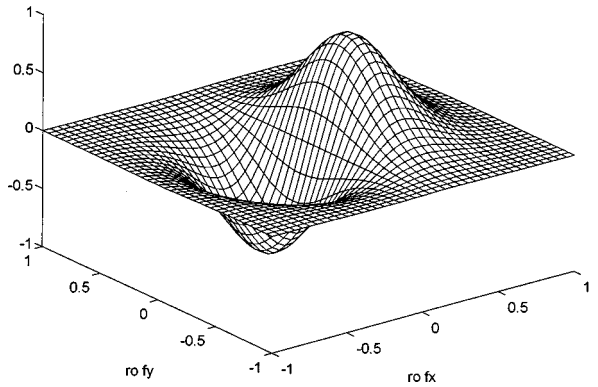


FIG. 2. Normalized transfer function for the knife-edge detector.

detected evanescent waves is limited only by the cross-sectional radius of the scanning laser beam. This frequency can be very high even though the temporal frequency of the sound is low. Thus the propagating acoustic wavelengths can be large, the acoustic temporal frequency (and hence the attenuation in the object) low, and at the same time the fine structure within the object can be highly resolved due to the detection of evanescent waves. Hence, greater sensitivity and higher resolution of the reconstructed image is possible with the RBD and there need be no Doppler shift in the detected frequency of the zero-order waves due to oblique insonification.

I. KNIFE-EDGE DETECTION IN SLAM

As illustrated in Fig. 1, the use of KED in SLAM involves transmitting sonic waves through an object and obtaining a dynamic ripple caused by these waves on a solid mirrorlike surface beyond the object. A scanning laser beam is then employed to probe the ripple and thereby read out the image information concerning the object, this information being contained in the transmitted sonic waves. The motion of the ripple causes the laser beam to be periodically deflected through an angle proportional to the slope of the surface displacement produced by the ripple. In its most sensitive operation, the deflected beam moves perpendicularly with respect to the knife edge causing the light collected by the photodiode behind the knife edge to be intensity modulated. The diode current is substantially proportional to the laser-beam deflection and therefore to the slope of the surface ripple.

Actually the laser beam as it hits the photodiode has been perturbed by two effects: (1) the periodic up-and-down movement of the surface giving rise to phase modulation in the beam and; (2) the periodic tilting of the surface resulting in the deflection modulation. The effect of the up-and-down movement is usually negligible and, as stated before, the magnitude of the signal detected by the photodiode can be taken to be proportional to the slope of the surface displacement in the ripple. This slope, in turn, is proportional to the strength of the sound field transmitted through the object.

As SLAM operates, the laser beam is caused to scan the surface containing the dynamic ripple. The spatial information modulates the beam and is converted to a corresponding temporal signal by means of the scanning and the knife-edge processing.

Based on the above description, the transfer function T_f for the KED using a laser beam with a Gaussian intensity profile is given by[4]

$$T_f = \exp\left(-\frac{\pi^2 r_0^2 (f_x^2 + f_y^2)}{2}\right) \mathrm{erf}\left(\frac{\pi r_0 f_x}{\sqrt{2}}\right), \tag{1}$$

where (f_x, f_y) are the spatial frequencies for sinusoidal variations in the transmittance function of the object along the axes (x,y), and r_0 is the effective beam radius. The knife edge is assumed to be oriented parallel to the x direction; and

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)\, dt. \tag{2}$$

A plot with three-dimensional perspective showing normalized T_f as a function of $r_0 f_x$ and $r_0 f_y$ is presented in Fig. 2.

From the plot we can see that insonification at normal incidence, which puts the carrier of the sound transmitted to the coverslip at the zero spatial frequency point, gives the carrier a null response. Obliquely incident insonification must be used to avoid this. As the laser beam moves across the surface to scan out the raster pattern, the waves produced by the sound transmitted through the object move away or toward the detecting beam. Since both are moving, the detected surface waves on the coverslip are Doppler shifted, appearing to have a higher or lower frequency. Thus even the zero-order sound is Doppler shifted, which would not be the case for insonification at normal incidence.

In order to stay away from the null response as much as possible, usually only the upper side band is used in the detection process. In spite of the above measures (oblique incidence and single side band detection) which are employed to widen the useful spatial spectrum, resolution is still not good. The useful spectrum for effective image reconstruction has to be limited to the positive part of the transfer function and must avoid the null at $f_x = 0$. This results in a correspondingly limited resolution in the reconstructed images.

II. REFERENCE-BEAM DETECTION FOR SLAM

The new detector we propose involves a reference beam, retarded or advanced by 90°, which is coherently mixed in a photodiode with the scanning laser beam that has been phase and deflection modulated by the acoustic waves on the coverslip. The transfer function for this reference-beam detector (RBD) will now be derived. In our treatment an object is represented in its most mathematically general form. For simplicity in the derivation we have assumed that the lenses of the system all have the same focal length.

The coverslip surface dynamic ripple can be expressed as

$$z(x,y,t)=\mathrm{Re}\{h(x,y)\exp(-j\omega_s t)\}, \tag{3}$$

where $h(x,y)$ is the complex amplitude distribution function of the surface excursion, (x,y) are the axes coordinates in this plane and $\omega_s=2\pi f_s$, the acoustical temporal frequency.

The coverslip is probed by a laser beam with incident power P and a Gaussian intensity profile centered at the point (x_0,y_0)

$$\Psi(x,y)=\sqrt{\frac{2P}{\pi r_0^2}}\exp\left(-\frac{(x-x_0)^2+(y-y_0)^2}{r_0^2}\right), \tag{4}$$

where Ψ is the complex amplitude distribution function of the beam. In order to simplify the equations, the parameters (x_0,y_0) will be suppressed (that is, assumed to be zero for the time being) and will be appropriately reintroduced in the final result.

The probing beam, after being reflected from the surface of the coverslip is phase and amplitude modulated by the dynamic ripple. Thus the reflected complex amplitude distribution function is given by

$$\Psi_r=\sqrt{\frac{P}{\pi r_0^2}}\exp\left(-\frac{x^2+y^2}{r_0^2}\right)[1+jkh(x,y)\exp(-j\omega_s t) + jkh^*(x,y)\exp(j\omega_s t)], \tag{5}$$

where $k=2\pi/\lambda$ and λ the light wavelength. In Eq. (5) only terms of the first order in $kh(x,y)$ have been retained due to the fact that the value of this product is very small compared with unity.

Lens L1 performs the Fourier transform of Ψ, giving at the plane of the detector

$$\Psi_p=\left(\frac{1}{j\lambda f}\right)\sqrt{\frac{P}{\pi r_0^2}}\left[\pi r_0^2\exp\left(-\frac{\xi^2+\eta^2}{\rho^2}\right) + jk\exp(-j\omega_s t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\left(-\frac{x^2+y^2}{r_0^2}\right)h(x,y)\exp\left(-j\frac{2\pi}{\lambda f}(x\xi+y\eta)\right)dx\,dy + jk\exp(j\omega_s t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\left(-\frac{x^2+y^2}{r_0^2}\right)h^*(x,y)\exp\left(-j\frac{2\pi}{\lambda f}(x\xi+y\eta)\right)dx\,dy\right], \tag{6}$$

where (ξ,η) are the axial coordinates for this plane.

As seen in Fig. 3, the complex amplitude function Ψ_p is coherently added to the complex amplitude function of the reference beam in the photodiode. This latter function corresponds to the Fourier transform of the complex amplitude function of the incident beam retarded or advanced 90°. Thus the total complex amplitude distribution function Ψ_F for the light at this plane will be

$$\Psi_F=\Psi_P+\left(\frac{1}{j\lambda f}\right)\sqrt{\frac{P}{\pi r_0^2}}\left[j\pi r_0^2\exp\left(-\frac{\xi^2+\eta^2}{\rho^2}\right)\right]. \tag{7}$$

The power received by the photodiode is given by

$$P_s=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Psi_F\Psi_F^*\,d\xi\,d\eta, \tag{8}$$

the limits of integration in Eq. (8) are chosen from $-\infty$ to ∞ because the photodiode is large compared with the size and excursions of the beams.

Evaluating the above integral we have

$$P_s=P\left[1+\frac{4k}{r_0^2}\mathrm{Re}\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}h(x,y)\times\exp\left(-2\frac{(x-x_0)^2+(y-y_0)^2}{r_0^2}\right)\times\exp(-j\omega_s t)dx\,dy\right\}\right], \tag{9}$$

where the parameters (x_0,y_0) has been reintroduced as indicated before.

The integral in Eq. (9) can be interpreted very simply as the convolution of the impulse response of the system $\mathfrak{T}_s(x_0,y_0)$ with the function $h(x,y)$. Thus we get

$$\mathfrak{T}_s(x_0,y_0)=\frac{4k}{r_0^2}\exp\left(-2\frac{x_0^2+y_0^2}{r_0^2}\right). \tag{10}$$

The transfer function T_s for the RBD is obtained by Fourier transforming $\mathfrak{T}_s(x_0,y_0)$. Thus

$$T_s(f_x,f_y)=2\pi k\exp\left(-\frac{\pi^2 r_0^2}{2}(f_x^2+f_y^2)\right). \tag{11}$$

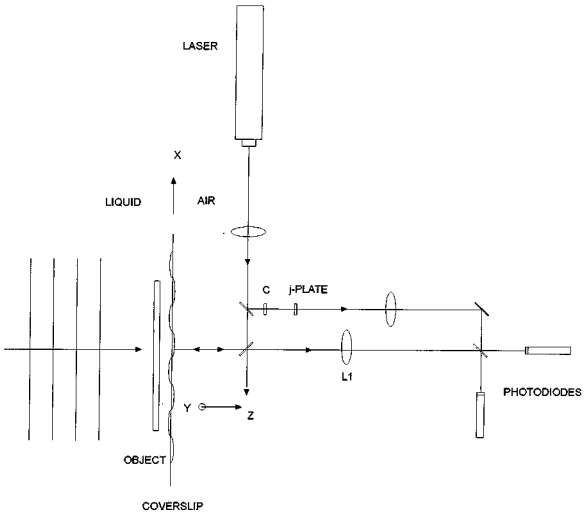


FIG. 3. Basic set up for the reference-beam detector. In the mathematical treatment the lenses are assumed to have the same focal length. C is an element to compensate the optical path of the reference beam. The outputs of both photodiodes are used in the signal processing so that no signal energy is lost.

Equation (11) shows that the transfer function for the reference-beam detector, plotted normalized in Fig. 4, is a Gaussian and therefore isotropic with maximum response at zero spatial frequency. This avoids having to insonify obliquely for best operation and, as we discuss below, permits evanescent-wave detection.

III. DISCUSSION AND CONCLUSIONS

The key component in this new detector is a reference beam advanced or retarded 90° which mixes coherently with the acoustic image-modulated beam and thus eliminates the need for a knife edge. The RBD is inherently superior to the KED used in a conventional SLAM because its transfer function is isotropic and maximum at the origin. These features give rise to high resolution, a reduction in the processing time required to analyze a given object and a simplification in the associated electronics. Furthermore, since the frequency response of the RBD has its maximum at the origin, there is no necessity for the oblique insonification that is

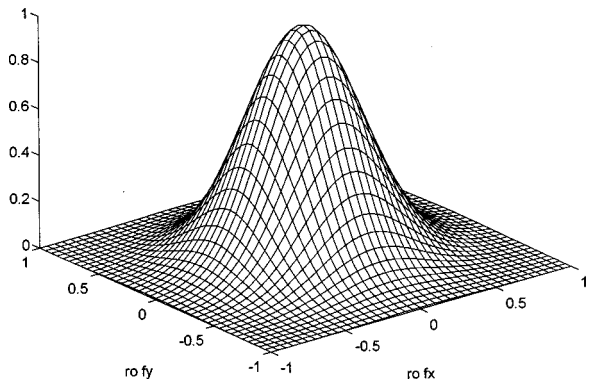


FIG. 4. Normalized transfer function for the reference-beam detector.

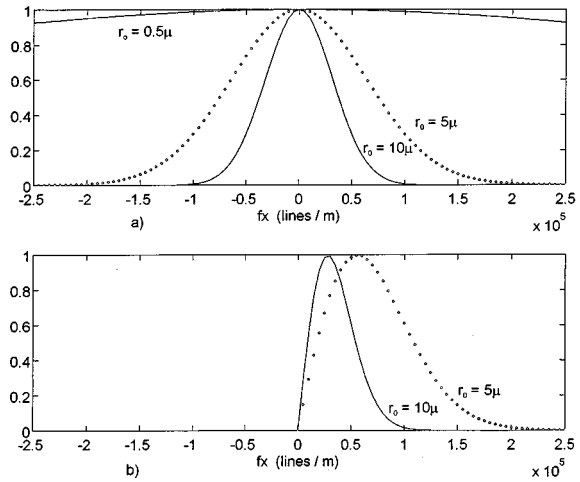


FIG. 5. Normalized one-dimensional transfer function: (a) for the RBD, and (b) for the KED.

required in the KED. This eliminates a Doppler shift in the detected frequency of the transmitted zero-order acoustic waves due to scanning.

We can see from the transfer functions of the two systems (Figs. 2 and 4) and the following arguments that the RBD is capable of much higher resolution than the KED in the direction perpendicular to the knife edge and is always higher in other directions. One reason is that the RBDs transfer function is isotropic and that of the KED is not. Figure 5 displays the normalized one-dimensional transfer functions for the RBD corresponding to three different spot sizes and the useful (that is, positive) part of the transfer functions of the KED for two of these spot sizes. In these plots both detectors appear to be more sensitive to the higher spatial frequencies for the smaller spot size.

However, for the KED (but not the RBD) the maximum response is moved to a higher spatial frequency as the spot size is lowered. The desired frequency for maximum response is that of the transmitted zero-order acoustic waves. Calculations show that when the spot size is cut from 10 μm to 5 μm, the response at the KED maximum drops to 70% of the previous value. Thus the optimal spot size for the KED is set by the angle of insonification needed to achieve the maximum response for the zero-order waves. This spot size is usually around 0.25–0.5 times the acoustical wavelength for the KED.

The RBD does not exhibit this problem. The maximum response remains at zero spatial frequency independent of the acoustic wavelength no matter what the spot size is. The limitation on broadening the spatial-frequency bandwidth is only on how tightly the laser beam can be focused. The smaller the spot size, the higher the spatial frequency that can be detected and the higher the object resolution that can be achieved.

The RBD is capable of detecting evanescent waves, that is, spatial wavelengths shorter than the acoustic wavelength. For example, a helium–neon laser beam can be focused to a radius of about half a micronmeter. If a 100-MHz acoustic frequency is used in water, the acoustic wavelength is 15 μm. However, if the object is mounted in the very near field,

the RBD could detect evanescent acoustic waves 15 times smaller than the acoustic wavelength. This provides a much higher resolution of the object that can be produced by conventional means at this acoustic wavelength. Furthermore, this resolution is independent of the actual acoustic wavelength. With evanescent-waves detection the acoustic wavelengths can be large with corresponding low attenuation in the object and at the same time the object's fine structure can be highly resolved.

ACKNOWLEDGMENTS

This work was supported in part by a CONACYT scholarship awarded to M. Cywiak and a 1996 UCMEXUS grant for Scanning-Laser Acoustic Microscopy to G. Wade.

[1] A. Korpel and P. Desmares, ''Rapid sampling of acoustic holograms by laser scanning techniques,'' J. Acoust. Soc. Am. **45**, 881 (1969).
[2] L. Whitman and A. Korpel, ''Probing of acoustic surface perturbations by coherent light,'' Appl. Opt. **8**, 1567 (1969).
[3] L. Kessler, ''Imaging with dynamic-ripple diffraction,'' in *Acoustic Imaging*, edited by G. Wade (Plenum, New York, 1976), pp. 229–239.
[4] R. Mueller and R. Rylander, ''New demodulation scheme for laser-scanned acoustical-imaging systems,'' J. Opt. Soc. Am. **69**, 407 (1979).
[5] Y. Kim, W. D. Hunt, Y. Liu, and C. Jen, ''Velocity surface measurements for ZnO films over {001}-cut GaAs,'' J. Appl. Phys. **76**, 1455 (1994).
[6] S. Samson, A. Korpel, and H. S. Snyder, ''Conversion of evanescent waves into propagating waves by vibrating knife edge,'' Int. J. Imaging Syst. Technol. **7**, 48 (1996).
[7] L. R. Sahagún-Ortiz, F. Mendoza-Santoyo, G. Wade, and S. Isakson, ''Detector de Espejo Piramidal Para Microscopio Acústico de Barrido Laser,'' in *Programa y Resúmenes*, Sociedad Mexicana de Física, XXXIV Congreso Nacional de Física (Oct. 1996), Oaxaca, Mexico, pp. 49.
[8] A. Meyyappan, S. Isakson, F. Mendoza-Santoyo, R. A. Duarte, L. R. Sahagun, and G. Wade, ''Detection systems for scanning laser tomoholographic acoustic microscopy,'' Int. J. Imaging Syst. Technol. **7**, 131–135 (1996).
[9] R. Y. Chiao and H. Lee, ''Recent advances in scanning tomographic acoustic microscopy,'' Int. J. Imaging Syst. Technol. **3**, 334–353 (1991).
[10] H. Lee, ''Scanning tomographic acoustic microscopy: A historical overview,'' in *Acoustical Imaging*, Vol. 22, edited by P. Tortoli and L. Masotti (Plenum, New York, 1996), pp. 645–650.

An adaptive noise reduction stethoscope for auscultation in high noise environments

Samir B. Patel
School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907-1285

Thomas F. Callahan and Matthew G. Callahan
University Research Engineers & Associates, Inc., Acton, Massachusetts 01720

James T. Jones, George P. Graber, and Kirk S. Foster
Hillenbrand Biomedical Engineering Center, Purdue University, West Lafayette, Indiana 47907-1296

Kenneth Glifort
Armstrong Laboratory, U.S. Air Force, San Antonio, Texas 78201

George R. Wodicka[a)]
School of Electrical and Computer Engineering and Hillenbrand Biomedical Engineering Center, Purdue University, West Lafayette, Indiana 47907-1285

(Received 17 January 1996; accepted for publication 23 January 1998)

Auscultation of lung sounds in patient transport vehicles such as an ambulance or aircraft is unachievable because of high ambient noise levels. Aircraft noise levels of 90–100 dB SPL are common, while lung sounds have been measured in the 22–30 dB SPL range in free space and 65–70 dB SPL within a stethoscope coupler. Also, the bandwidth of lung sounds and vehicle noise typically has significant overlap, limiting the utility of traditional band-pass filtering. In this study, a passively shielded stethoscope coupler that contains one microphone to measure the (noise-corrupted) lung sounds and another to measure the ambient noise was constructed. Lung sound measurements were made on a healthy subject in a simulated USAF C-130 aircraft environment within an acoustic chamber at noise levels ranging from 80 to 100 dB SPL. Adaptive filtering schemes using a least-mean-squares (LMS) and a normalized least-mean-squares (NLMS) approach were employed to extract the lung sounds from the noise-corrupted signal. Approximately 15 dB of noise reduction over the 100–600 Hz frequency range was achieved with the LMS algorithm, with the more complex NLMS algorithm providing faster convergence and up to 5 dB of additional noise reduction. These findings indicate that a combination of active and passive noise reduction can be used to measure lung sounds in high noise environments. [S0001-4966(98)02605-8]

PACS numbers: 43.60.Qv, 43.80.Cs, 43.80.Ev, 43.80.Vj [JLK]

INTRODUCTION

Auscultation is the process of listening to sounds emanating from the body. Medical personnel frequently make diagnoses pertaining to the health of the respiratory system using a standard binaural stethoscope. However, this stethoscope is highly prone to interference from ambient noise and thus becomes clinically useless in high ambient noise environments.

In a typical emergency, civil or military, an ambulance or aircraft (helicopter or fixed wing) is used to evacuate critically ill patients. Several sources of noise exist in such vehicles, and their interaction with the interior space or fuselage is complex and time variant due to the wide range of conditions and maneuvers encountered during a typical evacuation. For example, ambient noise levels in the range of 90–100 dB SPL (sound pressure level) exist in the cabin of C-130 aircraft used by the U.S. Air Force (USAF) for aeromedical evacuation. In contrast, normal breath sounds in an healthy adult are roughly 22–30 dB SPL in free space (excluding airflow generated noise radiating from the mouth) and 65–70 dB SPL within a stethoscope coupler (Zenk, 1994). Thus traditional auscultation in such a high noise environment is not feasible due to an inherently poor signal-to-noise ratio (SNR). In fact, even moderate background noise levels such as those found during average conversation are sufficient to hamper auscultation in the clinic. In general, the SNR can be improved either by increasing the signal strength (numerator) or by decreasing the interfering noise (denominator). However, the high noise levels and strong overlap of spectral content of lung sounds and the time-varying vehicle noise impose limitations on the successful employment of conventional passive noise reduction schemes.

Thus in this study, a digital scheme for adaptive noise reduction (ANR) was implemented in addition to passive noise reduction (PNR) at the transducer in an integrated approach. In this manner, the time varying characteristics of the noise are estimated and exploited to effectively extract the desired lung sounds. This manuscript details the design and preliminary testing of a Digital Active Noise Attenuation (DANA) Stethoscope, with a focus on the function of the

[a)]Electronic mail: wodicka@ecn.purdue.edu

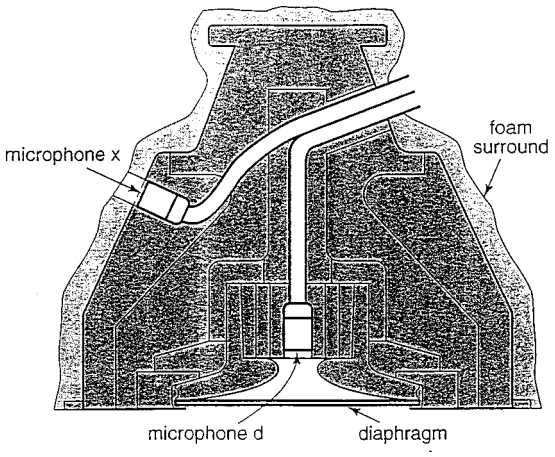


FIG. 1. Diagram of the stethoscope coupler. Diaphragm is 4 cm in diameter.

adaptive filters. The two algorithms that were implemented are the least-mean-squares (LMS) and normalized least-mean-squares (NLMS).

I. STETHOSCOPE COUPLER

A passively shielded coupler with two omnidirectional electret microphones (Telex ELM-22, Minneapolis, Minnesota) was constructed (Fig. 1). The microphone d is placed in a conical air cavity behind a standard stethoscope diaphragm to detect the lung sounds (Wodicka *et al.*, 1994). Such an arrangement allows significant lung sound energy to reach microphone d, with a frequency response that mimics that of the standard stethoscope. The other microphone, x, measures the noise reference that drives the adaptive filter. These sensors are positioned in the coupler so that microphone x detects very little lung sounds, yet measures sound that is highly correlated to the noise contaminating the lung sounds measured by microphone d. The stethoscope coupler is passively shielded from high frequency ambient noise via a foam surround since the effectiveness of adaptive noise reduction decreases with increasing frequency.

II. ADAPTIVE NOISE CANCELLATION

The conventional method of signal estimation from its noise corrupted version is to pass it through a filter that removes the unwanted noise without significantly affecting the signal. The filter employed can be fixed or self-adjusting, i.e., adaptive. If the signal and noise characteristics are known *a priori* the optimal fixed filter can be designed based on Wiener filter theory. In practice, *a priori* knowledge of noise characteristics is generally not available and hence a filter that adapts to the characteristics of the noise is desired. In the following, we present the basic theory of the adaptive filter algorithms considered in this study.

A. Least-Mean-Squares (LMS) algorithm

The LMS algorithm, developed by Widrow and Hoff in 1959 (Widrow *et al.*, 1985; Haykin *et al.*, 1986), has been studied in great detail and has found a myriad of applications (Widrow *et al.*, 1975; Darlington *et al.*, 1985; Rodriguez *et al.*, 1987; Harrison *et al.*, 1986; Pulsipher *et al.*, 1979; Elliott *et al.*, 1987; Poole *et al.*, 1984). Figure 2(a) depicts the basic construct for noise reduction in a linear system with additive noise. The primary input d consists of signal s and noise n_0 (for convenience the time index j is omitted). The reference input, i.e., the input of the adaptive filter, consists of noise x. The signal s is assumed to be uncorrelated with both n_0 and x, while n_0 and x are assumed to be correlated with each other in some unknown way. n_0 is often referred to as the primary noise and x as the noise reference. In this approach, x is adaptively filtered to produce an output y which is then subtracted from the primary input d to produce the system output e (the recovered signal) where

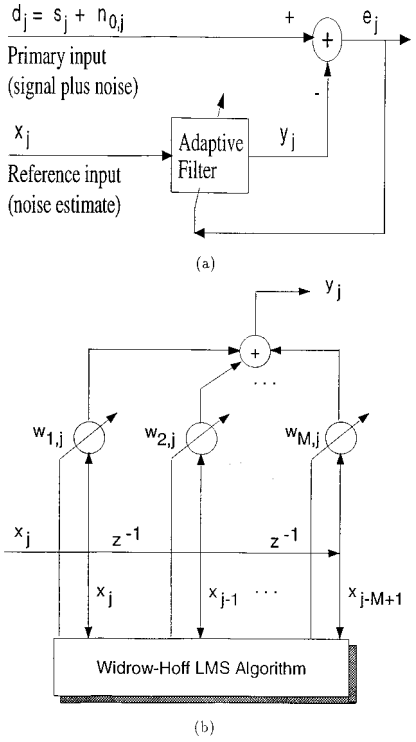


FIG. 2. Adaptive noise canceler: (a) block diagram; (b) tapped delay line.

$$e = s + n_0 - y. \tag{1}$$

From Eq. (1), it is observed that minimization of the mean-square error (MSE) $E[e^2]$ results from the minimization of $E[(n_0-y)^2]$, i.e.,

$$\min E[e^2] = E[s^2] + \min E[(n_0-y)^2] \tag{2}$$

and that smallest possible output power is $E[e^2]=E[s^2]$ when $E[(n_0-y)^2]=0$. In this case, the output y of the adaptive filter is a replica of the noise n_0 in the primary input and thus n_0 is completely eliminated and the output SNR is maximized.

The adaptive filter in Fig. 2(a) typically consists of a tapped delay line as depicted in Fig. 2(b). The output y_j, where j is the time index, is equal to the inner product of noise reference input vector X_j and filter coefficient vector W_j, each of length M, i.e.,

$$y_j = W_j^T X_j \tag{3}$$

and the general expression for the error signal as a function of weight vector is given by

$$e_j = d_j - y_j = d_j - W_j^T X_j \tag{4}$$

In the minimization of the MSE, when the true value of gradient (derivative of $\epsilon_j = E[e_j^2]$ with respect to filter coefficient vector W_j) is equated to zero, the steepest descent method (Widrow *et al.*, 1985) yields the form of the so-called Widrow–Hoff LMS algorithm

$$W_{j+1} = W_j + \mu e_j X_j, \tag{5}$$

where μ is the step size which controls the stability and rate of convergence. The larger the value of μ, the faster the algorithm converges and the more variable is the gradient estimate, and vice versa (Widrow *et al.*, 1976).

The magnitude-squared coherence function

$$\gamma_{pq}^2(e^{j\omega}) = \frac{|S_{pq}(e^{j\omega})|^2}{S_{pp}(e^{j\omega}) S_{qq}(e^{j\omega})} \tag{6}$$

is often used to measure the correlation between two stationary random processes p and q where S_{pq} is the cross-power spectral density of p and q, and S_{pp}, S_{qq} are the respective auto-power spectral densities.

When the filter output signal, e, is orthogonal to the noise reference data, x, the steepest-descent method converges to the optimal Wiener filter coefficients W^* and there are no further updates. On the other hand, in the steady state, the LMS algorithm continues to have updates. These fluctuations, called weight jitters, are about the optimal Wiener filter coefficients W^* and the LMS algorithm does not converge with unity probability. It can be shown that the mean of these updates converges to W^* under the assumption of stationarity of the input to the adaptive filter (Haykin, 1986). However, the MSE, ϵ, converges to a larger value than the optimal minimum mean square error, $\epsilon_{\min}$ (achieved with an optimal Wiener filter).

With the assumptions made in *independence theory* (Haykin, 1986), the necessary and sufficient condition for the convergence of LMS algorithm is that step-size parameter μ satisfies

$$0 < \mu < \frac{2}{\lambda_{\max}}, \tag{7}$$

where $\lambda_{\max}$ is the maximum eigenvalue of the autocorrelation matrix of the input to the filter, R_{MM}. In practice, the calculation of the individual eigenvalues of R_{MM} to find the upper limit of μ given by Eq. (7) is computationally very intensive. Under the stationarity assumption of the input to the filter, and positive definiteness of R_{MM},

$$\lambda_{\max} \leqslant \sum_{i=1}^{M} |\lambda_i| = \text{trace}(R_{MM}) = M\phi_x(0) = MP_x, \tag{8}$$

where P_x is the total power input to the adaptive filter, a generally known quantity. Thus in practice, a tighter bound is placed on the upper limit of μ and Eq. (7) becomes

$$0 < \mu < \frac{2}{MP_x}. \tag{9}$$

B. Normalized LMS (NLMS) algorithm

The LMS algorithm described in the previous section uses a constant step size μ. The stability, speed of convergence, and fluctuation of the adaptation process are governed by the step size μ and the input power to the adaptive filter P_x. The normalized LMS (NLMS) algorithm (Treichler *et al.*, 1987) represents a technique to improve the speed of convergence. This is accomplished while maintaining the steady-state performance independent of the input signal power. This NLMS algorithm uses a time variable convergence factor μ_j,

$$\mu_j = \frac{\alpha}{\xi + P_{x,j}}, \tag{10}$$

where α is the ''normalized'' step size chosen to be between 0 and 2, while ξ is a small positive term included to ensure that the update term does not become excessively large when the input average power at time j, $P_{x,j}$, becomes small. The NLMS algorithm gains its stability by normalizing the weight vector update with an estimate of the signal power. A computationally inexpensive way of obtaining this normalization is via recursive power estimation

$$P_{x,j+1} = (1-\beta)P_{x,j} + M\beta x_j^2, \tag{11}$$

where $0 < \beta \ll 1$ is a smoothing parameter and M is the filter length, with the NLMS algorithm weight vector update equation as

$$W_{j+1} = W_j + \mu_j e_j X_j. \tag{12}$$

C. Criteria for performance measures of adaptive algorithms

A study of the relationship between the speed of adaptation and performance of adaptive systems is complex since different choices of performance measures (often termed misadjustments) yield different ''optimized'' step sizes. To optimize the step size of the LMS algorithm and study its performance, the root-mean-squared (RMS) misadjustment, a performance measure similar to the temporal root-mean-squared MSE was used. This RMS misadjustment is defined as

$$\text{MRMS} = \left[\frac{\langle E[(s_j - e_j)^2] \rangle}{\langle E[s_j^2] \rangle} \right]^{1/2}, \tag{13}$$

where s_j is the lung sounds signal component and e_j is the output of the adaptive system (the recovered lung sounds). MRMS is a normalized root-mean-squared MSE and reflects the degree of the distortion of the desired signal introduced by the adaptive system.

Obviously, the optimal step-size parameter μ_* and the optimal filter length M_* that will yield a minimum MRMS is desired. In general, faster adaptation leads to a more noisy adaptive process. In a stationary environment, enhanced steady-state performance (smaller MRMS) results from a smaller value of the step size μ and thus a slower convergence. In a nonstationary environment however, a compromise has to be made between fast convergence necessary to track the changes in the input signal and slow adaptation needed to reduce the gradient noise (misadjustment). Thus

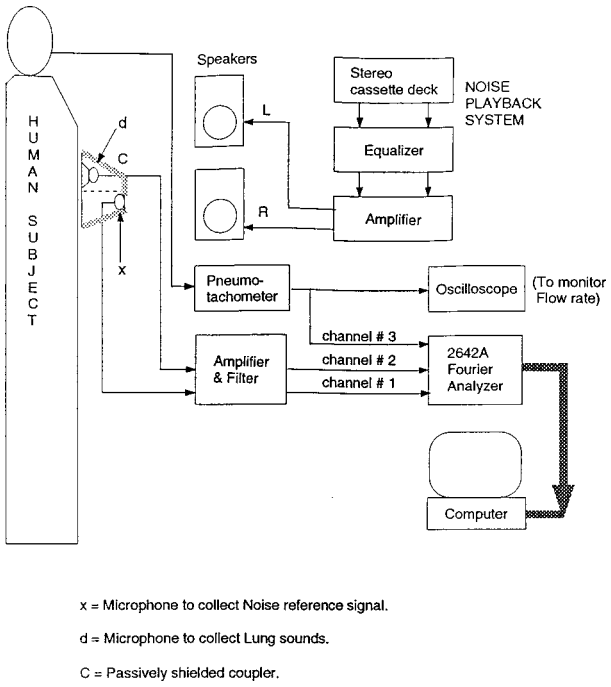


FIG. 3. Experimental apparatus.

the optimal value of μ is a compromise between a fast rate of convergence and a reduced misadjustment. Further, it can be shown that even for a stationary input, the misadjustment due to gradient noise, G, is proportional to μ (Haykin, 1986; Widrow *et al.*, 1976),

$$G = \mu M P_x, \tag{14}$$

and this relationship is used in this study to calculate the value of μ for given values of G, M, and P_x.

A performance measure that reflects the optimization of filter length M also must be chosen. As mentioned previously, the steady-state error signal e (the recovered breath sounds) and the reference data samples x used to obtain the minimum MSE (MMSE) estimate of the desired signal obeys orthogonality for the steepest descent method (and even for the optimal Wiener filter). Expanding upon this relationship, the cross correlation between the output error signal e and the reference noise x should attain some minimum value close to zero for the optimal filter length M_*. To quantitate this relationship, the correlation parameter COR was defined as

$$\mathrm{COR} \doteq \frac{1}{L} \sum_{j=1}^{L} e_j x_j. \tag{15}$$

As M increases, COR should decrease as relatively more and more uncorrelated data samples are taken into the estimation of the desired signal.

III. EXPERIMENTAL APPARATUS AND PROCEDURE

The primary goal of the experimentation was to record and process lung sounds in a simulated USAF C-130 aircraft environment. Figure 3 provides an overview of the experimental apparatus. The protocol was approved by the Purdue University Committee on the Use of Human Subjects. A subject was seated in a soundproof room that contained an audio system which played the recordings of USAF C-130 aircraft fuselage noise to simulate this environment. The subject breathed through a pneumotachograph (Fleisch #2) at target flow rates of 2 L/s. An experimenter held the stethoscope coupler on the subject's right anterior upper chest during the protocol. The signals from both microphones (d and x) were amplified by a factor of 32.5 and high-pass filtered using a fourth-order Butterworth filter at a cutoff frequency of 100 Hz. These signals and the pneumotachograph output were digitized at 5120 samples per second using a Tektronix Fourier Analyzer (2642A, with internal anti-aliasing filters) and stored on an IBM-PC compatible computer. Data segments of 20-s duration were recorded during breathing and apnea at aircraft noise levels of 80, 90, and 100 dB SPL, and also in the quiet room. In addition, noise measurements were made with the coupler placed on the subject's thigh. The stored digital epochs were subsequently processed with the adaptive noise reduction LMS and NLMS algorithms via MATLAB, where a detailed performance analysis was performed.

The signals d and x were also simultaneously input to a ''real-time'' implementation of one version of the LMS algorithm running on a digital signal processing board (AT-DSP2200, National Instruments) with a second IBM-PC compatible computer. The output of this algorithm was fed to an audio amplifier (DRA-345R, Denon) and presented to the experimenter within the chamber via a noise cancellation headset (AH-BG, Bose). Thus the experimenter could monitor the extracted lung sounds (from a specific realization of the algorithm) during the recording process, yet the focus of this study was on the off-line processing of the stored acoustic signals.

IV. RESULTS AND DISCUSSION

A. Characteristics of lung sounds and the interfering noise

Figure 4 depicts representative data collected in the quiet chamber. The pneumotachometer signal [Fig. 4(a)] is positive for inspiration and negative for expiration. Figure 4(b) is the corresponding lung sounds detected at microphone d. For all the plots of microphone measurements, 1 V corresponds to 3.1 Pa. The most clear and audible lung sounds occur for a flow rates above roughly 1 L/s. Figure 4(c) is the time record of the noise reference signal x, where some detection of breathing is observed during expiration due to airborne sounds from the mouth reaching the coupler. Figure 5 provides in the same format a representative data epoch in the presence of the high level noise. Here the measured (unprocessed) lung sounds are overwhelmed by the noise and hence are inaudible. The amplitude of the noise reference is less than that of the lung sounds detected in the presence of noise because the relatively high acoustic impedance of the air cavity in front of microphone d results in greater sound pressure levels measured as compared to microphone x which is essentially open to free space. An estimate of the power spectrum of both inspiratory lung sounds (from the microphone d signal detected during breathing in a quiet chamber) and aircraft noise (from the microphone d

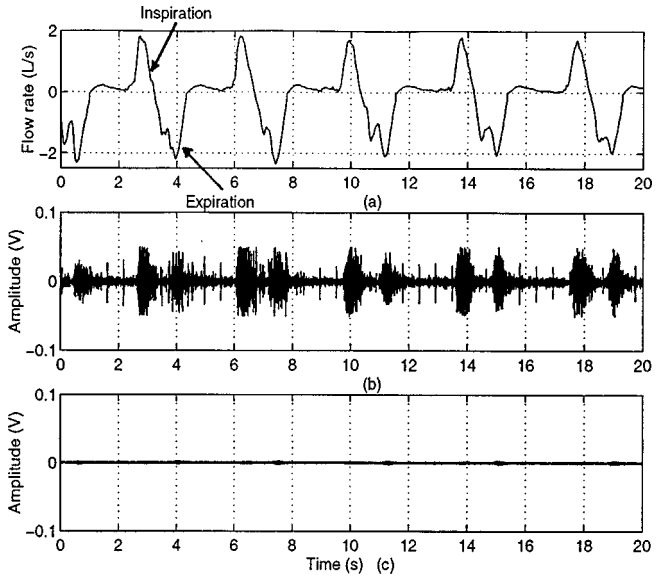


FIG. 4. Time record of (a) pneumotachometer, (b) microphone d (lung sounds), (c) microphone x (noise reference) signals in a quiet chamber.

signal during apnea in presence of 100 dB SPL of aircraft noise) is shown in Fig. 6. The majority of the energy in both signals is concentrated in the 100–600 Hz frequency range. This strong overlap in spectral content makes it particularly difficult to auscultate inside an aircraft. The coherence, as a measure of correlation between the noise reference x and the noise corrupting the lung sounds n_0, is depicted in Fig. 7. The valley in the coherence at just above 800 Hz corresponds to that in the aircraft noise spectra.

B. Performance analysis of the adaptive algorithms

To optimize and preliminary evaluate the processing schemes, the lung sound measurements (microphone d) made in the quiet chamber (signal, s) were first considered the pure desired signal. To this signal, s, the amplitude of the

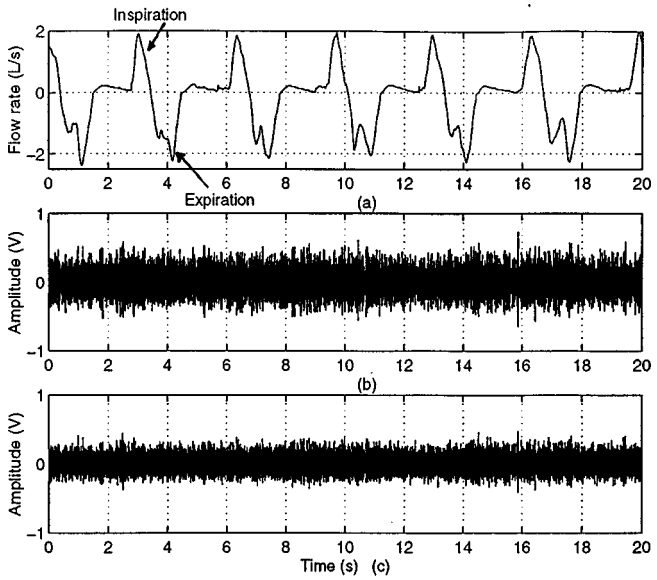


FIG. 5. Time record of (a) pneumotachometer, (b) microphone d (lung sounds), (c) microphone x (noise reference) signals in 100 dB SPL of aircraft noise.

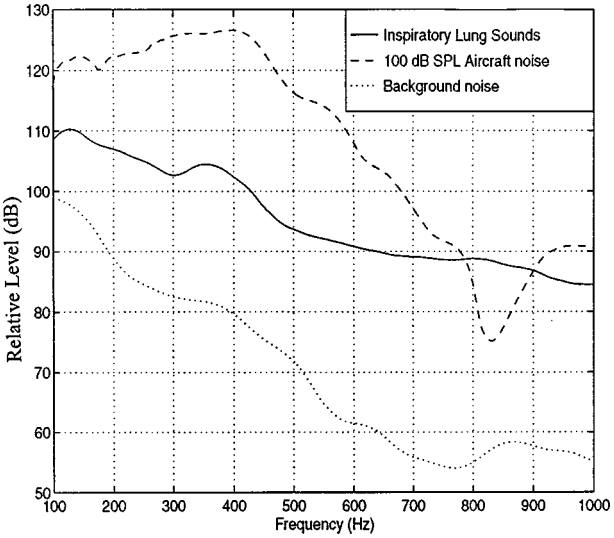


FIG. 6. Relative power spectra of aircraft noise, inspiratory lung sounds, and background noise in a quiet chamber.

aircraft noise detected by microphone d during nonbreathing with the stethoscope coupler on the right thigh was added to obtain a simulation of signal plus noise, $s+n_0$ for each of the respective noise levels. These generated $s+n_0$ data were input as the primary signal to the adaptive filter, with x as the noise reference signal, to analyze and optimize the performance of the algorithm. The resulting algorithms were then used to process actual experimental data. This initial data simulation stage was required since simultaneous recording of pure lung sounds and its corrupted version is not possible.

Some measures that were used to assess the algorithm were:

(1) Optimization of step size μ and filter length M: For the LMS algorithm, the values of μ were calculated using Eq. (14) for various values of gradient misadjustment G ranging from 0.01 to 1.0 and filter lengths M of 10, 20, and 40. With these values of M and the calculated values of μ,

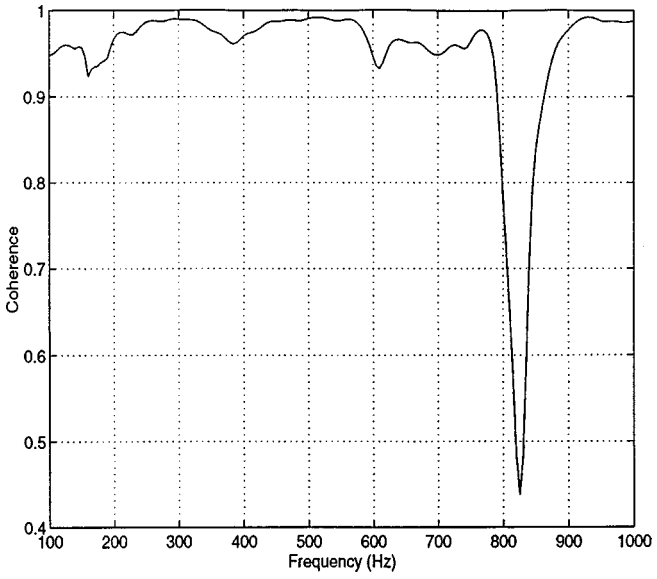


FIG. 7. Coherence estimate between the primary (microphone d) and the noise reference (microphone x) signals at 100 dB SPL of aircraft noise.

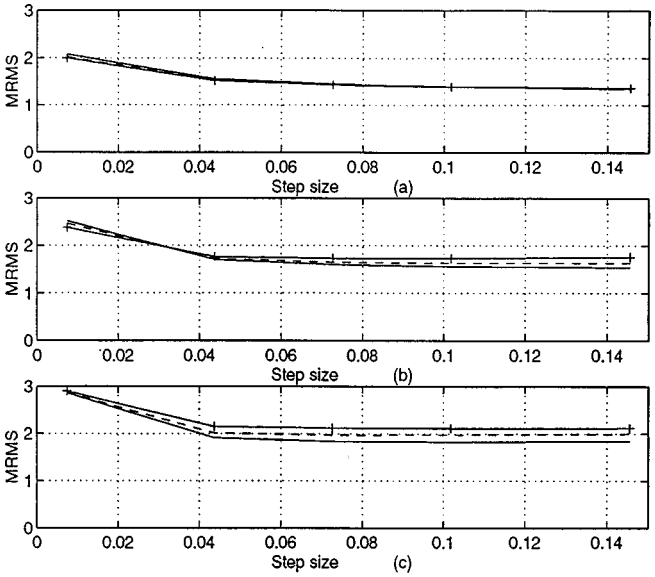


FIG. 8. For the LMS algorithm: Performance measure MRMS versus step size μ for filter length $M=10$ (———), $M=20$ (– – –), and $M=40$ (–+–) in (a) 80, (b) 90, and (c) 100 dB SPL of aircraft noise.

the LMS algorithm was implemented. For each case the performance measures of MRMS and COR were calculated. Figure 8 shows plots of MRMS versus μ for 80, 90, and 100 dB SPL noise levels which are generally consistent with theoretical predictions (Haykin, 1986; Widrow *et al.*, 1976). MRMS for 80 and 90 dB SPL of noise [Fig. 8(a), (b)] was a relatively weak function of μ, with increasing μ slightly increasing the MRMS as a reflection of G. The respective curves for 100 dB SPL of noise [Fig. 8(c)] exhibit the existence of ''optimal'' step-size parameter μ_* for which MRMS is a minimum. In the optimization of filter length M, the COR parameter was also taken into consideration. Plots of COR versus M (Fig. 9) show that in general a filter with a larger number of coefficients has a slightly lower COR.

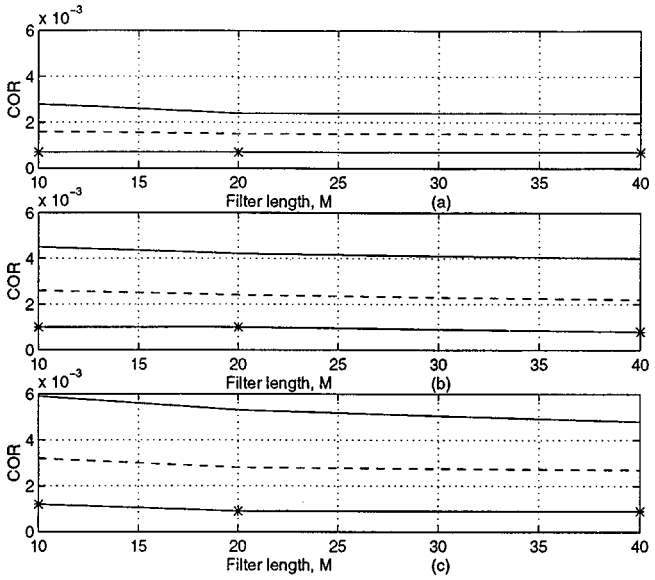


FIG. 9. For the LMS algorithm: Performance measure COR versus filter length M for step size $\mu=0.0073$ (———), $\mu=0.0146$ (– – –), and $\mu=0.0437$ (–*–) in (a) 80, (b) 90, and (c) 100 dB SPL of aircraft noise.

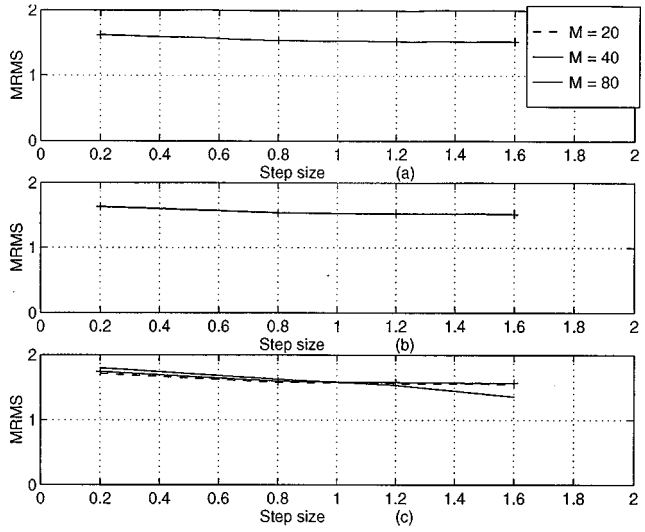


FIG. 10. For the NLMS algorithm: Performance measure MRMS versus step size α for filter length $M=20$ (– – –), $M=40$ (–+–), and $M=80$ (———). (a) 80, (b) 90, and (c) 100 dB SPL of aircraft noise.

When the filter length tends to an ''optimal'' value M_*, the value of COR does not change significantly thereafter. Thus from the perspective of implementation, the lowest value of M whereafter COR remains almost constant could be selected as the ''optimal'' value M_*. A similar evaluation of the NLMS algorithm was carried out for both the simulated and actual data for filter lengths M of 20, 40, and 80 with normalized step sizes α of 0.2, 0.8, 1.2, and 1.6. Plots of the performance measures for this case (Figs. 10 and 11) depict the existence of potentially optimal parameter settings. The MRMS for the NLMS algorithm (Fig. 10) is relatively lower than for the LMS algorithm because the use of a normalized dynamic step size drastically improves the tracking capability of the NLMS algorithm. Further, COR for the NLMS algorithm (Fig. 11) shows $M=40$ as an optimal filter length from an implementation perspective. This observation is similar to that made with the LMS algorithm.

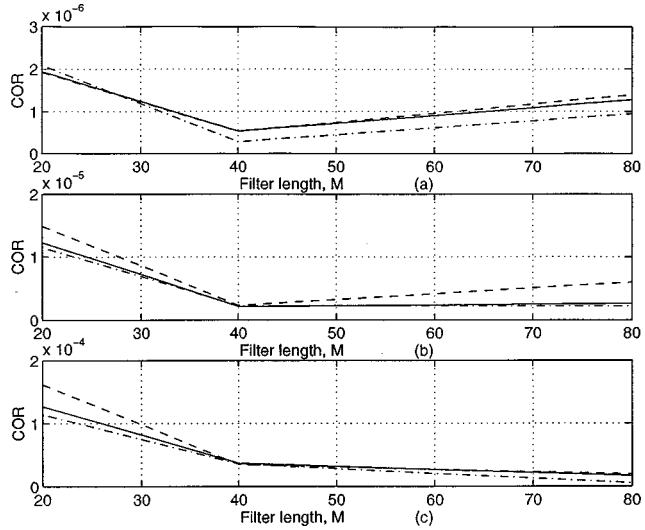


FIG. 11. For the NLMS algorithm: Performance measure COR versus filter length M for step size $\alpha=0.2$ (– – –), $\alpha=0.8$ (———), and $\alpha=1.2$ (–·–·–) *in* (*a*) 80, (*b*) 90, *and* (*c*) 100 *dB SPL of aircraft noise.*

TABLE I. SNR with the LMS algorithm ($M=40$ and $\mu=0.02$), and NLMS algorithm ($M=40$ and $\alpha=1.2$).

Noise level (dB SPL)	Simulated data		
	Input SNR	Output SNR	
...	...	LMS	NLMS
Quiet	3.0 : 1.0	3.0 : 1	3.0 : 1
80	1.1 : 1.0	1.7 : 1	2.8 : 1
90	1.0 : 2.1	1.7 : 1	2.9 : 1
100	1.0 : 7.7	1.8 : 1	2.7 : 1
Noise level (dB SPL)	Actual data		
	Input SNR	Output SNR	
...	...	LMS	NLMS
80	1.2 : 1.0	1.8 : 1	2.6 : 1
90	1.0 : 2.0	2.0 : 1	2.6 : 1
100	1.0 : 7.7	2.1 : 1	2.3 : 1

(2) Signal-to-noise ratio (SNR): In this study, the SNR (in linear form) at the input of the adaptive filter was conservatively defined as the square root of the ratio of variance of the microphone *d* signal when there is no aircraft noise for inspiratory flow rates >1 L/s, to that when the aircraft noise level is present at inspiratory flow rates <1 L/s. The SNR at the output of the adaptive filter was calculated as the square root of the ratio of variances of the recovered signal *e* during inspiratory flow rates >1 L/s to that during inspiratory flow rates <1 L/s. The difference of SNR at the input and output reflects a measure of the noise reduction capability of the configuration. Table I shows that higher SNR values are achieved at the output of the adaptive filter (both for the simulated and actual data) with NLMS algorithm than with the LMS algorithm, for essentially all cases.

C. Detailed performance comparison of LMS and NLMS algorithms

Figure 12(a) presents time records of typical lung sounds detected in a quiet chamber at microphone *d*. The lung sounds detected with microphone *d* and the reference noise detected at microphone *x* in the presence of 100 dB SPL aircraft noise are shown in Fig. 12(b) and (c), respectively. The recovered time series from the NLMS algorithm [Fig. 12(e)] is very similar to that of the pure lung sounds, with the misadjustment less than that of the LMS algorithm [Fig. 12(d)] due to the enhanced tracking ability of the NLMS algorithm. The speed of convergence is also much faster for the NLMS algorithm since it employs a normalized dynamic step size which renders the rate of convergence independent of the input signal power. The spectral features (Fig. 13) of the recovered lung sounds from the algorithms nearly overlap up to approximately 450 Hz, where the NLMS algorithm's recovered lung sounds more closely match the spectra of the pure lung sounds at higher frequencies. It can be inferred from Fig. 13 that roughly 15 dB of noise reduction is attained in 100 dB SPL of aircraft noise in the 100–600 Hz frequency range with LMS algorithm, while the NLMS algorithm provides up to 5 dB of additional noise reduction at frequencies above 450 Hz. The NLMS algorithm therefore affords higher noise reduction over wider bandwidth with a small amount of added computational complexity, and hence it is the algorithm of choice for the reduction of corrupting C-130 aircraft noise.

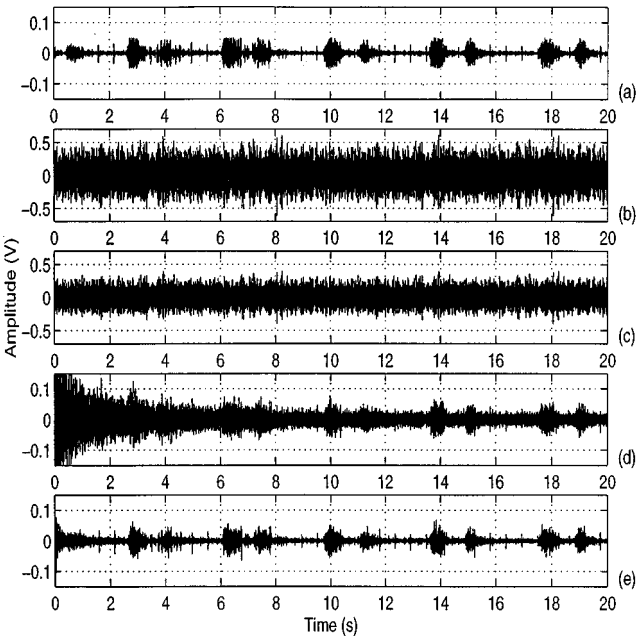


FIG. 12. For simulated data: Time record of (a) typical lung sounds in a quiet chamber, (b) noise corrupted lung sounds (microphone *d*) at 100 dB SPL of aircraft noise, (c) noise reference (microphone *x*) and recovered lung sounds with (d) LMS algorithm ($M=40$ and $\mu=0.02$), and (e) NLMS algorithm ($M=40$ and $\alpha=1.2$) ($1\,V=3.1$ Pa).

The recovered lung sounds for actual data are shown in Figs. 14 and 15, noting that the SNR is lower than that for the simulated case, as expected. This observation highlights the significance of the need for the reference noise (microphone *x*) to be highly correlated with the noise in the primary input (microphone *d*) to enhance the performance of the adaptive noise canceler. For the simulated data, the primary

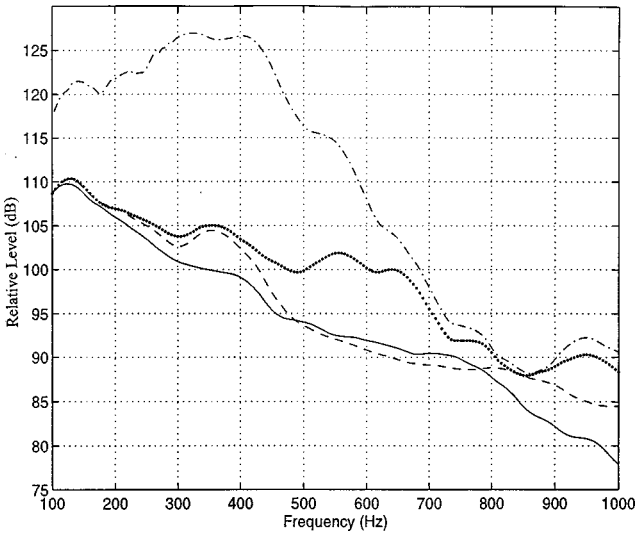


FIG. 13. For simulated data: Estimated relative power spectra of inspiratory lung sounds as measured in a quiet chamber (– – –), noise corrupted lung sounds at 100 dB SPL of aircraft noise (– · – · –), and recovered lung sounds from the LMS algorithm ($M=40$ and $\mu=0.02$) (···), and NLMS algorithm ($M=40$ and $\alpha=1.2$) (———).

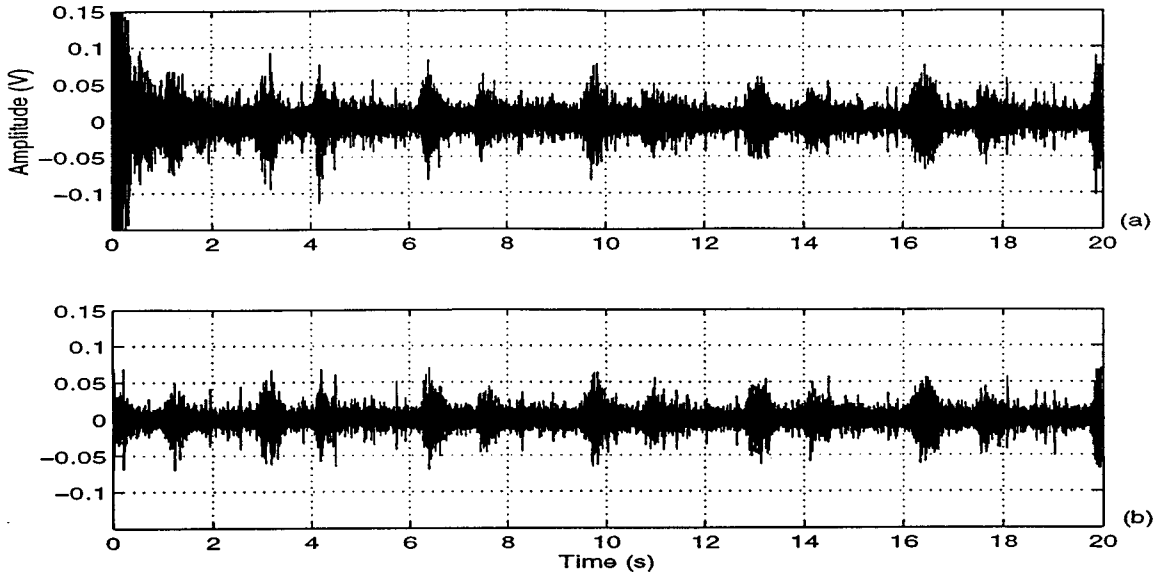


FIG. 14. For actual data: Time record of recovered lung sounds with (a) LMS algorithm ($M=40$ and $\mu=0.02$), (b) NLMS algorithm ($M=40$ and $\alpha=1.2$).

input was composed by adding the microphone d signal during apnea with the coupler placed on the thigh (in presence of the known level of the aircraft noise) to the pure lung sound signal (i.e., microphone d signal during breathing in a quiet chamber). Further, for processing of these simulated data, the same microphone d signal obtained during apnea in presence of the known level of the aircraft noise is used as the reference noise. Hence, for the simulated data, the correlation between the reference noise and the noise in the primary channel is perfect, whereas for the actual data, the primary input (microphone d) and reference noise (microphone x) signals are obtained using two different microphones. Here, the correlation between the primary input and the reference noise is governed by the transfer function between the physical locations of microphones d and x inside the stethoscope coupler. The trade-off in acoustic isolation of these microphones is to collect a reference noise signal that is least ''corrupted'' by the lung sounds and highly correlated with the noise component in the primary input. Thus the adaptive filter algorithm for the simulated data provides a better estimate of the noise in the primary input and hence results in a recovered lung sounds of higher fidelity as compared to those obtained for the actual data.

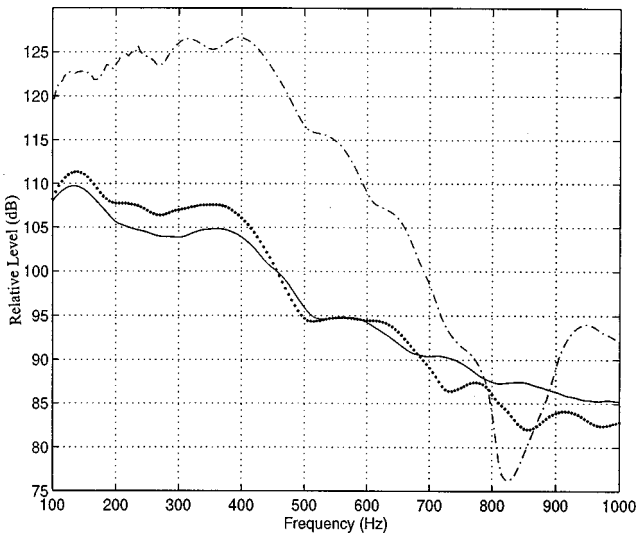


FIG. 15. For actual data: Estimated relative power spectra of noise corrupted lung sounds at 100 dB SPL of aircraft noise (– · – · –), and recovered lung sounds from the LMS algorithm ($M=40$ and $\mu=0.02$) (···), and NLMS algorithm ($M=40$ and $\alpha=1.2$) (———).

V. CONCLUSIONS

In patient transport vehicles, high ambient noise has significant spectral overlap with lung sounds, making lung sounds inaudible through a standard stethoscope and limiting the utility of conventional amplification and filtering techniques. Passive noise reduction techniques such as a coupler housing help to shield the lung sounds from high-frequency noise corruption. An adaptive noise reduction system implemented using a digital adaptive filter employing the LMS or NLMS algorithms significantly reduces the corruptive effects of low frequency noise. The proposed performance criteria of MRMS and COR are useful to determine algorithm parameters, such as convergence factor and filter length for the LMS approach. Approximately 15 dB of noise reduction over a 100–600 Hz frequency range was achieved by the LMS approach, with the somewhat more complex NLMS algorithm yielding faster convergence and up to an additional 5 dB of noise reduction. Further studies on patients during actual medical evacuation are required to refine this technology and determine the scope of its clinical utility under various breathing and noise conditions.

ACKNOWLEDGMENTS

This research was performed for the U.S. Air Force, Armstrong Laboratory under a phase II Small Business Innovation Research subcontract to University Research Engineers & Associates, Inc. The authors thank Saul B. Gelfand for his advice concerning the adaptive filter design and evaluation.

Darlington, P., Wheeler, P. D., and Powell, G. A. (**1985**). ''Adaptive noise reduction in aircraft communication systems,'' Proceedings of IEEE Conference on Acoustics, Speech and Signal Processing, Vol. 2, pp. 716–719.

Elliott, S. J., Stothers, I. M., and Nelson, P. A. (**1987**). ''A multiple error lms algorithm and its application to the active control of sound and vibration,'' IEEE Trans. Acoust., Speech, Signal Proc. **ASSP-35**, 1423–1434.

Harrison, W. A., Lim, J. E., and Elliott, S. J. (**1986**). ''A new application of adaptive noise cancellation,'' IEEE Trans. Acoust., Speech, Signal Proc. **ASSP-34**, 21–27.

Haykin, S. (**1986**). *Adaptive Filter Theory* (Prentice-Hall, Englewood Cliffs, NJ).

Poole, L. A., Warnaka, G. E., and Cutter, R. C. (**1984**). ''The implementation of digital filters using a modified widrow-hoff algorithm for the adaptive cancelation of acoustic noise,'' Proceedings of IEEE Conference on Acoustics, Speech and Signal Processing, 21.7.1–21.7.4.

Pulsipher, D. C., Boll, S. F., Rushforth, C., and Timothy, J. (**1979**). ''Reduction of nonstationary acoustic noise in speech using lms adaptive noise cancelling,'' Proceedings of IEEE Conference on Acoustics, Speech and Signal Processing, pp. 204–207.

Rodriguez, J. J., Lim, J. S., and Elliott, S. J. (**1987**). ''Adaptive noise reduction in aircraft communication systems,'' Proceedings of IEEE Conference on Acoustics, Speech and Signal Processing, pp. 169–172.

Treichler, J. R., Johnson, Jr., C. R., and Larimore, M. G. (**1987**). *Theory and Design of Adaptive Filters* (Wiley, New York).

Widrow, B., Glover, Jr., J. R., McCool, J. M., Kaunitz, J., Williams, C. S., Hearn, R. H., Zeidler, J. R., Dong, Jr., E., and Goodlin, R. C. (**1975**). ''Adaptive noise cancelling: Principles and applications,'' Proc. IEEE **63**, 1692–1716.

Widrow, B., McCool, J., Larimore, M. G., and Johnson, Jr., C. R. (**1976**). ''Stationary and nonstationary learning characteristics of the lms adaptive filter,'' Proc. IEEE **64**, 1151–1162.

Widrow, B., and Stearns, S. D. (**1985**). *Adaptive Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ).

Wodicka, G. R., Kraman, S. S., Zenk, G. M., and Pasterkamp, H. (**1994**). ''Measurement of respiratory acoustic signals—Effect of microphone air cavity depth,'' Chest **106**, 1140–1144.

Zenk, G. M. (**1994**). ''Stethoscopic detection of lung sounds in high noise environments,'' M.S. thesis, Purdue University, West Lafayette, IN.

Projection data error estimation and correction for high-resolution tomographic acoustic microscopy

Richard Y. Chiao,[a] Davis Kent, and Hua Lee
Department of Electrical and Computer Engineering, University of California, Santa Barbara, California 93106

(Received 30 June 1995; accepted for publication 15 January 1998)

In scanning tomographic acoustic microscopy (STAM), projection error correction is necessary for high-resolution tomographic reconstruction. In this paper, both phase and alignment errors are examined. These errors arise from several sources, including the quadrature receiver channels, the unknown initial phase term, and any misalignment in the rotational scan. The data acquisition process and the image formation algorithm for the STAM are reviewed, and a description of error estimation and correction is also presented. Experimental results from the STAM are included to demonstrate the capability and effectiveness of the error-removal techniques. © *1998 Acoustical Society of America.* [S0001-4966(98)02005-0]

PACS numbers: 43.60.Rw [JLK]

INTRODUCTION

The scanning tomographic acoustic microscope (STAM) is a scanning laser acoustic microscope (SLAM) enhanced with tomographic reconstruction capability. The SLAM system is a transmission mode acoustic microscope that employs continuous wave insonification.[1] It was developed for high-resolution, real-time imaging in nondestructive testing and biomedical tissue analysis.[2] At an operating frequency of 100 MHz, the SLAM is capable of producing high-resolution images of thin (relative to the acoustic wavelength) specimens. When thick specimens are imaged, resolution is degraded due to diffraction and the overlapping of different depth planes.[3]

In order to achieve high-resolution subsurface imaging for thick objects, the data acquisition system of the SLAM was modified by the addition of a quadrature receiver for the detection of the complex wave field distribution.[4] With the complex wave field, the effects of diffraction can be compensated through numerical holographic reconstruction,[5] and tomographic reconstruction can be performed by coherent superposition of holographic images.[6]

The detection of the phase information, combined with the tomographic summation process, gives the STAM a superior imaging capability as compared with the SLAM. During the tomographic image formation process, data errors occur in various forms at various stages of the operation. In this paper, we study the data errors from the quadrature receiver channels, illumination initial phase variation among projections, and alignment error associated with the rotational scan. Subsequently, we present error estimation and removal techniques for resolution enhancement.

In Sec. I we review the basic data acquisition process of the SLAM and the tomographic reconstruction algorithm. The quadrature data error is examined in Sec. II. The degradation due to the error terms is modeled and observed by experiments, and subsequently an algorithm is developed to estimate and remove the phase error. In Sec. III the illumination phase error is studied. With the phase errors corrected, tomographic reconstruction becomes effective and can be well demonstrated with experimental results of multiple-frequency tomography. In Sec. IV, the projection registration problem in tomographic superposition is discussed and correction schemes are implemented for multiple-angle tomographic reconstruction.

I. STAM DATA ACQUISITION AND RECONSTRUCTION

In this section, the data acquisition process of the SLAM is briefly reviewed and the tomographic image reconstruction algorithm is presented. Figure 1(a) and (b) shows the modified SLAM data acquisition system with the SLAM receiver replaced by the quadrature receiver. The object plane is located at z_0, the incident planar wave field is denoted by $u(x,y,z)$, and the transmitted wave field is $v(x,y,z)$. After being modulated by the specimen, the transmitted wave field causes dynamic surface ripples on the gold-plated coverslip.[7] A laser beam is raster scanned over the reflective surface to detect the wave field information. At each position in the raster scan, the laser beam is angularly deflected by the surface ripples. The instantaneous angular deflection is converted to an intensity signal by a knife-edge detector.[7] At this point the spatial wave field distribution is encoded in a temporal intensity-modulated laser signal. A photodetector placed after the knife edge converts the laser light signal into an electrical signal, which is then fed into the quadrature receiver.[4] The output of the quadrature receiver is digitized for computer processing.

The insonification wave can be written as

$$u(x,y,z)=a_0\ \exp(j2\pi(f_{xi}x+f_{yi}y+f_{zi}z)), \tag{1}$$

where a_0 is the amplitude of the plane wave and f_{xi}, f_{yi}, f_{zi} are the spatial frequencies of the plane wave. The spatial frequencies depend on the insonification angle of the plane wave as well as the acoustic wavelength. At the object plane $z=z_0$, the wave field is modulated by the object transmit-

[a]Currently with GE Corporate Research and Development, P.O. Box 8, Schenectady, NY 12301.

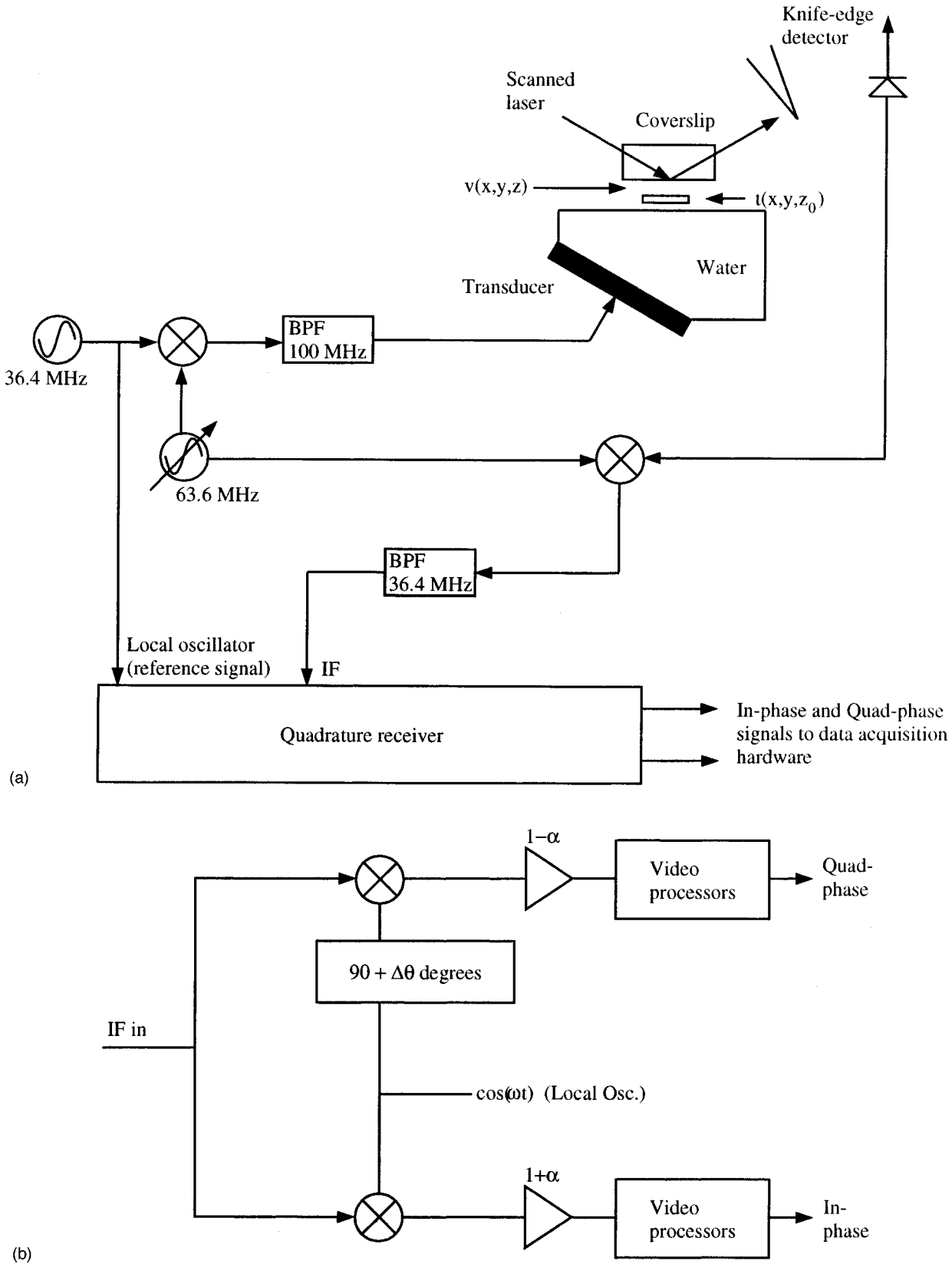


FIG. 1. (a), (b) SLAM data acquisition system.

tance $t(x,y)$, and the transmitted wave field is given by

$$v(x,y,z_0)=u(x,y,z_0)t(x,y). \tag{2}$$

In the Fourier domain Eq. (2) can be written as

$$V(f_x,f_y;z_0)=a_0\ \exp(j2\pi f_{zi}z_0)T(f_x-f_{xi},f_y-f_{yi}). \tag{3}$$

Thus the transmitted wave field is proportional to the object transmittance function modulated by f_{xi} and f_{yi}. The wave

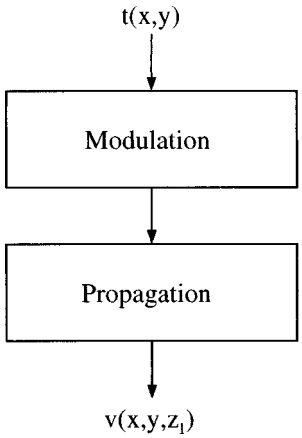


FIG. 2. Block diagram for the data acquisition process.

field $v(x,y,z_1)$ detected at the receiving plane is related to $v(x,y,z_0)$ by

$$V(f_x,f_y;z_1)=V(f_x,f_y;z_0)H(f_x,f_y;z_1-z_0), \tag{4}$$

where $V(f_x,f_y)$ is the Fourier transform of $v(x,y)$, and H is the propagation transfer function given by

$$H(f_x,f_y;z_1-z_0)=\begin{cases}\exp\left[j2\pi(z_1-z_0)\left(\frac{1}{\lambda^2}-f_x^2-f_y^2\right)^{1/2}\right], & f_x^2+f_y^2<\frac{1}{\lambda^2},\\ 0, \quad \text{otherwise,}\end{cases} \tag{5}$$

where λ is the acoustic wavelength. The propagation transfer function is a low-pass function because of the loss of evanescent waves. Figure 2 is the block diagram for the data acquisition process.

In tomographic processing, projections with different insonification angles or acoustic wavelengths are combined to construct the tomographic image. For multiple-angle tomography, f_{xi}, f_{yi}, and f_{zi} vary between projections, and for multiple-frequency tomography, λ varies between projections. Given a set of N projections $v_k(x,y,z_1)$, $k=1,2,...,N$ the subsurface plane $t(x,y)$ can be reconstructed using the backward propagation algorithm. The reconstruction equation for the algorithm is given by

$$\hat{t}(x,y)=\frac{\Sigma_{k=1}^{N}\, u_k^*(x,y,z_0)v_k(x,y,z_0)}{\Sigma_{k=1}^{N}\,|u_k(x,y,z_0)|^2}, \tag{6}$$

which is the least-squares estimate of the object transmittance given the incident wave field and the scattered wave field.[6,8]

To implement Eq. (6), it is necessary to obtain the incident wave field and the transmitted wave field at the plane of interest. One approach is to record the received wave field with and without the specimen. The transmitted wave field at the object plane, $v(x,y,z_0)$, can be estimated by backpropagating the received wave field with the specimen present

$$\hat{V}(f_x,f_y;z_0)=V(f_x,f_y;z_1)H(f_x,f_y;-(z_1-z_0)), \tag{7}$$

and the incident wave field at the object plane can be obtained by backpropagating the received wave field without the specimen. The above method requires two complex wave fields to be recorded for each projection. Alternatively, we can accurately approximate the incident wave field by a plane wave, and obtain

$$\hat{t}(x,y)=\sum_{k=1}^{N}\frac{1}{a_0^{(k)}}\hat{v}_k(x,y,z_0)\exp(-j2\pi(f_{xi}^{(k)}x+f_{yi}^{(k)}y+f_{zi}^{(k)}z_0)), \tag{8}$$

where the k subscripts and superscripts are for different projections. Since the transmitted wave field $\hat{v}(x,y,z_0)$ is modulated by $f_{xi}^{(k)}$, $f_{yi}^{(k)}$, and $f_{zi}^{(k)}$, multiplication by $\exp(-j2\pi(f_{xi}^{(k)}x+f_{yi}^{(k)}y+f_{zi}^{(k)}z_0))$ in Eq. (8) corresponds to demodulation. Consequently, the image formation algorithm can be implemented by superimposing the backpropagated, demodulated, and normalized projections, $w_k(x,y)$, such that

$$\hat{t}(x,y)=\sum_{k=1}^{N}w_k(x,y), \tag{9}$$

where

$$w_k(x,y)=\frac{1}{u_0^{(k)}}\hat{v}_k(x,y,z_0)\exp(-j2\pi(f_{xi}^{(k)}x+f_{yi}^{(k)}y)), \tag{10}$$

and $u_0=a_0\exp(j2\pi f_{zi}z_0)$. Using Eqs. (2), (4), and (7) in Eq. (10), we formulate it as

$$W_k(f_x,f_y)=\begin{cases}T(f_x,f_y), & (f_x+f_{xi}^{(k)})^2+(f_y+f_{yi}^{(k)})^2<1/\lambda_k^2,\\ 0, & \text{otherwise,}\end{cases} \tag{11}$$

which shows that each projection recovers spatial frequencies of the unknown object transmittance within a circular region centered at $(-f_{xi}^{(k)},-f_{yi}^{(k)})$. The superposition of different projections improves the signal-to-noise ratio as well as the resolution by extending the spectral coverage. In practice, there exists significant information redundancy between the projections which can be utilized for error estimation and correction.

II. QUADRATURE PHASE ERROR ESTIMATION AND CORRECTION

In tomographic acoustic microscopy, the first type of phase error occurs during data acquisition, in the quadrature receiver channels. The phase error can be due to either noise fluctuations or crosstalk in the receiver, and it causes significant degradation in the tomographic reconstruction. Here, we first model the phase error as a spatially varying quantity and derive observable effects of the error in a single tomographic projection. Subsequently, an algorithm is developed to estimate the phase error, and the data are then corrected. Finally, we apply the algorithm to experimental data and verify the reduction in phase error and image degradation.

A. Phase error model

From Eq. (4), we denote the received wave field as

$$v(x)=a(x)\exp(j2\pi f_{xi}x), \tag{12}$$

where we remove the y dependence for notational convenience, and $a(x)$ is given by

$$a(x)=F^{-1}\{u_0T(f_x)H(f_x-f_{xi};z_1-z_0)\}, \tag{13}$$

where F^{-1} is the inverse Fourier transform operator. Ideally, the quadrature receiver detects

$$p_r(x)=\text{Real}\{a(x)\exp(j2\pi f_{xi}x)\}$$
$$=|a(x)|\cos(2\pi f_{xi}x+\phi(x)) \tag{14a}$$

and

$$p_i(x)=\text{Im}\{a(x)\exp(j2\pi f_{xi}x)\}$$
$$=|a(x)|\sin(2\pi f_{xi}x+\phi(x)), \tag{14b}$$

where $p_r(x)$ and $p_i(x)$ are the real and imaginary components, respectively, of the wave field present on the coverslip. In practice, a relative phase error, $\Delta\theta(x)$, causes the two output components of the quadrature receiver to become

$$p_r'(x)=|a(x)|\cos(2\pi f_{xi}x+\phi(x)-\Delta\theta(x)/2) \tag{15a}$$

and

$$p_i'(x)=|a(x)|\sin(2\pi f_{xi}x+\phi(x)+\Delta\theta(x)/2). \tag{15b}$$

As a result, the actual recorded wave field is

$$p'(x)=p_r'(x)+jp_i'(x)$$
$$=\cos(\Delta\theta(x)/2)a(x)\exp(j2\pi f_{xi}x)$$
$$+\sin(\Delta\theta(x)/2)a^*(x)\exp(-j2\pi f_{xi}x). \tag{16}$$

Equation (16) shows that the phase error introduces an additional spectral component centered at $-f_{xi}$. The effect of the phase error can be observed by examining the intensity of the recorded wave field

$$|p'(x)|^2=|a(x)|^2+|a(x)|^2\ \sin(\Delta\theta(x))$$
$$\times\sin(4\pi f_{xi}x+2\phi(x)), \tag{17}$$

where the phase error introduces a background fringe pattern with spatial frequency $2f_{xi}$ in the detected wave field.

B. Phase error estimation

To obtain an estimate of the phase error $\Delta\theta(x)$, we first write Eq. (14) as

$$p_r'(x)=\tfrac{1}{2}e^{-j\Delta\theta(x)/2}a(x)\exp(j2\pi f_{xi}x)$$
$$+\tfrac{1}{2}e^{j\Delta\theta(x)/2}a^*(x)\exp(-j2\pi f_{xi}x) \tag{18a}$$

and

$$p_i'(x)=\frac{1}{j2}\,e^{j\Delta\theta(x)/2}a(x)\exp(j2\pi f_{xi}x)$$
$$-\frac{1}{j2}\,e^{-j\Delta\theta(x)/2}a^*(x)\exp(-j2\pi f_{xi}x). \tag{18b}$$

Since a relative phase error cannot be simply estimated from a pair of real numbers, it is not possible to use the outputs of the quadrature receiver to directly estimate the phase error. Instead, we take the single side band of the quadrature receiver outputs. Equation (18) then becomes

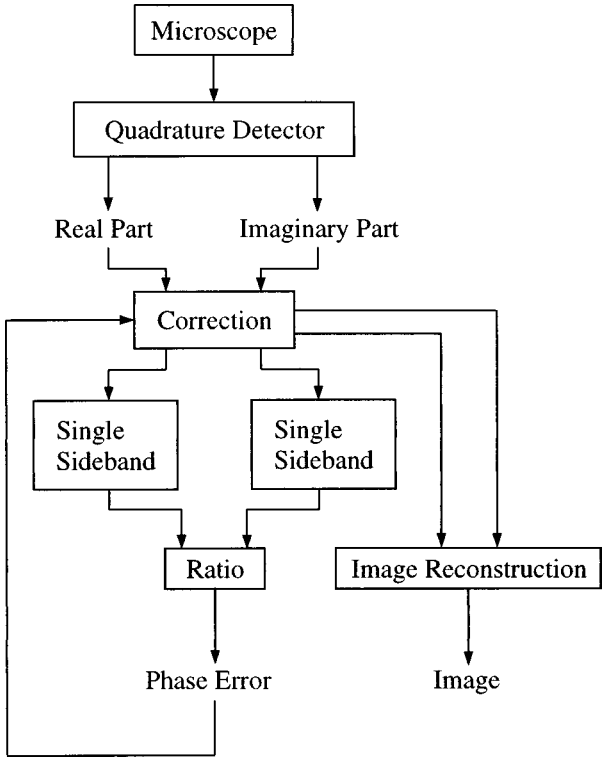


FIG. 3. Block diagram for the recursive quadrature error correction algorithm.

$$p_{r+}'(x)=\tfrac{1}{2}e^{-j\Delta\theta(x)/2}a(x)\exp(j2\pi f_{xi}x) \tag{19a}$$

and

$$p_{i+}'(x)=\frac{1}{j2}\,e^{j\Delta\theta(x)/2}a(x)\exp(j2\pi f_{xi}x), \tag{19b}$$

where the + subscript denotes the positive side band. Subsequently, the phase error can be estimated as

$$e^{j\Delta\theta(x)}=\frac{jp_{i+}'(x)}{p_{r+}'(x)}. \tag{20}$$

Normally, we use only the upper side band to estimate the phase error for better signal-to-noise ratio. In theory, the lower side bands can also be used to improve the phase error estimate.

C. Wave field correction

To derive the correction operation, we first consider the data pair

$$p_r'=A\ \cos(\alpha-\beta),\quad p_i'=A\ \sin(\alpha+\beta), \tag{21}$$

where α and β are constants. We wish to obtain the corrected data in the form of

$$\hat{p}_r=A\ \cos(\alpha),\quad \hat{p}_i=A\ \sin(\alpha). \tag{22}$$

Rearranging Eqs. (21) and (22), we obtain the matrix relationship

$$\begin{bmatrix}p_r'\\ p_i'\end{bmatrix}=\begin{bmatrix}\cos(\beta) & \sin(\beta)\\ \sin(\beta) & \cos(\beta)\end{bmatrix}\begin{bmatrix}\hat{p}_r\\ \hat{p}_i\end{bmatrix}. \tag{23}$$

The correction matrix operation is then in the form of

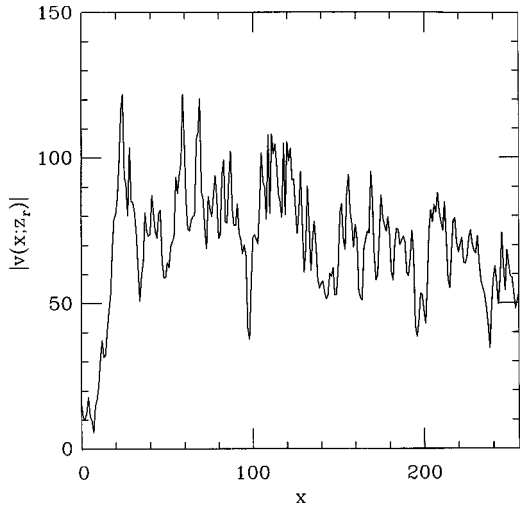


FIG. 4. Magnitude of a complex data sequence from the SLAM.

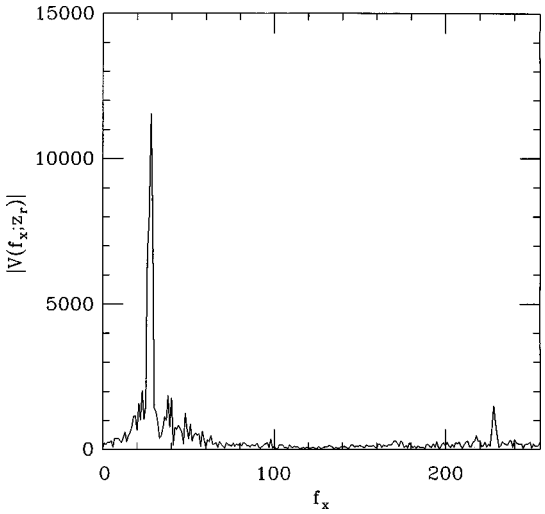


FIG. 5. Magnitude of the spectrum of the complex sequence.

$$\begin{bmatrix}\hat{p}_r\\ \hat{p}_i\end{bmatrix}=\begin{bmatrix}\cos(\beta) & \sin(\beta)\\ \sin(\beta) & \cos(\beta)\end{bmatrix}^{-1}\begin{bmatrix}p_r'\\ p_i'\end{bmatrix}$$
$$=\frac{1}{\cos(2\beta)}\begin{bmatrix}\cos(\beta) & -\sin(\beta)\\ -\sin(\beta) & \cos(\beta)\end{bmatrix}\begin{bmatrix}p_r'\\ p_i'\end{bmatrix}. \tag{24}$$

Letting $\beta=\Delta\theta(x)/2$, we can obtain the corrected signal for each position x

$$\begin{bmatrix}\hat{p}_r(x)\\ \hat{p}_i(x)\end{bmatrix}=\frac{1}{\cos(\Delta\theta(x))}\times\begin{bmatrix}\cos(\Delta\theta(x)/2) & -\sin(\Delta\theta(x)/2)\\ -\sin(\Delta\theta(x)/2) & \cos(\Delta\theta(x)/2)\end{bmatrix}\begin{bmatrix}p_r'(x)\\ p_i'(x)\end{bmatrix}. \tag{25}$$

As previously stated, the accuracy of the error estimate is governed by the variation of the phase error. For phase error variation with significant bandwidth, the error may not be completely removed in a single step. Therefore, for these cases, an iterative process is applied. The phase error is reduced at each iteration, and the algorithm terminates when the error reaches an acceptable level. Figure 3 shows the block diagram of the recursive correction algorithm.

D. General case

To simplify the derivation for the phase error estimator presented in Sec. II B, the amplitudes of the quadrature signals were assumed to be ideal. In practice, the amplitude is also contaminated by error and this will affect the phase error estimation. For the general case where both the magnitude and phase error are present, the signal model corresponding to Eq. (15) becomes

$$p_r'(x)=(|a(x)|\mp|\Delta a(x)/2|)\cos(2\pi f_{xi}x+\phi(x)-\Delta\theta(x)/2), \tag{26a}$$

$$p_i'(x)=(|a(x)|\pm|\Delta a(x)/2|)\sin(2\pi f_{xi}x+\phi(x)+\Delta\theta(x)/2), \tag{26b}$$

where $\Delta a(x)$ is the relative magnitude error between the real and imaginary components. The estimation and correction for the general case is similar to Eq. (25). With slight modification to the ideal magnitude case, the general case correction is

$$\begin{bmatrix}p_r(x)\\ p_i(x)\end{bmatrix}=\frac{1}{\det}\begin{bmatrix}\left(1+\frac{\alpha(x)}{2}\right)\cos(\Delta\theta(x)/2) & -\left(1-\frac{\alpha(x)}{2}\right)\sin(\Delta\theta(x)/2)\\ -\left(1+\frac{\alpha(x)}{2}\right)\sin(\Delta\theta(x)/2) & \left(1-\frac{\alpha(x)}{2}\right)\cos(\Delta\theta(x)/2)\end{bmatrix}\begin{bmatrix}p_r'(x)\\ p_i'(x)\end{bmatrix}, \tag{27}$$

where

$$\det=\left[1-\left(\frac{\alpha(x)}{2}\right)^2\right]\cos(\Delta\theta(x)). \tag{28}$$

E. Experimental results

The algorithm of Sec. D was applied to experimental data from the SLAM. Figures 4 and 5 show the magnitude of a complex data sequence from the SLAM and its corresponding spectrum, respectively. The small peak in the lower side band is due to the phase error, as predicted by Eq. (11). Figure 6 shows the estimate of the phase error distribution of the wave field data sample sequence, while Fig. 7 presents the magnitude of the corrected complex data sequence. Figure 8(a)–(c) shows the phase error after one, three, and five iterations, respectively, to demonstrate the effectiveness of the algorithm.

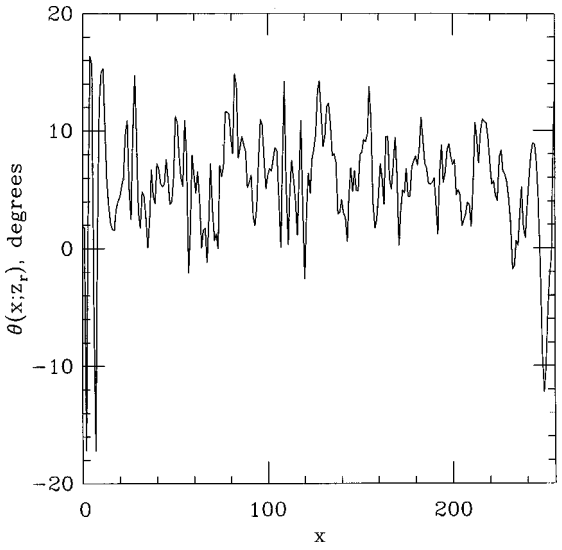


FIG. 6. Estimate of the phase error distribution.

III. INITIAL PHASE CORRECTION

An additional cause of degradation in the tomographic reconstruction is the presence of an unknown phase offset between projections. The phase offset can be due to the initial phase of the illumination wave field of each projection or velocity changes in the water coupler due to temperature fluctuations. In this section, we examine the initial phase error between projections. We first model the phase error in the received wave field, then derive an algorithm to estimate the phase error using the overlapping region of the spectrum of two projections. Finally, we present experimental results to demonstrate the effectiveness of the algorithm.

A. Initial phase model and estimation

To estimate the relative phase offset in tomographic projections, we first rewrite Eq. (11) as follows:

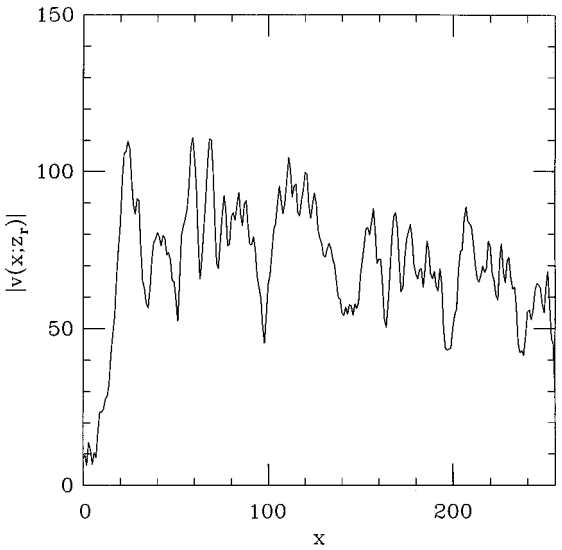


FIG. 7. Magnitude of corrected complex data sequence.

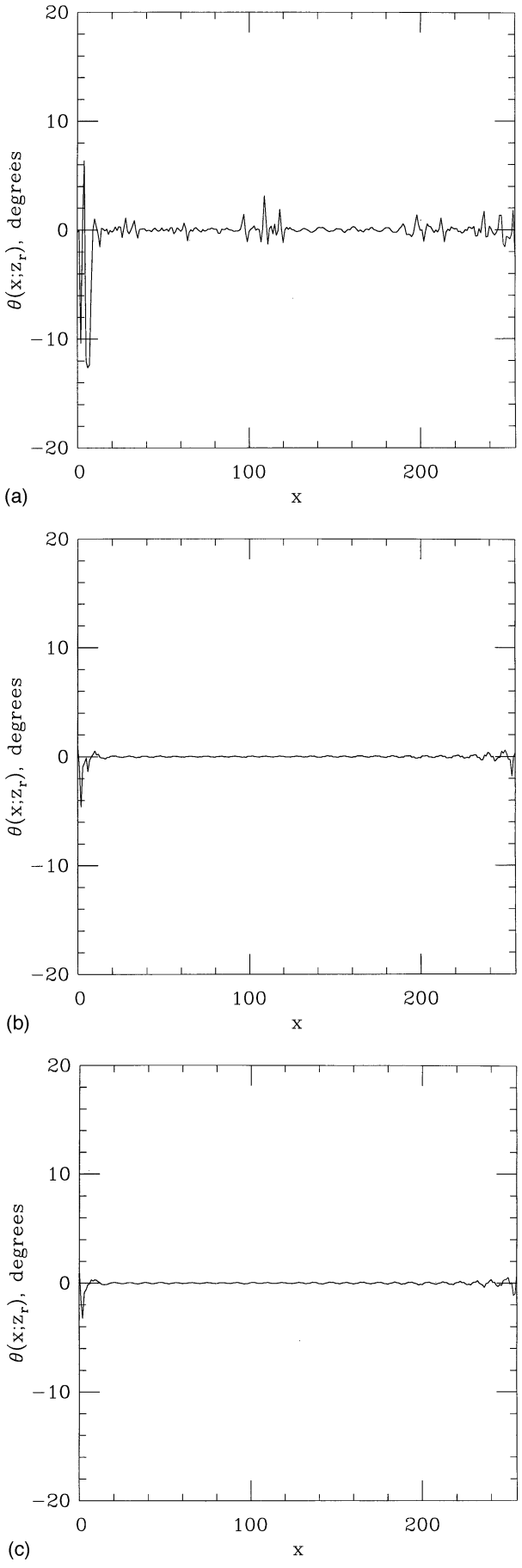


FIG. 8. (a)–(c) Phase error distribution after 1, 3, and 5 iterations.

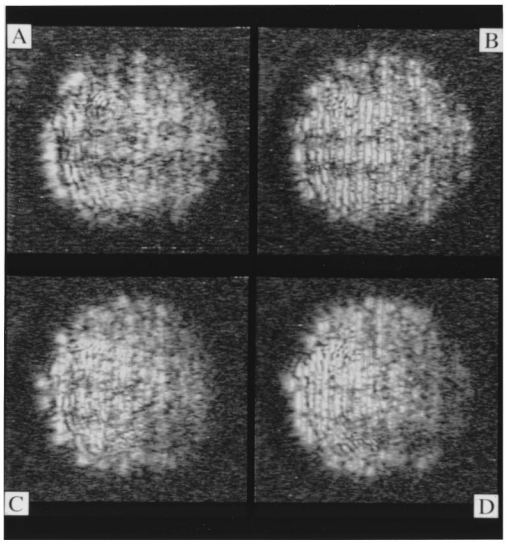


FIG. 9. (A)–(D) Magnitude images of even frequency projections.

$$W_k(f_x,f_y) = \begin{cases} \exp(j\Omega_k)T(f_x,f_y), & (f_x+f_{xi}^{(k)})^2+(f_y+f_{yi}^{(k)})^2<1/\lambda_k^2, \\ 0, \quad \text{otherwise} \end{cases} \tag{29}$$

where Ω_k is the unknown initial phase. Notice that $W(f_x,f_y)$ is nonzero only within a circular passband due to wave propagation, which is different for each projection. For multiple-angle projections, the passbands are centered at different locations, and for multiple-frequency projections, the radii of the passbands also change. The overlapping region of the passbands allows estimation of the phase offset between projections. Using Eq. (29), the phase offset between two projections m and n is expressed as

$$W_m(f_x,f_y)=\exp(j(\Omega_m-\Omega_n))W_n(f_x,f_y) \tag{30}$$

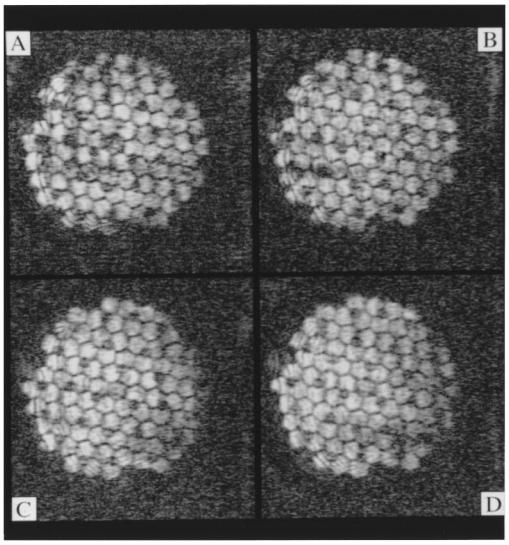


FIG. 10. (A)–(D) Magnitude images of backpropagated and demodulated projections.

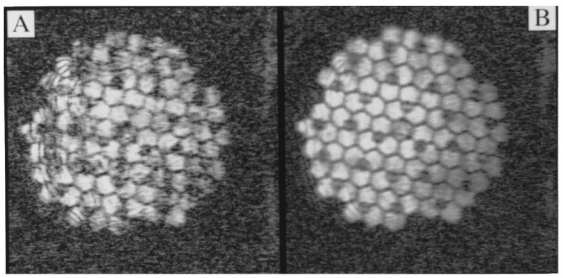


FIG. 11. (A), (B) Multiple-frequency tomographic reconstructions without and with phase correction.

which is valid for f_x and f_y in the overlapping region Γ of the passbands of $W_m(f_x,f_y)$ and $W_n(f_x,f_y)$. The least-squares estimate of $\exp(j(\Omega_m-\Omega_n))$ is computed from Eq. (30) as

$$\exp(j(\Omega_m-\Omega_n))=\frac{\int_\Gamma\int W_n^*(f_x,f_y)W_m(f_x,f_y)df_x\,df_y}{\int_\Gamma\int|W_n(f_x,f_y)|^2df_x\,df_y}. \tag{31}$$

Substituting Eq. (30) in the denominator of Eq. (29) yields as

$$\exp(j(\Omega_m-\Omega_n))=\frac{\int_{-\infty}^{\infty}\int W_n^*(f_x,f_y)W_m(f_x,f_y)df_x\,df_y}{\int_{-\infty}^{\infty}\int|W_n(f_x,f_y)||W_m(f_x,f_y)|df_x\,df_y}. \tag{32}$$

B. Experimental results

The algorithm of Sec. A is applied to four tomographic projections from the SLAM with insonification frequencies of 97 to 103 MHz, with a 2-MHz spacing. Figure 9(A)–(D) shows the magnitude images of the frequency projections. The field of view is 2 mm by 2 mm. The specimen is a honeycomb test grid located 0.5 mm beneath the detection plane ($z_1-z_0=0.5$ mm). Figure 10(A)–(D) shows the magnitude images of the backpropagated and demodulated wave fields. Figure 11(A) and (B) presents the tomographic superposition of the four images of Fig. 10 without and with initial phase correction, respectively.

IV. ALIGNMENT ERROR AND CORRECTION

Multiple-frequency and multiple-angle tomography require correction of the phase errors. Multiple-angle tomography, because of the rotation associated with the scan, requires additional processing to insure correct spatial registration and alignment of the projections. In this section, three alignment tasks are addressed:

(1) Sampling aspect ratio imbalance;
(2) Scanning skew; and
(3) Projection pose identification.

The scan is performed by rotating the specimen instead of the transducer, so estimation and compensation of these items becomes necessary. For multiple-angle tomographic imaging, each projection must be subsequently rotated to a common orientation. Failure to perform accurate compensation prevents a proper spatial alignment and results in degraded reconstructions.

A. Sampling ratio imbalance

The wave field impinging on the coverslip is detected by a laser scanning in a rectangular pattern with sample spacings of Δx and Δy in the x and y directions, respectively. Since the laser scan is not exactly adjusted, the sampling aspect ratio, $\alpha = \Delta x/\Delta y$, is not unity, and consequently the images appear expanded or contracted. If the projections do not involve rotation, this distortion does not seriously degrade the images because each projection is distorted consistently. With the subsequent rotation, the projections are expanded in different directions, preventing proper superposition.

The sampling aspect ratio can be estimated in the spatial domain using corresponding feature points in two projections.[9] Let $\mathbf{q}_i$ and $\mathbf{q}_i'$, $i = 1,2,...,N$ be a set of N corresponding feature points obtained from two projections related by the rotation matrix, $\mathbf{R}$, and the translation vector, $\mathbf{t}$. Then $\mathbf{q}_i$ and $\mathbf{q}_i'$ are related by

$$\mathbf{q}_i' = \mathbf{S}(\mathbf{R}\mathbf{S}^{-1}\mathbf{q}_i + \mathbf{t}), \tag{33}$$

where $\mathbf{S}$ is defined as

$$\mathbf{S} = \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix}. \tag{34}$$

The 2×2 covariance matrix, $\mathbf{Q}$, is defined as

$$\mathbf{Q} = \frac{1}{N}\sum_{i=1}^{N} (\mathbf{q}_i - \boldsymbol{\mu}_q)(\mathbf{q}_i - \boldsymbol{\mu}_q)^T, \tag{35}$$

where $\boldsymbol{\mu}_q$ is the centroid of the feature points of the first image. $\mathbf{Q}'$ is similarly defined. The two covariance matrices are then related by the equation

$$\mathbf{S}^{-1}\mathbf{Q}'\mathbf{S}^{-1} = \mathbf{R}\mathbf{S}^{-1}\mathbf{Q}\mathbf{S}^{-1}\mathbf{R}^T. \tag{36}$$

Equation (36) can be solved[9] for α, yielding

$$\alpha = \sqrt{\frac{q_{11}' - q_{11}}{q_{22} - q_{22}'}}, \tag{37}$$

where q_{ij} and q_{ij}' are the (i,j)th elements of $\mathbf{Q}$ and $\mathbf{Q}'$, respectively. Assuming the scanning system is stable, α remains unchanged and one estimate will suffice. Using this estimate, frequency domain interpolation techniques are employed to correct for the sampling error by expanding or contracting the image in the corresponding direction. After the compensation, the effective sample spacing becomes equal in both directions and numerical rotation can be performed without distortion.

B. Scanning skew

In addition to the sampling aspect ratio inaccuracy, the laser scanning pattern is not perfectly rectangular, as assumed. Instead, the laser scan forms a parallelogram with angles deviating slightly from 90 degrees. As a result, projections will be skewed, again preventing proper superposition after rotation. Several approaches for error estimation and correction are possible. One method matches corresponding feature points in two projections rotated 90 degrees relative to one another, so that simple spatial domain techniques can be used to estimate the error, which indicates the amount of skew. This method requires the use of two projections, accurate rotation of the specimen, and identification of corresponding feature points. A more robust method, which does not require rotation of the specimen or feature point correspondence, is employed instead. This method makes use of the Hough transform, and is described below.

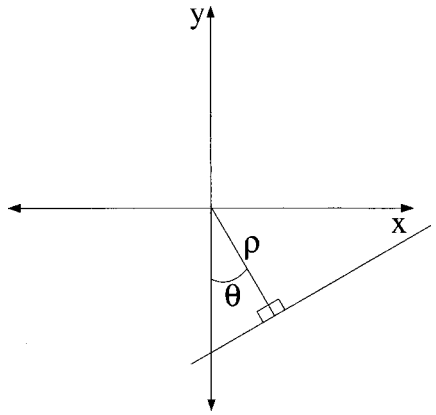


FIG. 12. Illustration of parameters ρ and θ used in the Hough transform for the detection of lines.

The STAM is equipped with an optical sensor that is in perfect registry with the acoustic sensor, permitting estimation of the skew in the acoustic data from the optical image. An optical target consisting of a square pattern is placed on the acoustic stage and illuminated by the laser scan. The optical image of the target is captured and used as the input to a Hough transform[10] which detects the lines corresponding to adjacent sides of the square target. The apparent angle, ϕ, between these lines is calculated. In the presence of a skewed scanning pattern, ϕ will deviate slightly from the correct value of 90 degrees. This angular error needs to be removed for accurate image formation. Since correction of the microscope hardware is complex, all acoustic data are corrected by software interpolation techniques.

The Hough transform may be used to locate any geometrical shape that can be described by an equation of the form

$$g(\mathbf{x},\mathbf{c}) = 0. \tag{38}$$

Here $\mathbf{x} = (x,y)$ are the space indices of the image, and $\mathbf{c}$ is a vector of coefficients or parameters. Since the current application requires detection of lines, the linear expression

$$g(\mathbf{x},\mathbf{c}) = x\,\cos(\theta) + y\,\sin(\theta) - \rho, \tag{39}$$

is used where $\mathbf{c} = (\rho,\theta)$ is defined in Fig. 12. To detect lines in the image, features are first extracted by applying an amplitude threshold. Each pixel greater than the threshold is considered to be a possible member of a line in the image. Since an infinite number of lines may be drawn through each of these pixels, there exist an infinite number of solutions to Eq. (38) for each pixel. The Hough transform represents these solutions in the $\mathbf{c}$-parameter space using an accumulator array, $A(\mathbf{c})$, where the elements of $A(\mathbf{c})$ are referenced by the components of $\mathbf{c}$. The accumulator array is used to form a histogram of the solutions of Eq. (38). After transformation of the image, the array element $A(\mathbf{c}^*)$ contains the

number of pixels in the image that are on the line described by $\mathbf{c}^*$. Pixels which are colinear result in large values, or peaks, in the accumulator array. The values of ρ and θ corresponding to these peaks are used to locate the adjacent sides of the square test target so that the angle between them can be calculated. The angular deviation indicates the amount of skew present in the laser scan.

Searching the accumulator array shows the angle between the horizontal edge and the vertical edge of the square to be $\phi = 92$ degrees, indicating an error of $\phi_e = 2$ degrees. As a result, a horizontal line will appear in the acoustic image with a downward slope of $-\tan(2\pi\phi_e)$. To compensate for the skew, the one-dimensional Fourier transform $V(x, f_y; z_r)$ of the acoustic data is formed in the vertical direction. A phase delay term is multiplied by each column, producing

$$\hat{V}(x, f_y) = \exp(j2\pi f_y x \tan(2\pi\phi_e)) V(x, f_y). \tag{40}$$

The inverse one-dimensional Fourier transform of $\hat{V}(x, f_y)$ is computed to obtain the undistorted image, $\hat{v}(x, y)$. The frequency domain technique is employed because of the automatic interpolation that occurs in the transformation process. In contrast, spatial domain techniques such as simply shifting the data by an integer number of pixels will not be sufficient. Therefore, the frequency domain technique is selected for its superior spectral characteristics.

C. Projection pose identification

Accurate knowledge of the rotation matrix, **R**, and the translation vector, **t**, is essential for high-resolution tomographic superposition. Incorrectly aligned multiple-angle projections can cause significant degradation in the reconstruction. Several methods have been studied for estimation of these imaging parameters. The first method utilizes the redundant information in the overlapping regions of the spatial frequency spectrums of each projection. Instead of extracting feature points in the Fourier domain, an algorithm has been developed to use the continuous distribution of the spectrum.[9] This method has been successfully employed, but some limitations exist. Using this method, the estimation of the parameters relies on the accuracy of the depth of the structures in the specimen. If the depth is incorrectly estimated,[11,12] the estimation of the rotation matrix and translation vector can be erroneous.

An alternative method has been developed for the estimation of the translation vector, or, equivalently, the center of rotation of the specimen using the optical image of the mirrored coverslip. The rotation angle is assumed known as a result of precise control of the rotational hardware.

In a typical configuration, the coverslip placed on the specimen is opaque. Therefore, the optical sensor detects only the image of the mirrored surface of the coverslip, and not the structure of the specimen under investigation. A small identifying mark is placed on this surface such that it is always in the optical sensor's field of view. As the specimen rotates, the mark rotates and optical images are captured at each angle of insonification. After data acquisition is complete, the optical images are superimposed to form one composite optical image. Since the coverslip mark traces out a circular path, the composite image contains a circular shape. Detection of the circular pattern is often complicated by the presence of noise caused by other features on the coverslip and ambient light conditions.

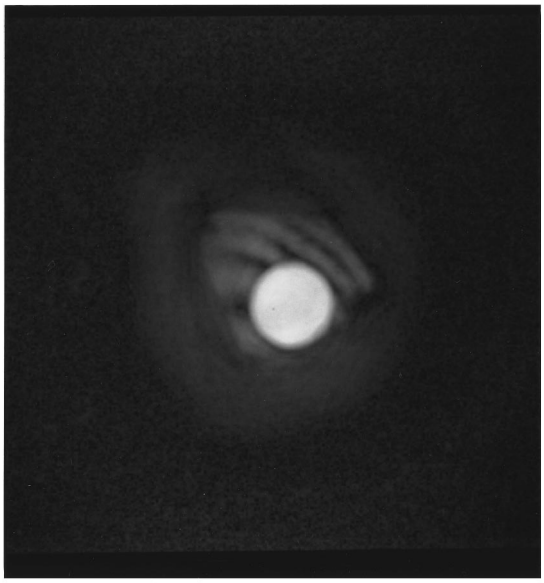


FIG. 13. Reconstruction of circular aperture using 36 multiple angle projections.

A two-step algorithm has been developed that is less sensitive to noise, and accurately detects the circular pattern and estimates the coordinates of the circle's center, which correspond to the center of rotation. The first step employs a Hough transform that detects circular shapes using the function

$$g(\mathbf{x}, \mathbf{c}) = (x - c_1)^2 + (y - c_2)^2 - c_3, \tag{41}$$

where $\mathbf{x} = (x, y)$ are the space indices of the image, c_1 and c_2 are the coordinates of the circle's center, and c_3 is the radius. The algorithm functions similarly to the line detection algorithm described earlier, but is more computationally intensive due to the larger dimension of the parameter space. The algorithm can be made computationally efficient by restricting image analysis to a small area.

After the circular Hough transform of the composite image has been obtained, the location, $\mathbf{c}_m = (c_{1,m}, c_{2,m}, c_{3,m})$, of the maximum value of this array provides an initial estimate of the center of rotation. To improve this estimate, a second step is employed. Using the known width of the circular shape, the values of the accumulator array are summed over the radius parameter, c_3, to form a two-dimensional array,

$$\hat{\mathbf{A}}(c_1, c_2) = \sum_{c_3 = c_{3,m} - W/2}^{c_{3,m} + W/2} A(\mathbf{c}), \tag{42}$$

where W is the expected width of the curve formed by the coverslip mark. The maximum value of $\hat{\mathbf{A}}(c_1, c_2)$ provides an improved estimate of the center of rotation. Without this second step, the estimate of the center may be in error by as much as W pixels, causing significant reconstruction degradation. Experimental results indicate that with this second

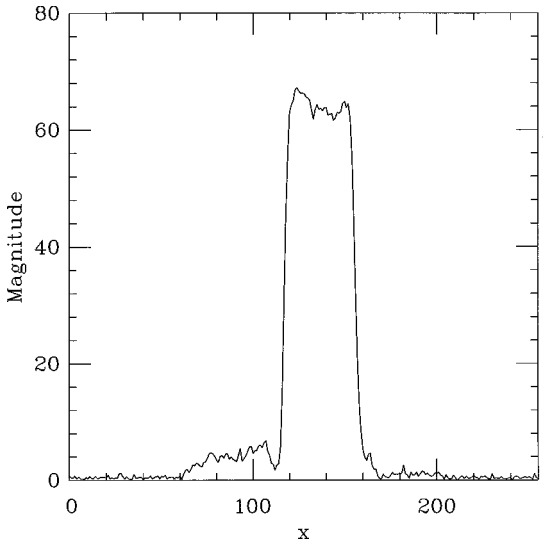


FIG. 14. Cross section through the center of the aperture shown in Fig. 13.

step, the error is reduced to less than one pixel.

To illustrate the effectiveness of this algorithm, a multiple-angle reconstruction is formed of a test specimen consisting of a 380-μm-diam circular aperture in a 3-mm-diameter copper test target. The specimen is placed 640 μm below the coverslip and 36 projections are acquired at 10-degree steps of rotation, using an insonification frequency of 104 MHz. The phase and sampling errors present in the data are first compensated. Using the two-step algorithm, the center of rotation is estimated to be at $x_c = -30$ μm and $y_c = 130$ μm relative to the center of the microscope's field of view. The resulting reconstruction of this specimen is shown in Fig. 13 with the circular aperture in the center of the image. The darker area around the aperture is a result of signal attenuation caused by the epoxy used to secure the specimen to the mounting hardware. Figure 14 shows a cross section taken through the center of the aperture. The sharp cutoff at the edges of the aperture is a result of properly aligned projections. Incorrect alignment would result in a slower transition from the region of the aperture to the surrounding copper material. Measurements taken from Fig. 14 indicate an aperture diameter of 390 μm, representing an error of 10 μm.

V. CONCLUSION

In scanning tomographic acoustic microscopy, projection phase error correction plays a pivotal role in the quality of the final reconstruction. In this paper, two aspects of projection error correction were examined: projection phase errors and projection alignment errors. Phase errors occur in the form of the phase error from the quadrature receiver channels and in the form of the initial phase error between different projections. Both phase errors were modeled and algorithms were developed to estimate and compensate the errors. Experimental results were provided to illustrate the effectiveness of the algorithms.

Three types of alignment errors due to incorrect spatial sampling were discussed: sampling aspect ratio imbalance, scanning skew, and projection pose identification. The error associated with inadequate registration between the projections was also examined. Algorithms for the estimation and correction of these errors were implemented and demonstrated with full-scale experiments.

Because of the successful development of these error estimation and removal techniques, the tomographic acoustic microscope system is now capable of high-resolution three-dimensional imaging at full capacity.

ACKNOWLEDGMENT

This research is supported by the National Science Foundation under Grant Nos. ENG-8451484 and ECS 8406511.

[1] L. W. Kessler, P. R. Palermo, and A. Korpel, ''Recent developments with the scanning laser acoustic microscope,'' in *Acoustical Imaging*, Vol. 5, edited by P. S. Green (Plenum, New York, 1974), pp. 15–23.
[2] L. W. Kessler and D. E. Yuhas, ''Acoustic Microscopy—1979,'' Proc. IEEE **67**, 526–536 (1979).
[3] Z. C. Lin, H. Lee, and G. Wade, ''Scanning Tomographic Acoustic Microscopy: A Review,'' IEEE Trans. Sonics Ultrason. **SU-32**, 168–180 (1985).
[4] H. Lee and C. Ricci, ''Modification of the scanning laser acoustic microscope for holographic and tomographic imaging,'' Appl. Phys. Lett. **49**, 1336–1338 (1986).
[5] Z. C. Lin, H. Lee, G. Wade, M. G. Oravecz, and L. W. Kessler, ''Holographic image reconstruction in scanning laser acoustic microscopy,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **UFFC-34**, 293–300 (1987).
[6] H. Lee, C. Schueler, G. Flesher, and G. Wade, ''Ultrasonic planar scanned tomography,'' in *Acoustical Imaging*, Vol. 11, edited by J. Powers (Plenum, New York, 1982), pp. 309–323.
[7] R. L. Whitman and A. Korpel, ''Probing of acoustic surface perturbations by coherent light,'' Appl. Opt. **8**, 1567–1576 (1969).
[8] Z. C. Lin, H. Lee, and G. Wade, ''Back-and-forth propagation for diffraction tomography,'' IEEE Trans. Sonics Ultrason. **SU-31**, 626–634 (1984).
[9] R. Y. Chiao, ''Signal processing and image reconstruction for scanning tomographic acoustic microscopy,'' Ph.D. thesis, University of Illinois, Urbana-Champaign, July 1990.
[10] R. Gonzalez and P. Wintz, *Digital Image Processing* (Addison-Wesley, Reading, MA, 1987).
[11] B. L. Douglas, S. D. Kent, and H. Lee, ''The use of image contrast and entropy criteria for the estimation of imaging parameters in acoustic microscopy,'' J. Acoust. Soc. Am. **95**, 306–312 (1994).
[12] B. L. Douglas, S. D. Kent, and H. Lee, ''Parameter estimation and the importance of phase in acoustic microscopy,'' Proceedings of the IEEE 1992 Ultrasonics Symposium, 1992.
[13] Z.-C. Lin and Glen Wade, ''On the resolution of planar ultrasonic tomography,'' J. Acoust. Soc. Am. **77**, 139–143 (1985).
[14] R. Y. Chiao, H. Lee, and G. Wade, ''Image restoration and wave-field error removal in holographic acoustic microscopy,'' Proceedings of 1989 IEEE International Conference on Acoustics, Speech and Signal Processing, 1989.

Spectral-ripple representation of steady-state vowels in primary auditory cortex

Huib Versnel[a] and Shihab A. Shamma
Institute for Systems Research and Electrical Engineering Department, University of Maryland, College Park, Maryland 20742

(Received 1 July 1997; revised 8 December 1997; accepted 30 January 1998)

Responses to various steady-state vowels were recorded in single units in the primary auditory cortex (AI) of the barbiturate-anaesthetized ferret. Six vowels were presented (/a/, /ε/, 2 different /i/'s, and 2 different /u/'s) in a natural voiced and a synthetic unvoiced mode. In addition, the responses to broadband stimuli with a sinusoidally shaped spectral envelope (called ripple stimuli) were recorded in each cell, and the response field (RF), which consists of both excitatory and inhibitory regions, was derived from the ripple transfer function. We examined whether the vowel responses could be predicted using a linear ripple analysis method [Shamma *et al.*, Auditory Neurosci. **1**, 233–254 (1995)], i.e., by cross correlating the RF of the single unit, and the smoothed spectral envelope of the vowel. We found that for most AI cells (71%) the relative responses to natural vowels could be predicted on the basis of this method. Responses and prediction results for unvoiced and voiced vowels were very similar, suggesting that the spectral fine structure may not play a significant role in the neuron's response to the vowels. Predictions on the basis of the entire RF were significantly better than based solely on best frequency (BF) (or ''place''). These findings confirm the ripple analysis method as a valid method to characterize AI responses to broadband sounds as we proposed in a previous paper using synthesized spectra [Shamma and Versnel, Auditory Neurosci. **1**, 255–270 (1995)].
[S0001-4966(98)03805-3]

PACS numbers: 43.64.Qh, 43.64.Sj [RDF]

INTRODUCTION

Timbre is the aspect of sound which enables us to distinguish the sounds of different musical instruments playing the same tone with the same loudness or to discriminate different vowels spoken with the same fundamental frequency and the same intensity (cf. Plomp, 1976). The spectral envelope is an important physical correlate of timbre: e.g., the more similar the spectra of two vowels the more likely they will be confused. There is growing evidence from recent experiments (Phillips *et al.*, 1985; Schreiner and Mendelson, 1990; Schreiner and Sutter, 1992; Schreiner and Calhoun, 1994; Shamma *et al.*, 1993, 1995; Shamma and Versnel, 1995; Versnel *et al.*, 1995; Kowalski *et al.*, 1996a, b) that the auditory cortex plays an important role in encoding the spectral shape, and thus in encoding of timbre.

Such evidence is found in the organization of the response areas in the primary field (AI): the bandwidth shows topographic organization (Schreiner and Mendelson, 1990; Schreiner and Sutter, 1992; Versnel *et al.*, 1995) and so does the asymmetry of the inhibitory sidebands (Shamma *et al.*, 1993). Secondary fields might widen the range of response areas already available in the primary field as Kowalski *et al.* (1995) found with respect to the bandwidth. The function of the topographic organization is apparent in the tendency of units to respond to spectral features according to their response area: e.g., a unit with a broad response area will preferably respond to a wide-band sound (Phillips *et al.*, 1985; Schreiner and Mendelson, 1990), or units with inhibition below the best frequency (BF) will respond best to sounds with most energy above BF (Shamma *et al.*, 1993). A recent series of experiments in which cortical responses to spectral ripples were examined provided a feasible method for cortical encoding of spectral features. We will refer to this method as ripple analysis (Shamma *et al.*, 1995), where ripple is defined as a sinusoidal modulation of the spectral profile of a broadband sound stimulus.

A. Ripple analysis and encoding of spectral profiles

Units in AI were found to be tuned to different spectral ripples, mostly in bandpass fashion, and the units were topographically segregated with respect to the ripple responses (Versnel *et al.*, 1995). An important consequence of this linear analysis was that responses to an arbitrary spectral profile could be predicted by superposition of responses to the separate ripple components of the stimulus (Shamma and Versnel, 1995). This finding led to the suggestion that cortical cells behave as ripple filters, and accordingly perform a Fourier-like analysis of the auditory spectrum (see for a detailed theoretical background Wang and Shamma, 1995). The ripple analysis appears to be analogous to the spatial frequency analysis in the visual cortex where simple cells are bandpass tuned to spatial frequencies of sinusoidal gratings (De Valois and De Valois, 1990).

In this paper we apply the ripple analysis method and ideas to characterize and predict AI responses to natural broadband stimuli. Steady-state vowels were chosen as natu-

[a]Present address: University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, UK, Electronic mail: hv@physiol.ox.ac.uk

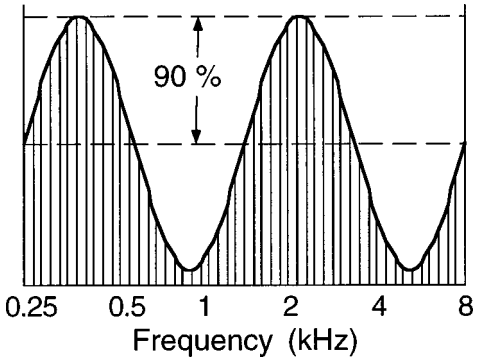


FIG. 1. Schematic of a rippled spectrum stimulus. It is composed of 101 tones equally spaced along the logarithmic frequency axis between 0.25 and 8 kHz. The envelope of the complex is sinusoidally modulated along the logarithmic frequency axis and the amplitude scale is linear. The ripple phase is defined relative to a sine wave starting at the left edge. For the ripple shown, the ripple frequency Ω is 0.4 cycles/octave, the ripple phase ϕ is 0, and the ripple amplitude is 90%.

ral stimuli. We examine whether the relative responses to different vowels can be predicted on the basis of the ripple responses and the spectral envelope of the vowel, analogous to the predictions in Shamma and Versnel (1995). The spectral profiles described in that paper were composed of 2–10 ripple components, and the carrier of this stimulus was the same as that of the rippled stimuli. Here, we use natural voiced vowels which are not artificially composed of ripples and whose carrier is the harmonic series of the natural voicing. Further, we test the ripple analysis on synthesized vowellike stimuli with a carrier as ripples and a spectral envelope as vowels, which can be considered a stage between testing with ripple combinations (Shamma and Versnel, 1995) and natural vowels.

I. METHODS

A. Surgery and animal preparation

The data presented here were collected in 6 young adult male ferrets (*Mustela putorius*) weighing 1.5 to 2.3 kg. A detailed description of the animal preparation is given in Shamma *et al.* (1993). The ferrets were intraperitoneally anesthetized with sodium pentobarbital (40 mg/kg) and atropine sulfate was administered subcutaneously (0.1 mg/kg). An areflexic level of anesthesia was maintained throughout the experiment by continuous intravenous infusion of sodium pentobarbital at a flow rate of about 5 mg/kg/h. The infusion fluids included 5% dextrose in Ringer's solution. The primary auditory cortex on the left hemisphere was exposed by craniotomy and the overlying dura was incised and reflected. The brain was covered in 2% agar in saline to reduce pulsations. The contralateral meatus was exposed, cleaned and partly resected, and subsequently a cone-shaped speculum containing a Sony MDR-E464 miniature speaker was sutured to the meatal stump.

B. Acoustic stimuli

Stimuli were monaurally presented in above described closed-field set up. Calibration of the sound delivery system was performed *in situ* over a range of 0.4–20 kHz using a 3.2-mm (1/8-in.) Bruël & Kjaer probe microphone (type 4170). The microphone was inserted into the ear canal through the wall of the speculum to 5–10 mm of the tympanic membrane. The speculum and microphone setup resembles closely that suggested by Evans (1979). Maximum tone intensities were limited to 85 dB SPL to avoid any distortion.

We presented pure-tone stimuli, complex stimuli with a rippled spectrum (ripple stimuli), natural voiced vowels, and synthetic unvoiced vowels. A Hewlett-Packard 9000/800 series minicomputer controlled stimulus delivery. Pure-tone stimuli of 200-ms duration were generated using a function generator, the vowels were taken from a data base and the other stimuli were computer synthesized. All stimuli were gated to obtain rise and fall times of 7 ms, and then fed

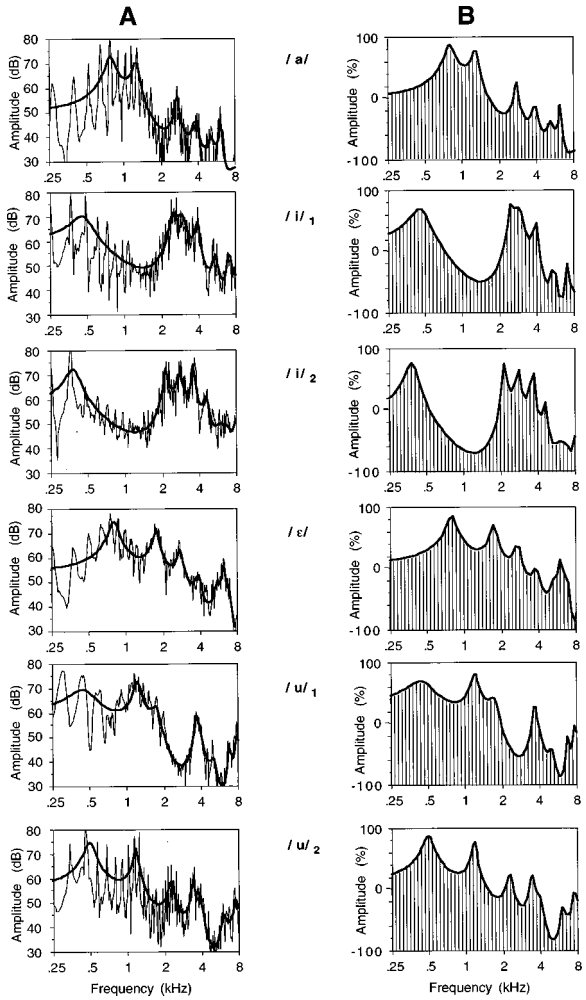


FIG. 2. (A) Frequency spectra of the six natural vowels used in the experiment, /a/, /ε/, 2 different /i/'s, and 2 different /u/'s. The thin line represents the Fourier spectrum and the thick line represents the smoothed spectral envelope which is obtained using linear predictive coding (LPC). (B) Frequency spectra of six unvoiced synthesized vowels. These sounds were composed of 101 logarithmically distributed frequency components between 0.25 and 8 kHz (as used in the ripple stimuli, see Fig. 1). The spectral envelope was derived from the LPC-smoothed spectral envelope of the six natural vowels shown in (A). The envelope was modified such that the maximum peak was at 90% and the minimum trough was at −90% comparable to the amplitude of the ripple stimulus. The vertical lines schematically represent the frequency components, and the thick line represents the envelope.

through an equalizer into the earphone. The interstimulus interval was 1.2 s for the natural vowels, and 1.0 s for the other stimuli.

A schematic of the ripple stimulus is shown in Fig. 1. It consisted of 101 tones that were equally spaced along the logarithmic frequency axis and spanning 5 octaves from 0.25 to 8 kHz. The envelope of the complex was modulated sinusoidally on a linear amplitude scale. The ripple amplitude or depth, ΔA, defined as the maximum percentage change in the component amplitudes, was 90%. The overall level of the ripple stimulus is given by the level of the corresponding flat complex ($\Delta A = 0$). Thus, if the sound pressure level for a flat complex with 101 components is L_1, the level of a single component is $L_1 - 10\log(101) \approx L_1 - 20$ dB. Usually the overall level was set 10–20 dB above the threshold to tones at BF. The ripple frequency (Ω) is measured in units of cycles/octave against the logarithmic frequency axis. The ripple phase (ϕ) is measured in degrees relative to a sine wave starting at the left edge (low-frequency edge) of the complex (Fig. 1). In a single test both Ω and ϕ are randomly varied, Ω from 0 to 1.4 cycles/octave with 0.2 steps, and ϕ over the full 360° cycle in steps of 45°. The duration of the complex stimulus bursts was 50 ms.

Six different vowels were taken from steady-state speech segments in the TIMIT database: /a/, /ɛ/, 2 different /i/'s, and 2 different /u/'s. Subsequently, these vowels were preemphasized (see Rabiner and Juang, 1993; Eq. 3.69, we applied $\tilde{a} = 0.9$) thereby enhancing energy at high frequencies. Figure 2(A) shows the amplitude spectra of the six vowels with their smoothed envelope obtained using linear predictive coding (LPC). The duration of the vowels was 80–140 ms. The vowels were presented at 1–3 levels typically in a range of 47–67 dB SPL.

Unvoiced versions of the natural vowels were synthesized using the smoothed spectral envelope of the natural vowels, and the carrier and other stimulus characteristics of the ripple stimuli [see Fig. 2(B)]. The spectrum was modified such that the maximum peak was at 90% and the minimum trough was at -90% according to the ripple-stimulus amplitude that was used, and the spectral range was 0.25–8 kHz. The transformation to the percentage scale is unlikely to significantly affect the response (Shamma *et al.*, 1995). The duration of the unvoiced vowels was 50 ms. The unvoiced vowels were presented at the level at which the ripple stimulus was presented, and often at one or two more levels in the range 45–65 dB SPL.

C. Recordings

Action potentials were recorded using glass-insulated tungsten microelectrodes with 5–6 MΩ tip impedances. Electrode penetrations were made orthogonal to the cortical surface, and in each penetration, one or two cells were studied. A search stimulus consisted of pure tones manually swept across frequency at various levels. Neural signals were amplified ($\times 10\,000$) and filtered (0.5–5 kHz bandpass) and led through a time-amplitude window discriminator (BAK electronics, model DIS-1). Single-unit spikes were separated and identified on the basis of their waveform and amplitude. During recording the waveform is monitored in order to check the stability of the spike waveform. The time of spike occurrence relative to stimulus delivery was stored using the HP 9000/800 computer. In a single test a type of stimulus (tone, ripple, voiced vowel, or unvoiced vowel) is presented and the stimulus parameters (including sound level) are varied in random order.

D. Data analysis for responses to tonal stimuli

We used the definitions for various response parameters as given by Kowalski *et al.* (1995). The best frequency (BF) is defined as the frequency which elicits the maximum response at a low level (10–20 dB above threshold). After manual assessment of the excitatory region, BF was determined by randomly presenting tones at a low stimulus level with a frequency resolution up to 1/8 octave. The rate-level function at BF was measured over the range of 35–85 dB SPL. The rate-level curve is considered nonmonotonic if the spike count at higher intensities has decreased by more than 25% of the maximum spike count. The best intensity is defined for nonmonotonic curves as the intensity which elicits maximum response, for monotonic curves as the highest level used (85 dB SPL). The threshold is the intensity at which the spike count reached 10% of the maximum. In determining the various parameters, corrections have been made for the cell's spontaneous rate. Spike counts were made over a time window from 10 to 60 ms after onset of the stimulus.

E. Data analysis for responses to ripple stimuli

Figure 3 shows responses of an AI unit to the ripple stimuli and it illustrates the analysis applied. Full details of the analysis are described in Shamma *et al.* (1995). In Fig. 3(A), the raster shows the responses to different phases of the ripple at $\Omega = 0.8$ cycles/octave. Spike counts, obtained from 15 repetitions of each stimulus, were made over a 60-ms time window starting 10 ms after the onset of the stimulus. The spike count curve shown in the inset plot to the right of the raster demonstrates the dependence of the response on the phase of the ripple. In Fig. 3(B), such spike counts from all ripple frequencies are combined in one display. The baseline at each ripple frequency was set equal to the spike count obtained from the flat spectrum ($\Omega = 0$). The best sinusoidal fit to the data (solid line) at each ripple frequency yield amplitude and phase values, and accordingly a transfer function, $AC_1(\Omega)$, is obtained. In order to account for statistical noise and the quality of the sinusoidal fit the amplitude of the transfer function is adjusted as follows. An eight-point Fourier transform is performed on the spike counts at each ripple frequency. The magnitude and phase of the first harmonic, $AC_1(\Omega)$, is extracted and weighted by the rms value of the response as follows (Shamma *et al.*, 1995):

$$T(\Omega) = AC_1(\Omega) \cdot \frac{|AC_1(\Omega)| - |AC_1(0)|}{\sqrt{\Sigma_{i=1}^{4} |AC_i(\Omega)|^2}} \quad \text{if } |AC_1(\Omega| - |AC_1(0)| \geq 0,$$
$$T(\Omega) = 0 \quad \text{if } |AC_1(\Omega| - |AC_1(0)| < 0, \tag{1}$$

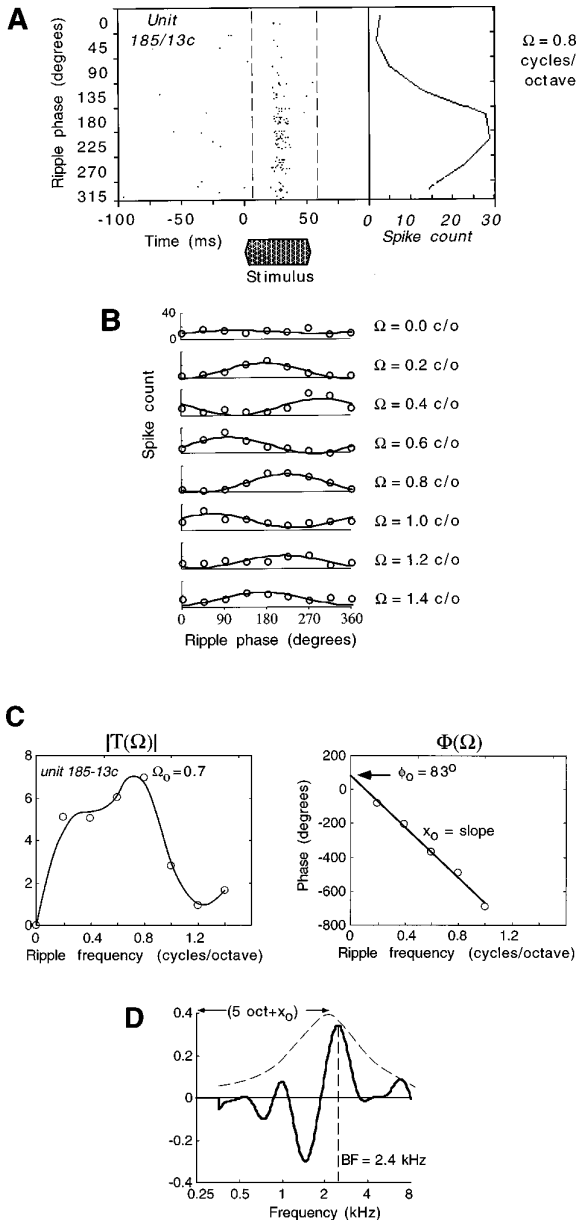


FIG. 3. The analysis of the responses to rippled stimuli. (A) Raster of the responses of an Al cell (185/13*c*; BF of 2.4 kHz) to a rippled spectrum stimulus with a fixed ripple frequency ($\Omega=0.8$ cycles/octave) and various ripple phases (0–315 degrees in steps of 45 degrees). The rise and fall time of the stimulus burst is 7 ms and its duration 50 ms. The stimulus is repeated 20 times for each ripple phase. Spike counts are computed over a time window 10–70 ms after the stimulus onset (indicated by the dashed lines), and are displayed in the inset plot to the right of the raster. (B) Spike counts as a function of ripple phase for various ripple frequencies Ω between 0 and 1.4 cycles/octave. At each Ω, spike counts are indicated by the circles, and the solid line is the best sinusoidal fit to the points (in the sense of mean square error). The baseline at each ripple frequency was set equal to the average spike count to the flat spectrum ($\Omega=0$). (C) The transfer function $T(\Omega)$ is derived from the magnitude and phase of the sinusoidal fitted curve shown in (B). The plot to the left shows the magnitude of the transfer function $|T(\Omega)|$, where the characteristic ripple frequency Ω_0 is the location of the maximum of $|T(\Omega)|$. The right plot shows the phase of the transfer function, $\phi(\Omega)$, where the solid line represents a linear fit to the data with intercept ϕ_0 and slope x_0. (D) The response field (RF) of the cell is derived by an inverse Fourier transform of the transfer function $T(\Omega)$. The dashed curve represents the envelope of the RF. The distance from the left edge of the rippled spectrum to the center of the RF envelope is given by $k2\pi/\Delta+x_0$, where Δ is the step size of the ripple frequencies Ω tested, and k is an integer $\geqslant 1$. The location of the maximum of the RF is defined as $\mathrm{BF}_{\mathrm{RF}}$.

where $|AC_i(\Omega)|$ is the magnitude of the ith Fourier component of the response. In general terms $T(\Omega)$ can be written as follows:

$$T(\Omega)=|T(\Omega|\cdot e^{j\phi(\Omega)}, \tag{2}$$

where $j=\sqrt{-1}$.

Figure 3(C) shows the ripple transfer function $T(\Omega)$: the magnitude $|T(\Omega)|$ (*left*) and the unwrapped phase $\phi(\Omega)$ (*right*). Note that the phase relation is approximately linear. The ripple transfer function can be inverse Fourier transformed to obtain the response field (RF) of the cell which is shown in Fig. 3(D). The RF can be interpreted as an isointensity frequency response curve with the positive peak representing the excitatory portion and the negative peak representing the inhibitory portion.

Several parameters characterize the ripple transfer function and the RF. Figure 4 shows data of three main param-

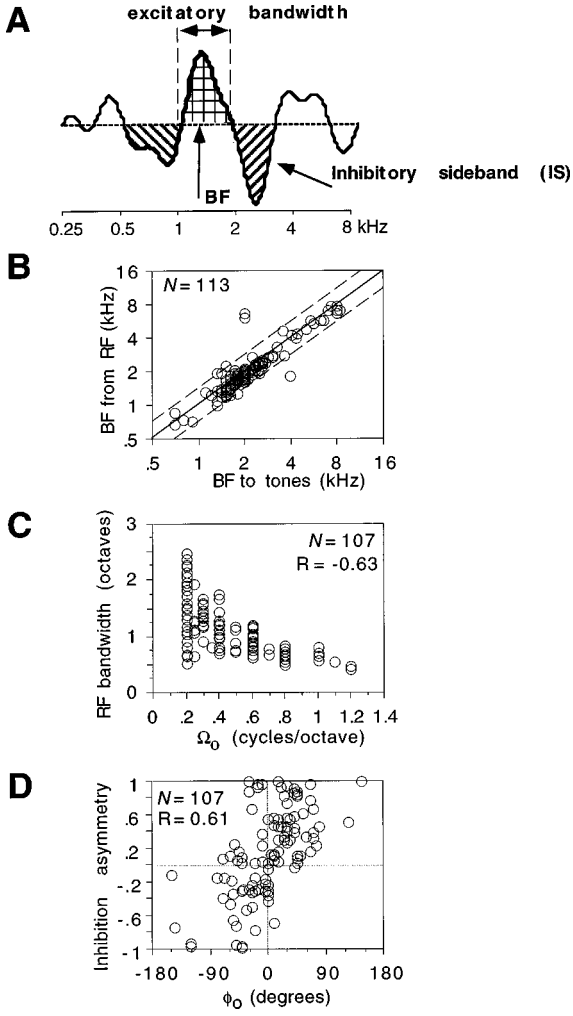


FIG. 4. Interpretation of ripple response parameters. (A) An example of an RF of an AI cell (unit 185/35*a*) illustrating that positive peaks of the RF might be associated to excitation and negative peaks to inhibition. The inhibitory sidebands are indicated with the oblique-line patterns and the excitatory peak with the cross pattern. (B) BF according to RF vs BF found with tones. (C) Excitatory bandwidth of RF vs Ω_0. The bandwidth is defined by the zero crossings below and above BF as illustrated in (A). (D) Asymmetry index of inhibitory sidebands vs ϕ_0. The measure of inhibition asymmetry, M, is defined as $M=(A_{is<\mathrm{BF}}-A_{is>\mathrm{BF}})/(A_{is<\mathrm{BF}}+A_{is>\mathrm{BF}})$, where $A_{is<\mathrm{BF}}$ is the area of the inhibitory sideband below BF and $A_{is>\mathrm{BF}}$ the inhibitory sideband area above BF.

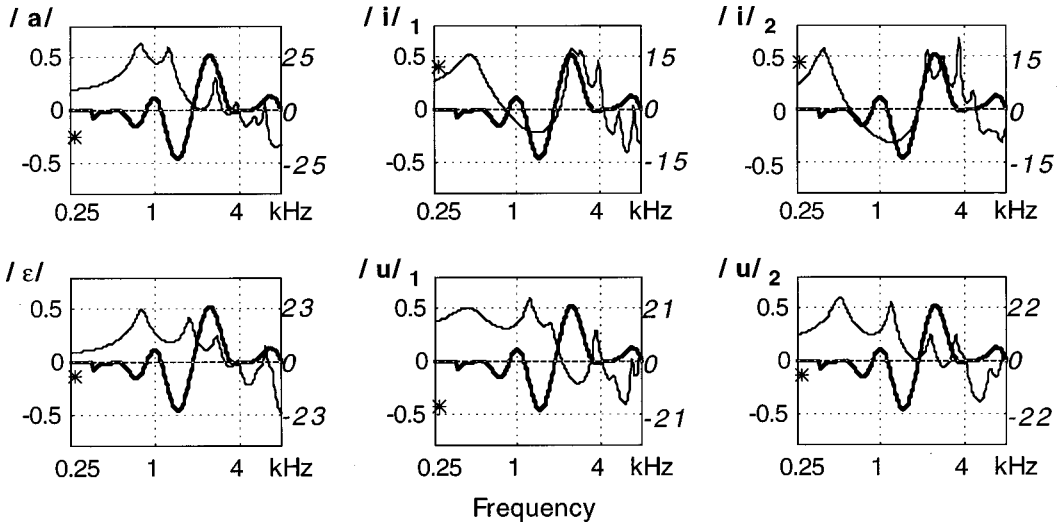


FIG. 5. Example of RF of AI unit (thick line) with vowel's spectral profiles (thin lines). The vowel profiles are scaled and aligned to the RF. The ordinate is given in RF units (spike count). The asterisk reflects the prediction of the response to the vowel which is the cross correlation of the RF and vowel's spectral profile [see Eq. (4)].

eters and illustrates their significance. The most obvious one and easily interpretable is the location of the maximum of the RF along the tonotopic axis. This generally corresponds very well to the tonal BF of the cell [Fig. 4(B)] and will be labeled accordingly as BF_{RF}. The second parameter is the characteristic ripple frequency, Ω_0, which is the ripple frequency where the magnitude, $|T(\Omega)|$, is maximum. Figure 4(C) shows that Ω_0 inversely relates to the width of RF: in general, the higher Ω_0 the narrower RF, and vice versa. Third, the characteristic ripple phase, ϕ_0, is the intercept of the phase function $\phi(\Omega)$. It is derived by fitting a line to the phase function [see Eq. (6), Shamma *et al.*, 1995]:

$$\phi(\Omega)=x_0\Omega+\phi_0, \tag{3}$$

where x_0 is the slope of the line.

The parameter ϕ_0 indicates the asymmetry of the RF about its center [Fig. 4(D)]. For instance, the RF is symmetric for $\phi_0\cong 0$, and strongly asymmetric for $\phi_0\cong -90°$ or $\phi_0\cong 90°$ [as in Fig. 3(D)].

F. Data analysis for responses to vowels

The responses to vowels were obtained from 40 repetitions of each vowel stimulus, and spike counts were made over a 60-ms time window starting 10 ms after the onset of the vowel. The responses were analyzed either across units or across vowels. The former type of analysis provided a population response, in which case the response was normalized by dividing by the maximum response found to BF tones (cf. Sachs and Young, 1979). In the other type of analysis the responses of a cell to the six different vowels are considered relative to each other, i.e., we examine whether the cell responds better to the one than to the other vowel. In this case normalization was not necessary. The relative vowel responses are compared to predictions.

The prediction model basically follows the linear ripple analysis applied in Shamma and Versnel (1995). In that paper the stimuli were profiles shifting along the logarithmic frequency axis, i.e., with a constant ripple frequency content (two or more ripple components) and a varying ripple phase, and the responses were predicted by superposition of the responses to individual components. This superposition can be computed either in the ripple domain by taking the inverse Fourier transform of the product of the cell's ripple transfer function and the stimulus' ripple spectrum or in the logarithmic frequency domain by crosscorrelating the cell's RF with the stimulus' spectral profile. In this paper the stimuli are six profiles which differed in ripple frequency content and ripple phase content. We predicted for each cell the relative responses to the vowels in the logarithmic frequency domain as follows [cf. Eq. (5) in Shamma and Versnel, 1995]:

$$r_p(k)=\frac{\Sigma_x w(x)\cdot p(x;k)}{\sqrt{\Sigma_x w^2(x)\cdot\Sigma_x p^2(x)}}, \tag{4}$$

where integer $k=1,2,...,6$ denotes the vowel, x is the logarithmic frequency axis, $w(x)$ is the RF and $p(x)$ the spectral profile. Strictly speaking, for each vowel k the center element of the cross correlation of $w(x)$ and $p(x)$ is computed whereas Shamma and Versnel computed the entire cross correlation vector in order to predict responses to shifting profiles. The right term of Eq. (4) can be thought of as a match between the profile and RF which is illustrated in Fig. 5. When the vowel's spectral profile matches the RF, e.g., because a formant peak coincides with an excitatory area, then the response is predicted to be large; this is demonstrated for both vowels /i/ in Fig. 5. When the match is poor, e.g., because a formant coincides with an inhibitory area, then the response is predicted to be weak; this is demonstrated for the vowel $/u/_1$. It is obvious from Eq. (4) that an absolute response amplitude or discharge probability is not predicted, instead one predicts whether the response to the one vowel is stronger than to the other. The value of $r_p(k)$ is between -1 and 1. For the prediction the profile $p(x)$ is dB scaled, which gave similar but slightly better results than linear scaling.

In order to examine whether the entire RF is needed to predict the vowel responses or instead information of the BF

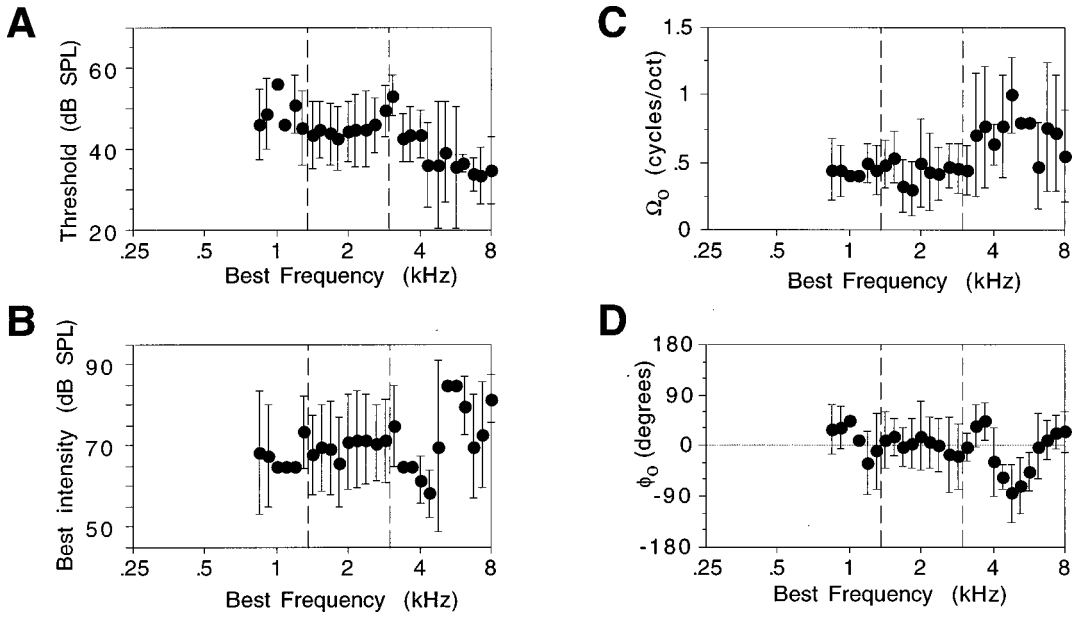


FIG. 6. Average response parameters as a function of BF. The parameter values are averaged over AI cells in 1/4 octave moving segments at 1/8 octave resolution. The BF is estimated from responses to tones. The dashed lines indicate the region where the number of units per segment is 10 or larger; outside this region of the numbers of units are typically between 3 and 5. The bars represent the standard deviation. (A) Threshold for BF tones. (B) Best intensity for BF tones. (C) Characteristic ripple frequency. (D) Characteristic ripple phase.

location alone might be sufficient, we also performed predictions assuming that RF only consists of a sharp excitatory peak at BF. This prediction was as follows:

$$r_p(k) = \frac{w(x = x_{\mathrm{BF}}) \cdot p(x = x_{\mathrm{BF}}; k)}{\sqrt{\Sigma_x w^2(x) \cdot \Sigma_x p^2(x)}}, \tag{5}$$

where x_{BF} is the location of BF along the logarithmic frequency axis.

For each cell a correlation coefficient R is computed using linear regression analysis of the predicted and measured responses, and R^2 is used as a measure of the accuracy of the prediction. For cases where $R<0$, R^2 was made 0 in statistical analyses and data presentation.

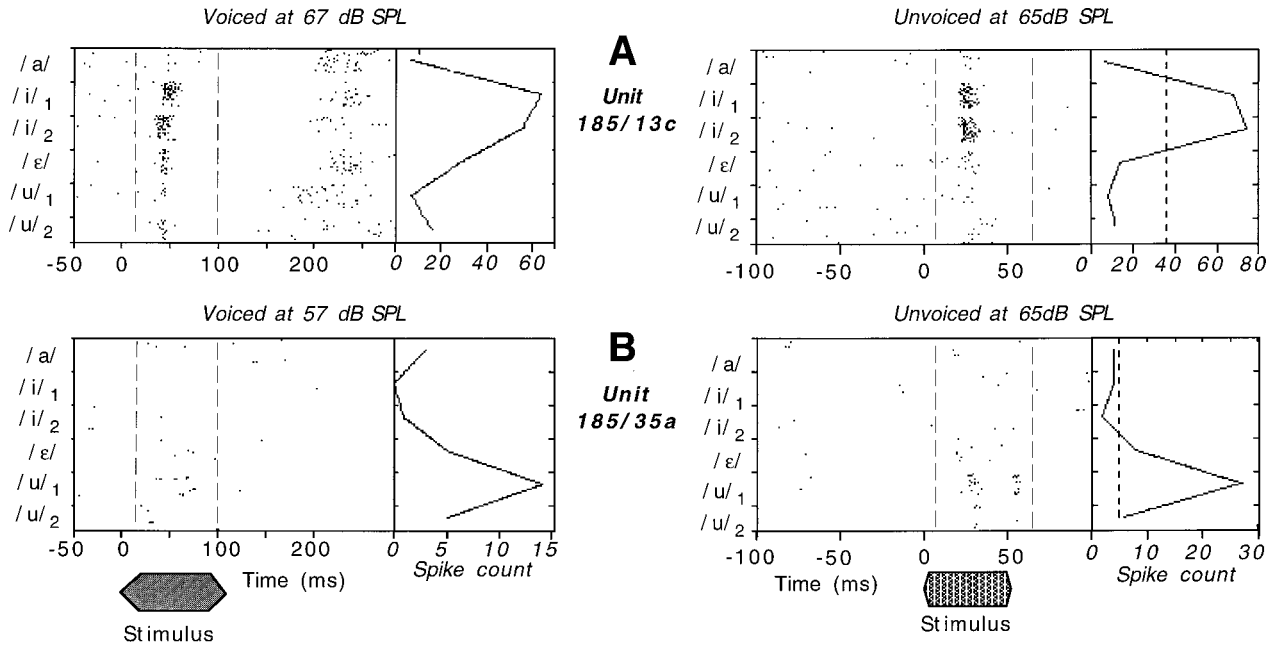


FIG. 7. Rasters of responses of two single units to 6 different vowel stimuli presented in voiced (left) or unvoiced mode (right). Each stimulus is repeated 40 times. Spike counts computed over a time window 10–100 or 10–70 ms after the stimulus onset (indicated by the dashed lines) are displayed in the inset plot on the right. The dashed line in the inset of the responses to unvoiced vowels represent the spike count to a flat spectrum presented at the same sound level. (A) Responses of unit 185/13*c* (BF: 2.4 kHz; $\Omega_0 = 0.7$ cycles/oct; $\phi_0 = 83°$; threshold: 47 dB SPL; best intensity: 65 dB SPL). (B) Responses of unit 185/20*a* (BF: 1.2 kHz; $\Omega_0 = 0.6$ cycles/oct; $\phi_0 = -69°$; threshold: 56 dB SPL; best intensity: 65 dB SPL).

II. RESULTS

We collected data from 113 single units in 6 ferrets. For all units the response to BF tones at the best intensity varied from 0.5 to 4.9 spike per stimulus presentation. The response to the optimal vowel varied from 0.1 to 5.4 spike, and could be twice as large as the response to BF tones at best intensity. A large majority of BFs were between 1.2 and 3 kHz (79%). For 109 units (96%) the BF estimated from the RF differed less than 0.5 octave from BF measured to tones [see Fig. 4(B)]. A BF difference $|BF_{tone} - BF_{RF}|$ can be caused by errors in the RF or tonal BF assessments, and/or by nonlinearities (cf. Shamma *et al.*, 1995). Therefore, BF difference can be used as a rough indicator of RF goodness, and we shall examine whether RF predictions of vowel responses [see Eqs. (4) and (5)], are related to this measure.

A. Basic response characteristics

Figure 6 plots threshold, best intensity, Ω_0, and ϕ_0, as a function of the unit's BF, parameters which characterize the unit's stationary response behavior to a great extent. The data are expressed as averages over 0.25 octave in order to be able to relate them to vowel response profiles which are shown in a later figure (Fig. 8). The parameters have average values expected on the basis of previous reports (Kelly *et al.*, 1986; Kowalski *et al.*, 1995; Shamma *et al.*, 1995). This report, however, covers mostly a lower BF range (below 3 kHz) than our previous reports on ripple responses. Figure 6 shows that in this range the parameters were fairly constant compared to the variations observed at higher BFs (3–8 kHz), partially due to the lower sample sizes.

B. Examples of responses to vowels

Figure 7 shows typical examples of responses of two different AI units to vowel stimuli. The responses are phasic as generally is the case, at least at normal levels of barbiturate anesthesia, for responses to other stationary stimuli such as tones. The two examples in Fig. 7 demonstrate two important response features. First, both units respond selectively, e.g., the unit in Fig. 7(A) responds much stronger to the /i/ vowels than to the other ones, and the unit in Fig. 7(B) prefers the /u/$_1$ vowel. Second, for each unit the responses to the unvoiced and voiced vowels show similar trends.

C. Population response

Figure 8 shows normalized spike counts which are averaged over several units in BF segments of a 1/4 octave as a function of BF, resulting in a population response profile. The response profiles are shown for both voiced and unvoiced vowels and for several levels, and they are compared to the vowel spectrum. Focusing on the BF region covered by the majority of units (1.4–2.8 kHz, indicated by dashed lines) one observes that the response profiles coarsely follow the stimulus spectrum, e.g., the opposite trends of the two vowel spectra are reproduced by the responses, in particular in the 1.4–2 kHz region. However, the responses do not follow the spectral details, e.g., between 2 and 2.8 kHz the response profiles for both vowels are virtually flat while the spectra are sloping. Although the sample size at high BFs is small (about 3 units per data point), the discrepancy between response profiles and vowel spectra, consistent over all sound levels, is striking. The threshold profile [Fig. 6(A)] does not show the variation with BF that might have accounted for this discrepancy.

D. Predictions: RF approach

For each unit we predicted relative responses to vowels on the basis of the RF of the unit and the spectral profile of the vowel [Eq. (4)]. Figure 9 illustrates for three different units how the RF predictions compare to the experimental responses to voiced vowels. In all three examples a response to the one vowel being significantly larger than to the other vowel is predicted well. The good prediction is reflected by the high value of R^2 (greater than 0.7).

Vowel responses vary with level and consequently, the goodness of the prediction as measured by R^2 varies with sound level of the vowel. Figure 10 shows for six different units the responses for three levels of the voiced vowel (A)

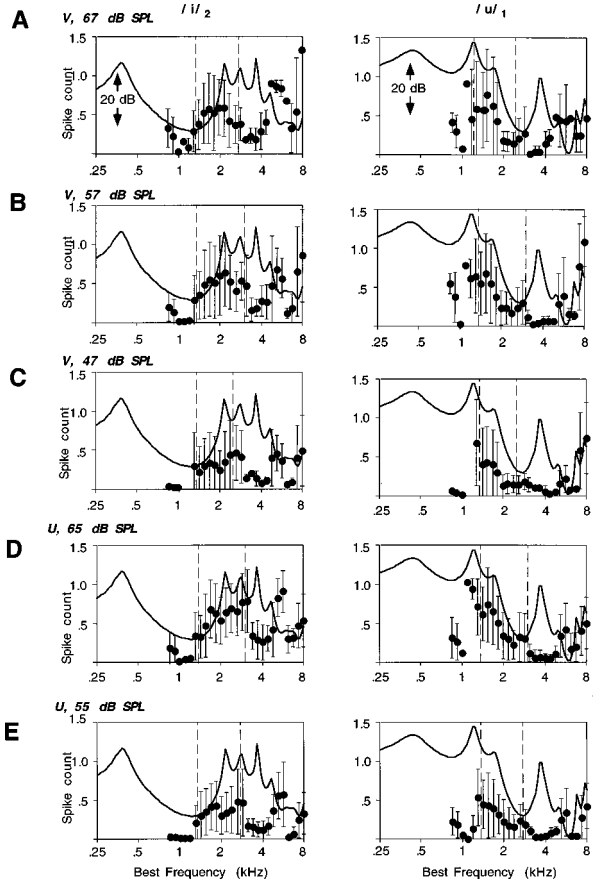


FIG. 8. Average spike count profiles for two different vowels (/i/$_2$ and /u/$_1$). The responses are normalized by dividing over the response to BF tone at best intensity and subsequently averaged over units in 1/4 octave moving segments at 1/8 octave resolution. The BF is estimated from responses to tones. The dashed lines indicate the region where the number of units per segment is 10 or larger; outside this region the numbers of units are typically between 3 and 5. The bars represent the standard deviation. The thick line represents the envelope of the vowel spectrum, which amplitude scale is indicated in the top plots. (A) Response profiles for voiced vowels at 67 dB SPL. (B) Response profiles for voiced vowels at 57 dB SPL. (C) Response profiles for voiced vowels at 47 dB SPL. (D) Response profiles for unvoiced vowels at 65 dB SPL. (E) Response profiles for unvoiced vowels at 55 dB SPL.

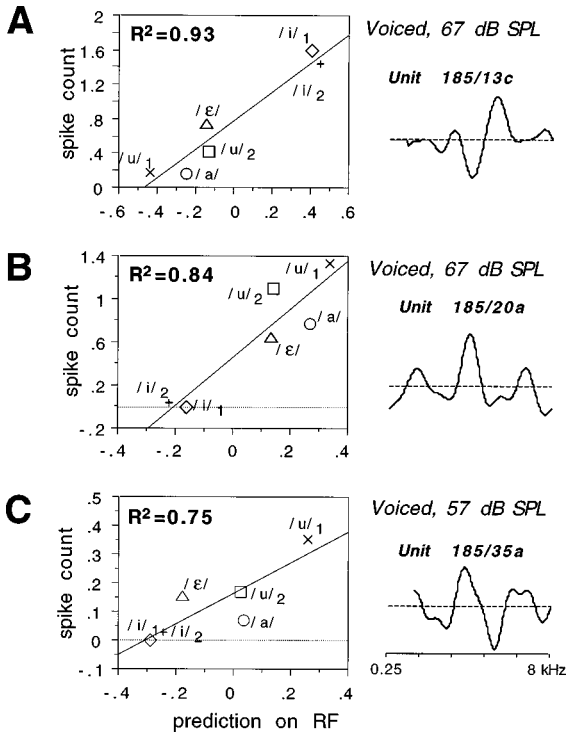


FIG. 9. Examples of spike count (per stimulus presentation) versus predicted response for voiced vowels, for three different AI neurons. The prediction on the basis of RF and spectral envelope [see Eq. (4)] is expressed in relative units. The line reflects the linear regression of the experimental versus predicted response with correlation coefficient R. The index R^2 is used as a measure of the accuracy of the prediction. The unit's RFs are plotted on the right. (A) unit 185/13*c*; (B) unit 185/20*a*; (C) unit 185/35*a*.

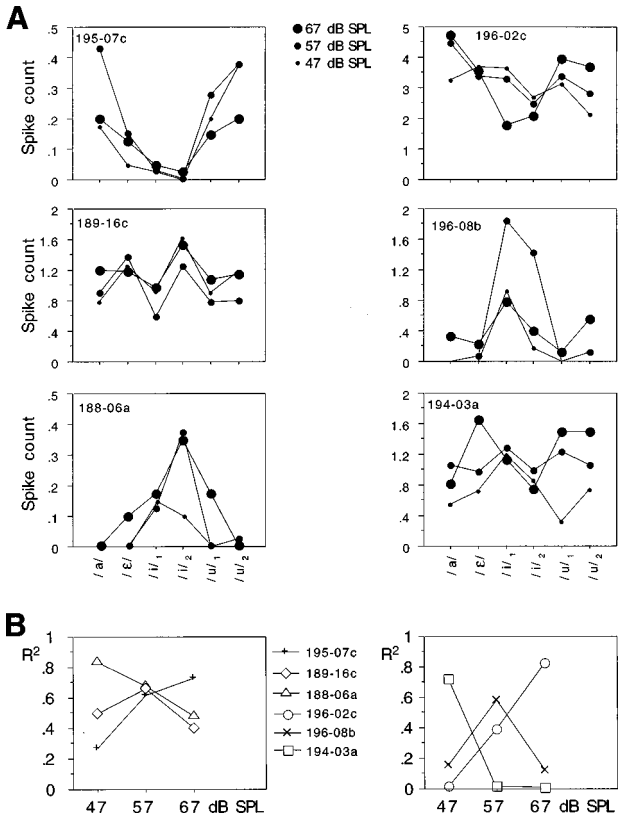


FIG. 10. Level dependence of voiced-vowel responses and their predictions. (A) The response curves at three different levels (47, 57, and 67 dB SPL, represented by different dot sizes) for six different units. The plots on the left show examples where response curves have similar shapes, the plots on the right show examples where the curves change with level. (B) The index $R^2_{\mathrm{RF}v}$ as a function of intensity for the vowel responses shown in (A). The index $R^2_{\max}$, is defined by the maximum. The plot on the left and right correspond to the left- and rightside plots in (A), respectively.

and the corresponding R as a function of level (B). In several cases the level dependence of the relative vowel responses was small which is illustrated by the three plots on the left in Fig. 10(A): all response curves have similar shapes. Consequently, R is stable with level [left plot in Fig. 10(B)]. However, the responses often varied considerably [Fig. 10(A), right-side plots] and consequently the prediction of the vowel responses could be virtually perfect ($R^2>0.7$) at one level and absolutely false ($R^2\approx 0$) at another level [Fig. 10(B), right plot]. In 30% of the units where we found a decrease of R^2 with increase of intensity the drop was more than 0.4 over 10 dB. These findings were similar for voiced and unvoiced vowels. In the following, the maximum correlation coefficient will be referred to as $R_{\max}$.

We examined whether the best-prediction level relates to the best intensity and/or threshold derived from the rate-level function measured for BF tones. If a change in relative vowel responses giving rise to the level dependence of R^2 is related to nonmonotonic response behavior then one would expect a relation to best intensity. However, multiple regression analysis, with threshold and best intensity as independent variables, indicates that the best-prediction level correlates with threshold but not with best intensity. The regression plots in Fig. 11 demonstrate that the intensity of the best prediction correlates well with threshold at BF for the unvoiced vowels ($r=0.63$) but very weakly for the voiced vowels ($r=0.17$).

The histograms in Fig. 12(A), (B) represent the distributions of the maximum squared correlation coefficient, $R^2_{\max,\mathrm{RF}v}$ across units for voiced (A) and $R^2_{\max,\mathrm{RF}u}$ for unvoiced vowels (B). In both histograms the data are split into three groups distinguished by the difference between BF estimates, $|\mathrm{BF_{tone}}-\mathrm{BF_{RF}}|$. In both the voiced and unvoiced case the percentage of units with $R^2>0.4$ is significantly larger for units where the smallest BF difference was found (χ^2 test, $P<0.05$). For voiced vowels [Fig. 12(A)], 44% of the units with $|\mathrm{BF_{tone}}-\mathrm{BF_{RF}}|<0.25$ octave have an $R^2_{\max}$ greater than 0.7 (as the examples in Fig. 8), and 71% of these units have an R^2 greater than 0.4 reflecting a fair to good prediction. For unvoiced vowels [Fig. 12(B)], $R^2_{\max}$ is greater than 0.7 for 51% of the units with $|\mathrm{BF_{tone}}-\mathrm{BF_{RF}}|<0.25$ octave, and greater than 0.4 for 78%. Table I shows the percentages of units with $R^2>0.4$ and the BF difference less than 0.25 octave at the various sound levels. Comparing the prediction results as just described between unvoiced and voiced vowels we see that the results are slightly better for unvoiced vowels, but $R^2_{\max}$ is not significantly larger for unvoiced than for voiced according to a nonparametric test (Wilcoxon test, $p>0.1$).

We examined whether the indices $R^2_{\max}$ vary with any of the characteristic response parameters threshold, best intensity, Ω_0, or ϕ_0. $R^2_{\max}$ was larger for units with a medium Ω_0 (0.4–0.8 cycles/oct) than for high or low Ω_0, which was

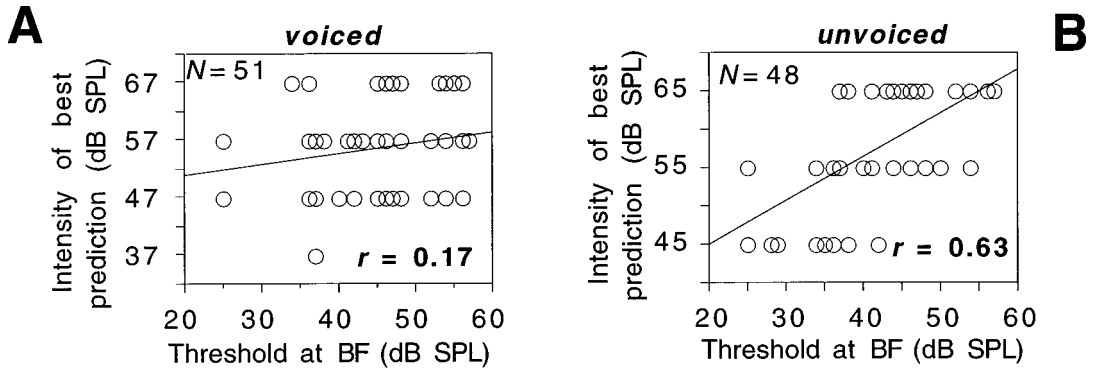


FIG. 11. Intensity of best prediction versus threshold at BF for voiced (A) and unvoiced (B) vowels. The intensity of best prediction is determined if data at three intensity levels are available and $R_{max}>0.4$, or if data at two intensity levels are available and $R_{max}>0.9$. For all AI units $|BF_{tone}-BF_{RF}|<0.5$ octave.

significant for both voiced and unvoiced vowels ($p<0.05$, ANOVA F test). There was no significant relation for other parameters.

The spectral envelope is used as the stimulus descriptor in the prediction [see Eq. (4)]. If indeed the response is determined by the spectral envelope (and not by, e.g., the harmonic fine structure), then the responses to the voiced and unvoiced vowels, which have the same envelope but different carriers (Fig. 2), would be expected to be very similar. Figure 12(C) demonstrates this similarity for a majority of neurons. Figure 12(C) shows the distribution of the index $R^2_{max,vu}$, which is a measure of the match between the responses to voiced and unvoiced for the sound levels with the best match: 61% of the neurons have R^2 greater than 0.7 (as in Fig. 7), and 80% have R^2 greater than 0.4.

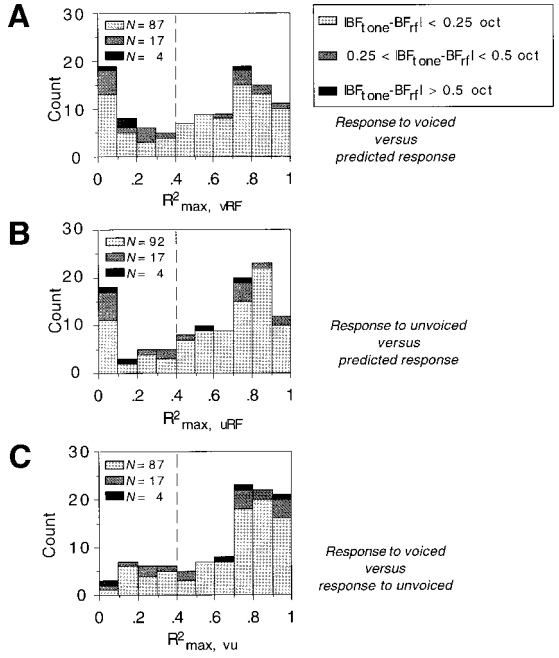


FIG. 12. Distributions of squared correlation coefficients at the level of best correlation, R^2_{max}. The neurons are distinguished on the basis of the difference in tonal and ripple BF, $|BF_{tone}-BF_{RF}|$, which is an indicator for the accuracy of the RF and BF measurement. If $R<0$, R^2 was made 0. (A) Distribution of $R^2_{max,RFv}$ for correlations between responses to voiced vowels and predicted responses. (B) Distribution of $R^2_{max,RFu}$ for responses to unvoiced vowels versus predicted responses. (C) Distribution of $R^2_{max,vu}$ for responses to unvoiced versus responses to voiced vowels.

E. Predictions: BF approach

We wanted to know whether the entire RF is needed to predict the vowel responses or instead information of the BF location alone might be sufficient. Thus, the vowel response was predicted assuming that the RF only consisted of a narrow peak at BF [Eq. (5)]. Figure 13 shows the experimental responses versus these BF-predicted responses for the same three units for which the RF-predicted responses are shown in Fig. 9. The predictions as judged from the correlation coefficients R are fair to good ($R^2>0.4$). However, BF predictions are worse than RF predictions, especially in one case (C) where the decrease in R^2 is considerable.

Focusing on the latter case we see that the responses to /a/, /u/$_1$, and /u/$_2$ are predicted to be similar according to the BF prediction while the response to /u/$_1$ is by far the largest according to the RF prediction [Fig. 9(C)]. The reason for the difference between the predictions is illustrated in Fig. 14 which shows the RF and vowel profiles. All three vowels, /a/, /u/$_1$, and /u/$_2$, have a formant peak around BF giving rise to excitation. However, unlike /u/$_1$, both vowels /a/ and /u/$_2$ have a formant peak at frequencies corresponding to the RF inhibitory sideband above the BF, which gives rise to significant inhibition canceling the excitatory response. In cases A and B the effects of inhibition are more subtle. For instance, examination of the RF of case A (shown in Fig. 5) reveals that both BF and RF predictions should be very similar (large response for the /i/'s, smallest response for /u/$_1$).

The difference between R^2 obtained from the RF prediction and BF prediction for the examples shown in Figs. 9 and 13 were 0.1–0.3. These differences are typical as illustrated in Fig. 15 which shows distributions of these differences, $R^2_{RF}-R^2_{BF}$, across all units for voiced (A) and unvoiced (B)

TABLE I. Percentage of units with $R^2>0.4$, BFdif<0.25 octave.

Vowel type	SPL (dB)	Percentage	N
voiced	67	45	64
voiced	57	51	86
voiced	47	46	70
voiced	best level	71	87
unvoiced	65	55	91
unvoiced	55	60	87
unvoiced	45	47	30
unvoiced	best level	78	92

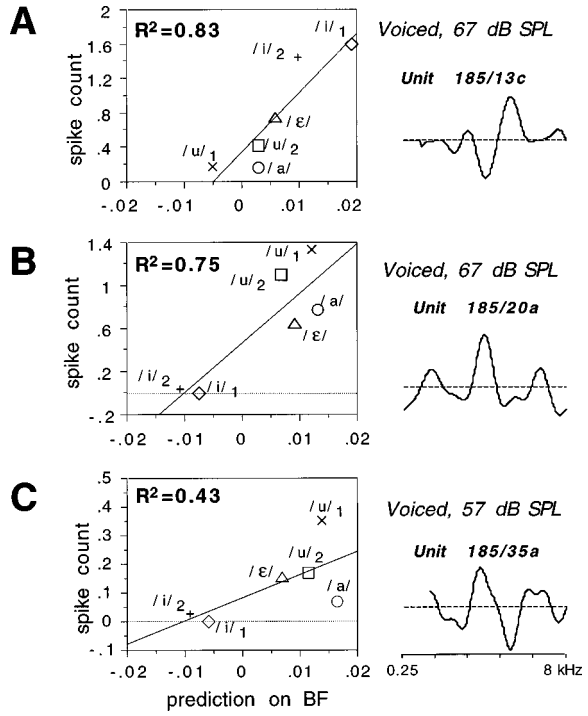


FIG. 13. Examples of spike count (normalized per stimulus presentation) versus responses predicted on the basis of BF and spectral envelope [see Eq. (5)], for voiced vowels, for the three units shown in Fig. 9. See Fig. 9 for further details. (A) unit 185/13*c*; (B) unit 185/20*a*; (C) unit 185/35*a*.

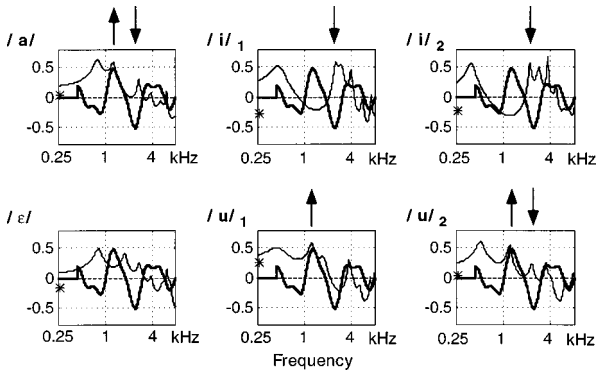


FIG. 14. RF of AI unit 185/35*a* (thick line) with vowel's spectral profiles (thin lines). The vowel profiles are scaled and aligned to the RF. The ordinate is given in RF units (spike count). The upward and downward arrows indicate significant excitatory and inhibitory responses, respectively, as expected from the RF prediction.

vowels. The distribution histograms reveal that for a majority of units the prediction is better, often to a small extent, when based on the entire RF than when based solely on BF. For units with $|\mathrm{BF_{tone}}-\mathrm{BF_{RF}}|<0.25$ octave the R^2 difference is larger than for units with $0.25<|\mathrm{BF_{tone}}-\mathrm{BF_{RF}}|<0.5$ octave, which is statistically significant for voiced vowels at 57 dB SPL [Fig. 15(A)] and for unvoiced vowels at 55 and 65 dB SPL ($p<0.05$, Mann–Whitney test). The results of the statistical analysis of the R^2 difference of units with $|\mathrm{BF_{tone}}-\mathrm{BF_{RF}}|<0.25$ octave are given in Table II and show that the R^2 difference is small but significant for all but one sound level ($p<0.05$, Wilcoxon test).

In Fig. 16 the distributions shown in Fig. 15 are plotted as a function of BF in the same mode as Fig. 8. Two features are evident: between 1.4 and 2.8 kHz where the response and vowel profiles are similar (Fig. 8) the RF prediction is better than the BF prediction, and around 4 kHz where the largest discrepancy between response and vowel profiles was found the RF prediction is worse than the BF prediction.

The difference between R's did not significantly vary with any of the parameters threshold, best intensity, Ω_0, or ϕ_0.

III. DISCUSSION

A. Encoding of vowels at central auditory level

Few data are known on vowel responses at high auditory centers. In contrast, numerous reports have addressed the neural representation of vowel or vowellike sounds in the peripheral auditory system (e.g., Sachs and Young, 1979; Young and Sachs, 1979; Delgutte and Kiang, 1984; Palmer *et al.*, 1986; see overview in Palmer, 1996). The main thread throughout these reports is the question of place and time representation. The auditory nerve and cochlear nucleus have the response properties to use either a rate-place or temporal-place coding scheme to transfer essential information on the vowel to the central system.

The auditory cortex probably plays a role in the processing of natural complex sounds. This role is evident in the bat (Suga, 1984), but not in other mammals in spite of several studies which used in particular animal vocalizations (Newman and Wollberg, 1973; Winter and Funkenstein, 1973; Sovijärvi, 1975). Since those early studies the use of complex stimuli for the study of the cortex in mammals other than the bat has been limited, and more commonly, simple stimuli and semi-complex stimuli as amplitude- and frequency-modulated tones were used. Recently, more use has been made of complex and natural stimuli to characterize cortical responses (Schreiner and Calhoun, 1994; Schreiner and Wong, 1996; Shamma *et al.*, 1995; Wang *et al.*, 1995; Kowalski *et al.*, 1996a; Ohl and Scheich, 1997).

Shamma and Versnel (1995) suggested that for a majority of cells in AI the response to a broadband sound stimulus (expressed in spike count) is based on a linear summation of responses to the ripple components of the spectral envelope of the stimulus. In that paper the sounds used were purposely composed of a few ripple components in order to have a direct test of a ripple superposition. It was still unclear, however, whether this concept, which we refer to as ripple analysis, would in general apply to broadband sounds with any spectral profile. In this paper we addressed this question by examining the responses to vowel sounds in AI in relation-

TABLE II. Difference between goodness of RF and BF prediction for units with BFdif<0.25 octave. The Wilcoxon test is applied for the statistical analysis.

Vowel type	SPL (dB)	R^2rf-R^2bf	P	N
voiced	67	0.01	NS	64
voiced	57	0.08	0.002	86
voiced	47	0.08	0.03	70
unvoiced	65	0.10	0.003	91
unvoiced	55	0.07	0.02	87
unvoiced	45	0.10	NS	30

ship to the ripple responses. The major findings are as follows. First, for most AI cells (71%) the relative responses to natural vowels could be predicted on the basis of the cell's response field (which consists of both excitatory and inhibitory regions) and the smoothed spectral envelope. Second, the responses could be predicted to a fair extent on the basis of the BF (rate-place coding), but predictions on the basis of RF were significantly better. These findings indicate that for most units in AI the ripple analysis applies to natural vowels and to stimuli with natural spectral profiles, thus confirming the ripple analysis as feasible method to characterize Al responses to broadband sounds.

Ohl and Scheich (1997) combined 2DG mapping of responses to synthetic 2-formant vowels with single unit recordings in gerbil AI. Their study demonstrated that spectral profiles are analyzed by features that characterize the entire RF such as bandwidth and inhibitory side bands, thus it supports our findings. Wang *et al.* (1995) found in marmoset AI that the major features of marmoset's vocalization spectra were reflected by the response profiles. This agrees with our finding that the vowels, to some extent, can be encoded on the basis of rate-place information, point-to-point. One might wonder why this straight forward BF coding scheme, although inferior to the ripple coding, holds reasonably well. First, according to the ripple analysis the response to the BF component of the stimulus is included in the total response. Second, when the vowel has a formant peak around BF giving rise to strong excitation, often the troughs of the spectrum lie at the inhibitory regions of the RF giving rise to a weak inhibitory contribution to the total response (e.g., Fig. 5, /i/$_1$), and vice versa, i.e., a weak excitatory component of the vowel response accompanies a strong inhibitory component (e.g., Fig. 5, /u/$_1$). Consequently, the inhibitory effects enhance the contrast between responses to different spectra that already exist due to the excitatory components. Thus it might be a direct consequence of ripple coding that place coding appears to exist.

B. Factors that influence the prediction

There are various factors which influence the prediction of the vowel response on the basis of the ripple model. There might be conditions under which the prediction model [Eq. (4)] is not valid. In testing the model we assume that the RF is well estimated, that the fine structure of the spectrum can be neglected, and that the vowel responses are measured correctly. In the following we discuss assumptions and conditions involved in the prediction.

1. Estimation of the response field

In applying the prediction model formulated in Eq. (4) it is assumed that an Al cell has one unique RF. However, it is not obvious that this is the case. In the primary visual cortex it appears that the receptive field is not static (Gilbert *et al.*, 1996). For instance, the response of a VI cell to stimuli within the receptive field appears to depend on the context in which those stimuli are presented. Thus the RF depends on the stimulus used. Is this also true for AI? We found that RFs measured with tones were generally similar to RFs especially with respect to BF. However, level dependence for the two

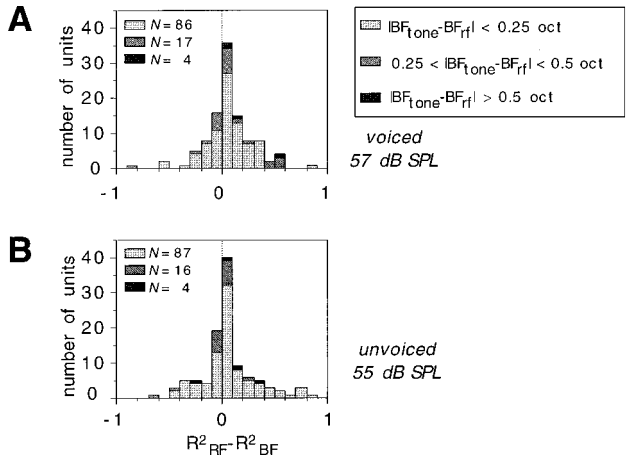


FIG. 15. Distributions of differences between R^2_{RF} reflecting the goodness of the RF prediction and R^2_{BF} reflecting the goodness of the BF prediction. Units are distinguished on the basis of the difference in tonal and ripple BF, $|BF_{tone}-BF_{RF}|$, which is an indicator for the accuracy of the RF and BF measurement. (A) Distribution of $R^2_{RF}-R^2_{BF}$ for voiced vowels at 57 dB SPL. (B) Distribution of $R^2_{RF}-R^2_{BF}$ for unvoiced vowels at 55 dB SPL.

measures is quite different as RFs measured by ripples are quite stable over a 20–30 dB range (Shamma *et al.*, 1995) while tuning to tones can vary dramatically over that same range. We ascribe this discrepancy to nonlinearities that exist along the auditory pathway and are more manifest for narrow band than for broadband stimuli (Shamma and Versnel, 1995). Therefore, it is expected that RFs derived from responses to broadband stimuli would be superior to tonal response measures in order to predict responses. Apart from this principle argument, there are more practical reasons to prefer broadband sounds above narrow-band sounds (see

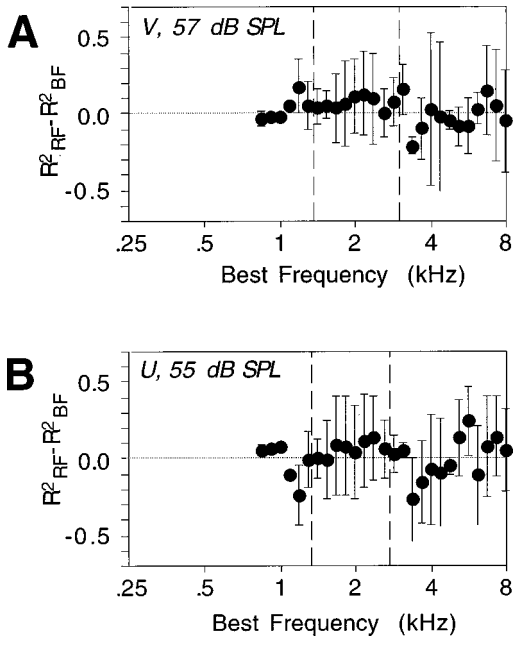


FIG. 16. Average differences between R^2_{RF} and R^2_{BF} vs BF for voiced (A) and unvoiced (B) vowels. The data correspond to those presented in Fig. 15, and all data are included (both units with small and units with large BF differences). The values of $R^2_{RF}-R^2_{BF}$ are averaged over units in 1/4 octave moving segments at 1/8 octave resolution. For further explanation of plots see Figs. 6 and 8.

also Shamma and Versnel, 1995). For instance, the lack of spontaneous activity causes a problem in measuring inhibitory regions using tones.

The ripple responses can be affected by nonlinearities as half-wave rectification and saturation (extensively discussed in Shamma *et al.*, 1995). At low levels the response is affected by rectification, and at high levels by saturation. Although ripple responses are quite robust with level over a 20–30 dB range, other nonlinear effects than half-wave rectification and saturation cannot be ruled out. In this experiment potential errors due to these nonlinearities were largely avoided by using medium levels, which were in the same range of the levels of the vowels.

We assumed that the RF integrates to zero, but it is well possible that either the excitatory or inhibitory regions dominate (which could vary with level, e.g., in nonmonotonic units the inhibitory portion of the RF might be larger at high levels). The dip in the response profile around 4 kHz (Fig. 8) which was unexpected considering the local peak in the vowel spectrum could be caused by inhibitory effects which were stronger than estimated. The relatively low best intensity around 4 kHz [Fig. 6(B)] suggests such strong inhibitory effects. Finally, the RF estimate is affected by statistical noise. A comparison with tonal responses, reflected by the BF difference $|BF_{tone}-BF_{RF}|$ was used as criterion to distinguish RF estimates effected by noise or nonlinearities.

2. Fine spectral structure

The predictions are based on the RF and smoothed spectral envelope, thus neglecting the carrier of the spectrum. A large effect of, e.g., the fine harmonic structure is not expected as this fine structure contains high ripple frequencies (above 2 cycles/octave) which will be filtered by most neurons [highest Ω_0 was 1.2 cycles/octave, Fig. 4(C)]. Similar results obtained for unvoiced and voiced vowels [Fig. 12(C)] confirm that the effect of the carrier is small. The responses to unvoiced vowels can be slightly better predicted [Fig. 12(A), (B); Table I]. This was expected since the carrier is the same as for the rippled stimuli which were used to derive the RF. Yet, the difference is small which underlines that the response is primarily determined by the global spectral structure.

C. Nonlinearities of vowel responses

Half-wave rectification and saturation will affect vowel responses as these nonlinearities affect ripple responses (see above). An example of possible half-wave rectification in vowel responses is the response to $/u/_1$ in Fig. 9(A): this is predicted to be smaller than that to /a/, but the response to /a/ is virtually zero, and consequently the response to $/u/_1$ cannot be smaller. Shamma and Versnel (1995) assessed the subthreshold and supersaturation regions of the responses to multiple-ripple profiles, thus reconstructing the measured response. The correlation between reconstructed and predicted responses was significantly larger than that found between measured and predicted (R increased by 0.11). The biophysical correlate of the reconstructed signal can be thought of as the synaptic potential. The synaptic potential which underlies the neural activity exhibits less saturation and negative responses, and thus is expected to behave more linearly. Jagadeesh *et al.* (1993) experimentally confirmed this by demonstrating that the membrane potential of visual cortical cells evoked by a complex stimulus (moving grating) can be accurately predicted assuming linearity whereas prediction of the corresponding neural response requires nonlinear components (Reid *et al.*, 1991). It would be of great interest to record intracellularly in AI, analogously to Jagadeesh's experiments, to test the ripple analysis without the distorting effects of threshold and saturation.

D. Level dependence of vowel responses

In a majority of the units recorded the relative vowel responses did not vary significantly with absolute level over a 20-dB range [Fig. 10(A)], which is consistent with our findings on multiple-ripple profiles (Shamma and Versnel, 1995). In several cases however, the vowel response often changed with level, sometimes to an extent that within 10 dB a response changed from well predictable to unpredictable [Fig. 10(B)]. These were changes that cannot only be ascribed to saturation or half-wave rectification. We conclude that in general there is an optimal level for which the vowel response can be predicted on the basis of the ripple model and that the linear ripple analysis is restricted to a certain level range. This leaves us with the question how the cortex can apply its ripple analysis to encode profiles over a large range of sound levels. One possibility is that the cortex utilizes the spread of optimal levels across units (Figs. 10 and 11) to encode by means of the ripple analysis at any level. It is likely that optimal levels in mammalian AI in general are represented over a large range given the correlation of optimal level with threshold (Fig. 11) and the large range of thresholds available along isofrequency contours (Schreiner *et al.*, 1992). The ripple analysis model as discussed in previous papers proposes AI to analyze a profile along three dimensions which are spatially organized across AI (BF, ripple frequency, ripple phase). The optimal level could be added as a fourth dimension.

Responses at levels outside the optimal range (not following the ripple analysis) remain puzzling as they seem to distort the ripple encoding of spectral profiles. One might have to assume that at a particular level the responses of units evoked at their optimal level dominate the responses of units responding outside their optimal levels.

IV. CONCLUSIONS

In this paper we have demonstrated that for most units in AI responses to natural vowels can be predicted by cross-correlating the RF and the smoothed spectral envelope of the vowel. In other words, the vowel responses are based on a linear summation of responses to the ripple components of the smoothed spectral envelope of the vowel. This confirms findings in our previous paper with respect to AI responses to synthetic spectra (Shamma and Versnel, 1995). Thus more evidence is provided that AI performs a Fourier-like analysis on the spectral profile of a broadband sound rather than a point-to-point analysis (as a rate-place representation). This

does not mean that for a full description of the responses of the system (which after all integrates several nonlinear processes), a linear model is adequate. For instance, we found that the superposition rule applies to a restricted sound level range, which for some units is quite narrow (less than 20 dB, see Fig. 10). In practical terms our findings imply that given the response field of the AI cell and given the spectral profile of the sound the response magnitude can be predicted relative to the response to other sounds of similar sound level.

ACKNOWLEDGMENTS

This work is supported by grants from the Air Force Office of Scientific Research, and from the Office of Naval Research. We would like to thank P. Gopalaswamy for his help in developing the data acquisition system, P.-W. Ru for his help in preparing the vowel stimuli, and A. L. Owens, D. A. Depireux, and N. A. Kowalski for their assistance in animal preparation and data recordings. The research was carried out at the Institute for Systems Research which is partially funded by an NSF grant No. (NSFD CD 8803012).

Delgutte, B., and Kiang, N. Y. S. (**1984**). ''Speech coding in the auditory nerve I. Vowel-like sounds,'' J. Acoust. Soc. Am. **75**, 866–878.

De Valois, R., and De Valois, K. (**1988**). *Spatial Vision* (Oxford U.P., New York).

Evans, E. F. (**1979**). ''Single unit studies of mammalian cochlear nerve,'' in *Auditory Investigation: The Technological and Scientific Basis*, edited by H. A. Beagley (Clarendon Press, Oxford), pp. 324–367.

Gilbert, C. D., Das, A., Ito, M., Kapadia, M., and Westheimer, G. (**1996**). ''Spatial integration and cortical dynamics,'' Proc. Natl. Acad. Sci. USA **93**, 615–622.

Jagadeesh, B., Wheat, H. S., and Ferster, D. (**1993**). ''Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex,'' Science **262**, 1901–1904.

Kelly, J. B., Judge, P. W., and Phillips, D. P. (**1986**). ''Representation of the cochlea in primary auditory cortex of the ferret (*Mustela putorius*),'' Hearing Res. **24**, 111–115.

Kowalski, N., Depireux, D. A., and Shamma, S. A. (**1996a**). ''Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra,'' J. Neurophysiol. **76**, 3503–3523.

Kowalski, N., Depireux, D. A., and Shamma, S. A. (**1996b**). ''Analysis of dynamic spectra in ferret primary auditory cortex. II. Prediction of unit responses to arbitrary dynamic spectra,'' J. Neurophysiol. **76**, 3524–3534.

Kowalski, N., Versnel, H., and Shamma, S. A. (**1995**). ''Comparison of responses in the anterior and primary auditory fields of the ferret cortex,'' J. Neurophysiol. **73**, 1513–1523.

Newman, J. D., and Wollberg, Z. (**1973**). ''Multiple coding of species-specific vocalizations in the auditory cortex of squirrel monkeys,'' Brain Res. **54**, 287–304.

Ohl, F., and Scheich, H. (**1997**). ''Orderly cortical representation of vowels based on formant interaction,'' Proc. Natl. Acad. Sci. USA **94**, 9440–9444.

Palmer, A. R. (**1996**). ''Low-level processing of speech sounds in the auditory nervous system,'' in *Workshop on the Auditory Basis of Speech Perception*, edited by W. Ainsworth and S. Greenberg (Keele University, UK), pp. 49–56.

Palmer, A. R., Winter, I. M., and Darwin, C. J. (**1986**). ''The representation of steady-state vowel sounds in the temporal discharge patterns of the guinea pig cochlear nerve and primarylike cochlear nucleus neurons,'' J. Acoust. Soc. Am. **79**, 100–113.

Phillips, D. P., Ormand, S. S., Musicant, A. D., and Wilson, G. F. (**1985**). ''Neurons in the cat's primary auditory cortex distinguished by their responses to tones and wide-spectrum noise,'' Hearing Res. **18**, 73–86.

Plomp, R. (**1976**). *Aspects of Tone Sensation* (Academic, London).

Rabiner, L. R., and Juang, B. H. (**1993**). *Fundamentals of Speech Recognition* (Prentice-Hall, Englewood Cliffs, NJ).

Reid, R. C., Soodak, R. E., and Shapley, R. M. (**1991**). ''Directional selectivity and spatiotemporal structure of receptive fields of simple cells in cat striate cortex,'' J. Neurophysiol. **66**, 505–529.

Sachs, M., and Young, E. (**1979**). ''Encoding of steady state vowels in the auditory nerve: representation in terms of discharge rate,'' J. Acoust. Soc. Am. **66**, 470–479.

Schreiner, C. E., and Calhoun, B. M. (**1994**). ''Spectral envelope coding in cat primary auditory cortex: properties of ripple transfer functions,'' Aud. Neurosci. **1**, 39–61.

Schreiner, C. E., and Mendelson, J. R. (**1990**). ''Functional topography of cat primary auditory cortex: Distribution of integrated excitation,'' J. Neurophysiol. **64**, 1442–1459.

Schreiner, C. E., and Sutter, M. L. (**1992**). ''Topography of excitatory bandwidth in cat primary auditory cortex: Single-neuron versus multiple-neuron recordings,'' J. Neurophysiol. **68**, 1487–1502.

Schreiner, C. E., Mendelson, J. R., and Sutter, M. L. (**1992**). ''Functional topography of cat primary auditory cortex: Representation of tone intensity,'' Exp. Brain Res. **92**, 105–122.

Schreiner, C. E., and Wong, S. W. (**1996**). ''Spatial-temporal representation of syllables in cat primary auditory cortex,'' in *Workshop on the Auditory Basis of Speech Perception*, edited by W. Ainsworth and S. Greenberg (Keele University, UK), pp. 127–132.

Shamma, S. A., and Versnel, H. (**1995**). ''Ripple analysis in ferret primary auditory cortex. II. Prediction of unit responses to arbitrary spectral profiles,'' Aud. Neurosci. **1**, 255–270.

Shamma, S. A., Fleshman, J. W., Wiser, P. R., and Versnel, H. (**1993**). ''Organization of response areas in ferret primary auditory cortex,'' J. Neurophysiol. **69**, 367–383.

Shamma, S. A., Versnel, H., and Kowalski, N. (**1995**). ''Ripple analysis in ferret primary auditory cortex. I. Response characteristics of single units to sinusoidally rippled spectra,'' Aud. Neurosci. **1**, 233–254.

Sovijärvi, A. R. A. (**1975**). ''Detection of natural complex sounds by cells in the primary auditory cortex of the cat,'' Acta Physiol. Scand. **93**, 318–335.

Suga, N. (**1984**). ''The extent to which biosonar information is represented in the bat auditory cortex,'' in *Dynamic Aspects of Neocortical Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 315–373.

Versnel, H., Kowalski, N., and Shamma, S. A. (**1995**). ''Ripple analysis in ferret primary auditory cortex. III. Topographic distribution of ripple response parameters,'' Aud. Neurosci. **1**, 271–286.

Wang, K., and Shamma, S. A. (**1995**). ''Spectral shape analysis in the central auditory system,'' IEEE Trans. Speech Audio Process. **3**, 382–395.

Wang, X., Merzenich, M. M., Beitel, R., and Schreiner, C. E. (**1995**). ''Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: Temporal and spectral characteristics,'' J. Neurophysiol. **74**, 2685–2706.

Winter, P., and Funkenstein, H. H. (**1973**). ''The effect of species-specific vocalization on the discharge of auditory cortical cells in the awake squirrel monkey (Saimiri sciureus),'' Exp. Brain Res. **18**, 489–504.

Young, E. D., and Sachs, M. B. (**1979**). ''Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers,'' J. Acoust. Soc. Am. **66**, 1381–1403.

Within-channel gap detection using dissimilar markers in cochlear implant listeners

Monita Chatterjee,[a] Qian-Jie Fu, and Robert V. Shannon
Department of Auditory Implants and Perception, House Ear Institute, 2100 W. Third Street, Los Angeles, California 90057

(Received 25 June 1997; revised 23 October 1997; accepted 15 January 1998)

Temporal gap detection thresholds were measured between perceptually dissimilar electrical markers in cochlear implant listeners. Both markers were presented to the same electrode pair. The amplitude and pulse rate of the first marker were fixed, and gap thresholds were measured as a function of either the pulse rate or the amplitude of the second marker. In either case, U-shaped functions were obtained, with lowest gap thresholds occurring when the two markers were similar in both amplitude and pulse rate. Because the two markers were presented to the same electrode pair, the data cannot be accounted for on the basis of across-channel interactions. It is hypothesized that when different markers are used, the perceptual discontinuity from the first marker to the second is similar to the sensation of a brief gap, and dominates the gap detection process. Thus, gap threshold functions with electrically dissimilar markers serve more as indicators of perceptual distance between the markers and less as measures of temporal resolution. © *1998 Acoustical Society of America.* [S0001-4966(98)01205-3]

PACS numbers: 43.66.Ba, 43.66.Mk, 43.66.Ts, 43.64.Me [RHD]

INTRODUCTION

In a typical gap detection experiment, the subject is required to detect the presence of a silent temporal gap between two markers. When the two markers are identical, the task can be thought of as a ''within-channel'' task. The detection of the gap in this case will be limited only by the ability of the system to encode the onset of the second marker while it is recovering from the first marker. When the two markers are different in frequency by an amount exceeding a critical bandwidth, the detection of the gap is more difficult. In normal-hearing listeners, one can think of such a task as an ''across-channel'' task because, in this case, the detection of the gap involves the detection of a silent interval between the offset of excitation in one frequency region, and an onset in a distant frequency region. Such a comparison may involve more processing in higher centers of the auditory system than in the ''within-channel'' situation. Another factor likely to influence the detection of a gap is the amplitude scaling performed by the auditory system. For instance, a compressive transformation will reduce the size of the amplitude change signaling the gap. Consequently, although gap threshold has often been considered to be a measure of temporal resolution, researchers have argued that gap detection combines analysis of temporal aspects of the stimulus with analysis of dynamic amplitude changes (Moore and Glasberg, 1988; Moore *et al.*, 1989).

Gap detection thresholds are shortest when the two markers flanking it are identical. When the markers are different, an apparent gap is heard even when there is no actual gap between them. We hypothesize that the primary cue is the system's response to the discontinuity in signal representation, not a discontinuity in time *per se*. Thus, in a two-interval forced choice gap detection task with different markers, both the gap and no-gap intervals would sound as though there is a discontinuity, and gap thresholds would be longer than when the two markers are similar.

In acoustic hearing, the frequency analysis performed by the cochlea, and the broad spread of excitation across different frequency regions, makes it difficult to separate ''within channel'' from ''across channel'' processes. In contrast, in electrically stimulated hearing, one can present two stimuli of different frequency to the same electrode pair, and be reasonably certain that the same neural population is responding to both markers. This makes it possible to explore ''within channel'' mechanisms of perception in a way that would not be possible in acoustic hearing.

The present research stems from work done by Hanekom and Shannon (1997) using ''across-channel'' gap detection as a measure of channel interaction in multielectrode cochlear implant listeners. In that study, the markers flanking the gap were identical signals, but presented to different electrode pairs (the amplitudes were loudness balanced). This created a discontinuity along the location of stimulation in the cochlea (the perceptual correlate of which may be pitch, or timbre, of the sound). Gap detection thresholds were measured as a function of the electrode separation between the first and second markers. When plotted as gap threshold versus electrode number, lowest thresholds were obtained when the two markers were presented to the same electrode, while thresholds rose on either side of this minimum, thus forming a ''tuning'' curve.

In this paper, we describe the results of a study of gap detection using different markers presented to the same electrode pair. It is anticipated that the results will provide an understanding of within-channel gap detection mechanisms, against which we can compare results from across-channel gap detection experiments.

[a]Electronic mail: monita@hei.org

Our experiments were performed with users of the Nucleus 22 multielectrode cochlear implant. Only one electrode pair was stimulated (BP+1 mode). The stimulus comprised a silent gap flanked by markers that were identical (experiment 1), different in amplitude (experiment 2), or different in pulse rate (experiment 3).

I. EXPERIMENTAL DESIGN AND METHODS

All three subjects were adult users of the Nucleus 22 cochlear implant device. They were all postlingually deafened. The subjects were highly trained in various psychophysical experiments.

Three experiments were performed. In experiment 1, the markers were identical in all respects. Gap threshold was measured as a function of marker amplitude, for three pulse rates: 125, 250, and 1000 pulses per second (pps). In experiment 2, both markers were presented at the same rate. The first marker amplitude was fixed at comfortable level, and gap thresholds were measured as a function of the amplitude of the second marker, for two pulse rates (125 and 1000 pps). In experiment 3, both markers were presented at the same subjective loudness. The rate of the first marker was fixed at one of three values (125, 500, or 1000 pps) and gap thresholds were measured as a function of the rate of the second marker. The second marker was always loudness balanced to the first, and both were presented at comfortable level.

The two markers were 100-ms-long, biphasic pulse trains (100 or 200 μs/phase) presented in BP+1 mode to the middle electrode pair (10–12). In this mode of stimulation, the active and indifferent electrodes are separated by one electrode. Thus, current flows from electrode 10 to electrode 12. Subjects N3 and N10 received 100 μs/phase pulse trains. Subject N4, however, required 200 μs/phase pulses to reach comfortable level.

Experiments were software controlled. Stimuli were presented using a custom Nucleus interface (Shannon *et al.*, 1990). The stimuli did not go through the implant processor, but were directly controlled through the interface. All experiments used an adaptive, three-down, one-up paradigm to measure gap thresholds in blocks of two-interval, forced choice trials. Randomly, one of the two intervals in each trial contained a silent gap between the markers. The subject pressed a mouse button to indicate which interval contained the gap. The average of the last 8 out of 12 reversals was computed as the mean gap threshold for each run. The initial (first three reversals) step size was generally twice the final step size, and both were set according to the anticipated threshold for each condition. Visual feedback was provided at the end of each trial. The three-down, one-up paradigm results in convergence at the 79.4% correct point (Levitt, 1971). The mean and standard deviation of two repetitions were calculated for each data point. When the two repetitions yielded widely different results, a third repetition was performed and the mean of all three was calculated.

Stimuli were loudness balanced using a two-up, one-down, 2IFC, double-staircase procedure (Jesteadt, 1980; Zeng and Turner, 1991). The standard was always the 1000-pps signal fixed at comfortable level. The subject heard the standard and a comparison signal in two successive intervals (random order) of a trial and indicated which interval contained the louder signal. The amplitude of the comparison signal was adjusted adaptively to the subject's responses. Each experimental run was composed of two tracks, in which the comparison signal was adjusted until it was heard to be louder 70.7% and 29.3% of the time, respectively. For each track, the last 8 out of 12 reversals were averaged to compute the mean, and the average of the means from the two tracks was computed as the point of subjective equality.

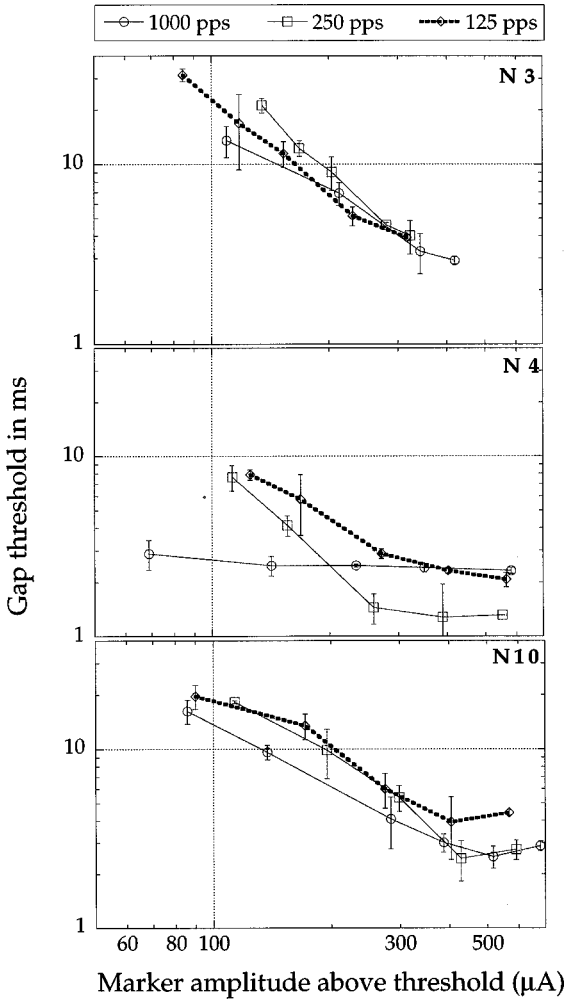


FIG. 1. Gap threshold as a function of marker amplitude in three subjects. The parameter is pulse rate. The two markers were always identical. The horizontal axis represents the level of the markers above threshold in μA.

II. RESULTS

A. Experiment 1: Gap detection as a function of level

The results of experiment 1 are summarized in Fig. 1. In agreement with previous work in acoustic stimulation (Buus and Florentine, 1985; Fitzgibbons, 1983; Plomp, 1964) and electrical stimulation (Shannon, 1986, 1989; Moore and Glasberg, 1988), gap detection thresholds decline with increasing marker level, reaching an asymptote in the range of 1–4 ms above an amplitude corresponding to moderate loudness. Below the asymptote, the function declines as a power function (straight line on a log–log scale). In one subject (N4), the 1000-pps function is quite different from the per-

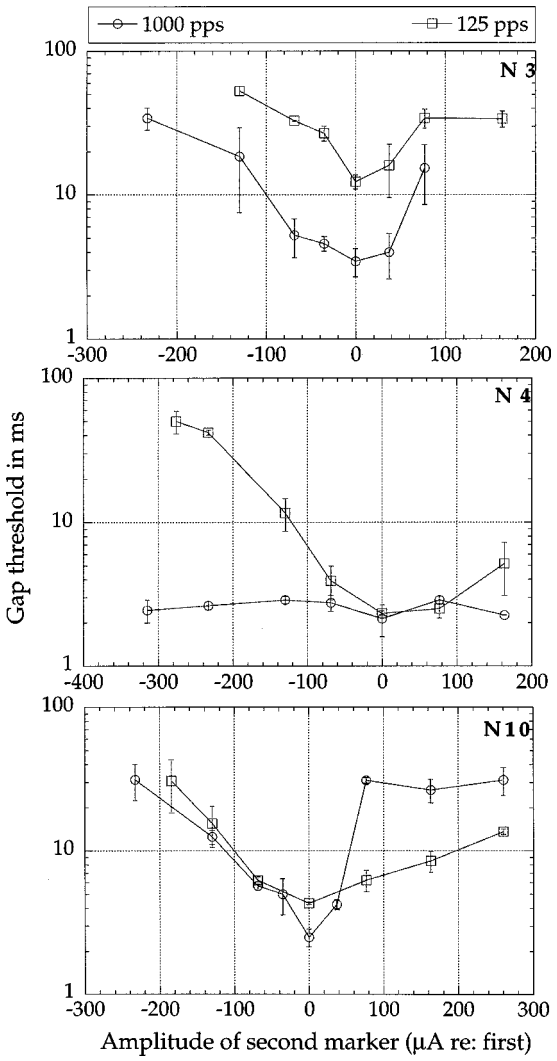


FIG. 2. Gap threshold as a function of the amplitude difference between the second and first markers (second marker−first marker) in μA. The first marker was fixed at comfortable level (zero along the abscissa). The independent variable is the pulse rate, which was identical for both markers.

formance functions for the other two pulse rates: at low current levels, gap thresholds are much lower at 1000 pps, and the slope of the function is much shallower than at 125 or 250 pps. Functions obtained in the other two subjects show little dependence on pulse rate.

B. Experiment 2: Gap detection with markers of unequal amplitudes

Results of experiment 2 are shown in Fig. 2. As in Fig. 1, the parameter is pulse rate (the same for both markers). For each pulse rate, the amplitude of the first marker was fixed at the most comfortable level. The amplitude of the second marker was fixed at different levels, ranging from below to above the level of the first marker. Gap threshold was measured for each level of the second marker.

The gap thresholds in Fig. 2 are plotted on a log axis as a function of the difference in amplitude (in μA) between the first and second markers. In general, gap threshold is minimum at the point where the amplitude of the second marker equals that of the first. Because the pulse rates of the two markers are equal, the markers are identical at the minimum threshold point. Above and below the point of equality, the perceptual discontinuity signaling the change in amplitude is confused with the perceptual discontinuity signaling the gap, and gap thresholds rise with increasing level difference.

It is evident that the shape of the function varies across subjects. The function also varies within subjects, as a function of rate and the level of the first marker. For example, for subject N4, there is little or no dependence on the level of the second marker when the pulse rate is 1000 pps. However, in the same subject and for the same electrode pair, there is considerable dependence on the level of the second marker when the pulse rate is 125 pps. Minimum gap thresholds are typically less than 5 ms, with the exception of subject N3, whose minimum threshold for the 125-pps pulse rate exceeded 10 ms.

C. Experiment 3: Gap detection with markers of unequal pulse rate

In this experiment, the pulse rate of the first marker was kept fixed at 125, 500, or 1000 pps. Gap thresholds were measured as a function of the pulse rate of the second marker, which was fixed at various values ranging from below to above the pulse rate of the first marker. At each point, the amplitude of each marker was loudness balanced to the comfortable level for the 1000-pps marker.

Results from this experiment are shown in Fig. 3. As in experiment 2, the functions show a broad minimum, usually less than or equal to 4 ms. The minimum threshold does not always correspond to the point at which the second marker equals the first in pulse rate. In some cases, the minimum occurs as much as an octave above the pulse rate of the first marker. This may be a reflection of poor pitch judgment among cochlear implant users. As in experiment 2, considerable between- and within-subject variability is observed in the shape of the functions.

III. DISCUSSION

Experiments that would be the direct counterpart in acoustic stimulation do not exist. Presumably, pairs of interrupted or modulated noise markers, presented at different rates or levels before and after the silent gap, would be most comparable to our electrical marker conditions. However, notwithstanding the differences in stimulus and perceptual characteristics, the pattern of our results is in qualitative agreement with similar studies of ''across channel'' gap detection in normal hearing (Formby and Forrest, 1991; Formby *et al.*, 1992; Phillips *et al.*, 1997). Whereas previous work with gap detection has been dominated by time domain analyses, Formby *et al.* (1996) account for their data using concepts of frequency analysis by the auditory system. Based on modeling efforts (Forrest and Formby, 1996; Heinz *et al.*, 1996), they concluded that across-channel gap detection threshold functions can only be accounted for by taking both peripheral filtering and central processing schemes into account.

Gap discrimination experiments of a similar nature performed by Divenyi and Danner (1977) and Divenyi and Sa-

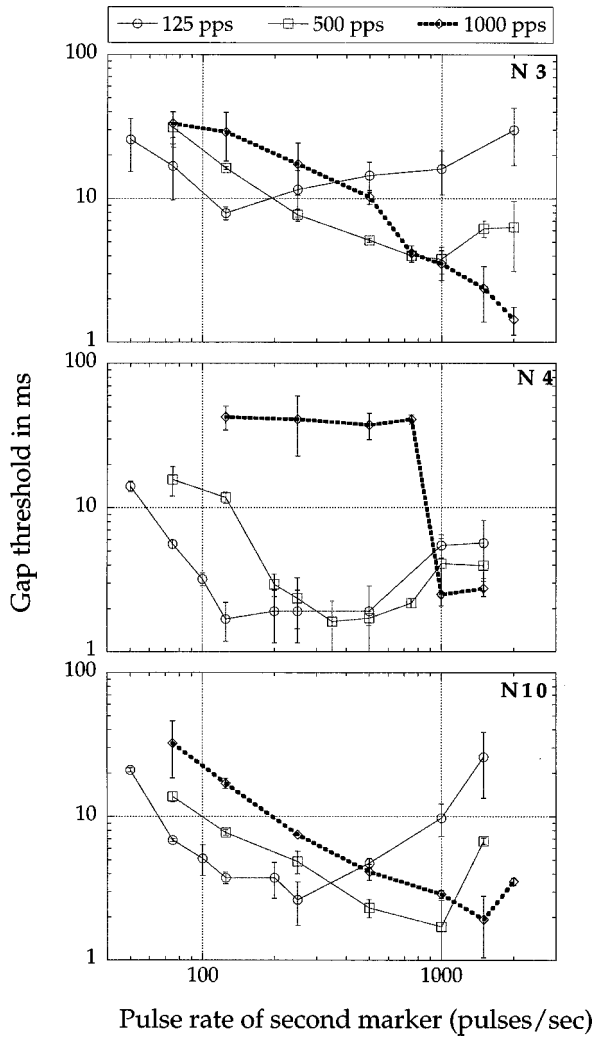


FIG. 3. Gap threshold as a function of the pulse rate of the second marker. The pulse rate of the first marker was kept fixed at 125, 500, or 1000 pps. The amplitudes of the two markers were individually balanced to comfortable level at 1000 pps.

chs (1978) have yielded similar results: the more spectrally similar the two markers, the more discriminable the silent temporal gap.

When experiments involving markers with different frequencies are performed in normal-hearing listeners, the frequency analysis performed by the cochlea maps the markers along the tonotopic axis. When the two markers fall within the same critical band, or ''channel,'' thresholds are lower than when their respective frequencies span more than one critical bandwidth. In experiment 3, the markers, although different in frequency, stimulate the same location in the cochlea. Therefore, we cannot attribute our results to within- versus across-channel mechanisms. The U-shaped functions measured using markers that are different in either frequency or amplitude, in single electrode pair stimulation of cochlear implant patients, indicate that similar functions observed in normal-hearing subjects may be at least partly derived from central mechanisms. However, peripheral capabilities will limit the perceptual distance perceived by the central detector: for example, in cochlear implant listeners, who can only perform temporal analysis on a sound presented to a single electrode, 500 and 1000 pps may sound roughly similar, whereas the normal-hearing listener can discriminate a far wider range of frequencies.

It is evident that gap detection measured with dissimilar marker pairs, which are asymmetric in level or rate, is heavily dependent on the magnitude of the perceptual discontinuity created by the asymmetry. This indicates that the task is not determined solely by peripheral mechanisms, but rather, that gap detection under these conditions is performed at a central stage where rapid changes in stimulus rate and amplitude have already been abstracted. From one perspective, the task of gap detection under these conditions is not a measure of temporal resolution but, rather, a measure of the perceived difference between the first marker and the second. The sharper the slopes of the function relating gap threshold to the difference between markers on either side of the minimum, the greater the subject's sensitivity to the difference. An interesting counterexample to this notion may be found in subject N4's data in Fig. 2: for the 1000-pps marker, his function shows the shallowest slope and the lowest gap thresholds. One may infer that the shallower the slopes of the function (the broader the U), the more the subject is able to attend to the temporal gap in spite of the confounding perceptual change. Thus, overall, the data show considerable differences within subjects under different parametric conditions that cannot be accounted for using either interpretation alone.

These data also emphasize the importance of loudness balancing two different markers in such an experiment as experiment 3. Small changes in perceived loudness will sharpen the slopes on either side of the minimum.

IV. SUMMARY

Gap detection thresholds were measured on a single electrode pair for electrical stimulation of cochlear implant patients, under conditions where the markers flanking the gaps were different in either pulse rate or amplitude. We found that gap detection thresholds increased as a function of the difference between the two markers. The functions generally showed a minimum in the region where the two markers were most similar. We conclude that gap detection under these conditions reflects the perceptual discontinuity between the markers, which presumably involves more than purely temporal processing mechanisms.

ACKNOWLEDGMENTS

Dr. Brian Moore and Dr. Craig Formby are thanked for their comments on the manuscript. We thank our subjects for their patient and enthusiastic participation in long hours of tedious experiments, and Alena Wilson for editing the manuscript. Research was supported by the National Institute on Deafness and other Communicative Disorders.

Buus, S., and Florentine, M. (**1985**). ''Gap detection in normal and impaired listeners: The effect of level and frequency,'' in *Time Resolution in Auditory Systems*, edited by A. Michelson (Springer-Verlag, Berlin).

Divenyi, P. L., and Danner, W. F. (**1977**). ''Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectral,'' Percept. Psychophys. **21**(2), 125–142.

Divenyi, P. L., and Sachs, R. M. (**1978**). ''Discrimination of time intervals bounded by tone bursts,'' Percept. Psychophys. **24**(5), 429–436.
Fitzgibbons, P. J. (**1983**). ''Temporal gap detection in noise as a function of frequency, bandwidth, and level,'' J. Acoust. Soc. Am. **74**, 67–72.
Formby, C., and Forrest, T. G. (**1991**). ''Detection of silent temporal gaps in sinusoidal markers,'' J. Acoust. Soc. Am. **89**, 830–837.
Formby, C., Morgan, L. N., Forrest, T. G., and Raney, J. J. (**1992**). ''The role of frequency selectivity in measures of auditory and vibrotactile temporal resolution,'' J. Acoust. Soc. Am. **91**, 293–305.
Formby, C., Sherlock, L. P., and Forrest, T. G. (**1996**). ''An asymmetric roex filter model for describing detection of silent temporal gaps in sinusoidal markers,'' Aud. Neurosci. **3**, 1–20.
Forrest, T. G., and Formby, C. (**1996**). ''Detection of silent temporal gaps in sinusoidal markers simulated with a single-channel envelope detector model,'' Aud. Neurosci. **3**, 21–33.
Hanekom, J., and Shannon, R. V. (**1997**). ''Gap detection as a measure of electrode interaction in cochlear implants,'' J. Acoust. Soc. Am. (submitted).
Heinz, M. G., Goldstein, Jr., M. H., and Formby, C. (**1996**). ''Temporal gap detection thresholds in sinusoidal markers simulated with a multi-channel, multi-resolution model of the auditory periphery,'' Aud. Neurosci. **3**, 35–56.
Jesteadt, W. (**1980**). ''An adaptive procedure for subjective judgments,'' Percept. Psychophys. **28**, 85–88.
Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.
Moore, B. C. J., Glasberg, E. D., McPherson, T., and Plack, C. J. (**1989**). ''Detection of temporal gaps in sinusoids by normally hearing and hearing-impaired subjects,'' J. Acoust. Soc. Am. **85**, 1266–1275.
Moore, B. C. J., and Glasberg, B. R. (**1988**). ''Gap detection with sinusoids and noise in normal, impaired and electronically stimulated ears,'' J. Acoust. Soc. Am. **83**, 1093–1101.
Phillips, D. P., Taylor, T. L., Hall, S. E., Carr, M. M., and Mossop, J. E. (**1997**). ''Detection of silent intervals between noises activating different perceptual channels: Some properties of central auditory gap detection,'' J. Acoust. Soc. Am. **101**, 3694–3705.
Plomp, R. (**1964**). ''Rate of decay of auditory sensation,'' J. Acoust. Soc. Am. **36**, 277–282.
Shannon, R. V. (**1986**). ''Temporal processing in cochlear implants,'' in *The Scott Reger Memorial Conference* (Iowa U.P., Iowa City, IA).
Shannon, R. V. (**1989**). ''Detection of gaps in sinusoids and pulse trains by patients with cochlear implants,'' J. Acoust. Soc. Am. **85**, 2587–2592.
Shannon, R. V., Adams, D. D., Ferrel, R. L., Palumbo, R. L., and Grandgenett, M. (**1990**). ''A computer interface for psychophysical and speech research with the Nucleus cochlear implants,'' J. Acoust. Soc. Am. **87**, 905–907.
Zeng, F.-G., and Turner, C. W. (**1991**). ''Binaural loudness matches in unilaterally impaired listeners,'' Q. J. Exp. Psychol. (A) **43**(3), 565–583.

Analysis of the performance of a model-based optimal auditory signal processor

Lisa C. Gresham[a] and Leslie M. Collins[b]
Department of Electrical and Computer Engineering, Box 90291, Duke University, Durham, North Carolina 27708-0291

(Received 19 August 1997; accepted for publication 9 February 1998)

Traditionally, psychophysical data have been predicted either by constructing models of the peripheral auditory system or by applying signal detection theory (SDT). Frequently, the theoretical detection performance predicted by SDT is greater than that observed experimentally and a nonphysiologically based ''internal noise'' source is often added to the system to compensate for the discrepancy. A more appropriate explanation may be that traditional SDT approaches either incorporate little or no physiology or make simplifying assumptions regarding the density functions describing the physiological data. In the work presented here, an integrated approach, which combines SDT and a physiologically based model of the human auditory system, is proposed as an alternate method of quantifying detection performance. To validate this approach, the predicted detection performance for a simultaneous masking task is compared to predictions obtained from traditional methods and to experimental data. Additionally, the sensitivity of the integrated method is thoroughly investigated. The results suggest that by combining SDT with a physiologically based auditory model, thereby capitalizing on the strengths of each individual method, the previously observed discrepancies can be partially explained as the result of physical processes inherent in the auditory system rather than unspecified ''internal noise'' and more accurate predictions of psychophysical behavior can be obtained.
[S0001-4966(98)05805-6]

PACS numbers: 43.66.Ba, 43.66.Gf [JWH]

INTRODUCTION

Data from psychoacoustic experiments have been used to relate physiological mechanisms of the human auditory system to the detection performance of a human listener. In order to gain a greater understanding of the relationship between perception and physiology, substantial research has been devoted to developing methods of predicting human detection performance as well as predicting the neural response of the auditory system to various stimuli. One approach has been to construct models which mimic the physiological mechanisms thought to be operating in the peripheral auditory system (e.g., Carney, 1993; Patterson *et al.*, 1995). Typically, these models simulate the functioning of the cochlea in two stages, one which performs a spectral analysis and one which simulates the neural activity pattern evoked by the stimulus (Yost *et al.*, 1996). A third stage, often added to model the processing performed in the central auditory system, performs temporal integration (Moore *et al.*, 1996). Since these models have been developed to predict peripheral auditory responses (often in terms of neural firing patterns), they have not been used to directly calculate theoretical limits on the discrimination performance achievable by the auditory system. In addition, since the models are in their infancy, their ability to predict the entire range of experimental results has not been fully validated.

Another approach has been to construct models based on the signal processing thought to occur in the peripheral auditory system, for example, an energy detector, an envelope detector, or a leaky integrator (Sherwin *et al.*, 1956; Pfafflin and Mathews, 1962; Jeffress, 1964, 1967, 1968; Green and Swets, 1974). More complex, multiple detector models have also been investigated (e.g., Gilkey and Robinson, 1986). Although some of these models provide a good fit to the general shape of some experimental data, their predictions usually exceed the performance of human subjects and the differences are often attributed to unspecified ''internal'' noise sources in the auditory system or uncertainty in the signal parameters.

Signal detection theory has also been used to generate estimates of the performance bounds by deriving the optimal receiver for a particular task and simulating its performance (Green and Swets, 1974), or by calculating a discriminability statistic, d'. When modeling some experimental tasks, such as detecting a tone in noise, the performance of these receivers far exceeds that of human subjects. Typically, this difference has been compensated for by incorporating an ''internal noise'' process into the receiver system whose variance is adjusted to match experimental data (Pfafflin and Mathews, 1962; Durlach *et al.*, 1986; Dau *et al.*, 1996). While some level of ''internal noise'' has been observed physiologically (Nuttall *et al.*, 1997), the magnitude of the discrepancy is too great to be attributed to ''internal noise'' alone (Green, 1970). Other studies (e.g., Teich and Khanna, 1985; Young and Barta, 1986; Relkin and Pelli, 1987; Viemeister, 1988) have used signal detection precepts to fit ROC curves to neural data. This approach eliminates the need to add ''in-

[a]Electronic mail: lcg@ee.duke.edu
[b]Electronic mail: lcollins@ee.duke.edu

ternal noise'' since it is inherent in the data. The ROCs can then be interpreted in terms of the parameters of an unequal variance Gaussian model. However, comparisons between physiological and psychophysical ROC curves, the former predicting better performance, reveal differences consistent with the idea that the auditory system does not function ideally (Relkin and Pelli, 1987). The second approach, in which a d' statistic is calculated directly from the response data, is based on the assumption that these data follow Gaussian distributions, which may be unrealistic in some cases. These approaches have not consistently modeled experimental data in an accurate manner nor have they provided a physiologically based explanation for the difference between theoretical and experimental results.

Each of the traditional approaches described above has been primarily concerned with either correctly modeling the underlying distributions of the signals or with incorporating the known physiological limitations of the auditory system. The work presented in this paper considers both physical limitations and signal characteristics by integrating a physiologically based auditory model and signal detection precepts. Doing so enables the difference between theoretical performance predictions and the experimental performance of humans to be linked to a physical process in the auditory system rather than attributed to unexplained ''internal noise.'' In order to evaluate the validity of the various approaches, a simple task was chosen for which experimental data and theoretical predictions have historically been mismatched, specifically, the task of detecting a 500-Hz tone in broadband, Gaussian white noise, or simultaneous masking.

In the subsequent sections, an existing model of the peripheral auditory system and the traditional methods of evaluating detection performance, described briefly above, are examined in more detail. It is then shown how the integrated method can be derived from a combination of the peripheral auditory model and signal detection theory. Using this integrated method, the performance of the ideal observer at the output of each stage of Patterson's model of the human auditory system (1995) is evaluated. In addition, the sensitivity of this method to mismatched assumptions is thoroughly investigated. Finally, the detection performance predicted by the integrated method, the energy-detector model, and classical signal detection theory (both the signal-known-exactly and signal-known-statistically cases) are compared to experimental data obtained by Watson and his colleagues (1964). The results suggest that a method combining signal detection theory and a physiologically based auditory model enables previously observed discrepancies between theoretical and experimental results to be partially explained as the result of physical processes inherent in the auditory system, and that more accurate predictions of psychophysical behavior can be obtained.

I. COMPUTATIONAL MODELS OF AUDITORY PROCESSING

Signal processing occurring in the peripheral human auditory system is commonly simulated in two stages. The functional path of the Auditory Image Model (AIM), developed by Patterson and his colleagues (Patterson and Holdsworth, 1994, and references therein) takes this approach and includes a third stage which attempts to link physiology and perception. The first stage, which models the acoustic to mechanical transduction of the basilar membrane, performs a spectral analysis. Experiments have shown that the basilar membrane behaves like a bank of overlapping bandpass filters with sharp high-frequency roll off and shallow low-frequency roll off (von Békésy, 1960). The basilar membrane motion (BMM) module of AIM models this stage of auditory processing as a bank of linear gammatone filters whose impulse response is a good fit to the measured impulse response of cats using the revcor technique (de Boer, 1978). The results from psychological and physiological experiments studying frequency selectivity indicate that the bandwidth of the filters varies with location on the basilar membrane. The gammatone filters have the characteristic of being narrow band at low frequencies with a bandwidth which increases proportionally with the center frequency. Despite some drawbacks (such as the linearity assumption), the gammatone filter was utilized in AIM since it appears to be a good first order approximation to the data, is easy to implement, and is computationally efficient (Patterson and Holdsworth, 1994).

The second stage of AIM encompasses the neural encoding phase of the auditory system in which the filtered output of the basilar membrane is converted into a neural activity pattern (NAP). Data have shown that the signal processing performed by the inner hair cells includes compression, rectification, and low-pass filtering (Patterson *et al.*, 1992). AIM models this stage by rectifying and compressing the output of the first stage (BMM) and applying a two-dimensional time and frequency adaptation, allowing small features to be suppressed in the region of a large feature (Patterson and Holdsworth, 1994).

The final stage of AIM does not directly model any signal processing which takes place in the peripheral auditory system. Rather, it is a purely functional module which attempts to represent an ''auditory image'' of the perceived sound (Patterson and Holdsworth, 1994). Proposed models of the central auditory system based on experimental data suggest that the ear performs temporal integration. One difficulty is that fine-grained temporal information must be retained despite this integration. The mechanism employed by AIM to model this is strobed temporal integration (STI). STI operates on the signal by maintaining an adaptive threshold for each channel. When the output of the NAP exceeds this threshold, the NAP signal is added point for point to the existing auditory image. The threshold then decays exponentially until the next peak is encountered. Thus for a periodic signal, integration will be triggered once in each period, emphasizing periodic information and causing cancelation of aperiodic information introduced by noise. One important feature of STI is that the trigger mechanism is independent across channels, resulting in phase alignment, or equivalently, the loss of phase information (Patterson *et al.*, 1995).

AIM calculates the responses for multiple channels, each of which corresponds to a single filter in the gammatone filterbank. It is reasonable to assume that, for most tasks, all channels will contribute to signal detection, requir-

ing information to be combined in some manner across channels. However, in the past, it has been commonly accepted that, to a first-order approximation, only the information in the channel with the best signal-to-noise ratio is used in detecting a tone in broadband noise (e.g., Dau *et al.*, 1996). Therefore, the following simulations and analyses of a simultaneous masking task utilize only the output of the filter with a center frequency of 500 Hz, the frequency of the tone.

II. TRADITIONAL SIGNAL PROCESSING MODELS OF AUDITORY PROCESSING

Historically, several different methods incorporating signal detection theory have been used to evaluate detection performance for various psychophysical tasks. One common link between these methods is the inability of each to satisfactorily explain the discrepancies between human performance and theoretical results. Although evaluating detection performance using these methods can be efficient and convenient, simplifying assumptions may lead to inaccurate conclusions about performance. The novel approach pursued and evaluated in this paper utilizes a rigorous application of signal detection theory and eliminates the need to make these incorrect assumptions.

The energy-detector model, consisting of a bandpass filter centered at the frequency of the tone followed by a square-law device and an integrator, has been used to model data from several experiments with some success. Pfafflin and Mathews (1962) found that this model accurately predicts the shape of the detection curve of a 1000-Hz sinusoid in noise, however, both internal and external ''noise'' sources were required to explain the discrepancies between the predicted and observed results. Sherwin *et al.* (1956) also added ''internal noise'' to reconcile the differences between human detector and electronic (energy) detector performance on a similar task. Finally, Kidd *et al.* (1989) found that the critical band energy-detector model could predict some, but not all, of the experimental results in a study of the detectability of tones in noise.

Another method commonly used to quantify detection performance in psychophysical experiments involves calculating a discriminability statistic, d', from the response data (Tanner and Birdsall, 1958; Yost, 1996). For the special one-dimensional case where the signal parameters are known exactly and the noise is additive, zero mean, and follows a Gaussian distribution, the theoretical d' can be calculated as

$$d' = \frac{\mu_1 - \mu_0}{\sigma},$$

where μ_1 and μ_0 are the means of the density functions of the signal plus noise and noise alone cases, respectively, and the standard deviation, σ, of both density functions is equal. In the multi-dimensional case, when the received signal is a sampled function of time, and the signal, **s**, is known exactly, then

$$(d')^2 = (\mathbf{\Delta m})^{\mathbf{T}}\mathbf{Q}(\mathbf{\Delta m}), \tag{1}$$

where $\mathbf{\Delta m} = \mathbf{s}$ and $\mathbf{Q}$ is a diagonal matrix for which $Q_{i,i} = [\sigma^2(t_i)]^{-1}$. If the variance is not a function of time, Eq. (1) reduces to

$$(d')^2 = \frac{\mathbf{s}^T\mathbf{s}}{\sigma^2},$$

which is the signal-to-noise ratio. When the statistics of the process are known to be deterministic, but are not known in advance, the statistics of the two density functions can be estimated across realizations and Eq. (1) can be used to calculate d' where $\mathbf{\Delta m} = \hat{\mathbf{s}}$ and $Q_{i,i} = [\hat{\sigma}^2(t_i)]^{-1}$ where $\hat{\mathbf{s}}$ is the estimated time-varying signal and $\hat{\sigma}^2$ is the estimated variance (Van Trees, 1968). In this case, Eq. (1) becomes

$$(d')^2 = \frac{\hat{\mathbf{s}}^T\hat{\mathbf{s}}}{\hat{\sigma}^2}. \tag{2}$$

A more appropriate method which has been used to evaluate the detection performance for a known signal, **s**, embedded in noise using signal detection theory is to generate distributions of the received signal given the following two hypotheses:

H_0: The received signal, **x**, is noise alone.

H_1: The received signal, **x**, is the known signal in noise.

From the density functions of the received data under each of the two hypotheses, $p(\mathbf{x}|H_0)$ and $p(\mathbf{x}|H_1)$, the likelihood ratio, defined as

$$\lambda = \frac{p(\mathbf{x}|H_1)}{p(\mathbf{x}|H_0)}, \tag{3}$$

can be formed. This ratio determines the form of the optimal processor, a monotonic function of the likelihood ratio, which can then be applied to the data in order to derive the Receiver Operating Characteristic (ROC) for a given signal-to-noise ratio (Tanner and Birdsall, 1958). Traditionally, the optimal detector has been derived assuming that signal parameters are known exactly and the distributions under H_0 and H_1 are both uncorrelated, equal-variance Gaussians, the classical Signal-Known-Exactly (SKE) detection problem (Tanner, 1960). Jeffress (1967) suggests that, although this processor does not accurately predict experimental data, it has been widely adopted since it is mathematically tractable. The resulting optimal detector, which will be referred to as the Gaussian–Gaussian processor due to the assumptions made regarding the two density functions, has the form

$$\mathbf{x}^T\mathbf{s} \underset{H_0}{\overset{H_1}{\gtrless}} \beta, \tag{4}$$

where **x** is the received signal, **s** is the known signal, and β is the threshold to which it is being compared (Van Trees, 1968). The performance of this processor is also quantifiable in terms of the d' described in Eq. (2) because the underlying assumptions are met.

Several studies (e.g., Marill, 1956; Egan *et al.*, 1959) have shown that experimental data are better fit when the signal is assumed to be known statistically (SKS) rather than exactly (SKE). In the context of AIM, this uncertainty is justified at the level of the SAI since, due to the strobed temporal integration, information about the phase of the signal is lost and, due to observer uncertainty, the exact frequency of the tone may not be known exactly. This uncertainty can be incorporated into the optimal ''signal-known-

except-for-frequency-and-phase'' (SKEFP) processor by assuming that the *a priori* distribution of the phase is uniform and that the *a priori* distribution of the frequency can be derived from the characteristics of the gammatone filter centered at 500 Hz. This processor, derived using classical signal detection theory methods, can then be applied to the signals outside the ear.

The approach taken in this paper is similar to the standard signal detection theory approach as far as the methodology. However, there are two key differences. First, in contrast to the previously described methods which analyze the signal at the input of the auditory system, this approach allows the analysis of data received after each stage of the model of auditory processing. Second, instead of assuming that the density functions of the data under each hypothesis are uncorrelated Gaussians, the true density functions are simulated at the output of each stage of AIM. This is achieved by generating 9000 realizations of the noise and signal plus noise, propagating them through one, two, or three stages of AIM, and forming histograms of the output signals at each point in time to approximate $p(\mathbf{x}|H_0)$ and $p(\mathbf{x}|H_1)$ and the corresponding statistics (e.g., the mean and variance in the case of Gaussian distributions). Using these density functions, the likelihood ratio is formed [Eq. (3)] and used to determine the corresponding processor. To generate ROCs, 9000 realizations of the signals under each hypothesis are generated and propagated through the model to the desired stage. The output signals are then used as the data upon which the processor operates to generate probability of false alarm and probability of detection values for each threshold. These values are then plotted on normal–normal axes to obtain the corresponding ROC curve. Since this method integrates signal detection theory and a model of the peripheral auditory system, it will be referred to as the integrated approach in the subsequent discussion. The ''optimal'' detector often requires substantial computation time, rendering it impractical for a real-time application. Therefore, the effect of making simplifying assumptions regarding the density functions under each hypothesis has also been investigated.

Each method described above is used to calculate ROC curves which can be used for two purposes. First, these curves allow the comparison of the detection performance of various proposed processors and methods. In addition, the sensitivity of the integrated method can be investigated by computing the ROCs for processors derived under various simplifying assumptions after each stage of the model. In this way, it can be determined which *a priori* assumptions about the received signal have a significant effect on detection performance and processor design. Second, the ROC curves can be used to validate or refute the model. If the results predicted by the model coincide with the experimental data, it can be concluded that the model may accurately represent the signal processing performed by the human auditory system for the given task. It should be noted, however, that this does not mean that the model necessarily specifies the mechanisms which actually exist.

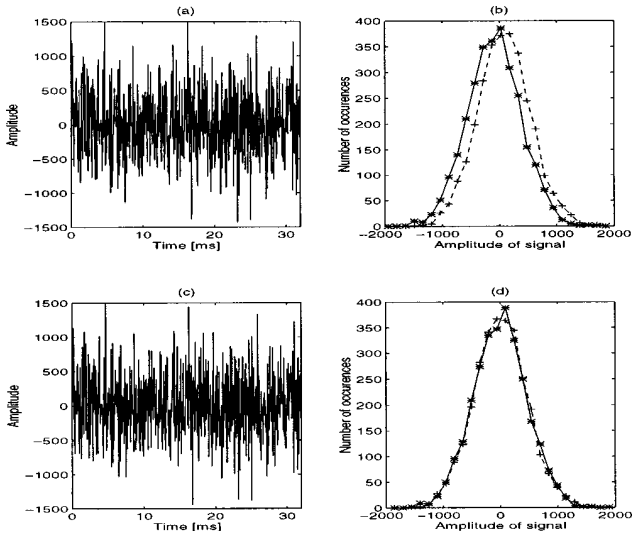


FIG. 1. Outside the ear: (a) Typical realization of 500-Hz tone in broadband, Gaussian noise, $E/N_0=14$; (b) Distribution of responses under H_1 at $t=20.7$ ms (+) and $t=21.7$ ms (∗); (c) Typical realization of broadband, Gaussian noise alone; (d) Distribution of responses under H_0 at $t=20.7$ ms (+) and $t=21.7$ ms (∗).

III. ANALYSIS OF THE INTEGRATED METHOD

A. Detection performance outside the ear

The task of detecting a 500-Hz tone in broadband, Gaussian white noise, known as simultaneous masking, was chosen for this analysis since theoretical and experimental results for this task have historically been mismatched, thereby making it a good test case for the integrated method. Since the integrated method is based on signal detection theory, the derived processor is equivalent to that of the SKE case outside the ear. However, the procedure for determining the optimal processor for the integrated method and simulating its performance was followed and applied to the signals outside the ear to illustrate the approach under a scenario where the correct result was known *a priori*.

In the first step, the density functions of the received data under H_0 and H_1 are determined. Outside the ear, these density functions are specified by the problem parameters for the simultaneous masking task. Namely, under H_0, $x(t_k)$ is a random variable described by a $\mathbf{N}(0,\sigma^2)$ distribution, under H_1, $x(t_k)$ is a random variable described by a $\mathbf{N}(A\sin 2\pi f t_k,\sigma^2)$ distribution, and the signal-to-noise ratio is defined as $KA^2/2\sigma^2$, where K is the total number of time samples. To determine the density functions of the simulated data, many realizations of the noise and signal plus noise were generated and a histogram was formed for each time slice. Figure 1(b) and (d) shows the histograms under H_1 and H_0, respectively, for $t=20.7$ ms, corresponding to the occurrence of a maximum in the input sinusoid, and $t=21.7$ ms, corresponding to the occurrence of a minimum in the input sinusoid. The histograms, and thus the density functions, under H_0 at each point in time are zero mean, uncorrelated, and Gaussian. Likewise, under H_1, at each point in time, the density functions are also uncorrelated and Gaussian, however, the mean of each density function varies sinusoidally in phase with the input tone. These density func-

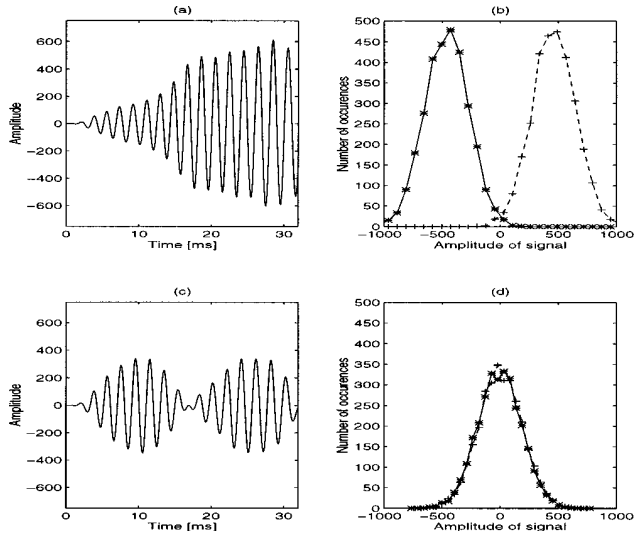


FIG. 2. Post-BMM (a) Typical realization of the received signal under H_1; (b) Distribution of responses under H_1 at $t=20.7$ ms (+) and $t=21.7$ ms (∗); (c) Typical realization of the received signal under H_0; (d) Distribution of responses under H_0 at $t=20.7$ ms (+) and $t=21.7$ ms (∗).

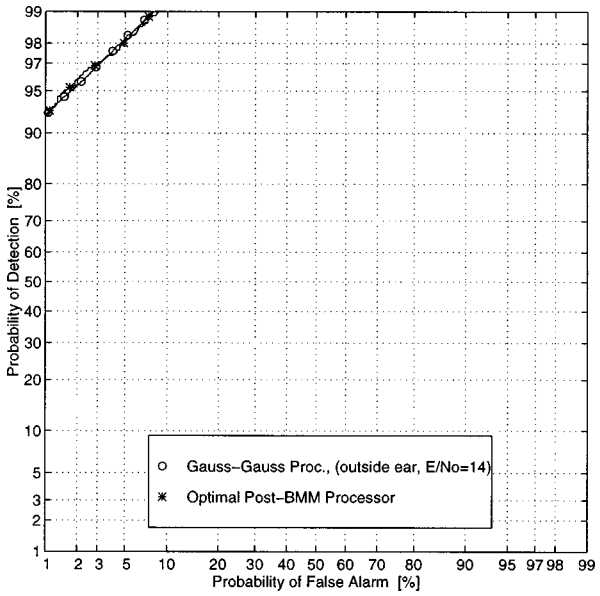


FIG. 3. ROC for the optimal post-BMM processor (∗). The Gaussian–Gaussian processor outside the ear (○) is included as a reference.

tions result in the classical signal-known-exactly (SKE) detection problem and the optimal processor, the Gaussian–Gaussian processor, is given by Eq. (4). Since the assumptions made by classical signal detection theory are met, the predictions made by the two methods are equivalent at this point. The ROC curve for the Gaussian–Gaussian processor corresponding to a value of $E/N_0=14$ outside the ear is included on all the ROC plots in this paper as a reference point to which detection performance of other processors at each stage of AIM can be compared. Since the most complete information about the signal is available outside the ear, and information can only be lost as the signal propagates through the ear, detection performance at successive stages of the model can never be greater than this reference curve for a given signal-to-noise ratio (Taylor and Forbes, 1969).

B. Detection performance post-basilar membrane

To determine the detection performance after the first stage of AIM using the integrated approach, the density functions under H_0 and H_1 were obtained by applying linear systems analysis. When the input to a linear filter is a sinusoid at a particular frequency, the output is a sinusoid at that same frequency whose amplitude and phase may be different. When the input is zero mean, uncorrelated, white Gaussian noise, the output is zero mean, correlated Gaussian noise with an autocorrelation function, $\mathbf{R}_y(\tau)$, equal to $\mathcal{F}^{-1}\{|\mathbf{H}(\mathbf{f})|^2\mathbf{S}_x(\mathbf{f})\}$, where $\mathbf{H}(\mathbf{f})$ is the Fourier transform of $h(t)$, the impulse response of the filter, and $\mathbf{S}_x(\mathbf{f})$ is the power spectral density function of the input signal (Oppenheim and Schafer, 1989). These analytic results were verified by generating multiple realizations of the uncorrelated Gaussian noise and the tone plus noise outside the ear and simulating the output of the first stage of AIM using an analytic form of the gammatone filter. Figure 2(a) and (c) illustrates typical realizations of the received signal under each hypothesis. Histograms were formed from the simulated responses in order to determine the density functions at each time slice. Figure 2(b) and (d) depicts the histograms under H_1 and H_0, respectively, for points in time corresponding to a maximum and a minimum in the sinusoid outside the ear. These two time slices were selected to illustrate the possible variations in the histograms since each is associated with an extreme point in the input. At each point in time, the density functions under H_0 are zero mean and Gaussian, as at the input of the model. However, unlike the input signal, there is a covariance structure inherent in the simulated response. Likewise, under H_1, the density functions are correlated and Gaussian with a time-varying mean.

Given these density functions, the likelihood ratio test can be formed. Simplification results in the optimal post-BMM processor, which incorporates the existing covariance structure,

$$\mathbf{x}^T\mathbf{C}_n^{-1}\mathbf{s} \underset{H_0}{\overset{H_1}{\gtrless}} \beta, \tag{5}$$

where $\mathbf{s}$ is the tone, $\mathbf{x}$ is the received signal, and $\mathbf{C}_n$, the covariance of the noise, is related to the autocorrelation of the output as $\mathbf{C}_n(i,j)=\mathbf{R}_y(i-j)$. To investigate the sensitivity of this method to simplifying assumptions, the Gaussian–Gaussian processor, given in Eq. (4), which ignores the covariance structure inherent in the data, was also applied at the output of the BMM and its performance was determined. Because the covariance matrix closely resembled an identity matrix, the performance for the two processors was essentially the same. A comparison between the ROCs for the optimal post-BMM processor and the Gaussian–Gaussian processor applied outside the ear reveals that the performance of these two processors is essentially the same as well (see Fig. 3). These results suggest that no information has been lost in the first stage of processing which is consistent with the theory of linear filters.

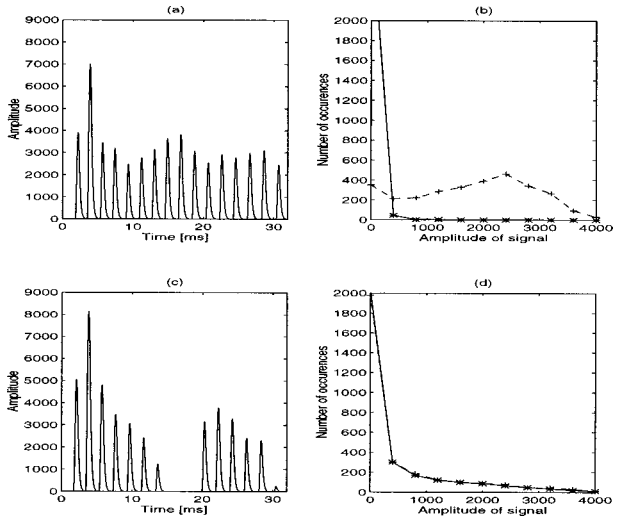


FIG. 4. Post-NAP: (a) Typical realization of the received signal under H_1; (b) Distribution of responses under H_1 at $t=20.7$ ms (+) and $t=21.7$ ms (∗); (c) Typical realization of the received signal under H_0; (d) Distribution of responses under H_0 at $t=20.7$ ms (+) and $t=21.7$ ms (∗).

C. Detection performance post-neural activity pattern

The linearity of the signal processing performed by the first stage of AIM results in a simple relationship between the input and the output signals. The second stage of AIM, however, is nonlinear. Therefore, it is instructive to first consider what form the responses are expected to take given the signal processing that is being modeled by AIM. First, consider the response to the tone alone. Since both the tone and AIM are deterministic, for each point in time the output of AIM will be identical across realizations. One of the functions of the NAP is to rectify the output of the BMM, thus the output of the NAP will be zero when the output of the BMM is negative. Next, consider the response to the noise alone. The noise at the output of the BMM [see Fig. 2(d)] follows a zero-mean Gaussian distribution. Therefore, the probability that the noise has a negative amplitude is equal to one-half. Thus half of the time, the response of the NAP to an input of noise alone will be zero, resulting in a density function that is heavily weighted toward zero and is zero for all negative values. Finally, consider an input of tone plus noise. Whenever the tone's amplitude is negative, the NAP responds as if to noise with a negative mean and, following the argument presented for noise alone, the density function will be even more heavily weighted toward zero than the density function under H_0. When the amplitude of the tone is positive, it will manifest itself as a positive shift in the mean of the density function. These observations suggest that the density functions are not Gaussian and plots of the actual density functions obtained by simulation (see Fig. 4) confirm this hypothesis.

As expected, when many realizations of the noise and tone plus noise are propagated through AIM, the density function under H_0 is identical at all time slices [see typical histograms in Fig. 4(d)] and can be approximated by an exponential density function with a parameter $\boldsymbol{\lambda}$, which can be derived from the data. The density function under H_1 varies across time, as predicted. For time samples when the BMM response to the tone is negative, such as when $t=21.7$ ms, the response of the NAP mimics its response to noise alone, resulting in similar exponential density functions [Fig. 4(b), [∗]]. When the response of the BMM to the tone is positive, such as when $t=20.7$ ms, the resulting NAP density function resembles a truncated Gaussian, however, there is a slight concentration of responses near zero [Fig. 4(b), [+]].

Under the assumptions that the noise follows an uncorrelated exponential distribution with parameter $\boldsymbol{\lambda}$ and the density function for the tone plus noise can be approximated by an uncorrelated Gaussian with mean $\boldsymbol{\mu}_{sn}$ and variance $\boldsymbol{\sigma}_{sn}^2$, the likelihood ratio can be formed. After simplification, the corresponding processor, to be referred to as the Gaussian–Exponential processor, can be expressed as

$$\boldsymbol{\lambda}^T\mathbf{x}-\frac{\mathbf{x}^T\mathbf{x}}{2\boldsymbol{\sigma}_{sn}^2}+\frac{\mathbf{x}^T\boldsymbol{\mu}_{sn}}{\boldsymbol{\sigma}_{sn}^2}\underset{H_0}{\overset{H_1}{\gtrless}}\beta. \tag{6}$$

The assumptions made when formulating the likelihood ratio that resulted in the Gaussian–Exponential processor yield a computationally intense processor. This gives rise to the issue of how sensitive the predicted performance is to the accuracy of the assumptions made regarding the density functions. An additional processor was derived from a likelihood ratio based on the assumptions that the density functions of the received data under H_0 and H_1 were both uncorrelated, equal-variance, nonzero-mean Gaussians. The processor corresponding to these assumptions, referred to as the Gaussian–Gaussian (nonzero-mean) processor, is

$$\mathbf{x}^T\boldsymbol{\mu}_{sn}-\mathbf{x}^T\boldsymbol{\mu}_n\underset{H_0}{\overset{H_1}{\gtrless}}\beta, \tag{7}$$

where $\boldsymbol{\mu}_{sn}$ and $\boldsymbol{\mu}_n$ are the means under H_1 and H_0, respectively. This processor simplifies the computations since it employs a Gaussian instead of an exponential density function. Finally, even the assumption of nonzero-mean noise can be eliminated, resulting in the standard Gaussian–Gaussian processor, given in Eq. (4).

Figure 5 shows the ROCs associated with each processor and indicates that there is some loss in performance predicted by each processor. In addition, these plots illustrate that the decrease in predicted performance is more substantial when simplifying assumptions are made. Thus one can conclude that incorporating more accurate *a priori* information regarding the output density functions increases the predicted performance. If only the results of the simplest processor, the Gaussian–Gaussian processor, were examined, the discrepancy between the theoretical prediction and the experimental data, whose ROC lies closer the the chance diagonal, would be mistakenly assumed to be smaller than that which actually exists and the observed decrease would be attributed to a loss of information in the auditory system rather than to the true cause, an imprecise application of signal detection theory.

As was alluded to above, the Gaussian–Exponential processor was based on a simplifying assumption regarding the density function of the received data under H_1. One final detector, derived by fitting polynomials to the density functions under H_1 at each time sample and assuming an expo-

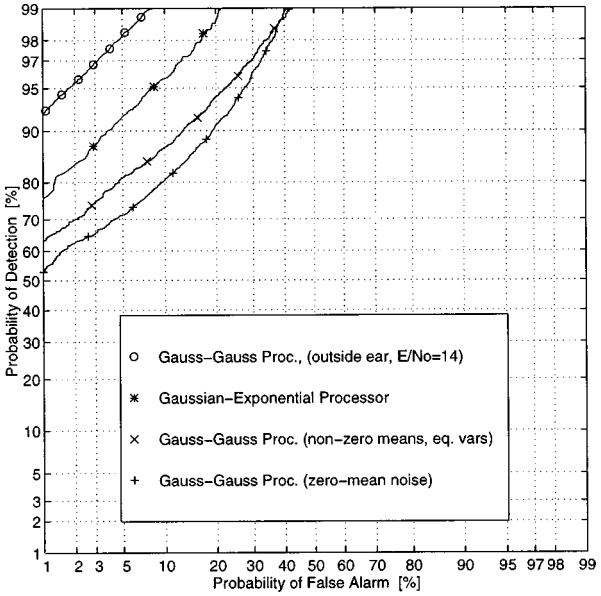


FIG. 5. ROC for the Gaussian-Exponential processor (∗), the nonzero-mean Gaussian–Gaussian processor (×), and the zero-mean noise Gaussian–Gaussian processor (+) applied to the output of the NAP. The Gaussian–Gaussian processor applied outside the ear (○) is included as a reference.

nential distribution for the noise, was evaluated. The performance was not significantly different from the Gaussian–Exponential processor, thus it can be concluded that the additional information contained in a more accurate representation of the density functions does not significantly increase the performance.

Finally, it should be noted that the covariance of the received data was ignored by all of the processors derived above for mathematical tractability. To derive the true ''optimal'' processor, this correlation should be incorporated, however, due to the complexity of the problem and computational constraints, this approach was not pursued. Thus it is possible that the loss in performance predicted by the Gaussian–Exponential processor may be at least partially attributed to this omission.

D. Detection performance post-stabilized auditory image

The purpose of the third stage of AIM is to generate a ''stabilized'' image by performing strobed temporal integration. The density functions under H_0 and H_1 were generated using AIM for a signal with K time samples (Fig. 6). Histograms are shown in Fig. 6(b) and (d), however, at this stage in AIM, they can no longer be considered ''typical'' since the histograms vary widely and unpredictably as a function of time. To obtain the most accurate approximation of $p(x_i|H_1)$ and $p(x_i|H_0)$, polynomials were fit to the density functions at each time slice i. Assuming statistical independence across time, these density functions were then used to calculate a likelihood ratio and corresponding processor of the form,

$$\frac{p(\mathbf{x}|H_1)}{p(\mathbf{x}|H_0)}=\prod_{i=1}^{K}\frac{p(x_i|H_1)}{p(x_i|H_0)}\underset{H_0}{\overset{H_1}{\gtrless}}\beta. \tag{8}$$

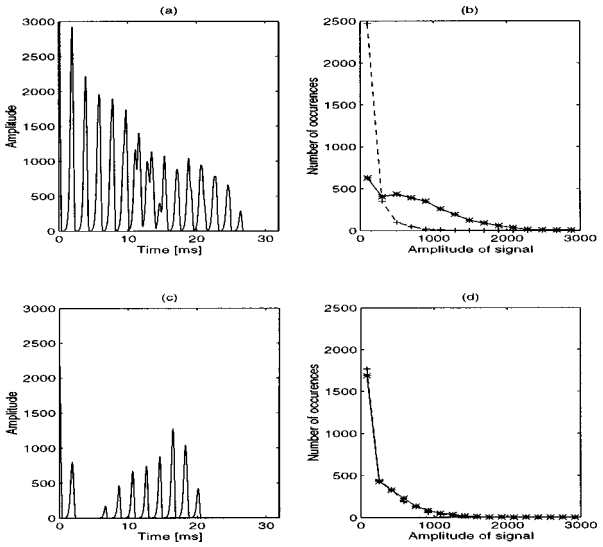


FIG. 6. Post-SAI: (a) Typical realization of the received signal under H_1; (b) Distribution of responses under H_1 at $t=20.7$ ms (+) and $t=21.7$ ms (∗); (c) Typical realization of the received signal under H_0; (d) Distribution of responses under H_0 at $t=20.7$ ms (+) and $t=21.7$ ms (∗).

Again, the sensitivity of this method was investigated by making simplifying assumptions similar to those described for the NAP. The nonzero-mean Gaussian–Gaussian processor, given in Eq. (7), and the classical Gaussian–Gaussian processor, expressed in Eq. (4), were also used to analyze the data. Figure 7 shows the resulting ROCs associated with each processor. Unlike the results after the BMM and the NAP, it appears that making simplifying assumptions does not greatly affect the predicted performance at this stage. However, there is a substantial decrease in performance predicted by all processors. The significance of this result, which is discussed further in the next section, is that white noise was only added to the signal outside the ear before

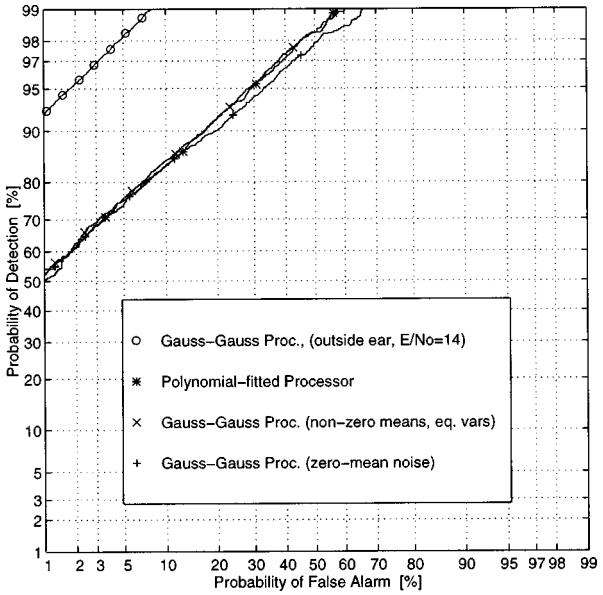


FIG. 7. ROC for the Polynomial-Fitted processor (∗), the nonzero-mean Gaussian–Gaussian processor (×), and the zero-mean noise Gaussian–Gaussian processor (+) applied to the output of the SAI. The Gaussian–Gaussian processor applied outside the ear (○) is included as a reference.

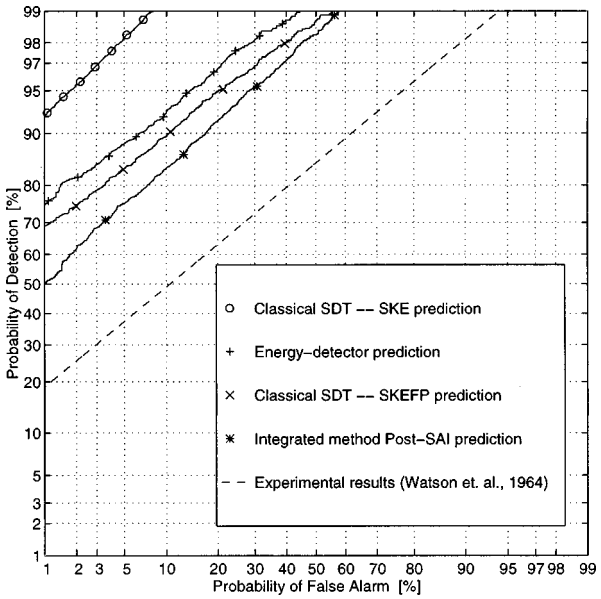


FIG. 8. ROC comparing the performance predicted by various methods, including the classical signal detection theory method (SKE case, ○; SKEFP case, ×), the energy-detector model (+), and the integrated method (*), to experimental results (- -). The experimental ROC was plotted assuming $d_s = 1.13$ and $\sigma_n/\sigma_{sn} = 0.82$, the values derived by Watson *et al.* from their 36-point ROC.

being processed by the auditory model in contrast to some other approaches (e.g., Dau *et al.*, 1996) which add white noise to the data after being processed by a model. As a result, external noise is processed in the exact same manner as the input signal and no assumptions have to be made regarding the internal distribution of the noise.

IV. COMPARISON OF THE PERFORMANCE OF INTEGRATED AND TRADITIONAL METHODS WITH EXPERIMENTAL RESULTS

The results of the integrated approach were compared to the predictions made by several traditional methods of quantifying detection performance and to experimental data obtained by Watson and his colleagues (1964) for a fixed-interval-observation masking experiment (see Fig. 8). In their experiment, the 500-Hz tone, when present, had a duration of 250 ms. The noise was white and the signal intensity was chosen so that $E/N_0 = 14$. Experimental ROCs were determined for three subjects using a 36-point rating method from which the average detectability index, d_s, and the average slope of the line on normal–normal axes were derived. As seen in Fig. 8, the integrated method yields the most accurate prediction of human performance. However, this prediction is still substantially greater than the experimental data. The classical signal detection theory prediction for the ideal observer in the SKE case far exceeds the human performance, yet this result is still useful as it gives an upper bound on performance. The energy-detector model provides an estimate that is closer to experimental data than the SKE case outside the ear, but this prediction exceeds even the integrated method's prediction. The ROC for the classical SDT problem assuming uncertain frequency and phase (SKEFP) yields a more accurate prediction than the SKE processor, however, the integrated method is still a better predictor of human performance. The detection performance predicted by the classical SKS processor (not shown), which only incorporates phase uncertainty, is slightly higher than that of the SKEFP case.

Table I summarizes the ROCs presented throughout this paper in terms of a detectability index, d_s, and a ratio of variances, σ_n/σ_{sn}, measurements suggested by Egan and his colleagues to quantify nonequal-variance ROCs (Egan *et al.*, 1961). Of the traditional methods evaluated outside the ear, the SKEFP processor predicts statistics closest to experimental values. This observation is supported by Marill (1956) and Egan and his colleagues (1959) who found that their experimental data could be best fit assuming the signal was known statistically and that the ratio of variances was less than 1. Regarding the predictions of the integrated method, the general trend is for d_s to become smaller as this method is applied to successive stages of AIM. As would be expected based on the ROCs presented in Fig. 8, the d_s value for the post-SAI integrated method is closer than any of the traditional methods to the experimental value of 1.13.

Examination of the ratio of variances allows additional comparisons to be made. It can be seen clearly that the

TABLE I. Values of the detectability index, d_s, and ratio of variances, σ_n/σ_{sn}, estimated from ROC curves. Experimental values were obtained by averaging the results of three subjects in an experiment conducted by Watson and his colleagues (1964).

Location	Method	Assumptions	d_s	σ_n/σ_{sn}
Outside the ear	Classical SDT	SKE	3.72	1.00
	Classical SDT	SKEFP	2.52	0.81
	Energy-detector		2.80	0.71
Post-BMM	Integrated method	Optimal Proc.	3.71	1.02
Post-NAP	Integrated method	Expon-Gauss. Proc.	3.09	1.02
	Integrated method	Nonzero means, equal variances	2.58	1.01
	Integrated method	Zero-mean noise, equal variances	2.45	1.10
Post-SAI	Integrated method	Poly-fitted Proc.	2.07	0.93
	Integrated method	Nonzero means, equal variances	2.08	0.91
	Integrated method	Zero-mean noise, equal variances	1.99	0.88
Experimental			1.13	0.82

TABLE II. Comparison of d' values calculated using Eq. (2) and actual or experimental values of d'. The post-SAI experimental value was obtained by determining $z(s|H_0)$ when the probability of detection was 0.50.

Location	Actual/experimental d'	Calculated d'
Outside the ear	14.0	14.0
Post-SAI	1.3	371.3

energy-detector method, the SKEFP, and the post-SAI integrated results have $\sigma_n/\sigma_{sn}<1$, similar to experimental data ($\sigma_n/\sigma_{sn}=0.82$). This phenomenon is not observed in the classical SDT results or the predictions made by the integrated method after the BMM or the NAP. This suggests that phase information is important to the ideal observer and leads to the conclusion that uncertainty in this parameter introduced by the auditory system may be one source of information loss partially explaining the deviation of human performance from that of the ideal receiver.

When estimated from the ROCs using Egan's method, d_s and σ_n/σ_{sn} are appropriate psychophysical measures. However, another frequently used method of generating an index of detectability, that of calculating a d' statistic using the raw data (for example, the input and the output data of AIM) and Eq. (2), is not an appropriate statistical measure in this case. Outside the ear, the actual value of d' used in simulations was 14. When Eq. (2) is used to compute d' based on the raw data at this point, the same value of 14 is obtained. Because the distributions of the noise and signal plus noise outside the ear are both uncorrelated Gaussians, the calculation of d' using Eq. (2) is an accurate statistical measure. However, a d' calculated in this manner from the model data (post-SAI) yields a value of 371.3 when the true value is 1.3 (see Table II). The conclusion is that d' is an inappropriate statistical measure, since, as is the case at the output of the model (post-SAI), the assumptions that the data are uncorrelated and Gaussian are not necessarily true.

V. DISCUSSION

The integrated method has been shown to predict psychophysical behavior more accurately than traditional methods and supports previous studies suggesting that phase uncertainty may partially explain why human performance is less than ideal. However, some discrepancy still exists between predictions generated using the integrated method and experimental performance. The success of the integrated method thus far suggests that, rather than default to the ''internal noise'' assumption for an explanation, the assumptions made by the model and the derived processors should be further investigated.

Despite evidence that certain parameters of the auditory system are stochastic (Swets *et al.*, 1961; Kiang, 1984; Teich and Khanna, 1985; Miller *et al.*, 1986), AIM itself is entirely deterministic. Specifically, Teich and Khanna have shown that the interspike time interval for a neural fiber follows a Poisson distribution and Relkin and Pelli (1987) have demonstrated that the pulse number distribution follows a Gaussian distribution. Patterson and Holdsworth (1994) acknowledged that reducing the activity of a fiber to a single number, the average firing rate in the case of AIM, probably results in a somewhat inaccurate representation of the signal at this stage in the auditory system.

In addition to being deterministic, the model does not incorporate certain aspects of the auditory system. For example, only the afferent activity of the auditory nerve has been included in this model. Studies have shown that efferent activity is also a significant factor and has been included in a model developed by Giguère and Woodland (1994). In addition, the filters have been modeled as linear filters, although the auditory system is known to behave in a nonlinear manner producing distortions such as harmonics and difference tones (Smoorenburg, 1972; Hall, 1974; Strube, 1985).

Changes to the model itself may be difficult to implement. However, in some cases, the changes can be effected by modifying the input signal. As discussed previously, a loss of phase information is consistent with experimental data. It has been observed that the incorporation of phase and frequency uncertainty into the SKEFP processor has a significant effect on the predicted performance when used to process the signals outside the ear, making this method better than the classical SKE approach. However, the predictions made by the integrated method are still closer to experimental data. These results suggest that phase and frequency uncertainty should be incorporated into the integrated approach and evaluated at the output of AIM rather than just at the outside of the ear.

VI. SUMMARY AND CONCLUSIONS

Prior to the application of the integrated approach, the difference between the theoretical performance predicted by traditional methods loosely based on signal detection theory and actual experimental performance had often been accounted for by adding internal noise with a level-independent variance after the signal had been processed by the model (e.g., Sherwin *et al.*, 1956; Pfafflin and Mathews, 1962; Dau *et al.*, 1996). While this enables theoretical and experimental results to be matched, there is no complete physiological justification for adding this level of noise. As Green (1970) noted, there are internal noise processes, such as spontaneous neural firings, which contribute to the mismatch, however, such a large discrepancy cannot be attributed to ''internal'' noise alone. Our work has shown that combining signal detection theory with a physiologically based auditory model allows the discrepancies to be partially explained as the result of physical processes inherent in the auditory system, specifically, the loss of phase information. However, there is still a discrepancy between theoretical and experimental results. Rather than add additional internal noise in an *ad hoc* manner, our work suggests that incorporating frequency and phase uncertainty into the integrated approach will likely result in an even closer agreement between theoretical and experimental performance.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Roy Patterson, Dr. Loren Nolte, Dr. Roger Miller, and Dr. David Smith for

useful discussions regarding this research. They would also like to thank the reviewers for their thoughtful comments and suggestions. Support for this work was provided by the National Institutes of Health (1R03-DC-03136-01) and the Lord Foundation.

von Békésy, G. (**1960**). *Experiments in Hearing* (McGraw-Hill, New York).

de Boer, E., and de Jongh, H. R. (**1978**). ''On cochlear encoding: potentialities and limitations of the reverse-correlation technique,'' J. Acoust. Soc. Am. **63**, 115–135.

Carney, L. H. (**1993**). ''A model for the responses of low-frequency auditory-nerve fibers in cats,'' J. Acoust. Soc. Am. **93**, 401–417.

Dau, T., Püschel, D., and Kohlrausch, A. (**1996**). ''A quantitative model of the ''effective'' signal processing in the auditory system. I. Model structure,'' J. Acoust. Soc. Am. **99**, 3615–3622.

Durlach, N. I., Braida, L. D., and Ito, Y. (**1986**). ''Towards a model for discrimination of broadband signals,'' J. Acoust. Soc. Am. **80**, 63–72.

Egan, J. P., Schulman, A. E., and Greenberg, G. Z. (**1959**). ''Operating characteristics determined by binary decisions and by rating,'' J. Acoust. Soc. Am. **31**, 768–773.

Egan, J. P., Greenberg, G. Z., and Schulman, A. I. (**1961**). ''Interval of time uncertainty in auditory detection,'' J. Acoust. Soc. Am. **33**, 771–778.

Giguère, C., and Woodland, P. C. (**1994**). ''A computational model of the auditory periphery for speech and hearing research. II. Descending paths,'' J. Acoust. Soc. Am. **95**, 343–349.

Gilkey, R. H., and Robinson, D. E. (**1986**). ''Models of auditory masking: A molecular psychophysical approach,'' J. Acoust. Soc. Am. **79**, 1499–1510.

Green, D. M. (**1970**). ''Application of detection theory in psychophysics,'' IEEE Proc. **58**, 713–723.

Green, D. M., and Swets, J. A. (**1974**). *Signal Detection Theory and Psychophysics* (Kreiger, New York).

Hall, J. L. (**1974**). ''Two-tone distortion products in a nonlinear model of the basilar membrane,'' J. Acoust. Soc. Am. **56**, 1818–1828.

Jeffress, L. A. (**1964**). ''Stimulus-oriented approach to detection,'' J. Acoust. Soc. Am. **36**, 766–774.

Jeffress, L. A. (**1967**). ''Stimulus-oriented approach to detection reexamined,'' J. Acoust. Soc. Am. **41**, 480–488.

Jeffress, L. A. (**1968**). ''Mathematical and electrical models of auditory detection,'' J. Acoust. Soc. Am. **44**, 187–203.

Kiang, N. Y.-S. (**1984**). ''Peripheral neural processing of auditory information,'' in *Handbook of Physiology: The Nervous System, Vol. 3, Sensory Processes*, edited by I. Darian Smith (Am. Physiol. Soc., Bethesda, MD), pp. 639–674.

Kidd, Jr., G., Mason, C. R., Brantley, M. A., Owen, G. A. (**1989**). ''Roving-level tone-in-noise detection,'' J. Acoust. Soc. Am. **86**, 1310–1317.

Marill, T. (**1956**). ''Detection theory and psychophysics,'' MIT Res. Lab. Electron. Rept. No. 319 (30 Oct.).

Miller, M. I., Barta, P. E., and Sachs, M. B. (**1986**). ''Strategies for the representation of a tone in background noise in the temporal aspects of the discharge patterns of auditory-nerve fibers,'' J. Acoust. Soc. Am. **81**, 665–679.

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.

Nuttall, A. L., Guo, M. H., and Ren, T. (**1997**). ''Basilar membrane noise in the absence of acoustic stimulation,'' Assoc. Res. Otolaryng. Abstracts, No. 41.

Oppenheim, A. V., and Schafer, R. W. (**1989**). *Discrete-Time Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ), pp. 63–67.

Patterson, R. D., Allerhand, M. H., and Giguère, C. (**1995**). ''Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform,'' J. Acoust. Soc. Am. **98**, 1890–1894.

Patterson, R. D., and Holdsworth, J. (**1994**). ''A functional model of neural activity patterns and auditory images,'' in *Advances in Speech, Hearing and Language Processing*, edited by W. A. Ainsworth (JAI, London), Vol. 3 (in press since 1991).

Patterson, R. D., Robinson, K., Holdsworth, J., McKeown, D., Zhang, C., and Allerhand, M. (**1992**). ''Complex sounds and auditory images,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford), pp. 429–446.

Pfafflin, S. M., and Mathews, M. V. (**1962**). ''Energy-detection model for monaural auditory detection,'' J. Acoust. Soc. Am. **34**, 1842–1853.

Relkin, E. M., and Pelli, D. G. (**1987**). ''Probe tone thresholds in the auditory nerve measured by two-interval forced-choice procedures,'' J. Acoust. Soc. Am. **82**, 1679–1691.

Sherwin, C. W., Kodman, Jr., F., Kovaly, J. J., Prothe, W. C., and Melrose, J. (**1956**). ''Detection of signals in noise: A comparison between the human detector and an electronic detector,'' J. Acoust. Soc. Am. **28**, 617–622.

Smoorenburg, G. F. (**1972**). ''Combination tones and their origin,'' J. Acoust. Soc. Am. **52**, 615–632.

Strube, H. W. (**1985**). ''The shape of the nonlinearity generating the combination tone $2f_1-f_2$,'' J. Acoust. Soc. Am. **79**, 1511–1518.

Swets, J. A., Green, D. M., and Tanner, Jr., W. P. (**1962**). ''On the width of critical bands,'' J. Acoust. Soc. Am. **34**, 108–113.

Tanner, Jr., W. P. (**1960**). ''Theory of signal detectability as an interpretive tool for psychophysical data,'' J. Acoust. Soc. Am. **32**, 1140–1147.

Tanner, Jr., W. P., and Birdsall, T. G. (**1958**). ''Definitions of d' and η as psychophysical measures,'' J. Acoust. Soc. Am. **30**, 922–928.

Taylor, M. M., and Forbes, S. M. (**1969**). ''Monaural detection with contralateral cue (MDCC). I. Better than energy detector performance by human observers,'' J. Acoust. Soc. Am. **46**, 1519–1526.

Teich, M. C., and Khanna, S. M. (**1985**). ''Pulse-number distribution for the neural spike train in the cat's auditory nerve,'' J. Acoust. Soc. Am. **77**, 1110–1128.

Van Trees, H. (**1968**). *Detection, Estimation, and Modulation Theory* (Wiley, New York), pp. 86–133.

Viemeister, N. F. (**1988**). ''Intensity coding and the dynamic range problem,'' Hearing Res. **34**, 267–274.

Watson, C. S., Rilling, M. E., and Bourbon, W. T. (**1964**). ''Receiver-operating characteristics determined by a mechanical analog to the rating scale,'' J. Acoust. Soc. Am. **36**, 283–288.

Yost, W. A., Patterson, R. D., and Sheft, S. (**1996**). ''A time domain description for the pitch strength of iterated rippled noise,'' J. Acoust. Soc. Am. **99**, 1066–1078.

Young, E. D., and Barta, P. E. (**1986**). ''Rate response of auditory nerve fibers to tones in noise near masked threshold,'' J. Acoust. Soc. Am. **79**, 426–442.

Beneficial effects of notched noise on intensity discrimination in the region of the "severe departure"

Christopher J. Plack
Laboratory of Experimental Psychology, University of Sussex, Brighton BN1 9QG, England

(Received 28 October 1997; accepted for publication 22 January 1998)

Intensity discrimination for a 6-kHz sinusoidal pedestal was measured in quiet and in the presence of a noise background. In the first experiment, the level of a 30-ms pedestal was fixed at 45 dB SPL and presented in the temporal and spectral center of a 110-ms notched noise. For a noise spectrum level of between 0 and 15 dB the noise produced a substantial reduction in the Weber fraction, i.e., an improvement in detectability, compared to the condition without the noise. The second experiment showed that, unlike the situation with notched noise, narrow-band noise produced no performance improvement, suggesting that the effect is dependent on noise frequency components outside the critical band of the pedestal. The third experiment showed that the improvement also occurred for a 6-ms pedestal presented in a 10-ms gap between two bursts of notched noise. The experiment rules out an explanation for the effect of the noise in terms of suppression on the basilar membrane. Finally, the effect was shown to decrease as the gap between the noise bursts was increased, in a manner at least broadly consistent with the decay of the temporal excitation pattern. It is suggested that the improvement in intensity discrimination in notched noise is due to an across-frequency comparison mechanism similar to ''profile analysis,'' perhaps operating on a temporally smoothed central representation of the stimulus. © *1998 Acoustical Society of America.* [S0001-4966(98)02505-3]

PACS numbers: 43.66.Ba, 43.66.Dc, 43.66.Fe [JWH]

INTRODUCTION

There are a growing number of examples of situations in which the auditory system can use across-frequency comparisons to aid signal detection. The most well-researched of these are comodulation masking release (Hall *et al.*, 1984) and profile analysis (Green, 1988; Green *et al.*, 1983). Profile analysis refers to the use of spectral changes to detect an intensity increment. In a classic intensity-discrimination task, a listener may be required to detect an intensity difference between sinusoids presented in two observation intervals. Over a series of such trials the standard, or pedestal, intensity is usually kept constant. On each trial the listener is required to choose the interval in which an intensity increment has been added to the pedestal. In a typical profile analysis experiment, however, pedestal intensity is randomized either *between* trials or between observation intervals *within* a trial. The former manipulation ensures that listeners cannot use a long-term memory representation of the standard to aid performance and the latter manipulation ensures that absolute intensity is not a reliable cue. In both cases performance is improved when a background stimulus, containing sinusoidal components with frequencies either side of the pedestal frequency and with intensities equal to the standard intensity, is added to both observation intervals. It appears that listeners can detect the change in spectral shape caused by the addition of the increment. Notice that when the standard intensity is randomized on each presentation, listeners must be basing discrimination on the *relative* difference between the pedestal plus increment and the background.

It has been suggested that measurements of spectral shape are more ''categorical'' and robust than other measures of intensity (Green, 1993). For example, there is very little effect of increasing the delay between the observation intervals when these stimuli are used (Green *et al.*, 1983) suggesting that the coding is highly resistant to memory decay. The idea is similar to the theory of ''context coding'' (Braida *et al.*, 1984; Durlach and Braida, 1969): The context code is assumed to be less susceptible to memory decay than the ''trace code'' that is used in the absence of a reference intensity.

More recently, evidence has been provided that across-frequency comparisons can produce a dramatic improvement in intensity discrimination even in the absence of intensity randomization. Plack and Viemeister (1992a, b) showed that the mid-level elevation in the Weber fraction under forward and backward masking (Plack and Viemeister, 1992b; Zeng *et al.*, 1991) could be reduced by presenting a notched noise with the pedestal. In backward masking, the notched noise caused a very large reduction in the Weber fraction of around 15 dB. The results were taken as evidence against the hypothesis (Zeng and Shannon, 1995; Zeng *et al.*, 1991) that the mid-level elevation in forward masking is due to the selective adaptation of low spontaneous rate (SR) auditory nerve fibers. It was argued that the effects of the notched noise could not be explained by processes at the level of the auditory nerve, although there is still some controversy over this issue (Zeng and Shannon, 1995).

Another situation in which intensity discrimination is poor at mid levels is for short tone bursts at high frequencies (Carlyon and Moore, 1984, 1986; Florentine, 1986). The phenomenon has been termed the ''severe departure'' from Weber's law (Carlyon and Moore, 1984). There have been several explanations for this effect, based on mechanisms at different stages of the auditory system.

(i) Mid-level compression on the basilar membrane (Ruggero, 1992) reduces the effective intensity difference between the pedestal and the pedestal plus increment (Klitzing and Kohlrausch, 1994).

(ii) The rate-level functions of auditory nerve fibers are such that the high-SR fibers are more sensitive to changes in low-level stimuli, whereas the low-SR fibers, with higher thresholds and wider dynamic ranges (Liberman, 1978), are more sensitive to changes in high-level stimuli. At mid levels neither fiber group is within its optimal range and performance is poor (Carlyon and Moore, 1984).

(iii) The increased variability of auditory-nerve firing at mid levels for short tone bursts impairs level discrimination (Oxenham and Moore, 1995; Raab and Taub, 1969).

(iv) At mid levels, the internal intensity reference points (or ''perceptual anchors'') of absolute threshold and discomfort threshold are perceptually distant, so that there is no proximal reference point for representing sound level, and hence coding is inaccurate (Braida *et al.*, 1984; Plack and Carlyon, 1995).

Carlyon and Moore (1984) found that gating a notched noise with the pedestal caused a large increase in the size of the mid-level elevation. However, Oxenham and Moore (1995) argued that this is a manifestation of the ''overshoot'' phenomenon (McFadden, 1989; Zwicker, 1965), whereby the threshold of a short-duration tone is raised near masker onset. It has been shown that overshoot reaches a maximum at mid levels (Bacon, 1990; Oxenham and Moore, 1995). In direct contrast to the results using gated notched noise, Plack and Viemeister (1992a) found that a notched noise with an onset 50 ms before the 30-ms 6-kHz pedestal, caused a substantial *reduction* in the Weber fraction at mid-levels, removing the severe departure completely. So it appears that notched noise can remove mid-level increases in the Weber fraction caused by either forward masking, backward masking, or high frequencies.

Just as the effects of notched noise on intensity discrimination under nonsimultaneous masking have implications for hypotheses regarding the origin of the mid-level elevation in these conditions, the finding that notched noise can remove the mid-level elevation in quiet has implications for explanations of the severe departure. For example, if performance could be improved by a mechanism located centrally to the auditory nerve, then this would raise doubts that discrimination is limited by the information present in the auditory nerve. There are, however, several possible hypotheses for the effect of the notched noise on the severe departure, some peripheral and some more central.

(i) Suppression. The notched noise may suppress the pedestal, thereby reducing the effective level of the pedestal so that it is no longer ''mid level,'' and/or reducing the on-frequency basilar-membrane compression. Palmer and Evans (1982) and Costalupes *et al.* (1984) have shown how notched noise can reduce auditory-nerve firing rate and steepen rate-level functions at mid levels, effects which may be a consequence of both of these processes.

(ii) Adaptation. Adaptation produced by the portion of the notched noise before the onset of the pedestal may reduce the onset response and hence reduce the associated variability (Oxenham and Moore, 1995). This basic argument has been used to explain certain characteristics of the overshoot phenomenon (Bacon, 1990; McFadden and Champlin, 1990). Auditory nerve adaptation is determined by the driven response to the adaptor in the fiber being recorded (Smith, 1979), and depends, therefore, on the adaptor and adaptee falling within the same physiological tuning curve (in psychophysical terms, the same *critical band*).

(iii) Temporal referential coding. Plack *et al.* (1995; Plack, 1996) showed how relative comparisons across time may improve intensity discrimination for tone bursts in the presence of forward or backward masking. These temporal comparisons seem to depend on the pedestal and the comparison tone being within the same critical band (Plack *et al.*, 1995). The same sort of across-time level comparisons, involving components of the noise before or after the pedestal, may improve coding accuracy in the severe departure condition.

(iv) Profile analysis. Similarly, relative comparisons across frequency may also improve discrimination, by providing a reliable level reference (Green, 1993).

Hypotheses (ii) and (iii) suggest that the improvement in discrimination should depend only on the frequency components of the noise within the critical band occupied by the pedestal. On the other hand, hypothesis (iv) suggests that the off-frequency components are more important. Hypothesis (i) suggests that components within the suppression region (Houtgast, 1972), which extends beyond the critical band, may have the most influence. Hypotheses (iii) and (iv) are linked to the context coding explanation of the severe departure, in that the notched noise may provide an intensity reference that is not present in quiet because of the perceptual distance of the internal anchors.

The aim of the experiments described here was to determine the likely mechanism for the effects of notched noise on intensity discrimination for short tone bursts at high frequencies. It was hoped that the data would be relevant both for explanations of the severe departure and for theories of referential intensity coding, such as profile analysis.

I. EXPERIMENT 1: EFFECTS OF NOTCHED NOISE LEVEL

A. Rationale

Experiment 1 was designed first, to replicate the basic finding of Plack and Viemeister (1992a), and second, to determine the optimum level of the notched noise relative to the pedestal. If the improvement in performance is the result of a process akin to profile analysis, it might be expected that the notched noise will be most effective when the excitation produced by the noise is similar to the excitation produced by the pedestal, as the increment would then produce a categorical ''bump'' in the excitation pattern. Theoretically, the accuracy of context coding is assumed to decline as the perceptual distance between the target and the reference is increased (Durlach and Braida, 1969).

B. Stimuli

Intensity discrimination was measured for a 45 dB SPL, 30-ms, 6-kHz pedestal presented either in quiet or in the temporal center of a 110-ms notched noise. Durations included 5-ms, raised-cosine, onset and offset ramps. The

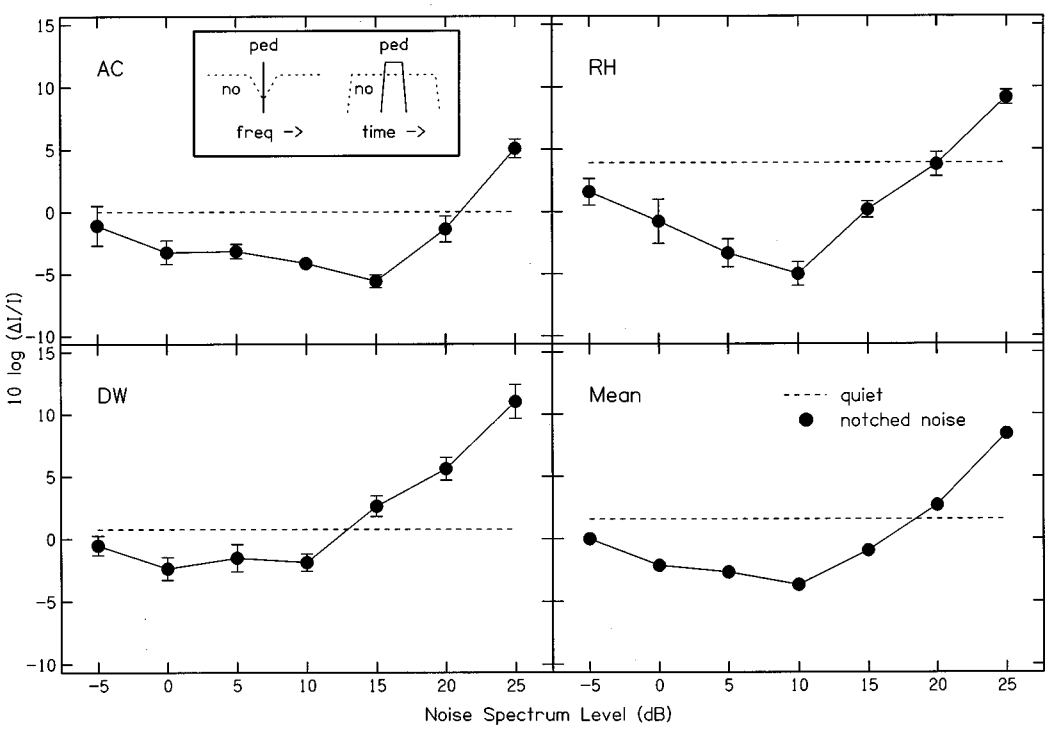


FIG. 1. The results of experiment 1 showing Weber fractions for a 45 dB SPL, 30-ms, 6-kHz pedestal presented in quiet (dashed line) or in the temporal center of a notched noise. The horizontal axis gives the spectrum level of the noise. Error bars are standard errors. A schematic illustration of the spectral and temporal characteristics of the stimuli are provided in the box.

notch had a bandwidth of 1 kHz, centered on 6 kHz. The spectrum level of the noise in the passband was varied between -5 and 25 dB in 5-dB steps.

Stimuli were generated digitally on a Silicon Graphics workstation with a sampling rate of 44.1 kHz. The noise was filtered using a digital FIR filter with a nominal attenuation of 90 dB per octave. Analog waveforms were generated using a built-in 16-bit DAC and were presented through Sennheiser HD414 headphones which were connected directly to the headphone output of the computer.

C. Procedure

A 2I, 2AFC paradigm was adopted throughout. The interstimulus interval was 500 ms. A two-down one-up adaptive tracking rule was used to estimate the 71% correct point on the psychometric function (Levitt, 1971). The level of the pedestal (I) in the standard interval was held constant and the level of the pedestal plus increment ($I+\Delta I$) in the comparison interval was varied adaptively, $10\log(\Delta I/I)$ being increased and decreased by 4 dB for the first four turnpoints, and by 2 dB thereafter. Sixteen turnpoints were recorded in each experimental block and the threshold estimate was taken as the mean of the values of $10\log(\Delta I/I)$ at the last 12 turnpoints. Eight such estimates were made for each listener (one estimate per condition in each experimental session) and the results averaged. Within a session each condition was tested in a random order.

Listeners were tested individually in an IAC single-walled sound-attenuating booth. They made their responses using the numeric keypad on the computer keyboard. ''Lights'' were presented in a graphical display on the computer monitor to delineate the observation intervals, and to provide feedback after each trial.

D. Listeners

Three normally hearing listeners took part. They were paid an hourly wage for their services. Listeners were given at least two hours practice on the task before data collection began.

E. Results and discussion

The results of experiment 1 are presented in Fig. 1. The Weber fraction in dB is plotted against the spectrum level of the notched noise. The horizontal dashed lines show the Weber fraction in the absence of the noise. A consistent pattern is seen across the three listeners. At high spectrum levels, the notched noise masks the pedestal and performance is worse than that in quiet. For noise spectrum levels below about 15 dB the noise produces a considerable improvement in intensity discrimination, confirming the results of Plack and Viemeister (1992a). The optimum spectrum level of the noise appears to be about 10 dB, and performance deteriorates as the noise level is increased or decreased away from this value.

Figure 2 shows the auditory excitation patterns of the notched noise alone (continuous line), the notched noise plus pedestal (dashed line), and the notched noise plus pedestal plus increment (dotted line). The pedestal level was 45 dB and the notched noise spectrum level was 10 dB (the level which produced the largest improvement in performance in experiment). The level of the pedestal plus increment was 48

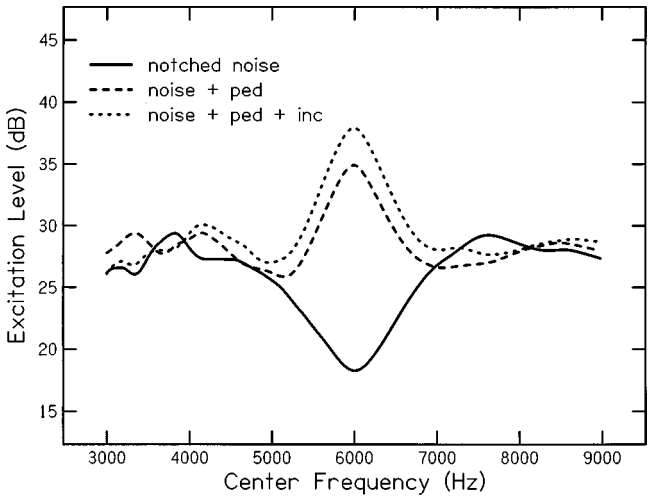


FIG. 2. Auditory excitation patterns for a notched noise with a spectrum level of 10-dB (solid line), a notched noise with a 45-dB, 6-kHz sinusoid (dashed line), and a notched noise with a 48-dB sinusoid (dotted line). Spectra were calculated using a 20-ms sampling window.

dB (Weber fraction=0 dB). The curves were generated from the spectra of actual stimuli used in the experiment, using the equations described by Glasberg and Moore (1990). For the notched noise alone curve, the spectrum was computed over a 20-ms portion of the stimulus before the pedestal onset. For the other two curves, the spectra were computed over the 20-ms portion of the stimulus during which the pedestal was at its steady-state level.

It can be seen from these curves that the excitation level of the noise plus pedestal at 6 kHz is about 8 dB above the level of the ''background'' (i.e., the excitation produced by the noise components away from the pedestal frequency). In a typical profile analysis experiment this would be equivalent to presenting a pedestal about 8 dB higher than the background sinusoids in the complex. This might appear to be inconsistent with the profile analysis/context coding account, since it would be expected from the theory that the noise would be most effective when it produced a flat excitation pattern when added to the pedestal, so that increment could be detected as a distinct ''bump.'' However, in a profile analysis experiment using a wide 20-component background (the configuration most similar to the notched noise used here), Green and Kidd (1983) found that the pedestal level which produced the largest effect was around 6 dB above the background, which is consistent with the results of experiment 1. They further showed that if they removed the adjacent low-frequency components from the background then the optimum relative level was 0 dB. They argued that thresholds might be elevated by the partial masking of the pedestal by these low-frequency components, which counteracts the advantages of a flat profile. The same argument could be applied here, since the 10-dB noise was close to the level at which the noise produced a deterioration, rather than an improvement, in performance. Whether or not this reasoning is valid, it can at least be stated that the effects of noise level are consistent with those seen in profile analysis experiments in which listeners were forced to measure intensity using *relative,* across-frequency, comparisons.

II. EXPERIMENT 2: NOTCHED NOISE VERSUS NARROW-BAND NOISE

A. Rationale

If the effect of the notched noise observed in experiment 1 is due purely to the on-frequency components in the excitation pattern of the noise (see Fig. 2), as implied by the adaptation and temporal referential coding hypotheses, then a narrow-band noise with an excitation level at the pedestal frequency similar to that of the notched noise should be equally effective in improving intensity discrimination. Calculations showed that for the notched noise the excitation level at 6 kHz was about 9 dB below the excitation level in the passband (see Fig. 2). For a narrowband noise with a bandwidth of 1 kHz (centered on 6 kHz) and with the same spectrum level, the excitation level at 6 kHz was about 3 dB less than the excitation level of the notched noise in the passband. In other words, a narrow-band noise with a spectrum level 6 dB below that of the notched noise should produce approximately the same excitation level at the 6-kHz center frequency. In experiment 2, three narrow-band noise spectrum levels were chosen covering a 10-dB range around this value relative to a notched noise spectrum level of 10 dB.

B. Stimuli and procedure

The temporal characteristics of the stimuli, and method of generation, were the same as in experiment 1. The 45-dB pedestal was presented either in quiet, in a notched noise, or in a narrowband noise with a bandwidth of 1 kHz. Notched noise spectrum levels of 10 and 15 dB were tested and narrow-band noise spectrum levels of 0, 5, and 10 dB were tested. The procedure was the same as in experiment 1 and the same three listeners were employed.

C. Results and discussion

The results of experiment 2 are shown in Fig. 3. The 10-dB notched noise once again produced an improvement in performance across all three listeners, although the effect was slightly less than that observed in experiment 1. The narrowband noise had no substantial effect on performance except at a spectrum level of 10 dB, where it produced a slight increase in the Weber fraction. It can be concluded that the on-frequency components of the notched noise are not responsible for the performance improvement. It follows that both the adaptation hypothesis and the temporal referential coding hypothesis should be rejected.

III. EXPERIMENT 3: EFFECTS OF NONSIMULTANEOUS NOTCHED NOISE

A. Rationale

Physiological suppression is a phenomenon believed to be a consequence of nonlinearities on the basilar membrane (Yates, 1995) in which the excitation at one characteristic frequency (CF) acts to reduce the excitation at neighboring CFs. Suppression has a very short time constant (Arthur *et al.*, 1971) which is consistent with the idea that the effect depends on the suppressor and the suppressee interacting on

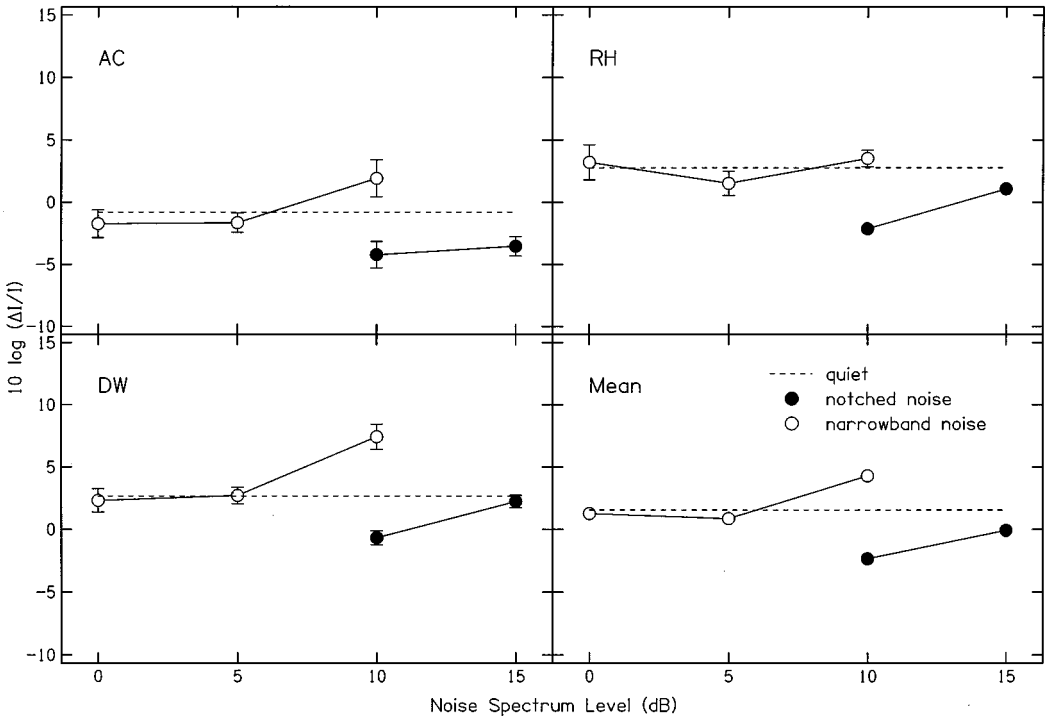


FIG. 3. The results of experiment 2, comparing the effects of a notched noise and a narrow-band noise on intensity discrimination for a 30-ms pedestal. The axes and error bars are as Fig. 1.

the basilar membrane. Following this reasoning, it was decided to determine if it is possible to obtain the beneficial effects of the notched noise without the pedestal and noise overlapping in time. The pedestal was presented in a 10-ms gap between two bursts of noise. At 6 kHz, where the temporal response of the auditory filter is very brief (Kiang *et al.*, 1965), it was reasoned that the gap would be sufficient to avoid significant interaction on the basilar membrane, and therefore to avoid suppression of the pedestal by the notched noise. Plack and Viemeister (1992a) found that a notched noise could reduce the elevation of the Weber fraction under forward masking even when the noise was terminated before the onset of the pedestal. They used this finding as evidence that the effects of the notched noise were not the result of suppression. Similarly, if nonsimultaneous notched noise can reduce the mid-level elevation in quiet, then this will be good evidence that the mechanism underlying this effect is not peripheral suppression.

B. Stimuli and procedure

The sinusoidal pedestal had a steady-state duration of 2 ms with 2-ms raised-cosine onset and offset ramps. The pedestal was presented either in quiet, in the temporal center of a notched noise (no gap), or in the temporal center of a 10-ms silent interval between two bursts of notched noise (gap). The spectral characteristics of the notched noise were the same as in experiment 1. In the no gap condition, the notched noise had a steady-state duration of 200 ms with 2-ms onset and offset ramps. In the gap condition, each noise burst had a steady-state duration of 100 ms with 2-ms onset and offset ramps. The spectrum level of the noise was always 40 dB below the pedestal level. In each of the three conditions, pedestal levels of 30–80 dB SPL were tested in 10-dB steps. The method of stimulus generation and the procedure were the same as in experiment 1.

C. Listeners

Three new listeners took part. They all had normal hearing and were paid an hourly wage for their services. Listeners were given at least two hours practice on the task before data collection began.

D. Results and discussion

The results of experiment 3 are shown in Fig. 4. The basic pattern of results is similar across all three listeners, although the Weber fractions for listener SM are much higher than the others, particularly in quiet. All listeners show a mid-level elevation in the Weber fraction in the absence of the noise (the severe departure). The Weber fraction is maximum for a pedestal level of 50 dB. Consistent with the earlier results, the simultaneous notched noise (no gap) improved discrimination performance at mid levels, removing the severe departure. For pedestal levels below 60 dB, the effects of the nonsimultaneous notched noise (gap) produced performance virtually identical to that observed with the simultaneous noise. For pedestal levels of 60 dB and above, however, the curves diverge. At the highest pedestal level, Weber fractions for the nonsimultaneous noise are substantially higher than those for both the simultaneous noise and quiet conditions.

It is possible that effects of the noise at high levels may be evidence for the beneficial effects of suppression at these levels. It has been shown previously that notched noise can impair performance at high levels, eliminating the ''near

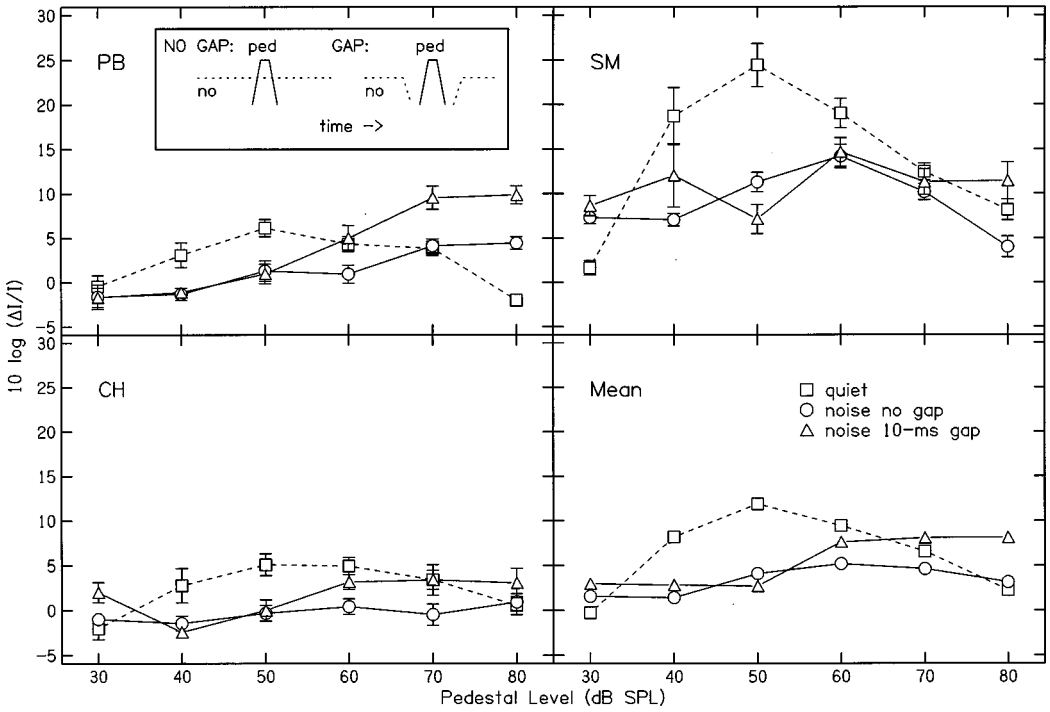


FIG. 4. The results of experiment 3 showing Weber fractions for a 6-ms pedestal presented alone (dashed lines), in the temporal center of a notched noise (circles), or in the temporal center of a 10-ms gap between two bursts of notched noise (triangles). The horizontal axis gives the level of the pedestal. The spectrum level of the noise was 40 dB below the pedestal level. Error bars are standard errors. The temporal characteristics of the stimuli in the region of the pedestal are illustrated in the box.

miss'' to Weber's law (Moore and Raab, 1974). The effect could be attributed to the removal of intensity information from the skirts of the excitation pattern. The present data show that a nonsimultaneous notched noise is even more effective in this regard.

The large dynamic range of hearing has long been a difficult theoretical problem (see Plack and Carlyon, 1995), since the majority of auditory nerve fibers saturate below about 60 dB SPL (Palmer and Evans, 1979). If performance at high levels is dependent on information from the skirts of the excitation pattern (where the neurons are not saturated) then masking the skirts should be very damaging to discrimination. It is conceivable that by suppressing the pedestal to a lower excitation level, where there is richer neural coding at CF, the simultaneous noise may not reveal the true detriment caused by the loss of off-frequency information (Plack and Viemeister, 1993).

Even if suppression by the notched noise is important at high levels, it appears that in the region of the severe departure it has little effect on performance. The results suggest that peripheral suppression is not the cause of the beneficial effects of the notched noise at mid levels.

It might be argued that the data also provide difficulties for the profile analysis account, since the effect has been shown not to rely on across-frequency comparisons between physically simultaneous sounds. However, temporal resolution experiments suggest that the internal representation of a sound stimulus is effectively blurred in time by the action of a central temporal integrator, or ''temporal window,'' with a time constant of about 10 ms (Moore *et al.*, 1988; Oxenham and Moore, 1994). Since the time interval between the end of the first masking burst and the center of the pedestal was only 5 ms in the gap condition, it is conceivable that at some higher stage of the auditory system the noise and the pedestal did in fact overlap in time and a simultaneous comparison might have been possible at this stage. If the argument is valid, then lengthening the gap should cause a reduction in the benefits of the noise, consistent with the decay of the temporal excitation pattern. The final experiment was designed to test this prediction.

IV. EXPERIMENT 4: EFFECTS OF GAP DURATION WITH NONSIMULTANEOUS NOTCHED NOISE

A. Rationale

Assuming that the benefit of the noise is the result of a simultaneous across-frequency comparison after the representation of the stimulus has been smoothed by a temporal window similar to that described by Moore *et al.* (1988), then a gap of around 20 ms should reduce the effective level of the noise at the temporal center of the pedestal by about 10 dB, and a gap of 40 ms should reduce the effective level of the noise by about 20 dB. On the basis of the results of experiment 1, this should be sufficient to substantially reduce the beneficial effects of the noise.

B. Stimuli and procedure

The basic stimuli and procedure were the same as in experiment 3. The level of the pedestal was 50 dB SPL, and the spectrum level of the notched noise was 10 dB. The pedestal was presented either alone or in the temporal center

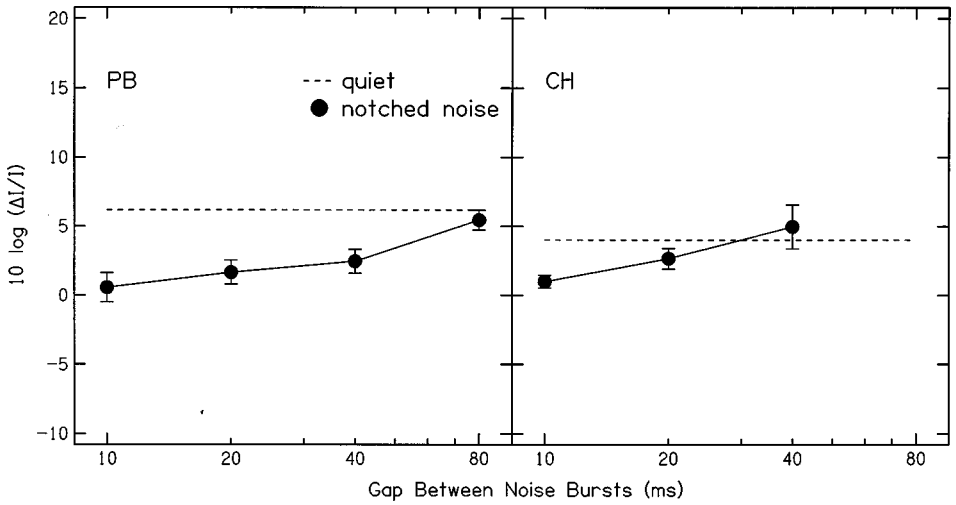


FIG. 5. The results of experiment 3 showing Weber fractions for a 6-ms pedestal presented alone (dashed lines) or in the temporal center of a gap between two bursts of notched noise (circles). The horizontal axis gives the duration of the gap.

of a silent interval between two 100-ms bursts of noise as before. Two of the listeners from experiment 3 were employed here. Gap durations of 10, 20, and 40 ms were tested for both listeners, and a gap duration of 80 ms was added for listener PB.

C. Results and discussion

The results are presented in Fig. 5. Both listeners show a deterioration in performance as the gap between the noise bursts is increased. In other words, the beneficial effects of the notched noise decrease as the gap increases. For listener CH the 40-ms gap eliminated the improvement entirely whereas for listener PB an 80-ms gap was required. In experiment 1, a 15-dB reduction in noise spectrum level almost eliminated the effects of the noise. However, the mean difference in the Weber fraction between noise spectrum levels of -5 and 10 dB in that experiment was 3.7 dB. The mean deterioration between a 10-ms gap and a 40-ms gap (corresponding to roughly a 20-dB attenuation at the output of the temporal window) in experiment 4 was 2.9 dB. If the Weber fraction increases at a steady rate as the noise spectrum level is reduced, then these results are not dramatically different from the predictions of the hypothesis presented in the previous section, that across-frequency comparisons are used after the internal representation of the stimulus has been smoothed by the temporal window.

V. GENERAL DISCUSSION

A. What is the mechanism underlying the effects of the notched noise?

In the introduction, four very different explanations for the performance improvements in notched noise were presented. Two of these, the adaptation and temporal referential coding hypotheses, were eliminated after experiment 2 showed that off-frequency noise energy is required for the effect. The suppression hypothesis appears to be untenable in the face of the results of experiment 3 in which a nonsimultaneous notched noise was shown to have an effect just as great as a simultaneous one. The only explanation which has survived this examination intact is the profile analysis hypothesis, which states that the improvements in intensity discrimination are due to the use of across-frequency, relative level comparisons.

Experiment 1 showed that the level of the notched noise that caused the greatest reduction in the Weber fraction was consistent with the relative level of the background sinusoids that produces the optimum performance in profile analysis experiments (Green and Kidd, 1983). Now, because in the latter case the level of the background is randomized between observation intervals, listeners must base discrimination on some form of across-frequency level comparison. The similarity between the results of the two experiments is further evidence, therefore, that across-frequency comparisons caused the performance improvement in the present experiments. The results of the final experiment raise the possibility that the comparison occurs after the representation of the stimulus has been smoothed in time by the action of a central temporal integrator.

It has been argued that profile analysis involves a ''categorical'' representation of intensity (for example, whether the spectrum is ''bumped'' or ''flat'') that may be simpler to code, less susceptible to memory decay, and hence more reliable than ''absolute'' representations of intensity (Green, 1993). It is suggested here that similar coding strategies may improve performance for intensity discrimination in notched noise.

B. Implications for explanations of the severe departure

Assuming for the moment that a profile analysis mechanism underlies the effects of the noise, it is now possible to consider the implications of this hypothesis for hypotheses relating to the severe departure. If the severe departure is the result of a *peripheral* limitation, then it is not obvious how a *central* mechanism could overcome this limitation. Consider, for example, that the severe departure is a consequence of mid-level basilar membrane compression (Klitzing and Kohlrausch, 1994). The effect of the compression will be to

reduce the level difference between the pedestal plus increment and the pedestal. Whatever happens at later stages will be limited by this process: A reduction in variability caused by a different coding strategy may improve performance overall, but it cannot change the fact that the input is impoverished at mid levels, and it might be expected that the hump should remain. The same basic logic applies to the hypothesis linking the hump to a discontinuity between the rate-level functions of the high-SR and low-SR fibers (Carlyon and Moore, 1984).

There are ways around this argument, although they are not entirely satisfactory. For example, it is conceivable that the across-frequency comparator can only operate effectively at mid levels, so that the peripheral mid-level impairment is followed by a central mid-level facilitation, and the consequence is a flat function. However, this is inconsistent with a study by Green *et al.* (1986) which suggests that profile analysis is not affected by overall level. Alternatively, it could be that there is a very nonlinear ''hard'' central limitation, that does not allow a Weber fraction of less than, say, -5 dB to be detected. The notched noise may only be able to improve performance this far at any level, again producing a flat response. However, this is not consistent with the much lower Weber fractions observed for longer pedestals (Florentine, 1986).

The strong conclusion that the severe departure is not the result of a peripheral mechanism does raise a problem, however. It is clear that the basilar membrane is more the compression remains at high levels (Murugasu and Russell, 1995; Ruggero, 1992; Ruggero *et al.*, 1997; Russell and Nilsen, 1997), and the compression should result in an effective reduction in the level difference at input. On the other hand, analyses of neural firing data suggest that a mid-level limitation is not present in the firing-rate information in the auditory nerve (Delgutte, 1987; Viemeister, 1988), the implication being that the richness of neural coding at mid levels may compensate for the reduction in the level difference.

VI. CONCLUSIONS

(i) For certain noise levels, the addition of notched noise reduced the Weber fraction for a 45-dB, 30-ms, 6-kHz pedestal. The spectrum level that produced the largest improvement in performance was 10 dB. The optimum relative level of the noise and pedestal was similar to the optimum relative level between background and pedestal in profile analysis.

(ii) A narrow-band noise did not produce any improvement in performance. The effect of the notched noise appears to depend on off-frequency components.

(iii) The notched noise was effective at reducing the mid-level elevation even when it was not simultaneous with the pedestal. The finding suggests that peripheral suppression is not involved in the improvement at mid levels, although there is some evidence that it may have an effect at high levels, perhaps improving the dynamic range by suppressing the pedestal to a lower level where there is richer neural coding.

(iv) For a pedestal presented in the gap between two notched noise bursts, the beneficial effect of the noise decreased as the gap was increased.

(v) The results are consistent with an across-frequency comparison mechanism similar to profile analysis. It is possible that this mechanism operates on the temporally smoothed representation of sounds at the output of a central temporal window.

ACKNOWLEDGMENTS

I am grateful to the two anonymous reviewers whose constructive criticisms improved the clarity of the manuscript. The research was supported by a Royal Society University Research Fellowship held by the author.

Arthur, R. M., Pfeiffer, R. R., and Suga, N. (**1971**). ''Properties of 'two tone inhibition' in primary auditory neurones,'' J. Physiol. (London) **212**, 593–609.

Bacon, S. P. (**1990**). ''Effect of masker level on overshoot,'' J. Acoust. Soc. Am. **88**, 698–702.

Braida, L. D., Lim, J. S., Berliner, J. E., Durlach, N. I., Rabinowitz, W. M., and Purks, S. R. (**1984**). ''Intensity perception. XIII. Perceptual anchor model of context coding,'' J. Acoust. Soc. Am. **76**, 722–731.

Carlyon, R. P., and Moore, B. C. J. (**1984**). ''Intensity discrimination: A severe departure from Weber's law,'' J. Acoust. Soc. Am. **76**, 1369–1376.

Carlyon, R. P., and Moore, B. C. J. (**1986**). ''Detection of tones in noise and the ''severe departure'' from Weber's law,'' J. Acoust. Soc. Am. **79**, 461–464.

Costalupes, J. A., Young, E. D., and Gibson, D. J. (**1984**). ''Effects of continuous noise backgrounds on rate response of auditory nerve fibers in cat,'' J. Neurophysiol. **51**, 1326–1344.

Delgutte, B. (**1987**). ''Peripheral auditory processing of speech information: Implications from a physiological study of intensity discrimination,'' in *The Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Nijhof, Dordrecht, The Netherlands).

Durlach, N. I., and Braida, L. D. (**1969**). ''Intensity perception. I. Preliminary theory of intensity resolution,'' J. Acoust. Soc. Am. **46**, 372–383.

Florentine, M. (**1986**). ''Level discrimination of tones as a function of duration,'' J. Acoust. Soc. Am. **79**, 792–798.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hearing Res. **47**, 103–138.

Green, D. M. (**1988**). *Profile Analysis* (Oxford U.P., New York).

Green, D. M. (**1993**). ''Auditory intensity discrimination,'' in *Human Psychophysics*, edited by W. A. Yost, A. N. Popper, and R. R. Fay (Springer-Verlag, New York).

Green, D. M., and Kidd, G. (**1983**). ''Further studies of auditory profile analysis,'' J. Acoust. Soc. Am. **73**, 1260–1265.

Green, D. M., Kidd, G., and Picardi, M. C. (**1983**). ''Successive versus simultaneous comparison in auditory intensity discrimination,'' J. Acoust. Soc. Am. **73**, 639–643.

Green, D. M., Zekiye, A. O., and Forrest, T. G. (**1986**). ''Frequency effects in profile analysis and detecting complex spectral changes,'' J. Acoust. Soc. Am. **81**, 692–699.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (**1984**). ''Detection in noise by spectro-temporal pattern analysis,'' J. Acoust. Soc. Am. **76**, 50–56.

Houtgast, T. (**1972**). ''Psychophysical evidence for lateral inhibition in hearing,'' J. Acoust. Soc. Am. **51**, 1885–1894.

Kiang, N. Y.-S., Wantanabe, T., Thomas, E. C., and Clark, L. F. (**1965**). ''Discharge patterns of single fibers in the cat's auditory nerve,'' Res. Monogr. No. 35, MIT, Cambridge, MA.

Klitzing, R. V., and Kohlrausch, A. (**1994**). ''Effect of masker level on overshoot in running- and frozen-noise maskers,'' J. Acoust. Soc. Am. **95**, 2192–2201.

Levitt, H. (**1971**). ''Transformed up-down methods in psychophysics,'' J. Acoust. Soc. Am. **49**, 467–477.

Liberman, M. C. (**1978**). ''Auditory-nerve response from cats raised in a low-noise chamber,'' J. Acoust. Soc. Am. **63**, 442–455.

McFadden, D. (**1989**). ''Spectral differences in the ability of temporal gaps to reset the mechanisms underlying overshoot,'' J. Acoust. Soc. Am. **85**, 254–261.

McFadden, D., and Champlin, C. A. (**1990**). ''Reductions in overshoot during aspirin use,'' J. Acoust. Soc. Am. **87**, 2634–2642.

Moore, B. C. J., and Raab, D. H. (**1974**). ''Pure-tone intensity discrimination: some experiments relating to the ''near miss'' to Weber's Law,'' J. Acoust. Soc. Am. **55**, 1049–1054.

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.

Murugasu, E., and Russell, I. J. (**1995**). ''Salicylate ototoxicity: The effects on basilar membrane displacement, cochlear microphonics, and neural responses in the basal turn of the guinea pig cochlea,'' Aud. Neurosci. **1**, 139–150.

Oxenham, A. J., and Moore, B. C. J. (**1994**). ''Modeling the additivity of nonsimultaneous masking,'' Hearing Res. **80**, 105–118.

Oxenham, A. J., and Moore, B. C. J. (**1995**). ''Overshoot and the ''severe departure'' from Weber's law,'' J. Acoust. Soc. Am. **97**, 2442–2453.

Palmer, A. R., and Evans, E. F. (**1979**). ''On the peripheral coding of the level of individual frequency components of complex sounds at high sound levels,'' in *Hearing Mechanisms and Speech*, edited by O. Creutzfeldt, H. Scheich, and C. Schreiner (Springer-Verlag, Berlin).

Palmer, A. R., and Evans, E. F. (**1982**). ''Intensity coding in the auditory periphery of the cat: Responses of cochlear nerve and cochlear nucleus neurons to signals in the presence of bandstop masking noise,'' Hearing Res. **7**, 305–323.

Plack, C. J. (**1996**). ''Temporal factors in referential intensity coding,'' J. Acoust. Soc. Am. **100**, 1031–1042.

Plack, C. J., and Viemeister, N. F. (**1992a**). ''The effects of notched noise on intensity discrimination under forward masking,'' J. Acoust. Soc. Am. **92**, 1902–1910.

Plack, C. J., and Viemeister, N. F. (**1992b**). ''Intensity discrimination under backward masking,'' J. Acoust. Soc. Am. **92**, 3087–3101.

Plack, C. J., and Viemeister, N. F. (**1993**). ''Suppression and the dynamic range of hearing,'' J. Acoust. Soc. Am. **93**, 976–982.

Plack, C. J., and Carlyon, R. P. (**1995**). ''Loudness perception and intensity coding,'' in *Hearing*, edited by B. C. J. Moore (Academic, New York), pp. 123–160.

Plack, C. J., Carlyon, R. P., and Viemeister, N. F. (**1995**). ''Intensity discrimination under forward and backward masking: Role of referential coding,'' J. Acoust. Soc. Am. **97**, 1141–1149.

Raab, D. H., and Taub, H. B. (**1969**). ''Click-intensity discrimination with and without a background masking noise,'' J. Acoust. Soc. Am. **46**, 965–968.

Ruggero, M. A. (**1992**). ''Responses to sound of the basilar membrane of the mammalian cochlea,'' Curr. Opin. Neurobiol. **2**, 449–456.

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.

Russell, I. J., and Nilsen, K. E. (**1997**). ''The location of the cochlear amplifier: Spatial representation of a single tone on the guinea pig basilar membrane,'' Proc. Natl. Acad. Sci. USA **94**, 2660–2664.

Smith, R. L. (**1979**). ''Adaptation, saturation, and physiological masking in single auditory-nerve fibers,'' J. Acoust. Soc. Am. **65**, 166–178.

Viemeister, N. F. (**1988**). ''Intensity coding and the dynamic range problem,'' Hearing Res. **34**, 267–274.

Yates, G. K. (**1995**). ''Cochlear structure and function,'' in *Hearing*, edited by B. C. J. Moore (Academic, New York), pp. 41–74.

Zeng, F.-G., and Shannon, R. V. (**1995**). ''Possible origins of the nonmonotonic intensity discrimination function in forward masking,'' Hearing Res. **82**, 216–224.

Zeng, F.-G., Turner, C. W., and Relkin, E. M. (**1991**). ''Recovery from prior stimulation II: Effects upon intensity discrimination,'' Hearing Res. **55**, 223–230.

Zwicker, E. (**1965**). ''Temporal effects in simultaneous masking by white-noise bursts,'' J. Acoust. Soc. Am. **37**, 653–663.

Auditory filter nonlinearity at 2 kHz in normal hearing listeners

Stuart Rosen, Richard J. Baker, and Angela Darling
Department of Phonetics & Linguistics, University College London, 4 Stephenson Way, London NW1 2HE, England

(Received 2 January 1997; revised 29 October 1997; accepted 26 January 1998)

Auditory filters broaden with increasing level. Using a recently developed method of fitting filter shapes to notched-noise masking data that explicitly models the nonlinear changes in filter shape across level, results at 2 kHz from 9 listeners over a wide range of levels and notch widths are reported. Families of roex(p,w,t) filter shapes lead to models which account well for the observed data. The primary effect of level is a broadening in the tails of the filter as level increases. In all cases, models with filter parameters depending on probe level fit the data much better than masker-dependent models. Thus auditory filter shapes appear to be controlled by their output, not by their input. Notched-noise tests, if performed at a single level, should use a fixed probe level. Filter shapes derived in this way, and normalized to have equal tail gain, are highly reminiscent of measurements made directly on the basilar membrane, including the degree of compression evidenced in the input–output function.
[S0001-4966(98)02205-X]

PACS numbers: 43.66.Ba, 43.66.Dc, 43.66.Sr, 43.64.Bt [WJ]

INTRODUCTION

A fundamental property of the peripheral auditory system is that it operates as a frequency analyzer. This processing mechanism can be conceptualized as a bank of overlapping bandpass filters, often referred to as *auditory filters*. Some of the applications exploiting the concept of auditory filters (and requiring accurate auditory filter shape characterization) include the design of realistic front ends for automatic speech recognition (Zwicker *et al.*, 1979), modeling the effects of normal and impaired frequency selectivity on the perception of speech or music (Patterson *et al.*, 1982; Rosen and Fourcin, 1986), and modeling loudness perception (e.g., Moore *et al.*, 1997). It has also been suggested that auditory filter shape measurement may be a useful diagnostic tool in the detection of preclinical noise-induced hearing loss (West and Evans, 1990).

The most well developed psychoacoustic technique for measuring auditory filter shapes is the notched-noise simultaneous masking method, supported by an explicit analysis algorithm (for a review, see Patterson and Moore, 1986). Almost always, thresholds are found for a tonal signal at a particular frequency in the presence of fixed-spectrum-level band-stop noises (*notched* noises), with the varying-width notch placed both symmetrically and asymmetrically around the probe frequency (e.g., Dubno and Dirks, 1989; Fidell *et al.*, 1983; Moore *et al.*, 1990; Niemiec and Yost, 1992; Patterson, 1976; Patterson *et al.*, 1982; Weber, 1977; Wright *et al.*, 1997; Shailer *et al.*, 1990, not to mention many others). The dependence of probe threshold on the notch configuration is exploited to estimate the auditory filter shape. The implicit assumption in such studies is that by fixing the spectrum level of the masker, the properties of the auditory filter being measured are also being fixed.

Of course, we could also fix the level of the probe, and vary the masker level for a particular notch. Were auditory filters linear, it would not matter which level was fixed and which was varied. But it is now well known that auditory filters are highly nonlinear, broadening with increasing level. Therefore it is crucial to decide whether to fix probe level or masker level in such experiments. At a more fundamental level, we need to determine what feature of the stimulus configuration controls filter shape. If filter shapes are somehow controlled by probe level, then fixing the masker level will distort considerably the shape of the derived auditory filters, and vice versa.

The question of which, if either of these parameters, actually controls filter shape can also be related to more complete models of auditory filtering. For example, Glasberg and Moore (1990) provide an algorithm for calculating excitation patterns based on the notion that auditory filter shape is determined by the level of the *input* to the filters. Their computations imply that fixing the level of the masker in a notched-noise experiment leads to an approximation of filter shape for a particular level (see below). On the other hand, Carney (1993) has proposed a model in which the output of the filter is rectified and smoothed, then used to control filter shape. With the additional assumption that the efficiency of the post-filtering detection process does not vary with level, such an *output-based* model implies that a fixed *probe* level leads to appropriate filter shapes.

For an *input-based* model it seems unlikely that the spectrum level of the masker *per se* controls the filter shape. Lutfi and Patterson (1984) suggested that the controlling factor is more likely to be the level of the masker in the region of the filter. However, if the ''region of the filter'' is restricted to one *equivalent rectangular bandwidth* (ERB) around the filter's center frequency and, given the above assumption about post-filtering detection, this ''restricted'' input is very closely correlated with filter output.

The excitation pattern model of Glasberg and Moore (1990; and Moore and Glasberg, 1987) implements another *input-based* model in which it is the level of an input com-

ponent integrated over an ERB that controls the amount of attenuation of that particular component by the filter concerned.[1] However, for a broadband signal such as those used in notched-noise masking, the ERB integrated input varies little over the width of the masking noise. For example the ERB is 111 Hz at 800 Hz and 241 Hz at 2 kHz. Accordingly, for a probe frequency of 2 kHz and a fixed masker spectrum level, the ERB integrated input level would vary only by 3.4 dB between masker components at an 800-Hz notch edge and those at the probe frequency in a no-notch condition. On the high-frequency side the filter shape is independent of signal level and so only signal components below and immediately around the probe frequency affect the filter shape in Glasberg and Moore's model. Thus in a notched-noise masking task, fixing the masker spectrum level would be a good approximation to the fixed ERB-integrated-input level that would be required to ensure a constant filter shape under Glasberg and Moore's assumptions.

While there is some debate as to which aspect of the signal controls the filter shape, it is the level of the probe tone or the spectrum level of the masker that are the ones most readily fixed in empirical studies (and in fact the only ones which have been). Given also that there are theoretical proposals to justify both, we focus solely on them.

It is, perhaps, a little surprising that this crucial issue has received relatively little attention until recently. Although two previous studies mentioned above have claimed to come to decisive conclusions (Lutfi and Patterson, 1984; Moore and Glasberg, 1987), we have recently shown that such claims cannot be upheld (Rosen and Baker, 1994). Instead, we have developed quite a different approach, which is to explicitly model the changes in filter parameters as a function of level, and to determine whether a dependence of these parameters on masker level (masker-dependent) or probe level (probe-dependent) better accounts for the data.

To illustrate this idea in a relatively simple case, consider a notched-noise masking experiment performed with symmetric notches only, at a fixed masker level of N_0 dB SPL/Hz. Let us assume the symmetric roex(p,w,t) filter shape to be appropriate (Patterson *et al.*, 1982), so off-place listening can be ignored:

$$(1-w)(1+pg)e^{-pg}+w(1+tg)e^{-tg}$$

where g is normalized frequency ($|f-f_0|/f_0$, f_0=filter center frequency), p and t are parameters determining the sharpness of the filter in the *p*assband and *t*ail, respectively, and w is a *w*eight determining where the tail slopes take over from the passband slopes. For such a shape, the tail slopes can only be shallower than those in the passband. Under certain simplifying assumptions (Rosen, 1989), the power spectrum model of masking (Patterson and Moore, 1986) leads to the level of the probe (P_s dB SPL) being given by

$$P_s = k + N_0 + 10\log(2f_0) + 10\log\left[\frac{(1-w)(2+pg)e^{-pg}}{p} + \frac{w(2+tg)e^{-tg}}{t}\right],$$

where g is now the particular symmetric notch width in question (normalized as above), and k is a measure of the detector efficiency at filter output (signal-to-noise ratio in dB).

Our *Poly*nomial *Fit*ting (PolyFit) procedure[2] might have the analogous equation:

$$P_s = (k_0+k_1N_0)+N_0+10\log(2f_0)+10\times\log\left[\frac{(1-(w_0+w_1N_0))(2+(p_0+p_1N_0)g)e^{-(p_0+p_1N_0)g}}{(p_0+p_1N_0)} + \frac{(w_0+w_1N_0)(2+(t_0+t_1N_0)g)e^{-(t_0+t_1N_0)g}}{(t_0+t_1N_0)}\right],$$

where each estimable parameter has been replaced by a linear function depending on masker level (N_0) with two coefficients to be estimated (the slope and intercept). Note the distinction made here between a *parameter* such as p or k whose single value determines the auditory filter shape, and a *coefficient* whose value is used in an equation to calculate the value of any particular filter parameter. Such an equation would correspond to a family of roex(p,w,t) filters whose shapes would be given by:

$$(1-(w_0+w_1N_0))(1+(p_0+p_1N_0)g)e^{-(p_0+p_1N_0)g} + (w_0+w_1N_0)(1+(t_0+t_1N_0)g)e^{-(t_0+t_1N_0)g}.$$

PolyFit allows the polynomial terms to be of any degree (although terms of a higher order cannot be included in the model without all the terms of lower order), for parameter values to be dependent upon the level of the masker (as above) or of the probe, and for parameters describing the high and low frequency sides of the filter to be specified independently. Thus instead of estimating coefficients for a parameter p, separate sets of coefficients are estimated for a lower p (p_l) and upper p (p_u), and similarly for w and t. The calculations also take into account the possibility of *off-place listening*, where a listener is supposed to attend to the filter which has the greatest signal-to-noise ratio at its output (Patterson and Moore, 1986). More generally, our implementation is a superset of the procedures described by Glasberg and Moore (1990). When analyzing a notched-noise experiment at a single fixed level, it is thus possible to select a model using PolyFit that will give identical fits to theirs.

In this way, it is possible to construct a single model, which, after appropriate coefficients are chosen, can predict thresholds for any combination of notch widths and level. While we have so far concentrated on whether the filter shape varies with N_0 or P_s, there is the separate, and somewhat unrelated, issue of whether P_s should be predicted from N_0 or vice versa when investigating the goodness of fit of a *given model* to the data. We have used the root-mean-square (rms) residual as our measure of goodness of fit. For example, we can take the value of N_0 from a given data point and use the model under investigation to predict P_s and then compare this predicted P_s to the empirical value given by the data point. Or we can, in the same way, predict N_0 given the value of P_s. Because the mathematical formula for predicting thresholds can be written as $P_s = N_0 + (\textit{other terms})$, or conversely as $N_0 = P_s - (\textit{other terms})$, predicting probe thresholds (P_s) from the masker spectrum levels (N_0), as in the first example, or masker levels from tone levels, as in the

TABLE I. Summary data concerning the nine listeners who participated in the study, including the range of conditions over which they participated. Maximum notch widths are shown separately for symmetric and asymmetric notches. Absolute thresholds were obtained using standard pure-tone audiometry.

Listener	Age	Absolute threshold at 2 kHz (dB HL)	Range of levels P_s	Range of levels N_0	Number of notch widths/level	Maximum notch width	Number of conditions	Number of thresholds
AMD	36	10	30–70	20–60	16	[0.5, 0.5], [0.4,0.6]	159	339
CT	21	0	40–70	30–60	14	[0.5, 0.5], [0.4, 0.6]	112	349
JD	25	0	30–70	20–60	16	[0.5, 0.5], [0.4, 0.6]	158	330
LS	26	10	40–60	30–50	13	[0.4, 0.4], [0.3, 0.5]	78	182
MB	35	20	40–60	30–50	13	[0.4, 0.4], [0.3, 0.5]	78	204
RC	21	−5	40–70	30–60	14	[0.5, 0.5], [0.4, 0.6]	110	291
RJB	30	0	30–70	20–60	16	[0.5, 0.5], [0.4, 0.6]	159	491
SK	21	−5	40–60	30–50	13	[0.4, 0.4], [0.3, 0.5]	78	185
WC	20	10	30–70	20–50	16	[0.5, 0.5], [0.4, 0.6]	143	354

second, leads to identical residuals. Therefore it does not matter which way around the prediction is made. In fact, because we use an additive power term to include the effects of absolute threshold (see below), it is intuitively more appealing to predict probe levels, even when the probe level itself is fixed in the experiment.[3]

Fitting the model to the data proceeds in essentially the same way as for the single level case. Instead of searching the space of the parameter values directly for the best-fitting values, the space of the coefficients making up the polynomials is searched.[4] Note that this fitting technique is quite different to the typical procedure used for studies across level, in which filter parameters are estimated separately for each of a number of masker levels, and polynomial fits are made to the values of the estimated parameters as a function of level (as in Moore and Glasberg, 1987). We have, for example, fitted as many as 158 mean data points in a single analysis, instead of doing ten separate analyses (five conditions in which the probe level is fixed, and five in which the masker level is fixed). Thus results from the fixed-probe and fixed-masker experiments can be analyzed together, putting stronger constraints on the way filter parameters can change with level.

A primary advantage of the PolyFit procedure is one of stability of the model fit, because such a large number of conditions can be fit by relatively few free coefficients. For example, in fitting one set of fixed-masker data (involving 13–16 data points) with a roex(p,w,t) filter shape in which p, w, and t are independently estimated on the upper and lower frequency side, seven free parameters must be estimated (including k). To describe the filter shape as a function of level requires multiplying the number of parameters by, for example, five levels of N_0 and five levels of P_s (70 parameters for 160 data points). Using PolyFit to fit the same data with a roex(p,w,t) model in which each of the seven parameters is allowed to take on the form of a linear function of N_0 (or P_s) results in a model with 14 coefficients (7×2) for the same number of data points. This large reduction in the number of free variables allows the stable fitting of complex models such as the roex(p,w,t) shape, something which was not possible using the ''standard'' techniques (Patterson *et al.*, 1982). PolyFit also provides a principled way of comparing results across fixed-masker and fixed-probe conditions. Individual fits leave these results incommensurable.

One further extension to typical roex fitting procedures was implemented. Because absolute thresholds can place a lower limit on masked thresholds in some conditions (for fixed noise maskers at low levels with wide notches), a single estimated parameter corresponding to the absolute threshold of the probe was incorporated. This was done by adding the value of the estimated threshold (in power terms) to the probe level predicted by the model for any particular condition. For probe levels more than 10 and 20 dB above the estimated absolute threshold, this term changes the initially predicted value by less than 0.5 and 0.05 dB, respectively.

Rosen and Baker (1994) report an application of PolyFit using a roex(p,r) shape [equivalent to a roex(p,w,t) shape with $t=0$] to a set of data obtained from two normal-hearing listeners. They showed that models which had filter parameters depending upon the level of the probe were considerably more successful than models in which filter parameters depended upon the level of the masker. They therefore argued that probe level should be fixed for simple measures of auditory filter shape at one level. More importantly, auditory filter shape appeared to be controlled by probe level, or something closely related to it. Here we extend those findings in a number of ways: (1) applying them to measurements made over a considerably wider range of notch widths and stimulus levels; (2) using the more complex roex(p,w,t) shape in place of the roex(p,r); (3) analyzing both group and individual results in a total of nine normally hearing listeners; (4) developing a technique for normalizing filter shapes so as to estimate changes in gain as well as shape across level, thus making our shapes comparable to direct measurements of basilar membrane vibration.

I. METHODS

A. Listeners

Table I presents summary data concerning the nine listeners who participated in the study. All had hearing thresholds within normal limits ($\leqslant$20 dB HL). All testing was

done monaurally, and in a single ear per listener. Mean results averaged over listeners LS and MB have appeared in a previous report (Rosen and Baker, 1994).

B. Threshold estimation

Masked thresholds were determined for sinusoidal probe tones of 2 kHz in the presence of notched-noise maskers with variable notch widths. The notches were placed both symmetrically and asymmetrically about the probe frequency and either the probe level or the noise level could be varied to determine the thresholds. A two-interval, two-alternative forced-choice paradigm with feedback was used to estimate the 79% point on the psychometric function. Listeners responded on a button box, with illuminated buttons indicating presentation intervals and providing feedback. From a starting level at which the probe was clearly audible, the varying sound, either probe or masker, was initially changed in 5 dB steps, with step-size decreasing by 1 dB after each turnaround. Once the step-size reached 2 dB, it remained constant for a further eight turnarounds, the mean of which was taken as the threshold. For each particular combination of notch-width and fixed probe or fixed masker level, two thresholds per listener were typically obtained. Threshold measurements where the standard deviation of the last eight turnarounds exceeded 3 dB were rejected and the measurement repeated. Also, where two measurements of the same condition in the same listener differed by more than 3 dB, a further measurement was taken and the average of all measurements used.

C. Stimulus configurations

The outside edges of the masker noise were fixed at $\pm 0.8 \times f_0$ (400 and 3600 Hz). A maximum of 16 different notch conditions were used, six symmetric and ten asymmetric. The frequencies of the edges of the notch are specified in normalized frequency units relative to the probe frequency (f_0) as given by $(|f - f_0|)/f_0$. In the symmetric conditions, both notch edges were placed at normalized values of 0.0, 0.1, 0.2, 0.3, 0.4, and 0.5. In the asymmetric condition one of the notch edges was set at a normalized value of 0.0, 0.1, 0.2, 0.3, and 0.4, while the other was set to 0.2 normalized units further away (0.2, 0.3, 0.4, 0.5, and 0.6). When the masker level was fixed, a subset of noise-spectrum levels (N_0) was chosen, ranging from 20 to 60 dB re: 10^{-12} W/Hz (the units hereafter referred to as SPL/Hz) in 10-dB steps. When the probe level was fixed, a subset of probe levels (P_s) was chosen, ranging from 30 to 70 dB SPL, again in 10-dB steps. Table I details the experimental conditions for each of the listeners.

D. Stimulus generation

All the stimuli were computer generated at a sampling frequency of 20 kHz. The time waveform of the probe was calculated independently of the masker and consisted of a steady state portion of 360 ms plus 20-ms raised-cosine onsets and offsets. The probe was temporally centered within the masker which consisted of a 460-ms steady-state portion with 20-ms raised-cosine-squared[5] onset and offsets. To generate the masker, the desired frequency spectrum was defined by setting all the spectral components (spaced at intervals of 0.61 Hz) within the appropriate frequency limits to have equal amplitudes while those outside were set to zero. Nonzero components had their phases randomized uniformly in the range of $0-2\pi$ radians. An inverse FFT was then applied to generate the time waveform. At the start of each threshold determination, a 3.2768-s buffer of noise was generated for use during that test. On each trial, a 500-ms portion of the buffer was chosen randomly for each of the two masker intervals within each trial.

The probe and masker were played out through separate channels of a stereo 16-bit D-A converter (PCLX™, Laryngograph Ltd.) and attenuated independently under computer control before being electrically mixed (PA4 and SM3 from Tucker-Davis Technologies). The signal was then sent via a balanced line to a final amplifier in a sound-treated room where it was presented monaurally to the right ear via Etymotic ER2 insert earphones. Calibrations were done using a B&K 4157 ear simulator [conforming to IEC 711 and ANSI S3.25/1979 (ASA 39/179)] with a B&K DB 2012 ear canal extension.

Because the experiments took place over a number of years, two distinct sets of apparatus were used. These differed in detail, but not in essentials. The later experiments (involving listeners AMD, JD, RJB, and WC) were run as described above. The earlier experiments (involving listeners CT, LS, MB, RC, and SK) had the following differences: (1) The 2-kHz sinusoidal probe was hardware generated, and gated by multiplying it with a computer-generated envelope with 10-ms raised-cosine onsets and offsets, and a steady-state portion of 380 ms. The probe was temporally centered within the masker, which consisted of a noise burst with the same gating envelope, but a 480-ms steady state. (2) A 500-ms portion of the buffer was chosen randomly for the masker burst on each trial, but the same masker burst was used for the two intervals of the trial. (3) The masker bursts were output from 12-bit D-A converters, simultaneously with the appropriate gating envelope for the probes. Probe and masker bursts were controlled independently in level by two digitally controlled attenuators (Charybdis), at the output of which they were mixed. We have no reason to suspect that any of these differences affect the results to any significant degree.

E. Analyses

All analyses were performed on means. When averaging over listeners, the contribution of each listener to the mean was kept equal by taking means within a listener before averaging across listeners. If there was no data for a particular condition for any one of the listeners in the mean, that condition was excised. In particular, it was occasionally not possible to present the masker at a sufficiently high level to mask the probe for the widest notch widths and higher probe levels.

Most of the analyses were done individually for each listener. However, in order to simplify the presentation, the first, and most extensive analyses concern the mean of the three listeners who experienced the greatest range of levels

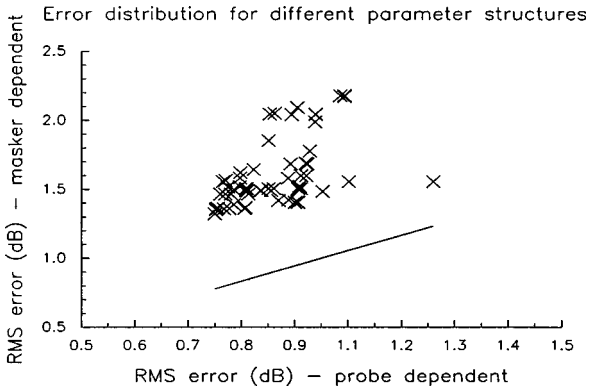


FIG. 1. A comparison of the goodness-of-fit of PolyFit models of identical structure, one of which has filter parameters depending upon probe level (abscissa), and one of which has filter parameters depending upon masker level. The goodness-of-fit measure is the root-mean-square residuals; hence smaller numbers indicate better fits. Only parameter structures resulting in models with a goodness-of-fit less than 2.5 dB are included (a total of 63 comparisons). The solid line indicates equal goodness-of-fit for the two possible dependences.

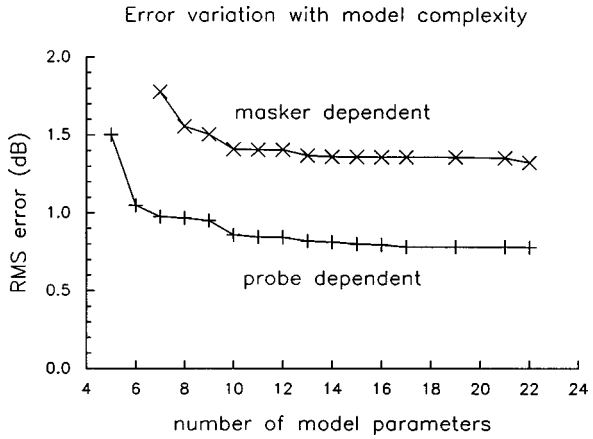


FIG. 2. Summary measures of the goodness-of-fit of PolyFit models which fit the data best with a given number of coefficients, for models in which filter parameters depend upon the level of the probe, and of the masker.

and notch widths (AMD, JD, and RJB). Comparisons are also made with the mean results of LS and MB, as used by Rosen and Baker (1994). Because listeners do vary somewhat in their performance, we only took means across listeners with a common set of conditions. Otherwise, we felt it would have been necessary to implement some kind of weighting to ensure listeners' responses were given equal weight across *all* conditions.

A variety of models were fitted to each data set, using Polyfit. All of the models were variants of the asymmetric roex(p,w,t) model. These included simplified models in which, for example, the upper half of the filter was described with a roex(p) shape whereas the lower half was a complete roex(p,w,t) shape. It is also necessary to estimate k, the signal-to-noise ratio necessary for detection at the output of the filter. All of these parameters can be arbitrary polynomial functions of the level of the masker or the probe, but we have never investigated models with more than a quadratic dependence on level (that is, three coefficients per parameter to be estimated). Finally, we also used Polyfit's ability to estimate an absolute threshold that is never allowed to vary with level (described above).

II. RESULTS AND DISCUSSION

A. Normal listeners: Mean of 3

This data set consists of mean thresholds obtained in 158 distinct conditions (as described above, excluding the [0.6, 0.4] notch at fixed probe levels of 60 and 70 dB SPL), based on a total of 1154 separate thresholds from the three listeners AMD, JD, and RJB.

We fitted 73 models, which differed in parameter structure, assuming filter parameters to vary with probe level or masker level (a total of 146 distinct models). Figure 1 shows that, for models that depend on level, probe-dependent models *always* fit the data better than masker-dependent models. Although the fits obtained from the masker-dependent models might be considered adequate in other circumstances, the rms residual is typically 70% larger than that obtained from the corresponding probe-dependent model (a statistic based on the median ratio of rms residuals).

Another way to demonstrate this is to compare the fits from the best-fitting probe- and masker-dependent models for fixed numbers of coefficients. Figure 2 shows, again, the consistently better fit obtained by making filter parameters depend upon probe rather than masker level.

Although it is quite clear that probe-dependent models give better fits than masker-dependent models, it is much more difficult to choose a particular probe-dependent model out of the many possible. Clearly, for a fixed number of coefficients, one would typically choose the best-fitting model. The difficult issue is the choice of the number of coefficients, and how this trades off against the goodness-of-fit.

To aid us in this task, we have adopted a heuristic approach based on those common in statistical model building (e.g., see Aitkin *et al.*, 1989). Starting off with a model with more coefficients than we think are necessary, we then determine which can be eliminated by looking at changes in the goodness-of-fit as they are excised. Unfortunately, for the type of nonlinear model employed here, there is no statistical theory which can assess the statistical significance of any given numerical change in the rms residual. Often, however, excising some coefficients hardly changes the rms residual, whereas excising others changes the fit of the model dramatically. In our earlier publication (Rosen and Baker, 1994), we had the good fortune for this to be true consistently. Here the results are not as clear cut, but as we shall see, models with similar goodness-of-fit lead to filter shapes that are very similar. Therefore it is not particularly important which model is chosen from the ''better-fitting'' ones. The relatively large number of good-fitting filter shapes is also an indication that the roex(p,w,t) shape may be *too* flexible. There are likely to be other adequate functional forms with fewer controlling parameters (e.g., Irino and Patterson, 1997; Lyon, 1996). For the moment, however, we concentrate solely on the roex family, as they are the only filter shapes which have been shown to be adequate for dealing with a wide variety of results.

First note that a quadratic dependence of all parameters

TABLE II. Goodness-of-fit for the best-fitting model containing 5–22 estimated coefficients. All are probe dependent models. The first seven columns contain the number of polynomial coefficients used for each filter parameter. The last column indicates the % change in the goodness-of-fit measure between the model given by that row, and the model on the row immediately above.

p_l	p_u	k	w_l	w_u	t_l	t_u	Number of coefficients	rms residual ($n=158$) (dB)	% increase
3	3	3	3	3	3	3	22	0.778	
3	3	2	3	3	3	3	21	0.780	0.3
3	3	2	3	3	2	2	19	0.780	0.0
3	3	2	2	2	2	2	17	0.780	0.0
3	2	1	3	2	2	2	16	0.794	1.8
3	2	1	2	2	2	2	15	0.808	0.7
2	2	1	2	2	2	2	14	0.812	1.5
2	2	1	2	1	2	2	13	0.818	0.7
2	2	1	2	1	1	2	12	0.842	3.0
2	3	1	2	x	2	x	11	0.845	0.4
1	3	1	2	x	2	x	10	0.859	1.6
1	3	1	2	x	1	x	9	0.950	10.6
1	1	2	2	x	1	x	8	0.968	1.8
1	0	2	2	x	1	x	7	0.976	0.8
1	0	1	2	x	1	x	6	1.049	7.5
1	0	1	2	x			5	1.501	43.2
1	0	1	1	x			4	5.109	240.3

on probe level leads to little improvement on the fits obtainable from a linear dependence. The quadratic model, with 22 coefficients to estimate (three for each of the six filter parameters and k, plus one for absolute threshold) leads to an rms residual of 0.778 dB, whereas the linear model, with 15 parameters, leads to a residual of 0.808 dB. Thus a loss of seven coefficients worsens the goodness-of-fit by only 0.03 dB.

Table II shows the best-fitting models for a particular number of estimated coefficients. Models are generally described by a letter (p or m indicating *p*robe or *m*asker level dependence) followed by a string of seven digits indicating the number of coefficients used for each of the parameters p_l, p_u, k, w_l, w_u, t_l, and t_u, respectively. An ''x'' indicates a parameter that is not needed, for example, when only a roex(p) shape is needed on the high-frequency side of the filter. Thus $p2212x1x$ indicates a model in which both lower and upper p depend upon probe level in a linear way, as does w_l, with a simple roex(p) shape on the high-frequency side. All other parameters are invariant across level. A ''0,'' used only for describing filter parameters for the upper half of the filter, indicates that the filter shape has identical values on its high- and low-frequency sides for that particular parameter. Thus $p1012x1x$ indicates a model with a simple roex(p) shape on its high-frequency side, but symmetric in its passband. Only one filter parameter varies with level (w_l).

Note that it is possible to account reasonably well for the data with as few as 6 coefficients. Of these, two concern absolute threshold and detection efficiency (k), while only four describe filter shape (assuming symmetry in the passband, by setting $p_u=p_l$, did not harm the fit and allowed the loss of one parameter). Only one filter parameter need depend on level (w_l). Also of note is the fact that k can generally be assumed to be invariant across level without affecting the fits obtained. This empirical fact supports a simpler interpretation of our results than would otherwise be the case (see the discussion below).

Figure 3 shows the filter shapes, as a function of level, derived from four of the models described in Table II. Although the goodness-of-fit varies across these four models by about 35%, there is little change evident on the low-frequency sides of the filter. Even the high-frequency sides (which are known to be somewhat difficult to pin down in notched-noise experiments as thresholds are largely determined by the shallower lower side of the filters, Glasberg and Moore, 1990) differ little unless they are not permitted to change with level. So, although there is some uncertainty in choosing a particular model, many of the models which fit the data reasonably well lead to similar conclusions about filter shapes.

We have chosen to focus on $p1312x2x$, as a further reduction in the number of coefficients leads to a relatively large increase (in terms of proportions) in the rms residual, considerably larger than reductions in models with more coefficients. This model appears to be a good compromise between the number of parameters used and the goodness-of-fit, but other choices would lead to conclusions that are little different. The quadratic term for p_u allows the upper slope of the filter to remain relatively constant for low levels, and then to become shallower with increasing level once the probe level reaches about 50 dB SPL. Models in which p_u is constant, or a linear function of level, lead to odd behavior in the upper tails of the filter. In particular, the tail starts just a few dB down from the peak. Clearly, the roex(p,w,t) was not designed with such a possibility in mind, so it seems far preferable to allow a quadratic dependence in one parameter.

That the models do fit the data quite well can be appreciated from Fig. 4, which shows the entire data set, plotted as growth-of-masking functions, along with the predictions from our chosen model. (Note that the ordinate here is signal-to-noise ratio as opposed to the probe level more usu-

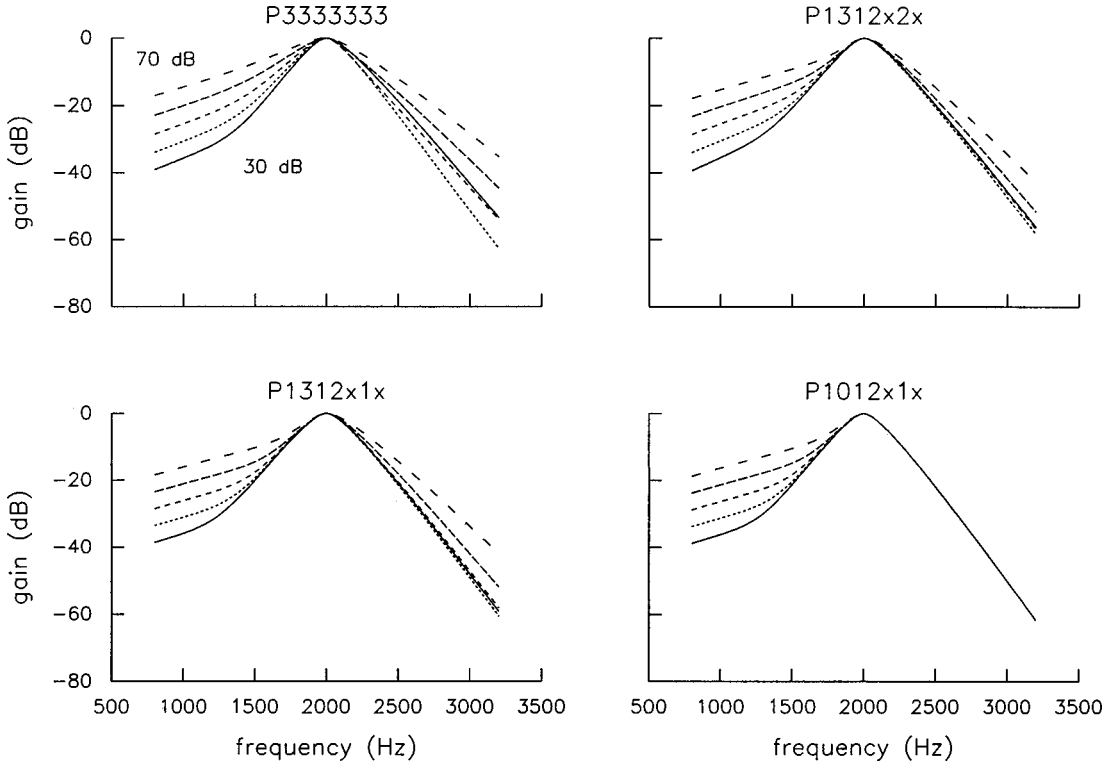


FIG. 3. Filter shapes for four different probe-dependent models. Each plot shows the filter shapes calculated for probe levels of 30–70 dB SPL in 10-dB steps.

ally plotted in such curves. The advantage of this measure is that departures from a linear growth of probe level with masker level, indicated by a horizontal line, are more easily assessed visually.) All predictions are within $-2.8/+2.0$ dB of the obtained data points and the rms residual is only 0.859 dB. Contrast this with the predictions from the best masker-dependent model with the same number of estimated coefficients in Fig. 5 ($m2222x1x$). Although in many situations this would be considered quite a good fit, note the many regions in which the data are consistently predicted poorly. The predictions are only within $-5.7/+3.3$ dB of the obtained data points with the rms residual being 1.41 dB, some 64% worse.

There is another way of plotting filter shapes which also points to probe-dependent models being superior. The filter shapes shown above are all normalized to have unity gain at their peak, as a consequence of the assumptions involved in fitting roex filter shapes. Yet we know from direct measurements of basilar membrane vibration that peak gain varies monotonically with level, being greatest at lowest levels (Ruggero *et al.*, 1992; Ruggero *et al.*, 1997). Such measurements also show basilar membrane response to be linear for frequencies sufficiently below the best frequency of the place on the membrane being investigated, as shown in Fig. 6. While Fig. 6 shows basilar membrane transfer functions for an isointensity stimulus paradigm, the same general properties (variation in peak gain and linearity below CF) can be seen in isoresponse measurements of basilar membrane motion (Ruggero *et al.*, 1997). Note too the compressive response at the peak of the filter, with a change in gain of about 20 dB for a 40 dB input range.

Working on the hypothesis that our behavioral results

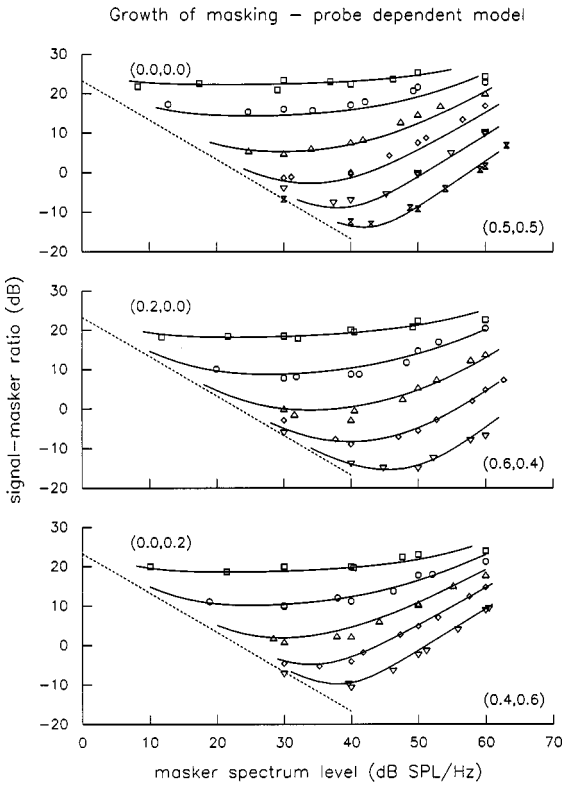


FIG. 4. Masked thresholds expressed as signal-to-noise ratios (probe level in dB SPL minus the masker level in dB SPL/Hz) as a function of masker level. Such curves are known as growth-of-masking functions (although more typically with the level of the probe as the ordinate). The results from symmetric notches and the two types of asymmetric notches are shown in separate graphs. Each symbol indicates a particular pair of notches. The lines are predictions from a model which assumes filter parameters to depend upon probe level. The diagonal line at left indicates absolute threshold.

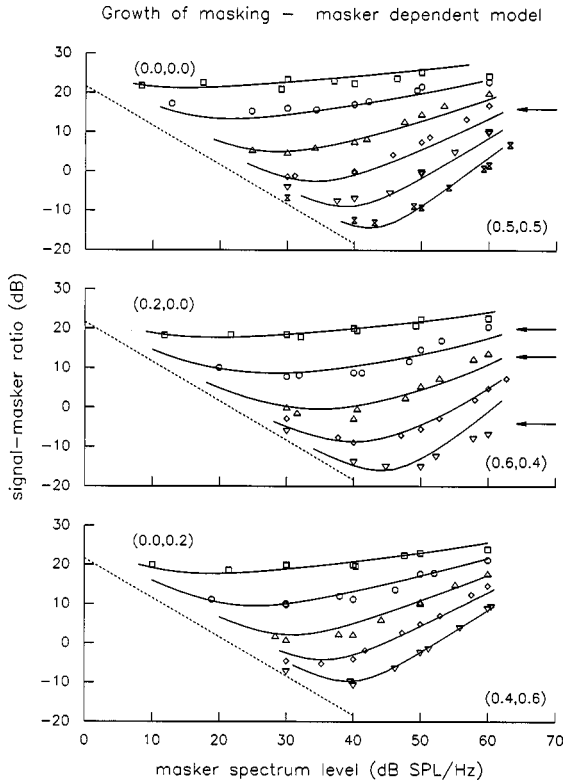


FIG. 5. As for Fig. 4, but with predictions from a masker-level-dependent model. Arrows indicate regions in which the fit of the model is consistently poor.

reflect basilar membrane filtering in a fairly direct manner, we assume that the auditory filter is linear a little more than one octave below its characteristic frequency, thus tacking together the shapes at this point. The resulting curves are highly reminiscent of filtering functions measured on the basilar membrane. In particular the filter sharpness and peak filter gain both increase with decreasing level (Fig. 7), and there is a tendency for the filters to become linear again at frequencies high above CF. The 2:1 compression ratio at the peak is also similar to that evidenced on the basilar membrane, although not too much should be made of this. Other basilar membrane studies indicate greater compression ratios (e.g., Nuttall and Dolan, 1996), but against this must be set the probability that compression ratios vary significantly

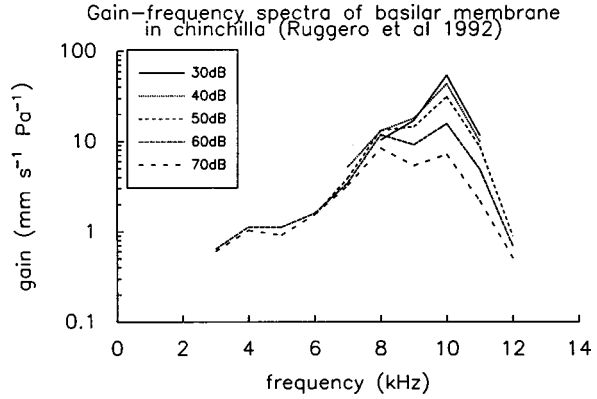


FIG. 6. The frequency response (isointensity) of a single place on the basilar membrane as a function of level (redrawn from Ruggero *et al.*, 1992).

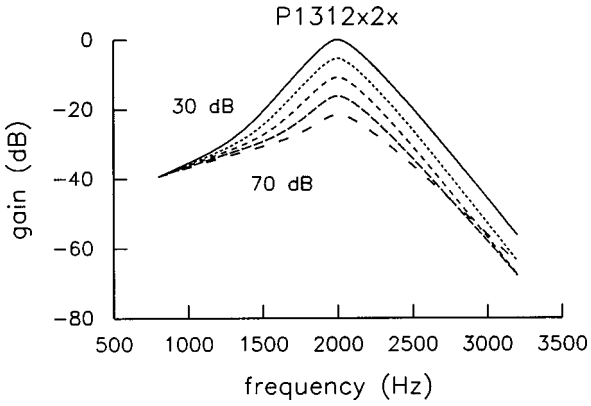


FIG. 7. Filter shapes for a $p1312x2x$ probe-dependent model, calculated for probe levels of 30–70 dB SPL in 10-dB steps, and normalized to have equal gain a little more than one octave below their center frequency.

with frequency, decreasing apically (i.e., at lower frequencies, Cooper and Yates, 1994; Wilson, 1992).

Note too that the filters plotted in Fig. 7 change shape right down to within 15 dB of absolute threshold. This is to say that the nonlinearity extends to levels as low as it is possible for us to measure. There is some controversy about this, with claims that the basilar membrane is linear even at its peak response for stimulus levels as high as 30–40 dB SPL. Recent measurements by Nuttall and Dolan (1996) support our view in showing that the response of the basilar membrane does indeed become linear at low enough levels, but only ‘‘ . . . for basilar membrane velocities below afferent neural thresholds based on discharge rates’’ (p. 1561).

The same manipulation leads to a much less orderly picture for the masker-dependent models. Figure 8 shows the filter shapes normalized in both ways. Although the filter does become shallower with level on its low-frequency side in the passband, the tail sharpness appears to *increase* with increasing level. Also, filter sharpness does not turn out to be linked to peak filter gain in the way suggested by basilar membrane experiments.

We have also found that the filter shapes obtained from masker-dependent models change much more with changes in parameter structure, as can be seen in Fig. 9. In the $m2212222$ model, the change in gain is even greater than the change in input level. Therefore according to this model, output level would *decrease* as input level increases.

In short, there are strong reasons to prefer models of auditory filtering which make filter parameters depend upon probe level rather than masker level: (1) The probe-dependent models predict the data more accurately over the entire stimulus space, with the root-mean-squared residuals 40%–70% larger for masker-dependent models; (2) Probe-dependent models lead to filter shapes much more in keeping with physiological measures. (3) Filter shapes derived from masker-dependent models change greatly with small changes in the parameter structure assumed, while filter shapes derived from probe-dependent models change little even with large changes in the assumed parameter structure.

B. Normal listeners: Mean of 2

Rosen and Baker (1994) reported similar analyses to those above on a set of mean data from two different listen-

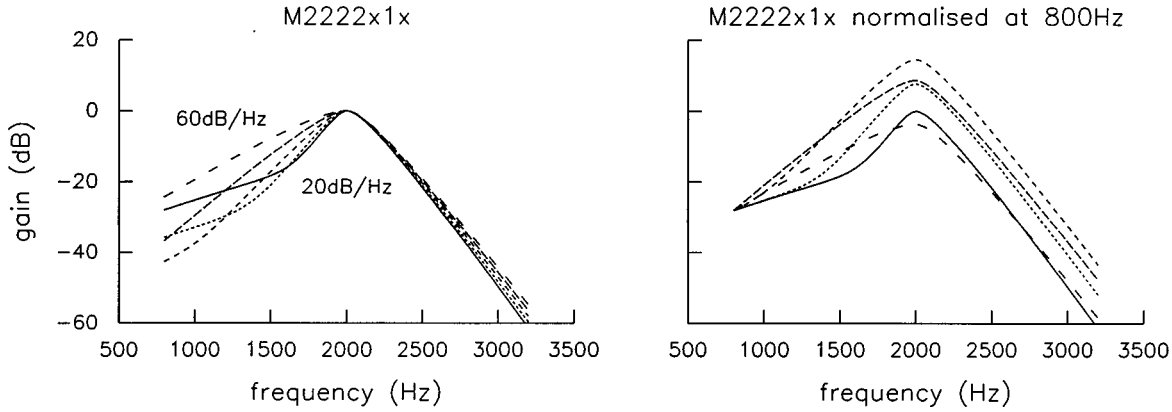


FIG. 8. Filter shapes for a *m*2222*x*1*x* masker-dependent model, calculated for masker levels of 20–60 dB SPL in 10-dB steps, normalized both to have equal gain at their tips, and about an octave below their center frequency.

ers. It is interesting to compare the filter shapes from the two sets of mean data given the differences in listener, apparatus, and particular conditions tested. Figure 10 (corresponding to Fig. 7) shows the filter shapes derived from the same probe-dependent model focused on above. It is clear that the two sets of data lead to highly similar derived filter shapes across level, in spite of all the differences between them.

C. Normal listeners: Individual results

Similar analyses were performed for the individual data sets obtained from each of the listeners. Of some interest is the relationship between the absolute thresholds obtained with standard audiometery (Table I) and the threshold estimated from the fits. Comfortingly, the estimated absolute threshold varied little with the parameter structure of the fitted model (e.g., at most 1.5 dB for the four probe-dependent models specified in Table III). Also, there was typically a close relationship between the two absolute thresholds (within 5.5 dB for six of the nine listeners). Only one listener (SK) had a large discrepancy (>8 dB), with the threshold estimated from the fits about 23 dB greater than her threshold measured audiometrically. Due to the fact that this listener did not participate in the lowest level fixed-masker condition, and that her signal detection efficiency (as measured by k) was the highest of any listener ($+7.5$ dB *re*: the mean level of the other eight normal listeners), no probe level even approached her actual absolute threshold. As it is thus clear

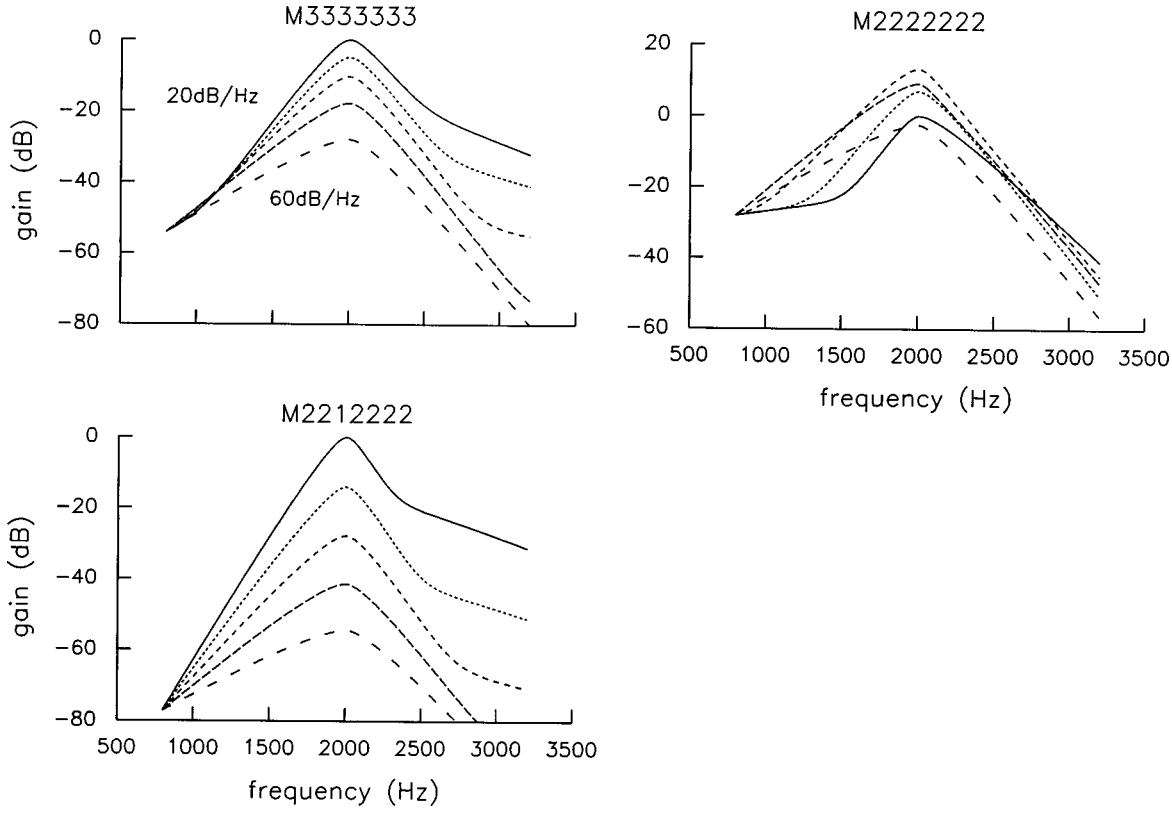


FIG. 9. Normalized filter shapes for three different masker-dependent models. Each plot shows the filter shapes calculated for masker levels of 20–60 dB SPL in 10-dB steps. Note the strong variation in shapes for different parameter structures.

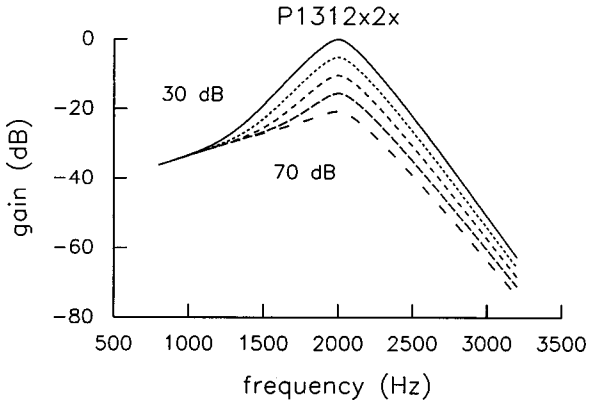


FIG. 10. Filter shapes from the data obtained by two normal listeners in Rosen and Baker (1994). The curves show a $p1312x2x$ probe-dependent model, calculated for probe levels of 30–70 dB SPL in 10-dB steps, and normalized to have equal gain at about one octave below their center frequency. Compare this to Fig. 7.

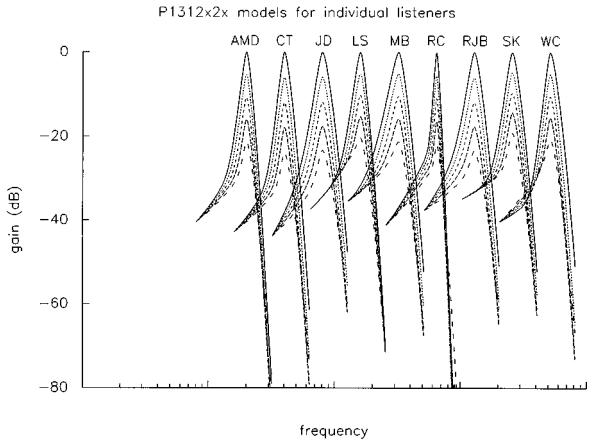


FIG. 11. Filter shapes for nine normal-hearing listeners using a $p1312x2x$ probe-dependent model, calculated for probe levels of 30–70 dB SPL in 10-dB steps, and normalized to have equal gain at just less than one octave below their center frequency. The curves have been shifted along the logarithmic frequency axis for clarity.

that the masked thresholds could not possibly constrain an estimate of her absolute threshold, no fits for SK incorporated an absolute threshold term.

Table III shows the goodness-of-fit measures for a selection of models for each individual data set. These measures vary significantly across listeners for at least two possible reasons. First, listeners differ in their inherent consistency, and we would expect more consistent listeners to have better fits. Second, given that there is some individual variation in filter shape, it may be that the roex(p,w,t) shape just happens to fit some listeners better than others.

Of more interest, in any case, is the pattern of results across conditions within an individual listener. Note first that probe-dependent models always fit the data better than masker-dependent models, again with the rms residual typically more than 50% greater for the masker-dependent models. In fact, we have never found a case in which a masker-dependent model fits the data better than a probe-dependent model of the same structure (in literally hundreds of comparisons). Second, for the probe-dependent model in which all parameters vary linearly with level, there is typically little loss of predictive power in assuming k (signal detector efficiency) to be constant across level.

Figure 11 shows the individual normalized filter shapes arising from the $p1312x2x$ model we used above. Again, however, there are a wide variety of parameter structures which lead to similar filter shapes. Clearly, individual listeners differ in their frequency selectivity, but the essential pattern is strikingly uniform.

In fact, for a number of listeners, there were ten parameter models that led to fits considerably better than the $p1312x2x$ used here. This was investigated further by finding the best ten-parameter model for each listener. If this model led to a goodness-of-fit that was more than 15% better than that from $p1312x2x$ (in terms of the sum of the squared residuals), a new model was chosen for that listener. (All the other listeners had improvements of fit ranging from 0% to 6.5%). One other constraint influenced the selection of the model. No model with a varying k was used, because the evidence from the Mean-of-3 data set so strongly supported the notion of a constant k. In those cases, another model was chosen, either one with fewer parameters, or with one additional parameter, depending upon the goodness of fit in each case. In all the listeners for whom a new model was chosen, the fit was improved by allowing p_l to vary with level, as shown in Table IV. Figure 12 shows the new set of filter shapes derived from the models. As compared to Fig. 11, the main change is seen in listener RC, whose filters were considerably narrower than those of the other listeners. This appears to arise from a poorly fitting model, and indeed, it was

TABLE III. Individual measures of goodness-of-fit from a number of Poly Fit models. Also shown are results from the Mean-of-3 and Mean-of-2 data sets described in the text.

Listener	$p2222222$	$p2212222$	$p1312x2x$	$p2222x1x$	$m2222222$	$m2212222$	$m1312x2x$	$m2222x1x$
Mean of 3	0.81	0.81	0.86	0.95	1.36	1.47	1.64	1.41
Mean of 2	1.17	1.17	1.24	1.19	1.65	1.65	1.81	1.72
AMD	1.12	1.13	1.16	1.19	1.97	2.02	1.97	2.09
CT	1.02	1.02	1.19	1.09	1.95	2.20	2.51	2.07
JD	1.31	1.33	1.38	1.45	1.79	1.87	2.02	1.83
LS	1.31	1.31	1.46	1.32	1.97	1.97	2.15	2.04
MB	1.30	1.30	1.31	1.30	1.68	1.72	1.74	1.72
RC	0.86	0.89	1.33	0.96	1.76	2.47	2.59	2.10
RJB	0.79	0.81	0.97	0.84	1.44	1.72	1.97	1.55
SK	1.40	1.41	1.49	1.48	2.01	2.04	2.10	2.01
WC	1.08	1.11	1.30	1.15	1.55	1.71	1.99	1.75

TABLE IV. Models chosen for listeners whose results were not fit adequately by the $p1312x2x$ model.

Listener	Chosen model	Number of parameters	rms residual	rms residual for $p1312x2x$
CT	$p2312x2x$	11	1.04	1.19
LS	$p2112x1x$	8	1.36	1.45
RC	$p2112011$	9	1.00	1.33
RJB	$p2312x1x$	10	0.82	0.97
WC	$p2312x2x$	11	1.20	1.30

the fit to RC's results that was improved the most by selecting a different model than our standard $p1312x2x$.

Contrast the filter shapes based on probe-dependent models with those that arise from the $m2222x1x$ masker-dependent model discussed above (Fig. 13). These are much more variable from listener to listener, both in the degree of compression (or even expansion) across level, and even in which level leads to the highest gain.

III. SUMMARY AND FINAL REMARKS

We have shown that it is possible to accurately account for the pattern of results across level in a notched-noise masking experiment with filter shape models that explicitly depend upon probe level. Such models have four advantages over models in which filter parameters depend upon the level of the masker: (1) They always fit the data better; (2) Filter shapes change relatively little with changes in the number of coefficients controlling each parameter, and in parameter structure; (3) Different listeners give results that are much more similar; (4) Appropriate normalization of tail gains lead to filter shapes that are highly reminiscent of vibration patterns observed directly on the basilar membrane. Therefore notched-noise measurements that are made only at one level should be performed with a fixed probe level. Fixing the masker level leads to a derived filter shape that is some kind of average of a number of shapes, caused by the change in probe level as notches are varied. In general, such filter shapes will be too narrow, simply because the filter is becoming sharper as the probe level decreases with increasing notch width.

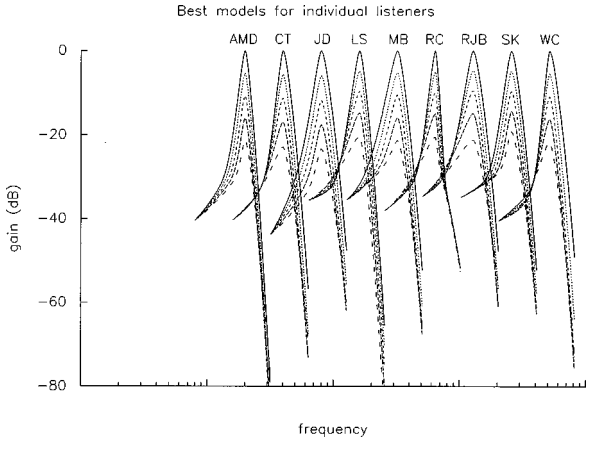


FIG. 12. Same as for Fig. 11, but with different filter shape models for the five listeners specified in Table IV.

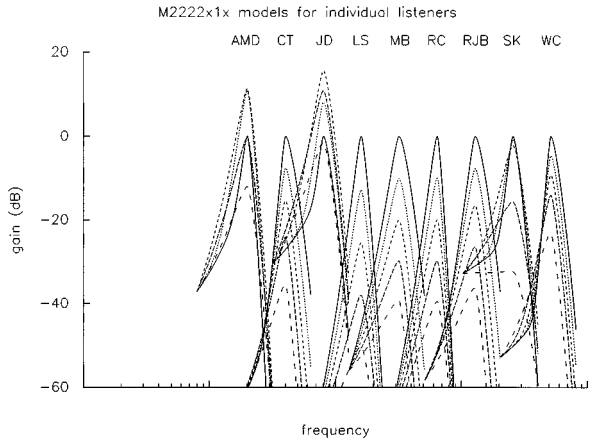


FIG. 13. Filter shapes for nine normal-hearing listeners using an $m2222x1x$ masker-dependent model, calculated for masker levels of 20–60 dB SPL in 10-dB steps, and normalized to have equal gain at just less than one octave below their center frequency (cf. Fig. 11).

Our technique is also likely to clarify the effects of hearing impairment on auditory nonlinearity, especially as it is currently believed that many of the concommitants of sensori-neural hearing loss can be attributed relatively straightforwardly to a loss of nonlinearity [see Moore (1995) for a review]. In particular, damage to the outer hair cells (where the ''motor'' for the nonlinearity is presumed to reside) would be expected to manifest itself in four ways relevant to the discussion here: (1) a loss of absolute sensitivity; (2) smaller or no changes in filter shape across level; (3) degraded selectivity at low levels; and (4) normal selectivity at sufficiently high levels.

Using the PolyFit procedure to analyze notched-noise masking data, we have found evidence for all of these four features in two hearing-impaired listeners with quite minor impairments (Rosen *et al.*, 1996). One had a hearing loss sloping from essentially normal at low frequencies to 35 dB HL at 8 kHz. Although she used hearing aids, her loss at 2 kHz was only 30 dB. Strikingly, her auditory filter shapes did not change at all with level. Filtering on the upper side of the filter was essentially normal, while that on the lower side could only be considered normal at relatively high levels. For low measurement levels, her auditory filter slopes were considerably shallower than normal on the low-frequency side. The other listener, with a subclinical impairment of 20 dB at 2 kHz attributed to frequent exposure to amplified music as a musician, also showed a reduced (but not absent) nonlinearity. Thus at least in these cases, hearing loss appears to consist of a loss of the nonlinear properties of the auditory periphery, most likely arising from damage primarily to the outer hair cells.

On the theoretical (and practical) side, there is still much to be done regarding the computational implementation of nonlinear filters, in order to produce a general-purpose nonlinear auditory filter bank. In terms of the distinction between ''input'' versus ''output'' control of filter shape (Lutfi and Patterson, 1984; Verschuure, 1981), our results clearly support the notion that the filter shape is controlled by its output level. This arises from the finding that k appears to be constant across level, so that fixing the probe level also fixes

the output level of the filter. On the other hand, it may well be that filter structures that are neither directly input- nor output-controlled lead to results consistent with our findings (e.g., the MBPNL model of Goldstein, 1988). Explicit models of auditory filtering will also be necessary for detailed comparisons of basilar membrane data with psychophysical results, preferably obtained in the same species.

ACKNOWLEDGMENTS

This work was supported by the Wellcome Trust (Grant Reference 046823/Z/96) and the Medical Research Council (Grant Reference G9225699N). Thanks to Sarah Kramer and Rosie Casson who collected some of the data reported here, and also to Alberto Recio who provided the data for Fig. 6.

[1]Such a model overcomes the problem associated with narrow- and wide-band maskers if the spectrum level of the masker is taken as the controlling factor. For example, a 1-Hz wide and a 500-Hz wide band of noise may have the same spectrum level, but their masking effect will be very different.

[2]The PolyFit program is available by anonymous ftp from ftp://pitch.phon.ucl.ac.uk/pub/polyfit/.

[3]Were we to make strong statistically based claims about the adequacy of the various models (which we cannot do in such a nonlinear model anyway), we would have to deal more explicitly with the fact that in one-half of the measurements it is the masker level that is fixed (i.e., known without sampling error), whereas in the other half, it is the probe level that is fixed. Here this distinction has little import.

[4]Best parameter values were found using the downhill simplex method (Nelder and Mead, 1965) as implemented by Press *et al.* (1988). In order to avoid the problem of local minima, a variety of strategies were used. First, every optimization actually consisted of two applications of the simplex procedure, in which the second was restarted from the best values found from the first. Second, many different models were fitted, most of which were supersets and/or subsets of other models. The goodness-of-fit was compared across models to ensure that a particular model fit the data at least as well as any model that was a subset of it. Optimizations were started from best fits obtained to supersets and subsets of the model in question, with the appropriate parameters deleted or added, respectively. Finally, the optimization procedure was run until no further improvement in a fit could be made, and with a minimum of ten restarts from the previously found best fit. Although no procedure can preclude local minima, we believe this combination of techniques makes it unlikely that a local minimum was retained.

[5]A 20-ms raised-cosine ramp was accidentally applied twice to the masker in the computer program. This is approximately equivalent to a single application of a 15-ms raised-cosine ramp.

Aitkin, M., Anderson, D., Francis, B., and Hinde, J. (**1989**). *Statistical Modelling in GLIM* (Clarendon, Oxford).

Carney, L. H. (**1993**). ''A model for the responses of low-frequency auditory-nerve fibers in cat,'' J. Acoust. Soc. Am. **93**, 401–417.

Cooper, N. P., and Yates, G. K. (**1994**). ''Nonlinear input-output functions derived from the responses of guinea-pig cochlear nerve fibres: variations with characteristic frequency,'' Hearing Res. **78**, 221–234.

Dubno, J. R., and Dirks, D. D. (**1989**). ''Auditory filter characteristics and consonant recognition for hearing-impaired listeners,'' J. Acoust. Soc. Am. **85**, 1666–1675.

Fidell, S., Horonjeff, R., Teffeteller, S., and Green, D. M. (**1983**). ''Effective masking bandwidths at low frequencies,'' J. Acoust. Soc. Am. **73**, 628–638.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hearing Res. **47**, 103–138.

Goldstein, J. L. (**1988**). ''Updating cochlear driven models of auditory perception: A new model for nonlinear auditory frequency analyzing filters,'' in *Working Models of Human Perception*, edited by B. A. G. Elsendoorn and H. Bouma (Academic, London).

Irino, T., and Patterson, R. D. (**1997**). ''A time-domain, level-dependent auditory filter: The gammachirp,'' J. Acoust. Soc. Am. **101**, 412–419.

Lutfi, R. A., and Patterson, R. (**1984**). ''On the growth of masking asymmetry with stimulus intensity,'' J. Acoust. Soc. Am. **76**, 739–745.

Lyon, R. F. (**1996**). ''The all-pole gammatone filter and auditory models,'' Paper presented at the Computational Models of Signal Processing in the Auditory System (Forum Acusticum '96), Antwerp, Belgium, 1–4 April 1996 (also available at http://www.pcmp.caltech.edu/anaprose/dick/Auditory/APGF/).

Moore, B. C. J. (**1995**). *Perceptual Consequences of Cochlear Damage* (Oxford U. P., Oxford).

Moore, B. C. J., and Glasberg, B. R. (**1987**). ''Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns,'' Hearing Res. **28**, 209–225.

Moore, B. C. J., Glasberg, B. R., and Baer, T. (**1997**). ''A model for the prediction of thresholds, loudness and partial loudness,'' J. Audio Eng. Soc. **45**, 224–240.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1990**). ''Auditory filter shapes at low center frequencies,'' J. Acoust. Soc. Am. **88**, 132–140.

Nelder, J. A., and Mead, R. (**1965**). ''A simplex method for function minimization,'' Comput. J. (UK) **7**, 308–313.

Niemiec, A. J., and Yost, W. A. (**1992**). ''Behavioral measures of frequency selectivity in the chinchilla,'' J. Acoust. Soc. Am. **92**, 2636–2649.

Nuttall, A. L., and Dolan, D. F. (**1996**). ''Steady-state sinusoidal velocity responses of the basilar membrane in guinea pig,'' J. Acoust. Soc. Am. **99**, 1556–1565.

Patterson, R. D. (**1976**). ''Auditory filter shapes derived with noise stimuli,'' J. Acoust. Soc. Am. **59**, 640–654.

Patterson, R. D., and Moore, B. C. J. (**1986**). ''Auditory filters and excitation patterns as representations of frequency resolution,'' in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London).

Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (**1982**). ''The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold,'' J. Acoust. Soc. Am. **72**, 1788–1803.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (**1988**). *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge U. P., Cambridge, England).

Rosen, S. (**1989**). ''Deriving auditory filter characteristics from notched-noise masking data: Modified derivations,'' Speech, Hearing and Language: Work in Progress **3**, 189–204.

Rosen, S., and Baker, R. J. (**1994**). ''Characterising auditory filter nonlinearity,'' Hearing Res. **73**, 231–243.

Rosen, S., Baker, R. J., and Darling, A. M. (**1996**). ''Auditory filter nonlinearity at 2 kHz in normal and impaired hearing,'' Speech, Hearing and Language: Work in Progress **9**, 161–186.

Rosen, S., and Fourcin, A. J. (**1986**). ''Frequency selectivity and the perception of speech,'' in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London).

Ruggero, M. A., Rich, N. C., and Recio, A. (**1992**). ''Basilar membrane responses to clicks,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford, UK).

Ruggero, M. A., Rich, N. C., Recio, N. C., Narayan, S., and Robles, L. (**1997**). ''Basilar-membrane responses to tones at the base of the chinchilla cochlea,'' J. Acoust. Soc. Am. **101**, 2151–2163.

Shailer, M. J., Moore, B. C. J., Glasberg, B. R., Watson, N., and Harris, S. (**1990**). ''Auditory filter shapes at 8 and 10 kHz,'' J. Acoust. Soc. Am. **88**, 141–148.

Verschuure, J. (**1981**). ''Pulsation patterns and nonlinearity of auditory tuning II. Analysis of psychophysical results,'' Acustica **49**, 296–306.

Weber, D. L. (**1977**). ''Growth of masking and the auditory filter,'' J. Acoust. Soc. Am. **62**, 424–429.

West, P. D. B., and Evans, E. F. (**1990**). ''Early detection of hearing damage in young listeners resulting from exposure to amplified music,'' Br. J. Audiol. **24**, 89–103.

Wilson, J. P. (**1992**). ''Cochlear mechanics,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford, U.K.).

Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., and Merzenich, M. M. (**1997**). ''Deficits in auditory temporal and spectral resolution in language-impaired children,'' Nature (London) **387**, 176–178.

Zwicker, E., Terhardt, E., and Paulus, E. (**1979**). ''Automatic speech recognition using psychoacoustic models,'' J. Acoust. Soc. Am. **65**, 487–498.

Modeling interaural-delay sensitivity to frequency modulation at high frequencies

Kourosh Saberi[a)]
Massachusetts Institute of Technology, Research Laboratory of Electronics, Cambridge, Massachusetts 02139

(Received 2 December 1996; revised 8 November 1997; accepted 20 January 1998)

Interaural-delay sensitivity to high-frequency ($\geqslant 3$ kHz) sinusoidal-frequency-modulated (SFM) tones is examined for rates from 25 to 800 Hz and depths of -12 to 18 dB. Comparison is made to thresholds obtained for sinusoidal-amplitude-modulated (SAM) tones for the same observers and modulation rates. Both SAM and SFM threshold-by-rate functions are U-shaped with optimum sensitivity to SFM tones occurring at higher rates ($f_m = 200-400$ Hz) compared to those for SAM tones ($f_m = 100-200$ Hz). Effects of modulation depth were examined for rates from 50 to 300 Hz. In all cases thresholds improved considerably with increasing modulation depth. It is also shown that a hybrid dichotic signal composed of an SFM tone presented to one ear and an SAM tone to the other, can perceptually fuse and be lateralized, with the contingency that both stimuli have equal modulation rates but not necessarily equal carrier frequencies. Using bandpass noise to restrict off-frequency listening, it was shown that for this stimulus, observers can use information from filters either below or above the carrier frequency. Consistent with FM-to-AM conversion from cochlear bandpass filtering, several important differences between the SAM- and SFM-tone data can be predicted from a nonstationary stochastic model of binaural interaction whose parameters are uniquely determined from the SAM-tone data.
[S0001-4966(98)03705-9]

PACS numbers: 43.66.Ba, 43.66.Pn, 43.66.Nm, 43.66.Lj [RHD]

INTRODUCTION

Sensitivity to interaural delays at high frequencies is documented extensively for amplitude-modulated (AM) sounds. The early studies of AM processing investigated the dependence of sensitivity on a wide range of simple stimulus parameters such as envelope rate, carrier frequency, and modulation type (Leakey *et al.*, 1958; David *et al.*, 1959; Harris, 1960; McFadden and Pasanen, 1974, 1975, 1976; Henning, 1974a, 1980, 1981; Hafter and DeMaio, 1975; Nuetzel and Hafter, 1976, 1981; Yost *et al.*, 1971). Later investigations extended to more complex binaural phenomena such as interference (McFadden and Pasanen, 1976; Zurek, 1985; Trahiotis and Bernstein, 1990; Buell and Trahiotis, 1993; Bernstein and Trahiotis, 1995), summation of interaural information (McFadden and Pasanen, 1976; Hafter *et al.*, 1990; Buell and Trahiotis, 1993; Saberi, 1995), testing models of cross correlation (Colburn and Esquissaud, 1976; Stern *et al.*, 1988a), masking (Henning, 1974b; Bernstein and Trahiotis, 1992; Kohlrausch *et al.*, 1995), motion discrimination (Saberi and Hafter, 1996), adaptation and precedence (Hafter *et al.*, 1988; Blauert and Divenyi, 1988; Saberi, 1996), all at high frequencies.

In contrast, there are only a handful of papers on binaural sensitivity to frequency-modulated (FM) sounds (Henning, 1980; Blauert, 1981; Hukin and Darwin, 1994). The types of FM stimuli to which binaural sensitivity has been shown include phase-jittered tones (Nordmark, 1976; Blauert, 1981), square-wave modulation (Green and Kay, 1975; Green *et al.*, 1976; Kay, 1982), noise-modulated phase (Bielek, 1975), and sinusoidal-frequency-modulated (SFM) tones (Henning, 1980). In spite of their completely flat envelopes, lateralization of high-frequency FM sounds is possible because of an FM-to-AM conversion by cochlear bandpass filtering (Zwicker, 1958; Henning, 1980; Blauert, 1981). The rate, depth, envelope shape, and phase of this AM-like response vary across filters and primarily depend on: (1) the frequency separation between the filter's resonant frequency and the stimulus carrier; (2) stimulus modulation rate, phase and depth; and (3) the impulse response of the filter, including its phase spectrum. The manner in which these factors interact is complex and requires careful and in-depth consideration.

With the exception of Henning (1980), all previous binaural studies on FM at high frequencies have used nonsinusoidal modulation. These stimuli have discontinuities in the slope of the carrier waveform and, therefore, transient byproducts. Sinusoidal modulation however has smoothly varying temporal properties with a power spectrum that is more confined and symmetric about the carrier frequency, and is therefore a useful stimulus for isolating the effects of frequency modulation. The one study that has examined binaural sensitivity to SFM tones at high frequencies was conducted by Henning (1980) who used only a single modulation rate (300 Hz) and carrier frequency (3.9 kHz). Henning reported that thresholds of 80–200 μs for that stimulus were comparable to those observed for quasi–frequency-modulation (QFM) and slightly higher than those for SAM tones with the same rate.

The goal of the present study is to: (1) investigate binaural processing of SFM tones at high frequencies for a wide

[a)]Current address: 216-76 Division of Biology, Caltech, Pasadena, CA 91125, Electronic mail: kourosh@caltech.edu

range of stimulus parameters and in comparison to SAM tones, and (2) evaluate for SFM tones the predictions of a model of binaural interaction whose parameters are uniquely determined from the SAM-tone data. The experiments begin with a measurement of interaural time difference (ΔITD) thresholds for SFM tones at several carrier frequencies and modulation rates. Thresholds are then compared at one carrier frequency to those obtained with SAM tones using the same observers. Data and observations are also reported on a variety of other psychophysical tasks that examine the effects of modulation versus carrier delay, and modulation depth. Finally, binaural interaction between an SFM tone presented to one ear, and an SAM tone to the other ear is examined. We begin with a description of the spectral characteristics of SFM tones.

I. FOURIER SPECTRUM OF SFM TONES

An SFM tone $X(t)=A_c\cos[2\pi f_c t+\beta\cos(2\pi f_m t)]$ has an instantaneous frequency $f_c+\Delta f\sin(2\pi f_m t)$ and a Fourier expansion (Inglis, 1988; Carlson, 1986) expressed as the harmonic series

$$X(t)=A_c\,J_0(\beta)\cos 2\pi f_c t-A_c\sum_{n=0}^{\infty}J_{2n+1}(\beta)\times\{\cos 2\pi[f_c-(2n+1)f_m]t-\cos 2\pi[f_c+(2n+1)f_m]t\}+A_c\sum_{n=0}^{\infty}J_{2n}(\beta)\{\cos 2\pi[f_c-2nf_m]t+\cos 2\pi[f_c+2nf_m]t\}, \quad (1)$$

where $\beta=\Delta f/f_m$, the frequency-modulation index, is the ratio of peak-frequency-deviation to modulation frequency, and

$$J_n(\beta)\triangleq\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{j(\beta\sin\lambda-n\lambda)}\,d\lambda \quad (2)$$

are Bessel functions of the first kind of order n (sideband number) and argument β which define the amplitude coefficients of the line spectra (Fig. 1). Equation (1) shows that the line spectra of SFM tones consist of a carrier frequency f_c and sidebands at $f_c\pm nf_m$ with the *odd*-order lower sidebands inverted in phase relative to the unmodulated carrier.

As β increases, so does the bandwidth of the SFM tone. Usually, the effective bandwidth of this stimulus is approximated from Carson's rule which defines the 98% energy bandwidth as $2(\beta+1)f_m$ (Cooper and McGillem, 1986; Carlson, 1986). While the bandwidth of the SFM tone increases with β, the amplitude A_c of the time waveform and therefore the stimulus power are constant and independent of β. Consequently, changing β must change the stimulus energy at the carrier frequency. For some values ($\beta=2.4$, 5.3) there is no energy at the carrier. This property is very different from that of SAM tones. The carrier of an SFM tone contains part of the modulation information while the carrier of an SAM tone contains no information about the modulation waveform (Couch, 1987; Carlson, 1986).

The amplitude of the sidebands of an SFM tone [Eq. (2)] cannot be expressed in nonintegral form, however, they have been numerically evaluated and tabulated for many values of β (Abramowitz and Stegun, 1972). The lower panel of Fig. 1 shows calculations of sideband attenuation relative to carrier amplitude in decibels for various values of $\beta=0.2$, 0.4, 0.6, 0.8, 1.0, 2.0. Note that these functions are dependent on only one stimulus parameter, β. The amplitude of the sidebands drop rapidly when the sideband number exceeds the value of β for $\beta>1$. For $\beta\leqslant 1$, the stimulus bandwidth from Carson's rule is confined to the frequency region bordered by the second sidebands. For most of the carrier frequencies and modulation rates used in the following experiments, the SFM-tone bandwidth is confined to frequencies considerably above 1.5 kHz. Nonetheless, lowpass noise and highpass filtering are used to eliminate possible low-frequency cues or intermodulation distortions.

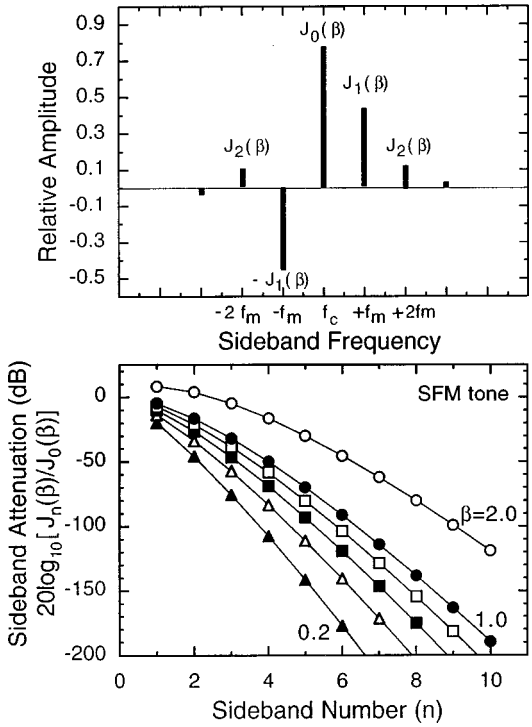


FIG. 1. Top panel shows the amplitude spectrum of an SFM tone with carrier f_c, modulation rate f_m, and a modulation index β of unity. The amplitudes are Bessel functions of the first kind of order n (harmonic number) and argument β. The lower odd harmonics are phase reversed relative to the remaining components [Eq. (1)]. The lower panel shows sideband attenuation in dB *re*: the amplitude of the carrier for $\beta=0.2$, 0.4, 0.6, 0.8, 1.0, and 2.0.

II. MODEL STRUCTURE

In this section a model of binaural interaction is described that is later used to examine the data obtained for both SAM and SFM stimuli. The important feature of the model is that its parameters are determined from the SAM data and predictions are subsequently made for SFM tones. Thus the extent to which the model captures performance reflects the extent to which amplitude envelope cues are used from an FM-to-AM conversion by bandpass filtering. A block diagram of the model structure is shown in Fig. 2. The stimuli are first filtered through a Gammatone filterbank (Holdsworth *et al.*, 1988). The filterbank consists of 30 filters

distributed logarithmically from 2 to 5 kHz. The Hilbert envelopes are then extracted and used to drive a nonstationary fractal point process.[1] The point process, which models the serial correlation of spike patterns from auditory afferents (Teich, 1989; Teich and Lowen, 1994; Lowen and Teich, 1995, 1996) is then subjected to a refractoriness model (Westerman and Smith, 1985; Carney, 1993) with a random dead-time distributed uniformly from $R_A = 0.25 - 0.75$ ms, and a recovery function described by the sum of two exponentials

$$Q(t) = \begin{cases} Q_{\max}\{c_0 \exp[-(t-t_1-R_A)/s_0] \\ \quad + c_1 \exp[-(t-t_1-R_A)/s_1], \quad \text{for } (t-t_1) \geqslant R_A; \\ 0, \quad \text{otherwise,} \end{cases} \tag{3}$$

where $t-t_1$ is the time since the last discharge, and $Q_{\max}$ is the maximum increase in threshold for the recovery function (normalized to unity in the current implementation), $c_0 = 0.55$ and $c_1 = 0.45$ are coefficients for the discharge history effect, and $s_0 = 0.8$ and $s_1 = 25$ ms are time constants. This first stage of the model is used to generate post-stimulus time histograms (PSTH) from 1000 sample waveform presentations; the PSTHs are then low-pass filtered to reconstruct the envelopes of the filtered waveforms. The primary effect of the point process and refractoriness was to enhance the waveform peaks, similar to an expansive nonlinearity used in psychophysical models (Colburn, 1973; Stern *et al.*, 1988a; Shear, 1987) and also reported for physiological recording in the auditory nerve in response to SAM tones (Joris and Yin, 1992). The resultant output of the point process is low-pass filtered at 150 Hz; the slope of the FIR low-pass filter is chosen to fit the SAM-tone data for modulation rates above 150 Hz.

The second stage of the model uses the reconstructed envelopes as inputs to an interaural cross-correlation function

$$\chi(t,\tau,f) = p(\tau) \int_{f_a}^{f_b} \int_{-\infty}^{t} X_l(t,f) X_r(t-\tau,f) \times \{1 + e^{-(t+\mu)/k}\}^{-1} \, dt \, df. \tag{4}$$

The running cross-correlation has a logistic decay with mean $\mu = 20$ ms and slope parameter $k = 0.002$, and a centrality-weighting function $p(\tau)$ that heavily weights τ values near zero and has the form of a sixth-order Butterworth bandpass filter that is maximally flat in the passband and monotonic overall with rolloffs at ± 0.6 ms.[2] The cross-correlation function of Eq. (4) is discretized at a sampling rate of 20 kHz and a deviate $\xi_{\tau,f} \sim \text{normal}(0,\sigma_\xi^2)$ is added, before frequency integration, to each delay-by-frequency sample of the $\tau - f$ plane to model the stochastic nature of the process. The lag (τ) associated with the peak of the cross-correlation output, $\chi_{\max}$, is then used to obtain a lateral position estimate, from which the index of detectability is calculated

$$d' = E[\tau(\chi_{\max}) + \epsilon_p] / (\text{var}[\tau(\chi_{\max})] + \text{var}[\epsilon_p])^{0.5}, \tag{5}$$

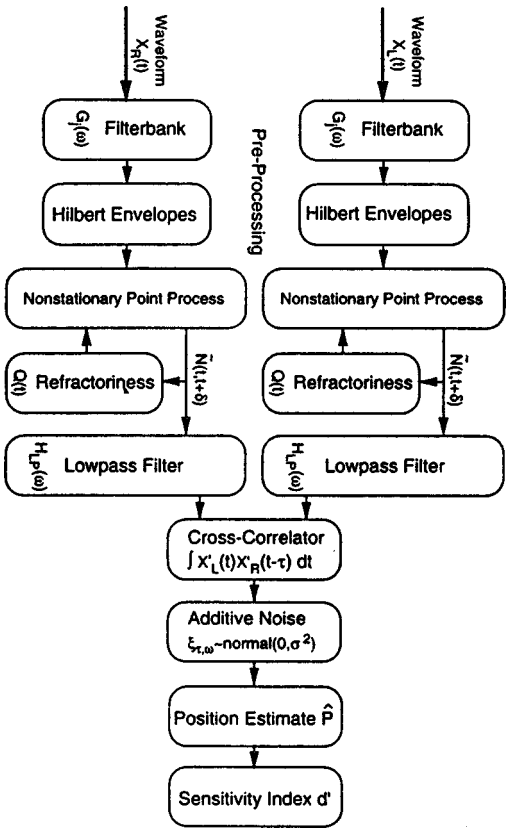


FIG. 2. Schematic diagram of the model structure.

where $\epsilon_p \sim \text{normal}(0,\sigma_p^2)$ is a position noise added to the estimate before calculation of d' to prevent the model from perfect discrimination in those cases where the waveform envelopes have a very high peak factor. Although it is, in principle, useful to maintain this parameter (ϵ), in practice the position noise may be chosen to be substantially smaller than ξ when using high-frequency stimuli, and therefore ϵ may be ignored or set to zero variance. In the model implementation the expected value of d' was estimated from computer simulations of 10 000 trials per each value of interaural delay. Between 2 and 4 values of interaural delay were used to obtain 2–4 values of d' that bracketed 0.78. A least-squares linear fit to log d' vs log ITD was used to obtain the interaural-delay threshold corresponding to 71% correct performance in a 2IFC task ($d' = 0.78$).

There are several reasons for choosing a stochastic approach. First, this approach predicts a reduction of interaural sensitivity at low modulation rates as previously reported for SAM-tone data (Nuetzel and Hafter, 1981). As the rate is reduced, the maximum slope of the envelope diminishes. As the slope of the envelope decreases, the addition of a constant-variance internal noise ξ increases the variance of estimating the peak of the cross-correlation pattern. As the modulation rate is increased, the slope of the envelope increases and the variance of the position estimate is reduced, thus a higher d'. As the rate is still further increased, the sidebands of the complex are increasingly filtered (due to increased component spacing) resulting in a reduced depth of modulation, increased variance of the position estimate, and thereby smaller d'. We next describe experiments that mea-

sure interaural-delay sensitivity to a variety of SFM stimulus conditions and compare performance with that from SAM tones. Model predictions are evaluated for several of these experimental conditions.

III. EXPERIMENT I

A. Lateralization thresholds for SFM tones as a function of modulation rate and carrier frequency

1. Procedure

In each interval of a two-interval, forced-choice (2IFC) task, a 400-ms dichotic SFM tone with a 20-ms cosine-squared rise-decay envelope was presented. The two intervals of the trial were separated by 300 ms. In one interval, the waveform ITD led to the left ear and in the other interval it led to the right ear (except for the gating envelopes which were not interaurally delayed). Waveform delay was achieved by appropriately phase shifting the carrier and modulation waveforms by $\phi_c = 2\pi f_c \mathrm{ITD}$ and $\phi_m = 2\pi f_m \mathrm{ITD}$, respectively. The subject's task was to identify the order of presentation of the tones (i.e., left-leading then right, or, right-leading then left). Each run consisted of 60 trials with visual feedback after each trial. A two-down, one-up procedure (Wetherill and Levitt, 1965; Levitt, 1971) was used to adaptively track the 0.707 probability of a correct response on the observer's psychometric function. The step change in interaural delay was 0.2 log μs up to the fourth reversal and 0.05 log μs for the remaining trials. The starting ΔITD value was 1300 μs (650 μs for each interval of the 2IFC). The first four reversals were discarded and the values of interaural delays at the remaining reversals were averaged to obtain one estimate of threshold.

Each subject completed a minimum of six runs per condition. The carrier frequencies were either 3, 4, or 8 kHz and the modulation rates were 100, 200, 250, 300, 350, 400, 450, 500, 600, and 700 Hz. The modulation index β was equal to unity. Each condition was defined by one value of modulation rate and one value of carrier frequency selected randomly between runs, but fixed within a run. All signals were computed before each trial of the adaptive procedure using an IBM PC and an array processor (Tucker-Davis Technologies TDT AP2). Signals were presented through 16-bit digital-to-analog converters (TDT DA2) at a sampling rate of 40 kHz and were low-pass filtered at 20 kHz. The sound pressure level for a continuous SFM tone was 60 dB SPL. Signals were highpass filtered at 1.5 kHz. In addition, continuous Gaussian noise, low-pass filtered at 1.5 kHz, was presented at a spectrum level of 32 dB to mask low-frequency intermodulation distortions (Henning, 1974a; Henning, 1980; Nuetzel and Hafter, 1981). Delays between left and right channels were checked for accuracy with a Philips dual-channel storage oscilloscope (model PM 3335). Signals were then led to a single-walled sound booth and presented to subjects through Sennheiser (HD-450) headphones.[3] All three subjects (EM, AN, AS) were experienced in lateralization experiments, and each was practiced on various conditions of the experiment until performance seemed to be stable. Two subjects were female and one male (EM) and their ages ranged from 18 to 22. The ear canals of all subjects were examined and cleaned at the University Infirmary. All subjects had normal hearing within 10 dB of ISO standards for frequencies between 125 and 8000 Hz as determined from a Bekesy audiometric test.

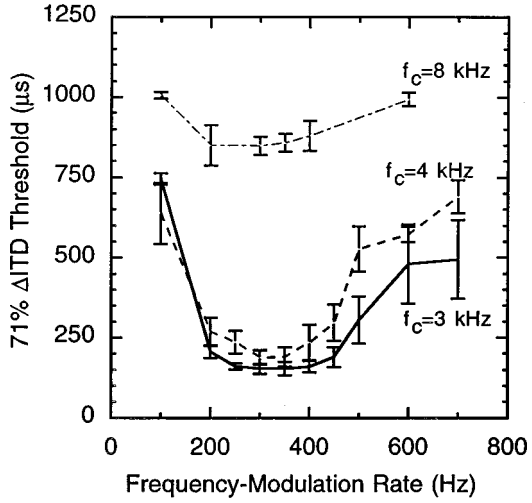


FIG. 3. Interaural-delay difference thresholds for SFM tones as a function of modulation rate. The parameter is carrier frequency. Note that each interval of the 2IFC task carried an interaural delay of ΔITD/2. Data are averaged for three observers and the error bars are one standard error of the mean.

2. Results

Results are shown in Fig. 3. Each line represents the averaged data for one carrier frequency. For the 8-kHz carrier, thresholds are elevated at all rates. At this carrier frequency, the signals are clearly audible and the decline in performance is not due to stimulus detectability. For the 3- and 4-kHz carriers, lowest thresholds are obtained at modulation rates of 200–400 Hz, with the 3-kHz carrier producing slightly better thresholds than the 4-kHz carrier. Best thresholds are on the order of 100–200 μs. One observer (AN) produced low thresholds even at a rate of 700 Hz (thus the large variability). For this one subject, performance degraded to chance as the modulation rate was increased to 1 kHz (not shown). The only previously available related data is from Henning (1980) who used a 300-Hz modulation of a 3.9-kHz carrier. The threshold of 200 μs reported by Henning is similar to the current results. Statistical analysis showed a significant effect of modulation rate for 3- and 4-kHz carriers ($F_{9,20}=2.66$, $p<0.05$; and $F_{9,20}=3.48$, $p<0.05$) but not for the 8-kHz carrier ($F_{5,12}=0.98$, n.s.). Carrier-frequency effects were significantly different between the 3- and 4-kHz conditions [$t(29)=2.36$ $p<0.05$].

B. Effects of carrier versus modulation delay

Studies on lateralization of SAM tones at high frequencies have shown that observers are only sensitive to the interaural delay in the envelope of the stimulus and not the carrier (Henning, 1974a, 1980; Henning and Ashton, 1981; Nuetzel and Hafter, 1976, 1981). If the SFM stimulus is converted into an AM-like response at the auditory periphery, then one may expect a similar dominance of modulation compared to carrier delay. It should be noted however that

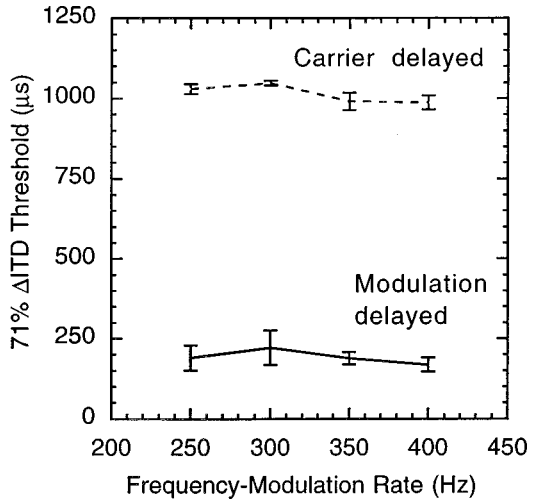


FIG. 4. The dashed line shows interaural-delay difference thresholds measured when the modulation waveforms of an SFM tone remained in phase at the two ears, and the 3-kHz carrier was interaurally phase shifted by $2\pi f_c$ITD. The solid line shows thresholds for the SFM tone when the modulation waveform was interaurally delayed by $2\pi f_m$ITD and the carrier phase was zero at the two ears. The data are averaged for three observers and the error bars are one standard error.

while phase shifting the carrier of an SAM tone produces a constant delay of the carrier, this is not true for SFM tones. A constant carrier phase shift of the SFM tone introduces a dynamic delay in the fine structure. For example, an interaural phase inversion would produce a smaller instantaneous interaural delay in the fine structure when the instantaneous frequency of the waveform is equal to the carrier frequency, compared to when it is at its minimum value $f_c-\Delta f$. While our results below show that these dynamic carrier effects are not important at high frequencies, they should be accounted for when low-frequency carriers are used, especially for large β. All subjects, procedures and stimuli for this experiment were the same as part A with the exception that for the carrier-delayed case the phase of the carrier to one ear was additionally delayed by $\phi_c=2\pi f_c$ITD while $\phi_m=0$ [see Eq. (6b)]. In the modulation-delayed case the modulation waveform was delayed by $\phi_m=2\pi f_m$ITD while $\phi_c=0$. Only one carrier frequency, 3 kHz, and four modulation rates of 250, 300, 350, and 400 Hz for which low thresholds had previously been observed were used ($\beta=1$). Averaged results for the three observers are shown in Fig. 4. The dashed line represents the carrier-delayed condition and the solid line shows the modulation-delayed condition. Error bars are one standard error of the mean. There is no significant carrier-phase effect for the range of parameters examined. Thresholds for the modulation-delayed case (solid line) are comparable to the whole-waveform delayed thresholds shown in Fig. 3.

IV. EXPERIMENT II

A. Comparison of SFM to SAM thresholds

To what extent is the AM induced by bandpass filtering of an SFM tone comparable to that produced by an SAM tone, and to what extent can performance for the SFM stimulus be predicted from data on SAM stimuli? In this section,

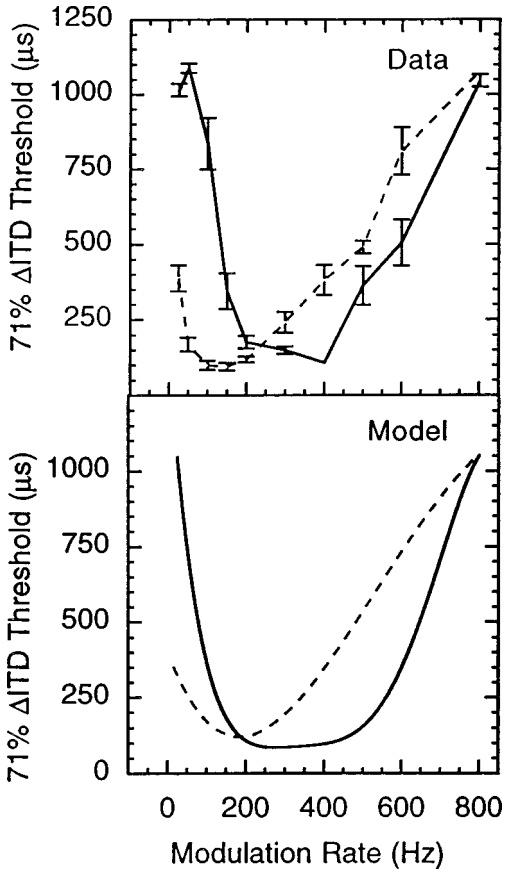


FIG. 5. ΔITD thresholds for SAM tones (dashed line) compared to SFM tones (solid line). The carrier frequency was 3 kHz. Data are averaged across three subjects. Error bars are one standard error of the mean. Lower panel shows model predictions.

ΔITD thresholds are measured for a 3-kHz carrier, either frequency or amplitude modulated at rates from 25 to 800 Hz. This is a slightly larger range of rates than those used in experiment I. The same controls and procedures as before were used in this experiment, except that the amplitude of the SAM tone was adjusted by $m^2/2$ (Viemeister, 1979) to equate the stimulus power with that of the SFM tone for which β was equal to one. Three new subjects were used in this experiment (JA,AW,KS).

1. Results

Top panel of Fig. 5 shows the results averaged for three observers. The bottom panel shows model predictions and will be discussed shortly. The dashed line shows thresholds measured for the SAM tone and the solid line, for the SFM tone. Error bars are one standard error. There is good agreement between the SFM data from experiment I (at 3 kHz) and the current data, in spite of using different observers. It is clear that the SFM data are shifted to the right by about 200 Hz compared to the SAM data. A second observation is that the same observers produced somewhat similar minimum thresholds for SFM and SAM tones, although the minimums occur at different rates (there is a slight edge for SAM thresholds). Previously published data on SAM tones has shown some variability in the minimum of this function, but generally, best sensitivity is observed anywhere from rates of

125–300 Hz (Henning, 1974a, Figs. 7, 8; Nuetzel and Hafter, 1981, Fig. 1). The data reported here show a minimum for the SAM function at a rate of 100–150 Hz for all observers, and a minimum of the SFM function that is centered around 200–400 Hz (Figs. 3, 5). There are statistically significant differences between SAM and SFM functions ($F_{1,40}=21.52$, $p<0.05$), rate effects ($F_{9,40}=23.42$, $p<0.05$), and interaction between condition (FM vs AM) and rate ($F_{9,40}=13.33$, $p<0.05$).

The bottom panel shows model predictions. It is critical to note that the model parameters were determined from the SAM-tone data of experiment II. The following procedure was used to determine these parameters. The noise parameter ξ was adjusted to force the model performance to match that of the single SAM-tone datum at a modulation rate of 150 Hz (average value in Fig. 5). The low-pass FIR filter was then adjusted to best fit the slope of the increasing thresholds as the modulation rate is increased (dashed line in top panel of Fig. 5 from 200 to 800 Hz). No parameter adjustments were made for the low modulation rates below 150 Hz. The model was then presented with the SAM stimuli at all rates (low and high) and with the SFM stimuli. The final output of the model was smoothed with a polynomial function. The following are the important features of the model prediction for experiment II.

(1) An increase in ΔITD threshold with SAM tones is observed as the modulation rate is decreased below 150 Hz. This feature is also observed in the model and is a result of its stochastic nature and the reduction in the slope of the envelopes at the filter outputs as modulation rate is decreased (approaching a dc function as modulation rate approaches zero). Because the cross-correlation function also becomes less peaked, for a constant variance noise ξ, the position estimate becomes less reliable.

(2) At low rates of modulation, thresholds for SFM tones also increase for the same reason, however, thresholds increase much faster than those for SAM tones. This is also observed in the model behavior and is due to the following. For a constant index of frequency modulation ($\beta=\Delta f/f_m=1$) as the modulation rate (f_m) is decreased, the peak frequency deviation must also decrease. Consequently, the instantaneous frequency of the stimulus sweeps through a progressively smaller region of an auditory filter, hence an output which will have a lesser depth of modulation. This is not true of an SAM tone whose amplitude is always fully modulated when $m=1$. Thus thresholds for an SFM tone increase not only because of a decrease in envelope slope (as for SAM tones) but additionally because of a reduction in peak-frequency-deviation and therefore the depth of modulation.

(3) As the modulation rate is increased above 150 Hz, thresholds increase for both SAM and SFM conditions, but they increase more quickly for the SAM stimuli (i.e., for a given modulation rate above 200 Hz, SFM stimuli produce lower thresholds). The model also shows this feature. Here, the reason is similar to that described in the previous point except in the opposite direction. As the rate increases, so does the peak-frequency-deviation. The output of a filter presented with a moderately high rate of frequency modulation (200–500 Hz) is highly peaked since the instantaneous frequency of the stimulus sweeps rapidly through the filter's passband. The highly peaked nature of the outputs of such filters provides a cross-correlation pattern that is also highly peaked. Therefore, the position estimates have a low variance in the SFM case.

(4) The predictions for SFM conditions show a wider region of ''good'' performance at both high and low rates of modulation. This difference is small, but consistent. The width of this region is primarily affected by the slope of the central lowpass filter and the bandwidth of the filterbank model.

(5) Thresholds for both SAM and SFM stimuli converge at the highest rate of modulation (800 Hz), both in the data and in the model predictions. This is not surprising in either case. In the experiments, the adaptive procedure had a ceiling ΔITD of 1300 μs (650 for each interval of the 2IFC task). The procedure produces a maximum threshold estimate of about 1050 μs, even for an undetectable signal, because correct responses are recorded by chance on half the trials (i.e., the probability of two successive chance correct responses is 0.25). For this same reason, we limited the model performance to a high threshold of 1050 μs. Both in the data and predictions, a negative curvature from this feature is observed as the modulation rate approaches 800 Hz.

B. Thresholds as β is varied

Interaural-delay thresholds for SAM tones improve with increasing depth of modulation (Nuetzel and Hafter, 1981). For an SFM tone, as the depth of frequency modulation increases, its instantaneous frequency sweeps through increasingly wider frequency regions and its spectrum widens. As the tone sweeps through wider frequency regions, the AM induced from bandpass filtering may also increase in depth, but not necessarily in all conditions. Because the SFM tone has an infinite number of sidebands and a constant power, increasing the depth of frequency modulation may in fact distribute the spectral energy such that components falling within a filter's passband would produce smaller envelope depths. The widening of the spectrum may also be problematic if leakage through filters with low resonant frequencies is suspected. In the current experiment, the carrier frequency was 3 kHz and three modulation rates of 50, 100, and 300 Hz were used. For the 300-Hz rate, β was restricted to less than 1.5. For the lowest rate of 50 Hz, a maximum value of $\beta=8$ was used (a peak-frequency-deviation of 400 Hz). Note that all signals were additionally highpass filtered at 1.5 kHz and low-pass noise with the same cutoff frequency was continuously presented at $N_0=32$ dB.

Top panel of Fig. 6 shows the averaged data from this experiment for three observers. The error bars are one standard error of the mean. The parameter is modulation rate. Modulation depth is expressed as $20\log(\beta)$ to be consistent with the expression of amplitude modulation depth $20\log(m)$ in decibels. As before, there is an upper bound of $\approx$1050 μs on performance measurement due to the ceiling imposed on the adaptive procedure, thus the negative curvature for the 50-Hz rate. Clearly, a significant monotonic effect of β is observed. It is noteworthy that thresholds at a modulation

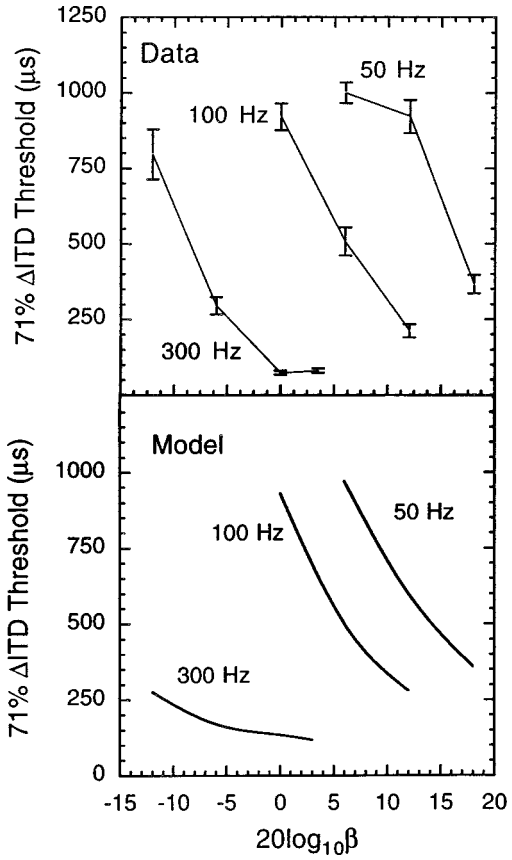


FIG. 6. Top panel shows the effects of changing the frequency-modulation depth on interaural-delay difference thresholds. The stimulus was an SFM tone with a 3-kHz carrier frequency and the parameter is modulation rate. There is an upper bound of $\approx 1050\ \mu$s on performance measurement due to the ceiling imposed on the adaptive procedure. Data are averaged across three observers. Error bars are one standard error of the mean. Lower panel shows model predictions.

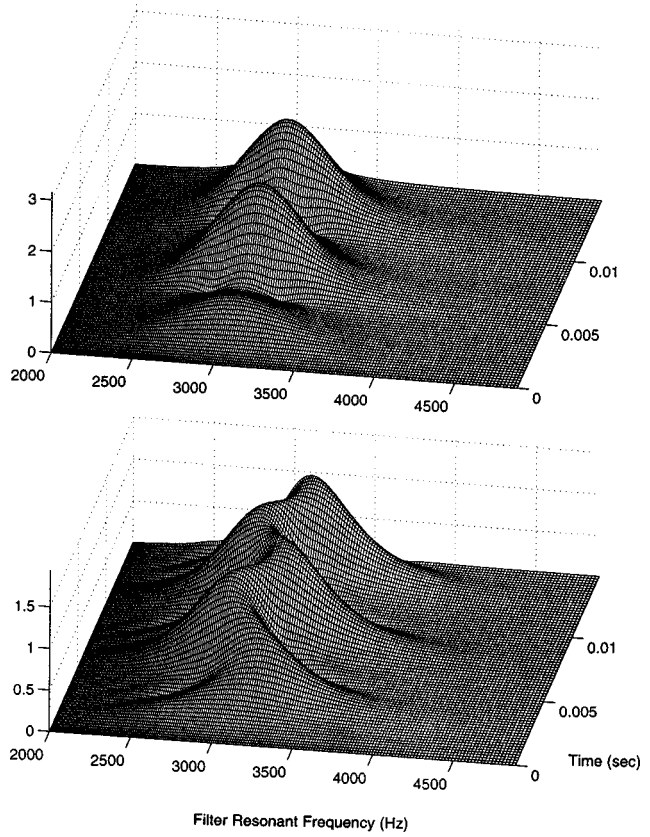


FIG. 7. Output of a Gammatone filterbank in response to an SAM tone (top panel) and an SFM tone (bottom panel). Both stimuli had a carrier frequency of 3 kHz, a modulation rate of 200 Hz, and a rise/fall time of 5 ms.

rate of 50 Hz, where virtually no sensitivity was previously observed, have noticeably improved as β was increased to 8. This indicates that the limiting factor is not rate *per se* but the depth of the induced AM.

The lower panel shows model predictions. Predicted thresholds improve with increasing β because for a constant rate of modulation the peak-frequency-deviation (Δf) increases, producing envelopes at the filter outputs that have a higher peak factor. For $f_m = 300$ Hz the model substantially underestimates the rate of change in threshold as β varies. The reason for this discrepancy is not clear. It should be noted that for these predictions, the noise parameter ξ was adjusted to best match the middle datum at a rate of 100 Hz (0-dB depth). If however, ξ is adjusted individually at each rate to best fit the datum at the middle value of β for that rate, considerably better fits are obtained even at 300 Hz (not shown).[4]

V. EXPERIMENT III

A. Binaural interaction between SFM and SAM tones

It is natural to pose the question that if an FM-to-AM conversion occurs in the auditory periphery, then would an SFM tone presented to one ear, interact with an SAM tone of identical parameters presented to the other ear? Consider a comparison of the outputs of a filterbank model in response to SFM- and SAM-tone stimulation (Fig. 7). The filterbank consisted of 300 Gammatone filters whose resonant frequencies were spaced logarithimically from 2 to 5 kHz. The top panel shows this response to an SAM tone and the bottom panel to an SFM tone, both modulated at 200 Hz and centered at 3 kHz. The SAM tone produces a modulation pattern whose peak amplitude is at the filter centered on the carrier. The rate of modulation is constant and independent of the filter placement. An SFM tone on the other hand produces a somewhat more complicated pattern. Both the depth and rate of modulation vary across filters. The output of the filter centered on the carrier produces a small AM response whose rate is twice that of the stimulus modulation rate. For the filters centered slightly away from the stimulus carrier frequency, the response is an envelope with asymmetrically occurring double-peaks during each cycle of modulation. There is, consequently, substantial spectral energy at the fundamental frequency of modulation at the outputs of even these off-frequency filters (see also Fig. 10). For filters centered further than the first sideband, the filter response is at the modulation rate, however, the response peaks occur at different times for filters centered above and below the carrier frequency. How do the different responses for different filters combine at a later processing stage? One may address some of these questions by examining the interactions that occur when an SFM tone is presented to one ear and an SAM tone to the other. The first part of the data reported here was

presented in summary form in an earlier paper (Saberi and Hafter, 1995). Here individual data are presented in addition to newly collected data and model predictions.

B. Procedure

The left- and right-ear stimuli were SAM and SFM tones, respectively,

$$X_L(t) = A'_C \sin 2\pi f_c t[1 + m \sin 2\pi f_m t], \quad (6a)$$

$$X_R(t) = A_C \sin[2\pi f_c t + \phi_c + \beta \sin(2\pi f_m t + \phi_m)]. \quad (6b)$$

Both modulator waveforms were defined in sine phase, were fully modulated $m = \beta = 1$, had equal modulation rates (250 Hz) and carrier frequencies (3 kHz). The SAM-tone modulation waveform remained in sine phase throughout the experiment and an interaural delay was introduced by shifting the phase of the SFM-tone modulation. Both waveforms had 20-ms cosine-squared ramps that were not interaurally delayed. The task was 2IFC with feedback. In one interval $\phi_m = 0$, and in the other interval $\phi_m = 2\pi f_m$ ITD. As before, observers were instructed to indicate if the auditory image was perceived left–right or right–left. Most observers reported that for the homophasic case, the images were not symmetric about the median plane, but shifted toward the right ear which carried the SAM tone. Because this was the referent condition, the level of the SAM tone was adjusted (A'_c) so that when the modulators were homophasic at the two ears, the intracranial image was centered. All other procedures and apparatus were the same as those described in experiment IA, except that a blocked psychophysical design with fixed stimulus parameters was used instead of the adaptive procedure and the measure of performance was proportion of correct responses in runs of 100 trials, with one run per subject per interaural delay.

C. Results

Subjects reported that a single intracranial image was perceived whose position varied as the interaural modulation phase was varied. Figure 8 shows proportion of correct responses for three observers (filled symbol) as a function of interaural phase disparity. The open symbols are the averaged data and the solid line is model predictions and will be discussed later. When the modulation of the SFM tone was antiphasic relative to the SAM phase (π radians or a 2-ms ITD) discrimination performance was at its maximum. As the phase difference approached a full cycle, performance dropped to chance. One observer reported that for the conditions where $\phi_m > \pi$, the images reversed position and she had to reverse her response pattern (i.e., the correct feedback light became correlated with a reversal of the order of the images). For this case, all observers were instructed to adopt a response strategy to maximize the number of correct responses. Such a reversal of image positions is consistent with the cyclic nature of the modulation waveform.

Figure 9 shows results of a second experiment of this type. As before, an SAM tone with a carrier frequency of 3 kHz and a modulation rate of 250 Hz was presented to the right ear. The left ear, however, received an SFM tone whose

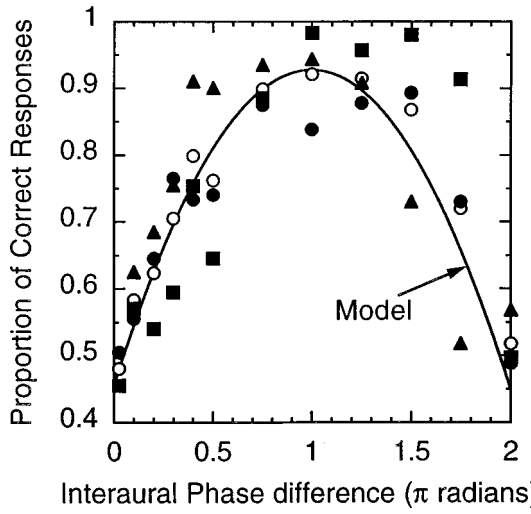


FIG. 8. Results of experiment III. An SFM tone was presented to the right ear and an SAM tone to the left ear. Proportion of correct responses are shown for three subjects (filled symbols). The task was to discriminate homophasic modulation from a delayed-modulation condition (gating envelopes were not delayed). The open symbols are the average for the three subjects and the line shows model prediction.

carrier frequency and modulation rate were parameters of the study. On each run of 100 trials, a single modulation rate and a single carrier frequency were selected for the SFM tone and percent-correct detectability of a 2-ms delay imposed on the SFM-tone modulator relative to the SAM-tone modulator were examined. Two features of these data are important. First, performance is above chance only when the modulation rate of the SFM tone is equal to that of the SAM tone. Other modulation rates do not binaurally interact with the 250-Hz SAM tone.

It is instructive to note a lack of interaction for an SFM-tone rate of 125 Hz with an SAM tone of 250 Hz. An SFM tone with a rate of 125 Hz produces an envelope with a rate of 250 Hz at the output of a filter centered on the carrier. However, this envelope, as observed from Figs. 7 and 10, has a very small depth and its effects are apparently easily overwhelmed by responses from off-frequency filters. Further support for off-frequency listening is provided by Fig. 8. If only the information from the filter centered on the stimulus carrier was used, a bimodal function may be expected with peaks at 0.5 and 1.5 rads.

A second feature of Fig. 9 is that binaural interaction is maintained for a wide range of interaural carrier-frequency disparities, as large as 1 kHz. This result is consistent with that reported for high-frequency SAM tones. Nuetzel and Hafter (1976, 1981) and Henning (1974a) have shown that some observers maintain interaural sensitivity to an SAM tone with interaural carrier-frequency disparity of 0.3–1 kHz. McFadden and Pasanan (1975) have also shown binaural beats with two-tone complexes for relatively large interaural-frequency differences between the complex to the left and right ears (≈ 1 kHz). Naturally, the skirts of cochlear filters extend beyond one critical band. If we consider a Gammatone filter as an approximation, the level of a 3-kHz carrier component of an SAM tone is 30 dB attenuated at the output of a filter centered at 4 kHz, and it is about

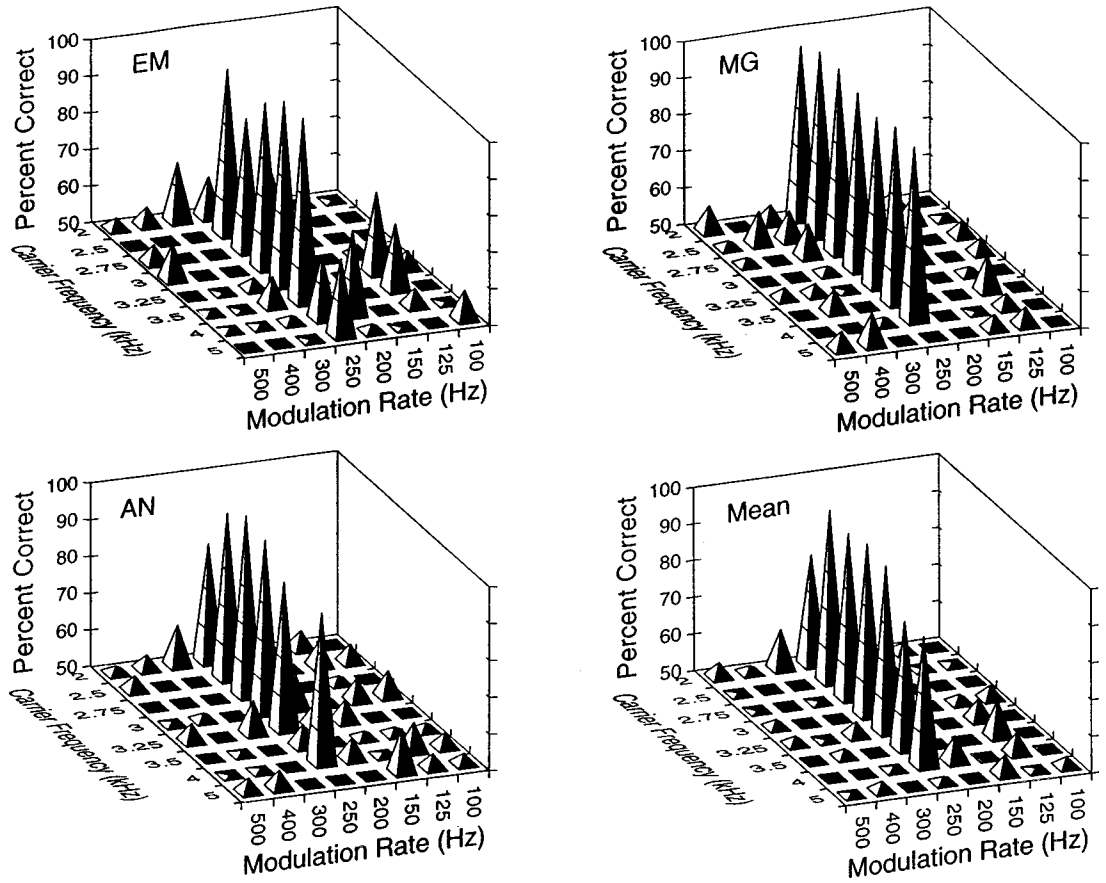


FIG. 9. Interaural-delay sensitivity for unequal carriers and modulation rates. The left ear received an SAM tone with a carrier of 3 kHz and a modulation rate of 250 Hz. The right ear received an SFM tone, the carrier and modulation rate of which were parameters (x and y coordinates). Data were collected for all 64 combinations of carrier versus rate. Each panel shows data from one observer. The lower right panel shows the average data for the three observers.

50 dB attenuated at the output of a filter centered at 2 kHz. These are only moderate attenuations and apparently insufficient to prevent binaural interactions.

Data were also collected at other modulation rates. Some preliminary listening showed that different observers showed best performance at different modulation rates, but generally rates between 100 and 250 Hz were optimal. For one observer, performance was measured at a modulation rate of 100 Hz as β was varied. The number of correct responses out of 100 trials per run were 100 for $\beta=1.0$, 92 for $\beta=0.2$, 88 for $\beta=0.1$, and 72 for $\beta=0.05$. Even at a modulation depth of -25 dB ($\beta=0.05$) where the peak-frequency-deviation is only ± 5 Hz this observer showed above chance performance. The frequency deviation in this case is too small to produce a noticeable modulation depth at the filter centered on the carrier (≈ 0.1 dB if calculated from a Gammatone filter). Listening must therefore occur at more remote filters. We show further data of this type from more observers in the discussion section when considering the off-frequency process underlying this interaction.

For simplicity, the data of experiment III on binaural interaction between an SFM and SAM tone are analyzed by a statistical approach since the data were collected at a fixed modulation rate (the stochastic aspects of the model do not significantly affect this analysis). Figure 11 shows a surface plot of the SFM–SAM cross-correlation function (after bandpass filtering) for the two extreme stimulus phase conditions, that is when the modulation waveforms of the two stimuli were either in-phase (left panels) or out-of-phase (right panels) relative to each other. The surface plots are a top-down view of the τ-f plane; the lighter regions represent the areas of greatest activity. The lower panels are the same as the upper panels except that a logistic frequency-weighting function was used to attenuate the activity above 3 kHz. The upper plots show nearly identical patterns that are mirror images along the frequency axis. It is easy to see that if one were to integrate across frequency, the in-phase and out-of-phase conditions would produce similar outputs and would therefore be the least discriminable conditions.[5] This is opposite to what was observed in the data. However, if observers were to listen off frequency either to auditory filters above or below the carrier (but not both), then these two phase conditions would be maximally discriminable as observed in the data. As the lower panels show, the inclusion of a frequency-weighting function allows discrimination between homophasic and antiphasic conditions because the peak activity is at negative delays for the homophasic conditions and at positive delays for the antiphasic conditions. To evaluate model predictions for this stimulus condition, the cross-correlation function was frequency-weighted prior to frequency integration as shown in Fig. 11. The position estimate ($\hat{p}$) was then calculated as the interaural delay corresponding to the peak cross correlation. Index of detectability ($d'=\hat{p}/\sigma$) was then determined from the position esti-

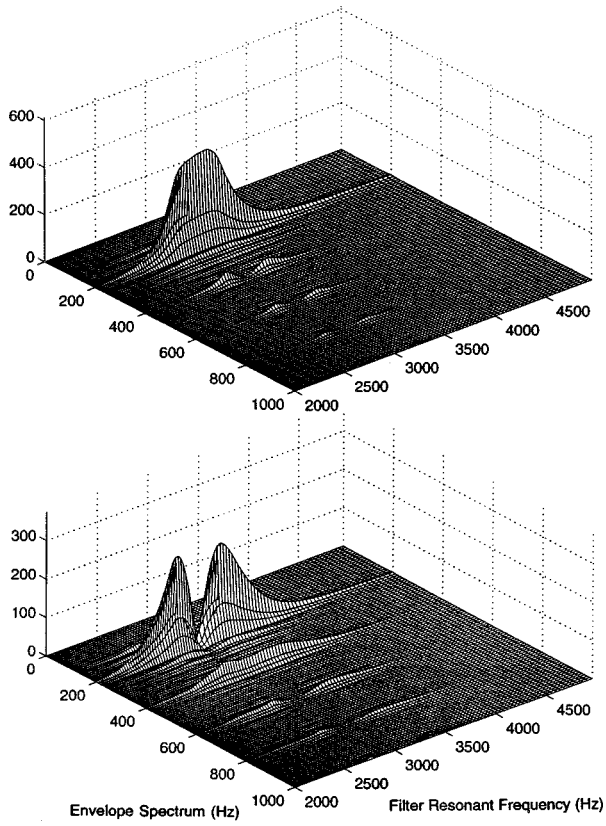


FIG. 10. Spectrum of the envelopes at the output of a Gammatone filterbank for an SAM tone (upper panel) and an SFM tone (lower panel). The stimulus was a 3-kHz carrier modulated at 200 Hz.

mate where σ is the single free parameter of the model. Predictions are plotted in Fig. 8 (solid line) together with the individual data (filled symbols) and averaged data (open symbols) of experiment III. Thus de-emphasizing binaural cross-correlation activity at the higher frequency regions captures the trend of the data.

To further test the validity of this analysis and the off-frequency explanation, an additional control experiment was run in which filtered noise was used to mask either upper, lower, or middle frequency regions of the stimuli. The stimulus conditions and procedures were the same as described for experiment III. The modulation rate and carrier frequency of both stimuli (SFM to one ear and SAM to other) were 250 Hz and 3 kHz, respectively. The task was to discriminate a homophasic from antiphasic condition. The three masking conditions included gated Gaussian noise with a passband of either 2.0–2.9 kHz, 2.8–3.2 kHz, or 3.1–4.0 kHz. No trial-by-trial feedback was provided.

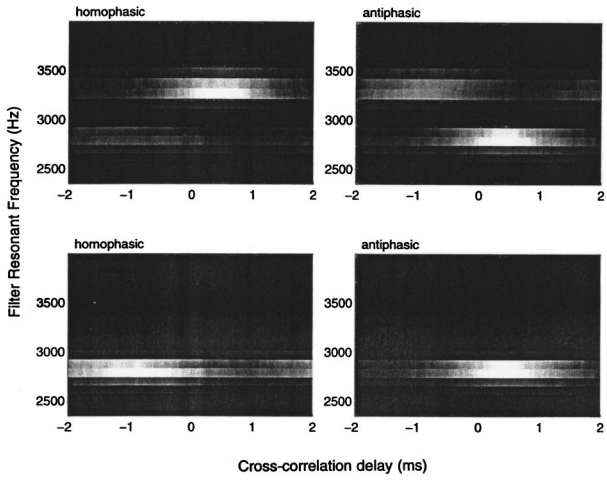


FIG. 11. Surface plots of interaural cross-correlation functions for four conditions (top-down view of the τ-f plane). The lighter regions show the areas of greatest activity. Left panels show the case where the SFM and SAM modulation waveforms were homophasic, and the right panels show antiphasic conditions. The upper panels are unweighted and the lower panels are frequency weighted.

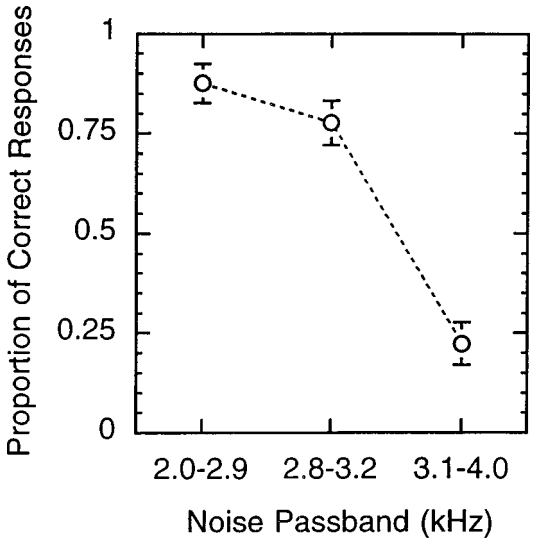


FIG. 12. Proportion of correct responses of detecting a homophasic from antiphasic modulation waveform at the two ears, as a function of masking either lower, upper, or middle frequency regions of the stimuli. Chance performance is 0.5. Left ear received an SAM tone and right ear an SFM tone, both with carriers of 3 kHz and modulation rates of 250 Hz. The subject's task was to respond right if the image orders were left–right and to respond left if the orders were right–left. Correct response was arbitrarily assigned to the homophasic–antiphasic, and incorrect to antiphasic–homophasic order of stimulus presentation in the 2IFC task. Data are averaged for three observers and error bars are one standard error.

Averaged data for three observers are shown in Fig. 12. Correct responses were arbitrarily assigned to one of the two orders of presentation (i.e., homophasic-antiphasic=correct; antiphasic-homophasic=incorrect), although as before, no trial-by-trial feedback was provided to the subjects. A few features of these data are worth emphasizing. First, a reversal of performance is observed when opposite frequency regions of the stimulus are masked, consistent with off-frequency listening and predicted from the cross-correlation analysis discussed above (note that chance performance is 0.5). Second, when the frequency region near the carrier is masked, observers seem to focus on the upper frequency region (i.e., similar performance for 2.0–2.9 and 2.8–3.2 kHz conditions). Third, unmasked performance (Fig. 8) is consistant with the low-frequency masker condition, suggesting that in the unmasked case, observers also attend to the frequency region above the carrier frequency, even though information below the carrier may also be available.[6]

VI. DISCUSSION

We would like to address three issues in this section: (1) Why are threshold-by-rate functions different for SAM and

SFM tones; (2) What are the envelope cues that promote off-frequency listening to SFM tones; and (3) How does the phase response of auditory filters alter the envelope of the AM-like response. The data and predictions both show that interaural-delay thresholds follow U-shaped functions for both SAM and SFM tones as modulation rate is varied (Fig. 5). Why do the minima of these functions occur at different modulation rates? First, consider the stimulus properties that limit interaural-delay sensitivity for SAM tones. At low modulation rates, eventhough the carrier is fully modulated, performance is limited by the inability of the binaural system to track envelope synchrony for tones whose envelopes are changing very slowly (i.e., a dc effect). This is of course not a characteristic of high-frequency fibers *per se* but of envelope slope since highpass filtered clicks are easily lateralized with a precision of better than 100 μs (Hafter *et al.*, 1983, 1988). Performance at low rates may also be affected by a reduction in the number of samples or envelope peaks that the binaural system receives for a fixed stimulus duration (McFadden and Moffitt, 1977). While we do not exclude this factor, it should be noted that the experiment on frequency-modulation depth (Fig. 6) shows that thresholds improve with increasing β, which in turn increases the effective slope of the envelope at a filter's output without changing the number of envelope-peak samples. For the higher modulation rates, SAM thresholds are affected both by cochlear filtering and a possible central low-pass filter (Viemeister, 1979). As the modulation rate is increased, the SAM sidebands are increasingly removed from the carrier and are thereby attenuated by the skirts of the filter centered on the carrier, resulting in a lower envelope depth and, thus, poorer performance (Nuetzel and Hafter, 1981).

What factors limit performance at high and low modulation rates for SFM tones? If the index of modulation is kept constant as the modulation rate is decreased, there is by definition a linear decline in peak-frequency-deviation, Δf. The smaller Δf yields an AM response with a smaller envelope depth. Thus for an SFM tone, not only is performance at low rates degraded by reduced envelope slope (as was the case for SAM tones), but there is an additional reduction of envelope depth due to smaller Δf. Therefore, thresholds degrade more rapidly for SFM tones compared to SAM tones as the modulation rate is reduced. Note that even if Δf for an SFM tone is kept constant as the rate is lowered, there is no reason to believe that the lower end of the rate-by-threshold functions would be identical for SAM and SFM stimuli since the envelope shapes would not be identical (Edwards and Viemeister, 1996). At higher modulation rates of about 300–400 Hz, the SFM tone produces lower thresholds than the SAM tone. One reason for this is an increase in the peak-frequency-deviation with increasing rate and a corresponding increase in depth of the AM-like envelope. At even higher rates (700 Hz) neither stimulus can be lateralized, either because the components are not resolved or because of limitations imposed by a central low-pass filter.

What are the cues that promote off-frequency listening? The data of experiment III suggest that observers do not heavily rely on ITDs from filters on or near the carrier frequency of an SFM tone. If an AM-like cue is the primary information from an SFM tone, then the rate associated with on-frequency listening would be twice the stimulus modulation rate and its depth would be considerably smaller than the response from off-frequency filters. Off-frequency filters produce an envelope rate that either has single peaks at the modulation rate or multi-modal peaks with energy at the fundamental frequency of modulation. The reliance on off-frequency listening is not at a major cost. One can estimate the available cues at different rates of frequency modulation at the outputs of a filterbank model from the Fourier spectrum of their envelopes. Figure 10 shows this output for both an SFM and SAM tone with a 3-kHz carrier and a 200-Hz modulation rate. Clearly, the predominant information is at the fundamental periodicity for both stimuli and the loss in information from neglecting the on-frequency filter is very small for the SFM tone since there is very little envelope-spectral-energy at the output of that filter to begin with. Consistent with these observations, others have provided evidence from monaurally presented stimuli that observers rely on off-frequency filters where the change in excitation pattern in response to frequency modulation is maximal (Moore and Glasberg, 1986; Moore and Sek, 1992; Hartmann and Hnath, 1982).

Finally, a concern throughout this study was how the phase response of auditory filters affect the processing of the FM stimulus. Equation (1) shows that the odd harmonics below the carrier of the SFM tone are inverted in phase relative to the remaining components. The phase characteristic of an auditory filter, of course, would affect the phase relations of this harmonic complex and thereby alter the time envelope at the output of that filter. Unless the filter has a linear phase (i.e., a delay) the envelope of any filtered waveform that contains more than two components within the filter passband will be affected. What then are the phase-by-frequency functions of auditory filters; how do they compare with the phase response of Gammatone filters; and, are changes in envelope from phase effects significant compared to external (stimulus) envelope modulation or that caused by a filter's magnitude response?

Kohlrausch and Sander (1995) suggest that the phase of an auditory filter is consistant with the phase pattern of a negative Schroeder-phase signal. A Schroeder-phase signal is a harmonic complex that has a quadratic phase-by-frequency function

$$\phi_n = \phi_1 - \pi n^2/N, \tag{7}$$

where n is harmonic number, N is the total number of components in the complex, and the sign in front of π defines the sign of the Schroeder phase (Schroeder, 1970). The Gammatone filter used in our analysis has an antisymmetric phase that is increasing in slope below the point of inflection (at the filter's resonant frequency) and decreasing above it. The curvature of this filter's phase (its second derivative) changes signs at the resonant frequency, whereas the curvature of a Schroeder-phase signal is constant. Kohlrausch and Sander suggest that a Gammatone antisymmetric phase would result in similar masking period patterns for negative- and positive-Schroeder-phase maskers, which is not what they have observed in their data (i.e., they suggest that the

TABLE I. Effects of filter phase on modulation depth (dB).

Filter CF (kHz)	Phase: Antisymmetric	+ Schroeder	− Schroeder
2.75	−1.57	−2.22	−2.21
3.00	−19.5	−18.5	−18.4
3.25	−3.3	−4.3	−4.3

Gammatone phase is not the appropriate model for auditory-filter phases). It is not however clear to what extent the filter's amplitude response affects the predicted differences. The models used by Kohlrausch and Sander for comparison were the Gammatone and Strube's (1985, 1986) basilar-membrane (BM) model. The BM model has a phase function with a negative curvature that is constant throughout most of the filter's passband, but has an amplitude response that simulates only the passive properties of the basilar membrane and therefore shows poor frequency selectivity compared to psychophysical and physiological data. The Gammatone on the other hand shows frequency selectivity comparable to psychophysical data, but has the antisymmetric phase response.

Positive- and negative-Schroeder-phase harmonic complexes such as those used by Kohlrausch and Sander may be phase shifted by an allpass filter with an appropriate negative curvature such that one signal would have a high peak factor and the other signal a low peak factor. If these two signals are then filtered with the amplitude response of a Gammatone filter, one observes little difference in the envelopes of the resultant waveforms at the output of the filter (i.e., the Gammatone amplitude response is substantially more significant than its phase response in determining the modulation envelope). Even if the Gammatone filter had a negative quadratic phase, similar predictions would be observed for the positive- and negative-Schroeder-phase maskers.

To observe how these phase-by-frequency functions affect the envelope of an SFM tone, we made comparisons under three conditions: (1) Gammatone antisymmetric phase; (2) negative-Schroeder-phase; (3) positive-Schroeder-phase. All stimuli were defined by summation of 13 primary harmonics of the SFM tone, and then filtered with an allpass filter with the assigned phase condition by phase shifting each component. The stimulus was then filtered with the amplitude response of a Gammatone filter by attenuating each component before summation. Thus, the filtering properties of a hypothetical Gammatone filter with a phase response defined by one of the three conditions was modeled. An SFM tone with a modulation rate of 250 Hz and a carrier of 3 kHz was the stimulus. The filter had a center frequency of 2.75, 3.0, or 3.25 kHz. The Hilbert envelopes of the filtered signals were then used to calculate an amplitude modulation depth in units of $20 \log_{10}[(\max-\min)/(\max+\min)]$.

Results are shown in Table I. The negative- and positive-Schroeder-phase filters produced identical modulation depth, but with slightly asynchronous envelope minima. The difference between the antisymmetric phase and Schroeder phases are generally small, approximately 0.6 dB for filters centered at 2.75 kHz, 1 dB at 3 and 3.25 kHz. The smallest envelope modulation, as expected, occurred at the filter centered on the carrier of the stimulus. Why are the results so similar for the various phase conditions? The primary components of an SFM tone are the carrier, the first and perhaps the second sidebands; further sidebands are not likely to contribute significantly to the envelope of the filtered waveform. As shown in Fig. 1, the first and second sidebands are approximately 5 and 16 dB lower in amplitude than the carrier. For a filter centered on the carrier, the first and second sidebands are further attenuated by 6 and 17 dB, respectively (assuming a Gammatone filter). The first and second sidebands for the modulation rate of 250 Hz are contained within the frequency region of 2.5–3.5 kHz. For components in this region, the filter produces a phase shift that is well modeled by a linear function of frequency for all three phase conditions ($r^2>0.999$). A linear phase shift is of course only a waveform delay, leaving its envelope unaffected; this nearly linear phase also explains the shift in the envelope minima when using filters with negative versus positive Schroeder phases. If we include the effects of the second through fourth sidebands (2.0–4.0 kHz), the deviation from linearity increases, but not substantially (Gammatone: $r^2=0.96$; Schroeder: $r^2=0.98$). For an off-frequency filter centered at 2.75 kHz, the Gammatone-phase curvature becomes more pronounced and the fit to linearity is reduced ($r^2=0.94$ for the 2- to 4-kHz region) while it remains the same for the Schroeder-phase filters ($r^2=0.98$). Although it is difficult to generalize these results (since both the auditory filter phase and amplitude functions are level dependent: Allen, 1983; Ruggero *et al.*, 1992) it seems that the magnitude response of the Gammatone filter and not its phase response is the primary determinant of envelope depth for a filtered SFM tone, at least for the stimulus conditions examined here.

VII. SUMMARY

(1) The binaural system shows optimum interaural-delay sensitivity when SFM tones are modulated at rates of 200–400 Hz, which is somewhat higher than those rates that are optimum for SAM tones (100–200 Hz). Lowest thresholds of about 100–200 μs for SFM tones at $\beta=1$ are not as good as the 80–100 μs lowest thresholds observed for SAM tones.

(2) An SFM tone presented to one ear interacts interaurally with an SAM tone presented to the other ear, but only for equivalent modulation rates. Interaction is maintained, however, for interaural carrier-frequency disparities of up to 1 kHz. Strongest interactions occur for modulation rates between 100 and 250 Hz. Using bandpass masking noise it was shown that observers can use information from auditory filters either above or below the stimulus carrier frequency, although when conflicting information is present, they attend to the high-frequency regions.

(3) SAM-tone data were used to set the parameters of a stochastic interaural cross-correlation model whose predictions were then obtained for SFM tones of various rates and depths. Several differences between the SAM data and SFM data were well predicted by the model whose main features were an FM-to-AM conversion by bandpass filtering, and an

internal-noise limitation on estimating the peak of the cross-correlation function, and thereby, discrimination of lateral position.

ACKNOWLEDGMENTS

Work supported by NIH Grants Nos. R29 DC03648-01, R01 DC00177, and R01 DC00100. I thank Drs. David M. Green, Ervin R. Hafter, Louis D. Braida, Patrick M. Zurek, Brent W. Edwards, and Malvin C. Teich for many helpful comments. I also thank members of the Sensory Communication Group and the Research Laboratory of Electronics at MIT.

[1]A fractal renewal process (FRP) is generated from sample times between adjacent events that are independent and selected from the same fractal probability distribution (Lowen and Teich, 1995). The survival function for this renewal process decays as a power law

$$t_i \sim \text{power}\{c/t^D - c/\zeta^D\} \quad \text{for} \quad t < \zeta$$
$$0 \quad \text{otherwise},$$

where t_i are independent samples from a power-law density, D is the fractal dimension set at 0.75 in current implementation (Teich and Lowen, 1994), $\zeta = 0.4$ sec is a cutoff parameter for the waveform duration, and c is a normalization constant that ensures the area under the density integrates to unity. The FRP exhibits fractal behavior over time scales under the value of ζ. The process consists of a set of points on the time axis, but it may be recast as a real-valued Bernoulli process that alternates between two states, for example zero and unity. This alternating function would then start at zero, switch to unity at the time corresponding to the first event, switch back to zero at the second event and so on. A number of identical and independent alternating FRPs may be summed to yield a binomial process with the same fractal dimension as the single alternating FRP (Lowen and Teich, 1995) which may then be scaled by the instantaneous level of the stimulus to be used as the driving function for the Poisson point process

$$\text{Poisson}\{\mu_t = X(t)\Sigma\text{FRP}(t)\} \quad \text{over} \quad [\widetilde{N}(t, t+\delta) = 0,1,2,...],$$

where μ_t is the time-dependent parameter of the nonstationary process, $X(t)$ is the instantaneous pressure waveform, and $\widetilde{N}$ is the number of events within a time specified by the discrete waveform sampling period (δ). The result is a fractal-binomial-noise-driven Poisson point process, a stimulus-level dependent doubly stochastic point process (DSPP) that models fractal spike generation by nerve fibers.

[2]The centrality weighting function, $p(\tau)$ implemented here has the shape of a Butterworth bandpass filter. Previous models have used either a function with a Gaussian density shape (Shackleton *et al.*, 1992; Saberi, 1996), or a hybrid function (Stern *et al.*, 1988b; Saberi, 1995) consisting of a uniformly flat maximum between $\tau = \pm 150$ μs and an exponential rolloff for $|\tau| \geqslant \pm 150$ μs. The Butterworth function was selected because it has the advantage of a relatively flat region near $\tau = 0$, a steep rolloff, but without discontinuities in its derivative at the rolloff transition point (i.e., as does the exponential model at ± 150 μs).

[3]The headphone transfer function, measured with a probe-tube microphone (Etymotic ER-7) inside the ear canal was flat within 0.8 dB from 2750 to 3250 Hz (the frequency region where many of the experiments were focused on). In some experiments additional controls were used to ensure that envelope cues resulting from headphone properties were not significant. For example, in some control cases, inverse digital filtering of the SFM signal inside the observer's ear canal was used to flatten the headphone transfer function. It should also be noted that any cues from the filtering of the SFM tone by the headphone transfer function is not pure AM but rather an FM–AM cue (an envelope modulation of a carrier that is changing in frequency).

[4]It is also worth documenting that attempts were made to measure masking-level differences (MLD) for high-frequency SFM tones in continuous noise. However, as has also be reported for SAM tones, no MLDs were observed for SFM stimuli.

[5]Previous models have used one of two measures of frequency coding. One model uses the straightness (rms error) of peak trajectories (Stern *et al.*, 1988b; Trahiotis and Stern, 1994) and others use frequency integration (Stern and Colburn, 1978; Shackleton *et al.*, 1994; Saberi, 1996). Both measures produce identical results in the analysis described here.

[6]By using a similar masking method, Moore and Sek (1994) have shown that monaural discrimination of SFM from SAM tones, at least at low rates of modulation (10 Hz), is also based on listening to off-frequency filters and that the specific listening region (above or below the carrier) is not critical to the discrimination task.

Abramowitz, M., and Stegun, I. (**1964**). *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables* (Dover, New York).

Allen, J. B. (**1983**). ''Magnitude and phase-frequency response to single tones in the auditory nerve,'' J. Acoust. Soc. Am. **73**, 2071–2092.

Bernstein, L. R., and Trahiotis, C. (**1992**). ''Discrimination of interaural envelope correlation and its relation to binaural unmasking at high frequencies,'' J. Acoust. Soc. Am. **91**, 306–316.

Bernstein, L. R., and Trahiotis, C. (**1995**). ''Binaural spectral interference in detection and discrimination paradigms,'' in *Advances in Hearing Research: Proceedings of the 10th International Symposium on Hearing*, edited by G. A. Manley, G. M. Klump, C. Koppl, H. Fastl, and H. Oekinghaus (World Scientific, Singapore), pp. 354–361.

Bielek, K. H. (**1975**). ''Spektrale Dominanz bei der interauralen Signalanalyse,'' Diploma thesis, Ruhr-University Bochum.

Blauert, J. (**1981**). ''Lateralization of jittered tones,'' J. Acoust. Soc. Am. **70**, 694–698.

Blauert, J., and Divenyi, P. L. (**1988**). ''Spectral selectivity in binaural contralateral inhibition,'' Acustica **66**, 267–274.

Buell, T. N., and Trahiotis, C. (**1993**). ''Interaural temporal discrimination using two sinusoidally amplitude-modulated, high-frequency tones: Conditions of summation and interference,'' J. Acoust. Soc. Am. **93**, 480–487.

Carlson, A. B. (**1986**). *Communication Systems: An Introduction to Signals and Noise in Electrical Communication*, 3rd ed. (McGraw-Hill, New York).

Carney, L. H. (**1993**). ''A model for the responses of low-frequency auditory-nerve fibers in cat,'' J. Acoust. Soc. Am. **93**, 401–417.

Colburn, H. S. (**1973**). ''Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination,'' J. Acoust. Soc. Am. **54**, 1458–1470.

Colburn, H. S., and Esquissaud, P. (**1976**). ''An auditory-nerve model for interaural time discrimination of high-frequency complex stimuli,'' J. Acoust. Soc. Am. **59**, S23(A).

Cooper, G. R., and McGillem, C. D. (**1986**). *Modern Communications and Spread Spectrum* (McGraw-Hill, New York).

Couch, L. W. II (**1987**). *Digital and Analog Communication Systems* (Macmillan, New York), 2nd ed.

David, E. E., Guttman, N., and van Bergeijk, W. A. (**1959**). ''Binaural interaction of high-frequency stimuli,'' J. Acoust. Soc. Am. **31**, 774–782.

Edwards, B. W., and Viemeister, N. F. (**1996**). ''Masking of a brief probe by sinusoidal FM,'' J. Acoust. Soc. Am. (submitted).

Green, G. G. R., and Kay, R. H. (**1975**). ''On the localisation by humans of frequency-modulated tones,'' J. Physiol. (London) **252**, 59–60.

Green, G. G. R., Heffer, J. S., and Ross, D. A. (**1976**). ''The detectability of apparent source movement effected by interaural phase modulation,'' J. Physiol. (London) **260**, 49.

Hafter, E. R., and De Maio, J. (**1975**). ''Difference thresholds for interaural delay,'' J. Acoust. Soc. Am. **57**, 181–187.

Hafter, E. R., and Dye, R. H. (**1983**). ''Detection of interaural differences of time in trains of high-frequency clicks as a function of interclick interval and number,'' J. Acoust. Soc. Am. **73**, 1708–1713.

Hafter, E. R., Buell, T. N., and Richards, V. M. (**1988**). ''Onset coding in lateralization: it's form, site, and function,'' in *Auditory Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 647–676.

Hafter, E. R., Dye, Jr., R. H., Wenzel, E. M., and Knecht, K. (**1990**). ''The combination of interaural time and intensity in the lateralization of high-frequency complex signals,'' J. Acoust. Soc. Am. **87**, 1702–1708.

Harris, G. C. (**1960**). ''Binaural interactions of impulsive stimuli and pure tones,'' J. Acoust. Soc. Am. **32**, 685–692.

Hartmann, W. M., and Hnath, G. M. (**1982**). ''Detection of mixed modulation,'' Acustica **50**, 297–312.

Henning, G. B. (**1974a**). ''Detectability of interaural delay in high-frequency complex waveforms,'' J. Acoust. Soc. Am. **55**, 84–90.
Henning, G. B. (**1974b**). ''Lateralization and the binaural masking-level difference,'' J. Acoust. Soc. Am. **55**, 1259–1262.
Henning, G. B. (**1980**). ''Some observations on the lateralization of complex waveforms,'' J. Acoust. Soc. Am. **68**, 446–454.
Henning, G. B., and Ashton, J. (**1981**). ''The effect of carrier and modulation frequency on lateralization based on interaural phase and interaural group delay,'' Hearing Res. **4**, 185–194.
Holdsworth, J., Nimmo-Smith, I., Patterson, R., and Rice, P. (**1988**). ''Implementing a Gammatone filter bank,'' in *Annex C of the SVOS final report (Part A: The auditory filter bank)*, APU Report 2341.
Hukin, R. W., and Darwin, C. J. (**1994**). ''Lateralization of a frequency-modulated 500-Hz tone,'' J. Acoust. Soc. Am. **95**, 2895–2896.
Inglis, A. F. (**1988**). *Electronic Communications Handbook* (McGraw-Hill, New York).
Jeffress, L. A., Blodgett, H. C., Sandel, T. T., and Woods III, C. L. (**1956**). ''Masking of tonal signals,'' J. Acoust. Soc. Am. **28**, 416–426.
Joris, P. X., and Yin, T. C. T. (**1992**). ''Responses to amplitude-modulated tones in the auditory nerve of the cat,'' J. Acoust. Soc. Am. **91**, 215–232.
Kay, R. H. (**1982**). ''Hearing modulation in sounds,'' Physiol. Rev. **62**, 894–975.
Kohlrausch, 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.
Kohlrausch, A., Par, S.v.d., and Houtsma, A. J. M. (**1995**). ''A new approach to study binaural interaction at high frequencies,'' in *Advances in Hearing Research: Proceedings of the 10th International Symposium on Hearing*, edited by G. A. Manley, G. M. Klump, C. Koppl, H. Fastl, and H. Oeckinghaus (World Scientific, Singapore), pp. 343–353.
Leakey, D. M., Sayers, B. McA., and Cherry, C. (**1958**). ''Binaural fusion of low- and high-frequency sounds,'' J. Acoust. Soc. Am. **30**, 222.
Levitt, H. L. (**1971**). ''Transformed up-down methods in psychophysics'' J. Acoust. Soc. Am. **49**, 467–477.
Lowen, S. B., and Teich, M. C. (**1995**). ''Estimation and simulation of fractal stochastic point processes,'' Fractals **3**, 183–210.
Lowen, S. B., and Teich, M. C. (**1996**). ''The periodogram and Allan variance reveal fractal exponents greater than unity in auditory-nerve spike trains,'' J. Acoust. Soc. Am. **99**, 3585–3591.
McFadden, D., and Pasanen, E. G. (**1974**). ''High frequency masking level differences with narrow-band noise signals,'' J. Acoust. Soc. Am. **56**, 1226–1230.
McFadden, D., and Pasanen, E. G. (**1975**). ''Binaural beats at high frequencies,'' Science **190**, 394–396.
McFadden, D., and Pasanen, E. G. (**1976**). ''Lateralization at high frequencies based on interaural time differences,'' J. Acoust. Soc. Am. **59**, 634–639.
McFadden, D., and Moffitt, C. M. (**1977**). ''Acoustic integration for lateralization at high frequencies,'' J. Acoust. Soc. Am. **61**, 1604–1608.
Moore, B. C. J., and Glasberg, B. R. (**1986**). ''The role of frequency selectivity in the perception of loudness, pitch and time,'' in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London).
Moore, B. C. J., and Sek, A. (**1992**). ''Detection of combined frequency and amplitude modulation,'' J. Acoust. Soc. Am. **92**, 3119–3131.
Moore, B. C. J., and Sek, A. (**1994**). ''Discrimination of modulation type (amplitude modulation or frequency modulation) with and without background noise,'' J. Acoust. Soc. Am. **96**, 726–732.
Nordmark, J. O. (**1976**). ''Binaural time discrimination,'' J. Acoust. Soc. Am. **60**, 870–880.
Nuetzel, J. M., and Hafter, E. R. (**1976**). ''Lateralization of complex waveforms: Effects of fine structure, amplitude and duration,'' J. Acoust. Soc. Am. **60**, 1339–1346.
Nuetzel, J. M., and Hafter, E. R. (**1981**). ''Lateralization of complex waveforms: Spectral effects,'' J. Acoust. Soc. Am. **69**, 1112–1118.
Ruggero, M. A., Rich, N. C., and Recio, A. (**1992**). ''Basilar membrane responses to clicks,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford), pp. 85–92.
Saberi, K. (**1995**). ''Lateralization of comodulated complex waveforms,'' J. Acoust. Soc. Am. **98**, 3146–3156.
Saberi, K., and Hafter, E. R. (**1995**). ''A common neural code for frequency- and amplitude-modulated sounds,'' Nature (London) **374**, 537–539.
Saberi, K. (**1996**). ''Observer weighting of interaural delays in filtered impulses,'' Percept. Psychophys. **58**, 1037–1046.
Saberi, K., and Hafter, E. R. (**1996**). ''Experiments on auditory motion discrimination,'' in *Binaural and Spatial Hearing*, edited by R. H. Gilkey and T. R. Anderson (Erlbaum, New York).
Schroeder, M. R. (**1970**). ''Synthesis of low-peak-factor signals and binary sequences with low autocorrelation,'' IEEE Trans. Inf. Theory **16**, 85–89.
Shackleton, T. M., Meddis, R., and Hewitt, M. J. (**1992**). ''Across-frequency integration in a model of lateralization,'' J. Acoust. Soc. Am. **91**, 2276–2279.
Shear, G. D. (**1987**). ''Modeling the dependence of auditory lateralization on frequency and bandwidth,'' master thesis, Carnegie Mellon University.
Stern, R. M., and Colburn, H. S. (**1978**). ''Theory of binaural interaction based on auditory-nerve data. IV. A model for subjective lateral position,'' J. Acoust. Soc. Am. **64**, 127–140.
Stern, R. M., Shear, G. D., and Zeppenfeld, T. (**1988a**). ''Lateralization predictions for high-frequency binaural stimuli,'' J. Acoust. Soc. Am. **84**, S80.
Stern, R. M., Zeiberg, A. S., and Trahiotis, C. (**1988b**). ''Lateralization of complex binaural stimuli: A weighted-image model,'' J. Acoust. Soc. Am. **84**, 156–165.
Strube, H. W. (**1985**). ''A computationally efficient basilar-membrane model,'' Acustica **58**, 207–214.
Strube, H. W. (**1986**). ''The shape of the nonlinearity generating the combination tone $2f_1-f_2$,'' J. Acoust. Soc. Am. **79**, 1511–1518.
Teich, M. C. (**1989**). ''Fractal character of the auditory neural spike train,'' IEEE Trans. Biomed. Eng. **36**, 150–160.
Teich, M. C., and Lowen, S. B. (**1994**). ''Fractal patterns in auditory nerve-spike trains,'' IEEE Eng. Med. Biol. Mag. **13**, 197–202.
Trahiotis, C., and Bernstein, L. R. (**1990**). ''Detectability of interaural delays over select spectral regions: Effects of flanking noise,'' J. Acoust. Soc. Am. **87**, 810–813.
Trahiotis, C., and Stern, R. M. (**1994**). ''Across-frequency interaction in lateralization of complex binaural stimuli,'' J. Acoust. Soc. Am. **96**, 3804–3806.
Viemeister, N. F. (**1979**). ''Temporal modulation transfer functions based upon modulation thresholds,'' J. Acoust. Soc. Am. **66**, 1364–1380.
Wetherill, G. B., and Levitt, H. (**1965**). ''Sequential estimation of points on a psychometric function,'' Br. J. Math. Stat. Psychol. **18**, 1–10.
Westerman, L. A., and Smith, R. L. (**1985**). ''Rapid adaptation depends on the characteristic frequency of auditory nerve fibers,'' Hearing Res. **17**, 197–198.
Yost, W. A., Wightman, F. L., and Green, D. M. (**1971**). ''Lateralization of filtered clicks,'' J. Acoust. Soc. Am. **50**, 1526–1531.
Zurek, P. M. (**1985**). ''Spectral dominance in sensitivity to interaural delay for broadband stimuli,'' J. Acoust. Soc. Am. Suppl. 1 **78**, S18.
Zwicker, E. (**1952**). ''Die Grenzen der Horbarkeit der Amplitudemodulation und der Frequenzmodulation eines Tones,'' Acustica **2**, 125–133.

Forward masked excitation patterns in multielectrode electrical stimulation

Monita Chatterjee[a] and Robert V. Shannon
Department of Auditory Implants and Perception, House Ear Institute, 2100 W. Third Street, Los Angeles, California 90057

(Received 20 June 1997; revised 21 October 1997; accepted 22 January 1998)

Across-channel interactions in multi-electrode cochlear implant patients may be critical to their performance in the natural auditory environment. One measure of channel interaction can be obtained using forward masking. The patterns of threshold shift were obtained as a function of the separation between masker and probe electrode pairs in four Nucleus-implanted patients with varying levels of speech recognition. In the three subjects with better speech recognition (N4, N7, and N13), the masking patterns showed the least parameter dependence. In the subject with the lowest speech scores (N3), the masking patterns showed the greatest dependence on masker level and on probe delay. Masking as a function of masker level also reflected these differences: N3's functions showed the greatest changes for different probe delays. Similar rates of recovery were observed in the three good performers (time constant approximately 70 ms). In contrast, subject N3's recovery function shows a more rapid recovery. These data indicate that channel interaction is individually variable and parameter dependent, both of which may play a role in the perception of dynamic stimuli in the natural auditory environment. © *1998 Acoustical Society of America.* [S0001-4966(98)01105-9]

PACS numbers: 43.66.Dc, 43.66.Ts, 43.66.Mk, 43.64.Me [RHD]

INTRODUCTION

In acoustic hearing, masking patterns have been used as a measure of the spread of excitation across different neural populations, or frequency channels. Psychophysical masking patterns appear to reflect some basic characteristics of cochlear excitation patterns. Thus, as the level of the masker increases, the pattern of threshold shift broadens in extent across different frequencies. This spread of excitation occurs primarily toward the base (higher frequencies), with a sharp cutoff toward the apex (lower frequencies). In electric hearing, cochlear frequency analysis is replaced by an array of electrodes stimulating afferent auditory neurons at different locations along the cochlea. Stimulation of a particular electrode pair creates an electric potential distribution and a corresponding current path, resulting in stimulation of excitable tissue in the vicinity. As the current strength increases, the potential distribution changes, spreading to more distant locations along the cochlea. The space constant in scala tympani (where electrodes are typically inserted) is relatively long (2–3 mm; Black and Clark, 1980), resulting in considerable current spread. Apart from the current spread, the distribution of surviving neurons in damaged cochleae may be individually variable. Changes in the pattern of excitation created by a stimulus will therefore depend on two factors: (a) cochlear electro-anatomy and (b) the nerve survival pattern.

In multi-electrode cochlear implant patients, channel or electrode interaction is expected to play an important role in auditory perception. Such interaction can occur at various levels of processing. For instance, in the case of simultaneous electrical stimulation of two electrodes, the fields of the two stimuli will interact instantaneously. In nonsimultaneous stimulation, interaction can occur due to charge summation at the capacitive nerve membrane (time scale of the order of microseconds). Another kind of interaction can occur between the response of a neuron to two stimuli presented successively in time: Because the neuron takes time to recover from having responded to the first stimulus, its response to the second stimulus is depressed if it is not given a large enough interstimulus interval. This lowered responsivity is one form of masking. Psychophysical experiments reveal a similar increase in threshold (lowered sensitivity) of a signal (probe) following a prior stimulus, or masker. This effect is known as forward masking.

The amount of threshold shift due to a forward masker is presumed to be related to the degree to which the system responded to the masker, and the temporal course of recovery from the masker. If the masker and the probe activate the same ''channel'' (neural population) in the system, we expect to find a larger threshold shift due to the masker than if the masker and the probe activate different, or independent, channels (neural populations). Assuming this to be true, we can obtain a measure of channel interaction as a function of cochlear distance of the masker electrode pair from the probe electrode pair (Shannon, 1983; Tong and Clark, 1986; Lim *et al.*, 1989). Similar experiments have been performed in acoustic stimulation of normal-hearing listeners in order to estimate the spread of excitation across different frequency regions in the cochlea, in response to a tonal masker (Munson and Gardner, 1950; Zwislocki and Pirodda, 1952; Ehmer and Ehmer, 1969).

The perception of dynamic stimuli such as speech, with parameters changing continuously in time, may be heavily

[a]Electronic mail: monita@hei.org

TABLE I. Duration of implant use and sentence recognition scores for each subject.

Subject	Duration of implant use	Sentence recognition scores (CUNY, sound only)
N3	4 years	68%
N4	4 years	95%
N7	6 years	92%
N13	Single channel: 10 years Multichannel: 7 years	96%

influenced by the way in which the perceptual ''channels,'' and their interaction, change with stimulus parameters. The present study examines the dependence of the measured masking patterns on the masker amplitude and the time interval between the masker and the probe (recovery time).

We present forward masked patterns from four postlingually deafened adult users of the Nucleus-22 cochlear implant. Three of the subjects (N4, N7, and N13) performed well (scoring above 90% correct, sound only) on the CUNY sentence tests (Boothroyd, 1988) whereas the fourth (N3) scored below 70% correct on the same test. Forward masking was measured as a function of cochlear location of the electrode pair, masker level, masker-probe delay, and masker-probe separation.

I. METHODS

A. Subjects

Subjects were adult users of the Nucleus-22 device. All subjects were postlingually deafened. Their sentence recognition scores varied, but overall thresholds in quiet were similar. Sentence recognition scores and duration of implant use are shown in Table I for each subject. Calibration tables for stimulus amplitude for each subject's device were obtained from the manufacturer.

B. Experimental procedure

Stimuli were delivered to the subjects using a custom interface (Shannon *et al.*, 1990). Stimuli were software generated. Both masker and probe were trains of biphasic pulses, 200 μs/phase, presented at 1000 pulses/s. Masker and probe durations were 300 and 20 ms, respectively. The delay between masker offset and probe onset (probe delay) was changed by the experimenter as desired. All stimuli were presented in the bipolar-plus-one (BP+1) mode of stimulation: when electrode (10,12) is selected, current flows from electrode 10, the active electrode, to electrode 12, the indifferent electrode.

In order to measure masking patterns, the masker was presented to a fixed electrode pair, while the probe was presented to different electrode pairs. Threshold shift was plotted as a function of probe location. The masker electrode pair was located in the basal [electrode pair (5,7)], middle [electrode pair (10,12)] or apical [electrode pair (14,16)] regions of the cochlea. Subject N13 was stimulated only on electrode pairs numbered (8,10) and less (basal) in order to avoid uncomfortable/nonauditory sensations.

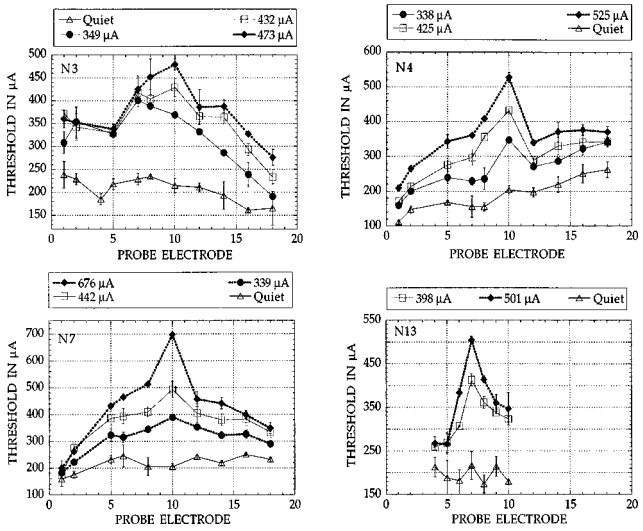


FIG. 1. Examples of masked threshold in μA as a function of probe electrode pair for the four subjects. For subjects N3, N4, and N7, the masker electrode pair was fixed at (10,12). For subject N13, the masker electrode pair was (7,9). The parameter is masker level. The triangles represent probe threshold in quiet for each electrode pair. Solid lines with squares represent the masker amplitude yielding comfortable loudness level for each subject.

Thresholds were measured using a two-interval forced-choice (2IFC), three-down, one-up paradigm. Using this method, the probe level converges at the 79.4% correct point (Levitt, 1971). The mean of the last 8 of a maximum of 13 reversals was calculated as the masked threshold. The step size was 2 dB for the first three reversals, and 0.5 dB for the remaining reversals. Stimuli were presented in two intervals. The intervals were accompanied by visual markers on the computer monitor. The subject heard the masker in both intervals. Randomly, the probe was present in only one of them. The subject indicated which interval contained the probe by pressing the appropriate mouse button. Feedback was provided visually. Threshold shift was measured as (masked threshold - quiet threshold). Generally, an average of two to four repetitions was taken for each data point.

II. RESULTS

A. Dependence on masker amplitude

The effects of masker amplitude (level) on forward masked excitation patterns were studied in individual subjects, for a probe delay of 1 ms. Figure 1 shows examples of raw data obtained in the four subjects. The standard deviations shown are typical. Quiet thresholds were subtracted from the masked thresholds to obtain patterns of threshold shift. Figure 2 shows such patterns obtained in subjects N3, N4, and N7 for maskers located on electrode pair (14,16) (apical location). It is evident that,other than a multiplicative shift, changes in masker amplitude (shown in the legend above each graph) create little or no change in the shape of the pattern for N4 and N7. However, in subject N3, a larger amplitude dependence is observed. This is due to a relatively larger increase in masked threshold around the masker electrode pair at the higher masker amplitude, which results in a more peaked or bandpass masking pattern than at the lower levels.

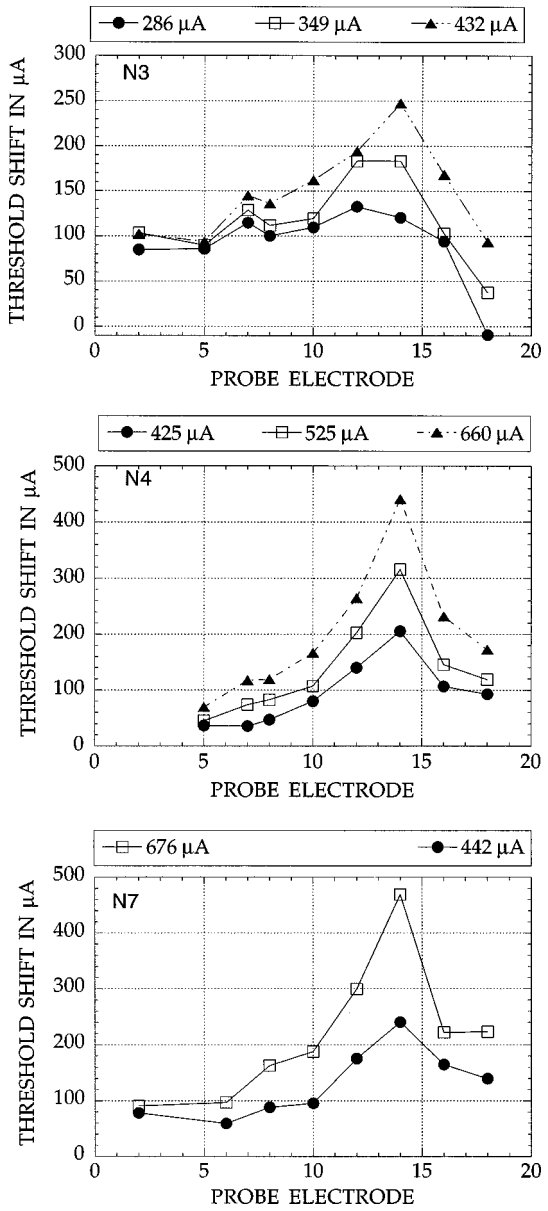


FIG. 2. Threshold shift in μA as a function of probe electrode pair, measured for different masker amplitudes in three subjects. The masker was on electrode pair (14,16) in each case. The open squares represent the masker amplitude yielding comfortable loudness for each subject (subjects N3, N4, and N7).

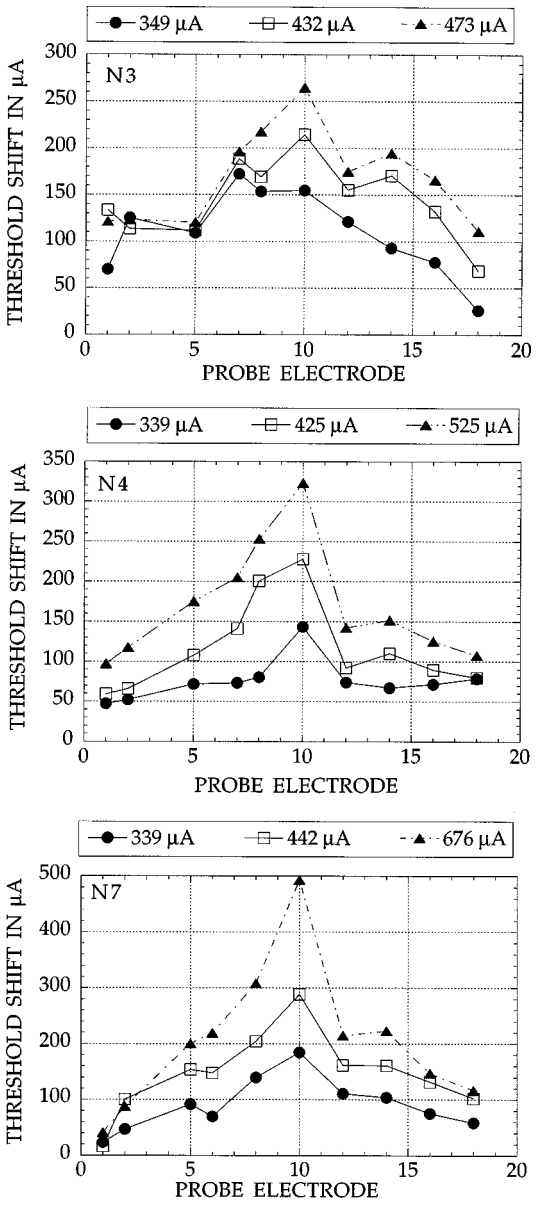


FIG. 3. The same as in Fig. 2: masker on electrode pair (10,12), subjects N3, N4, and N7.

Figures 3 and 4 show further examples of similar patterns obtained with maskers presented to electrode pairs (10,12) (middle location in the cochlea: subjects N3, N4, and N7), and (5,7), (basal location: subjects N3, N4, and N13). Note that the shape of subject N3's patterns show minimal dependence on masker amplitude when the masker is on electrode pair (5,7).

B. Dependence on probe delay

The effect of changing probe delay (Δt) was measured keeping the masker amplitude fixed at a comfortable level for each electrode pair in all subjects. In Fig. 5, patterns of threshold shift are plotted for subjects N4, N7, and N13 for probe delays of 1 and 10 ms. Apart from a multiplicative shift, patterns in these subjects show little change in shape as the probe delay is changed.

Figure 6 shows examples of patterns obtained in subject N3 at probe delays of 1 and 10 ms, at three locations of the masker electrode pair. In this subject, the amount of masking and the shape of the masking pattern both change with probe delay. For the 10-ms probe delay, the masking pattern is relatively flat. For the 1-ms probe delay, the pattern develops a peak around the masker electrode, thus assuming more of a bandpass characteristic. There are also variations with masker electrode location along the cochlea. Thus, note that the change in shape with probe delay is minimal when the masker is on the most basal electrode pair studied [electrode pair (5,7)]. This parallels the minimal change of shape with increasing masker level measured at this electrode pair in this subject (Fig. 4).

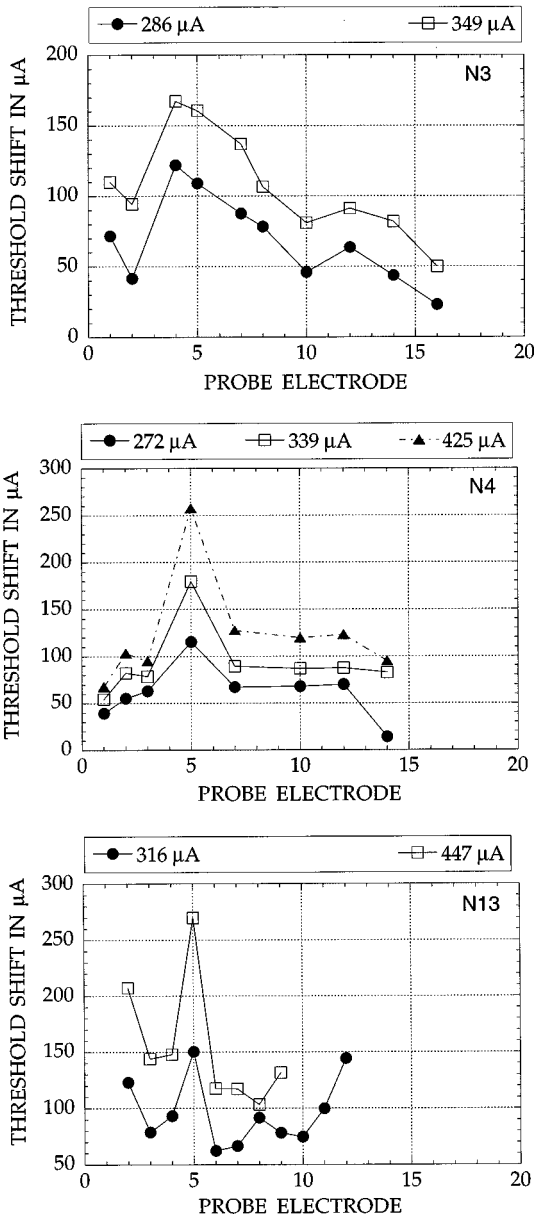


FIG. 4. Same as in Figs. 2 and 3: masker on electrode pair (5,7), subjects N3, N4, and N13.

C. Recovery from forward masking: Dependence on parameters

1. Effects of masker amplitude

Figure 7 shows examples of recovery functions obtained at two different masker amplitudes (corresponding to ''soft'' and ''comfortable'' loudness levels) in subjects N3 and N7. The actual threshold shifts were normalized to the peak in each case, to facilitate comparisons. Masker and probe were on the same electrode pair. It appears that the shape of the recovery function is independent of masker amplitude. Note that N3, the poorer performer, shows complete recovery from masking at less than 200 ms, while N7, who scores well on the sentence recognition test, is still not recovered at 200-ms delay.

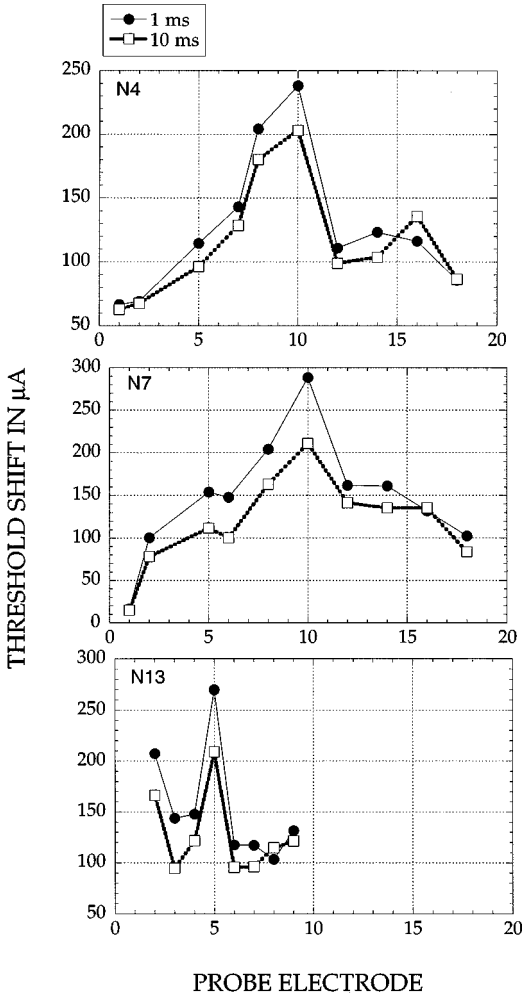


FIG. 5. Threshold shift in μA as a function of probe electrode pair, measured for probe delays 1 and 10 ms. Top: subject N4, masker on electrode pair (10,12). Middle: subject N7, masker on electrode pair (10,12). Bottom: subject N13, masker on electrode pair (5,7). Masker at comfortable level in each case.

2. Effects of probe location

Figure 8 shows recovery functions measured at a comfortable level at two electrode locations in subjects N3 and N7. In each case, masker and probe were presented to the same electrode pair. The threshold shifts have been normalized to the maximum in each case. There is no apparent dependence of the shape of the recovery function on cochlear location.

The shape of the recovery function may, however, change when the masker and the probe are on *different* electrode pairs. Figure 9 shows examples of recovery functions measured in subjects N3 and N7 under such conditions. Each function is normalized to the peak threshold shift to facilitate comparisons. In the case of subject N3 (top), the masker was fixed at electrode pair (14,16) at a comfortable level (432 μA). Normalized recovery functions measured on probe electrode pairs (8,10), (10,12) and (14,16) are shown. For subject N7 (bottom), the masker was fixed at electrode pair (10,12) at comfortable level (442 μA). Normalized recovery functions measured on probe electrode pairs (5,7), (8,10), (10,12), (12,14), and (14,16) are shown. In these two subjects it is evident that the functions change from one probe

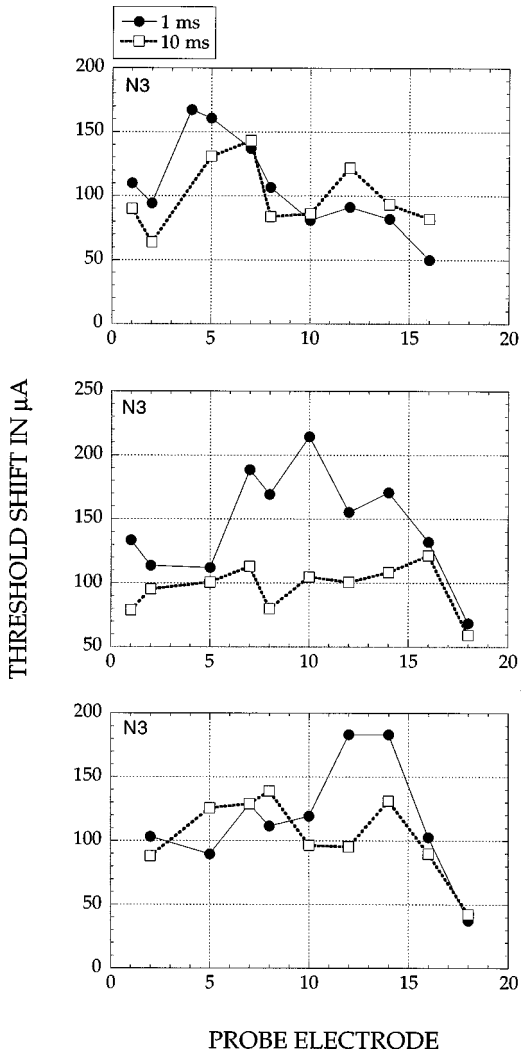


FIG. 6. Same as in Fig. 5: subject N3. Top, middle, and bottom panels: masker on electrode pairs (5,7), (10,12), and (14,16), respectively. Masker at comfortable level in each case.

location to the next. Occasionally, a non-monotonicity, or a bump, is evident in the early part of the recovery function. Such non-monotonicities could be due to the presence of multiple mechanisms in the recovery process.

3. Time constants of recovery: Subject-dependence

Figure 10 shows normalized recovery functions for four subjects. The functions were measured with masker and probe on the same electrode pair in each case. The functions are normalized to the peak threshold shift. The dashed lines represent single exponential curve fits to the normalized data. Time constants calculated from the fits are displayed in the inset in ms. Note that all the subjects, with the exception of subject N3, have very similar recovery functions, with a mean time constant of 72 ms. The time constant obtained for subject N3 is 1.85 times shorter than the mean of subjects N4, N7, and N13. Although the single exponential provides a good fit to the data of N4, N7, and N13, it underestimates the early recovery evident in N3's data.

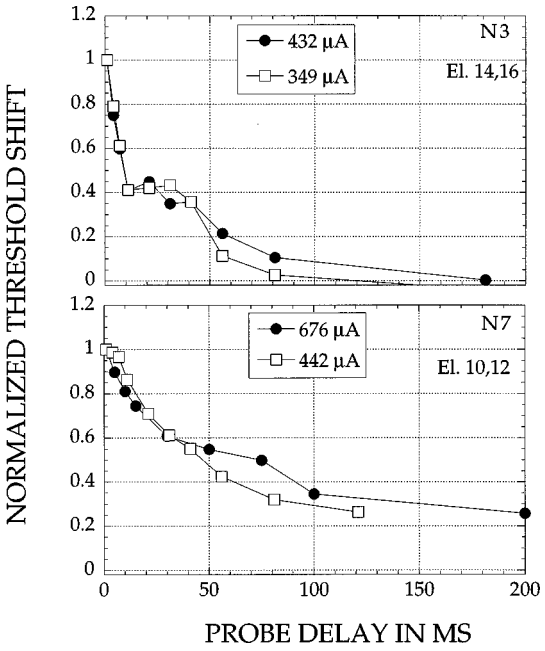


FIG. 7. Recovery from forward masking with masker level as a parameter in two subjects. Masker and probe presented to the same electrode pair in each case. Top: subject N3. Bottom: subject N7.

D. Growth-of-masking functions: Dependence on probe delay

In Fig. 11, growth-of-masking functions are plotted for subjects N3, N4, and N7, at probe delays of 1 and 10 ms. Masker and probe were presented to electrode pair (10,12) in all cases. In subjects N4 and N7, the functions are very similar at the two probe delays, the function for the longer delay being slightly shallower in slope. In these subjects, the

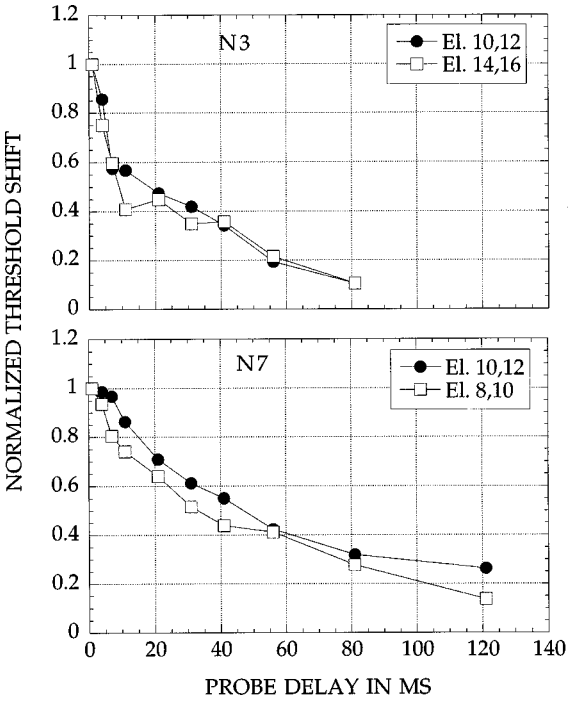


FIG. 8. Recovery from forward masking with cochlear location as parameter: data from two subjects. Masker and probe presented to the same electrode pair in each case. Top: subject N3. Bottom: subject N7.

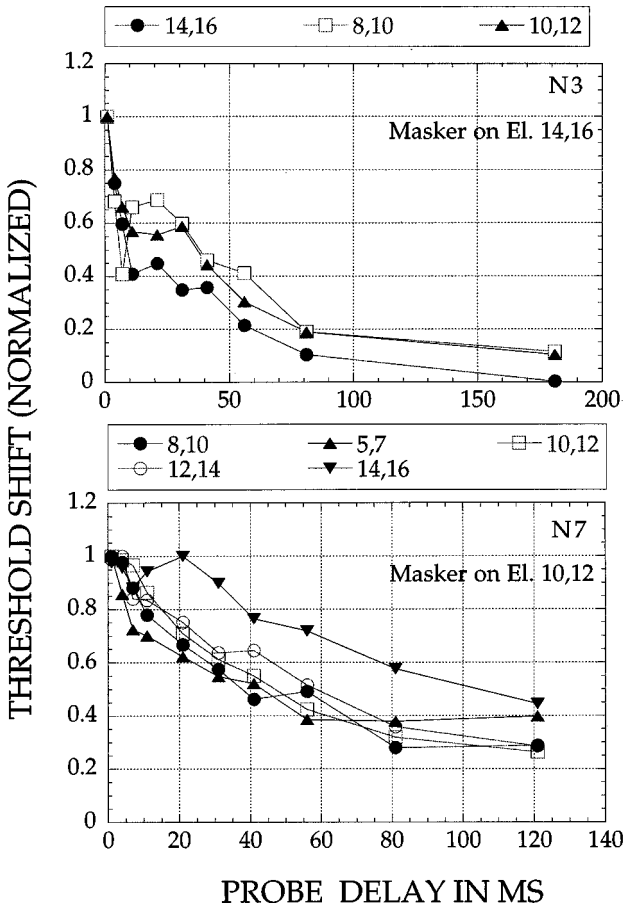


FIG. 9. Recovery from forward masking. Masker and probe on different electrode pairs. Top: subject N3, masker on electrode pair (14,16). Bottom: subject N7, masker on electrode pair (10,12).

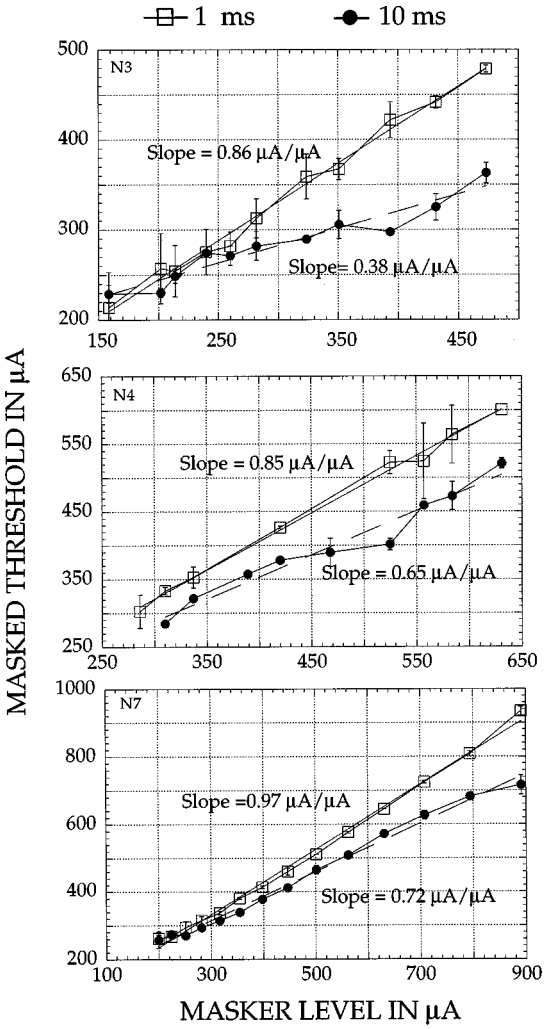


FIG. 11. Growth-of-masking functions in subjects N3, N4, and N7 for probe delays=1 ms (open squares) and 10 ms (filled circles). The slopes are derived from linear curve fits to the data (dashed lines).

change in the slope of the growth-of-masking function is consistent with the multiplicative shift observed at the two probe delays in Fig. 5. In the case of subject N3, consistent with the changes observed in Fig. 6, the growth-of-masking function changes dramatically when Δt is changed to 10 ms: for $\Delta t = 1$ ms, the function is quite linear, with a slope of about 0.85 (close to the slope obtained in the other subjects); when $\Delta t = 10$ ms, the slope changes to 0.38. It is important to recall that in this subject, the spatial masking patterns change from one masker level to another for a fixed probe delay,indicating that, for a fixed masker location, the slope of the growth-of-masking function changes with probe electrode location. In the other subjects, we can expect the slope differences to be much smaller at different electrode locations, given the relatively unchanging shape of the masking pattern with changes in masker amplitude.

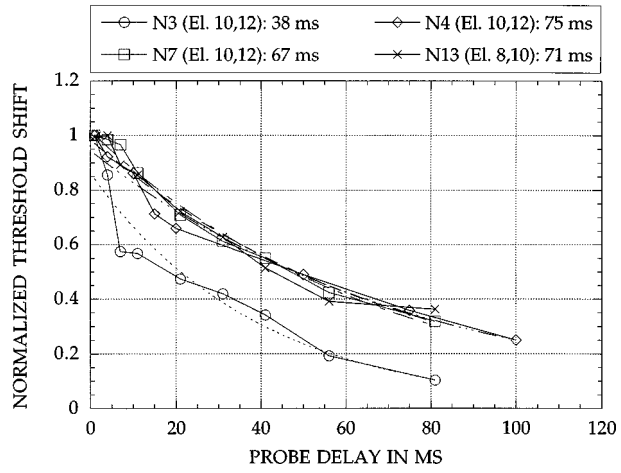


FIG. 10. Recovery from forward masking at most comfortable level in four subjects. The dashed lines show exponential curve fits to the four functions. The inset shows the time constants derived from the curve fits.

III. DISCUSSION

A. Forward masked excitation patterns: General observations

In broad agreement with previous work (Shannon, 1983; Tong and Clark, 1986; Lim *et al.*, 1989) forward masked excitation patterns in all subjects show the following general characteristics:

(1) a spatial band pass characteristic, with a peak in the region of the masker electrode.
(2) Intersubject variability is evident in the actual shape of

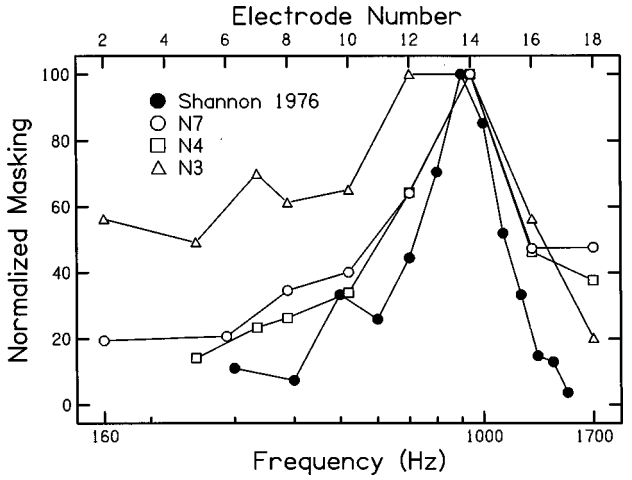


FIG. 12. Masked excitation patterns compared in acoustic and electric stimulation (see text for details).

the pattern. This variability may result from differences in etiology of deafness, nerve survival patterns, electrode insertion depth, etc.

(3) A general shift in symmetry of the excitation pattern is observed when the masker electrode is moved from the basal electrode pair (5,7) to the apical electrode pair (14,16). This is evident in most cases (see Figs. 4–6). For electrode pair (5,7), the pattern is skewed toward the base. For electrode pair (10,12), the pattern is generally somewhat symmetrical, whereas for electrode pair (14,16), there is a skew toward the apex.

B. Comparison with results obtained in normal-hearing listeners

Figure 12 shows a direct comparison of masked patterns obtained in our present study with the forward masking pattern obtained by Shannon (1976) in a normal-hearing listener. All four functions were measured at comfortable loudness levels. The log frequency axis for the normal-hearing listener's data is scaled to correspond to the linear cochlear distance mapped by the electrodes of a Nucleus-22 cochlear implant. The functions are normalized to aid comparisons.

It is evident that, at comfortable loudness, masking patterns in normal acoustic hearing and in electric stimulation are surprisingly comparable in extent as well as in shape. Data from subjects N4 and N7 in particular are very similar to the normal pattern, whereas subject N3's masking pattern is the broadest. It is important to keep in mind when making such comparisons that (a) the dynamic range in normal hearing is large compared with that in electric hearing and (b) that excitation patterns in normal hearing show changes in tuning and shape as the stimulus level is changed. As the compressive nonlinearity present in normal hearing cochleae is absent in electric hearing, phenomena such as the ''upward spread of masking'' are absent.

In agreement with previous work by Shannon (1990), time constants of recovery from masking are comparable with those obtained in normal-hearing listeners. We find a change in the slope of the growth of masking function with a change in probe delay. It may be worthwhile to note that Jesteadt *et al.* (1982, Fig. 1) found a similar trend in normal-hearing listeners.

C. Dependence of masking patterns on stimulus parameters

Results from subjects N4, N7, and N13, who have better sentence recognition scores, show masked thresholds that grow linearly with masker amplitude for all electrodes, with approximately the same slope, resulting in a relatively constant masking pattern. Results from subject N3, who has the poorest sentence score, indicate a growth-of-masking function that has a steeper slope when the probe is on electrodes close to the masker electrode than when the probe is on distant electrodes. This results in a masking pattern that changes from being relatively flat at lower levels of masking to being more peaked around the masker electrode at higher levels. This is in contrast to normal acoustic hearing, where the decay of threshold shift with probe delay is relatively independent of the frequency separation between the masker and the probe.

Similar to the results with changing masker amplitude, patterns obtained in subjects N4, N7, and N13 remained relatively unchanged when probe delay was increased. Subject N3's patterns on the other hand were flat for longer probe delays, but became peaked around the masker electrode for short probe delays. Consistent with these data, subject N3 was found to have a more rapid recovery at and around the masker electrode pair.

Fu (1997) has shown that electrode location plays a significant role in information transfer. Information about electrode location can be conveyed only by the pattern of excitation in the periphery. If the pattern of excitation in the periphery plays a role in coding properties of the stimulus, the shape of the pattern should be robust to changes in stimulus parameters. In the case of N3, information about electrode location, as it is preserved in the neural trace, is lost rapidly, within 10 ms of stimulus offset. This may adversely influence his performance in recognizing features of speech such as formant transitions, which the speech processor translates into electrode transitions. Such tasks would presumably require rapid comparisons between successive patterns in time.

D. Recovery from forward masking

The present data were collected in a small group of subjects, and examined effects of masker amplitude as well as masker-probe distance. We did not find significant changes in recovery functions with masker level, when masker and probe were presented to the same electrode pair. However, when the probe electrode pair was moved to another location, the functions sometimes changed in shape. As we have not found evidence of changes in the shape of the recovery function on the amount of masking (i.e., on the level of the masker), these changes may not be simply due to masked thresholds being lower when the masker and probe are presented to different electrode pairs.

When masker and probe were on the same electrode, recovery functions compared across subjects showed an almost identical shape, with the exception of data obtained from subject N3, which showed a much more rapid recovery. The data were well approximated by an exponential function. Time constants calculated from the curve fits were 75, 67, and 71 ms for subjects N4, N7, and N13, and 38 ms for subject N3. The more rapid recovery in subject N3 is consistent with the large change in the amount of masking at the masker electrode location when Δt was changed from 1 to 10 ms (Fig. 6).

We also found that, when masker and probe were on the same electrode, the slope of the growth-of-masking function decreased with the probe delay, or recovery time, in all three subjects. The change was the greatest in subject N3.

In the normal-hearing system, psychophysical recovery from forward masking has been hypothesized (Smith, 1977, 1979; Harris and Dallos, 1979; Zeng *et al.*, 1991) to be related to peripheral mechanisms of neural adaptation and recovery. Models of single auditory nerve adaptation require a functioning hair cell transducer mechanism at the input stage of a diffusion-limited synaptic transfer (Smith and Brachman, 1982). In electric stimulation, hair cells and synapses can safely be assumed to be nonfunctional. Therefore, the fact that recovery from forward masking follows comparable time courses in electric and acoustic hearing indicates that peripheral synaptic mechanisms are not necessary for psychophysical forward masking (Shannon, 1990). In fact, single-unit recordings in electrically stimulated cats do not show the dramatic adaptation seen in acoustic stimulation (Javel, 1990). On the other hand, Brown and Abbas (1996), in a two-pulse forward masking experiment in patients using the Ineraid device, found similar rates of recovery in the electrically evoked compound action potential and psychophysical forward masked thresholds. We infer from the above that (a) adaptation in the auditory nerve is not related to psychophysical forward masking and (b) other auditory nerve mechanisms of recovery must contribute to forward masking phenomena.

E. Significance of the results

The psychophysical measures studied in this investigation examine perceptual effects of prior stimulation. The relationships of these measures to speech perception are still ambiguous. Subject N3's data are both qualitatively and quantitatively different from the others. However, without further data from a larger subject pool, it is not clear whether the lower sentence recognition scores of subject N3 are in any way related to the greater parameter dependence observed in his forward masked excitation patterns. If similar results are obtained in the future with other subjects, we may hypothesize that robustness to changing stimulus parameters plays a role in speech processing by the electrically stimulated auditory system.

The results presented here serve to underscore the importance of detailed parametric studies of channel interaction. It is clear that in the future, designing stimuli with realistic, time-varying parameters will help further our understanding of the large individual variability in speech recognition by cochlear implant subjects.

ACKNOWLEDGMENTS

Some of the results were presented at the 1997 Midwinter Meeting of the Association for Research in Otolaryngology. We thank the reviewers, particularly Dr. Søren Buus, for helpful comments on the manuscript. Dr. Qian-Jie Fu programmed the software used in these experiments. We thank Dr. Fan Gang Zeng for useful suggestions. Our subjects are thanked for their enthusiastic participation. J. Heller and Cochlear Corporation kindly provided us with the amplitude calibration tables for each subject. Research supported by NIDCD.

Black, R. C., and Clark, G. M. (**1980**). ''Differential electrical excitation of the auditory nerve,'' J. Acoust. Soc. Am. **67**, 868–874.

Boothroyd, A. (**1988**). *CUNY Topic Related Sentence Test Laser Videodisk* (CUNY, New York)

Brown, C. J., and Abbas, P. J. (**1996**). ''Electrically evoked whole nerve action potentials in Ineraid cochlear implant users: responses to different stimulating electrode configurations and comparison to psychophysical responses,'' J. Speech Hear. Res. **39**, 453–467.

Ehmer, R. H., and Ehmer, B. J. (**1969**). ''Frequency pattern of residual masking by pure tones measured on the Bekesy audiometer,'' J. Acoust. Soc. Am. **46**, 1445–1448.

Fu, Q. J. (**1997**). ''Speech perception in acoustic and electric hearing,'' Ph.D. dissertation, University of Southern California.

Harris, D. M., and Dallos, P. (**1979**). ''Forward masking of auditory nerve fiber responses,'' J. Neurophysiol. **42**(4), 1083–1107.

Javel, E. (**1990**). ''Acoustic and electric encoding of temporal information,'' in *Cochlear Implants. Models of the Electrically Stimulated Ear*, edited by J. M. Miller and F. A. Spelman (Springer-Verlag, Berlin).

Jesteadt, W., Bacon, S. P., and Lehman, J. R. (**1982**). ''Forward masking as a function of frequency, masker level, and signal delay,'' J. Acoust. Soc. Am. **71**, 950–962.

Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**(2, Suppl. 2), 467.

Lim, H. H., Tong, Y. C., and Clark, G. M. (**1989**). ''Forward masking patterns produced by intracochlear electrical stimulation of one and two electrode pairs in the human cochlea,'' J. Acoust. Soc. Am. **86**, 971–980.

Munson, W. A., and Gardner, M. B. (**1950**). ''Loudness patterns—A new approach,'' J. Acoust. Soc. Am. **22**, 177–190.

Shannon, R. V. (**1976**). ''Two-tone unmasking and suppression in a forward-masking situation,'' J. Acoust. Soc. Am. **59**, 1460–1470.

Shannon, R. V. (**1983**). ''Multichannel electrical stimulation of the auditory nerve in man. II. Channel interaction,'' Hearing Res. **12**, 1–16.

Shannon, R. V. (**1990**). ''Forward masking in patients with cochlear implants,'' J. Acoust. Soc. Am. **88**, 741–744.

Shannon, R. V., Adams, D. D., Ferrel, R. L., Palumbo, R. L., and Grandgenett, M. (**1990**). ''A computer interface for psychophysical and speech research with the Nucleus cochlear implants,'' J. Acoust. Soc. Am. **87**, 905–907.

Smith, R. L. (**1977**). ''Short-term adaptation in single auditory nerve fibers: some poststimulatory effects,'' J. Neurophysiol. **40**(5), 1098–1112.

Smith, R. L. (**1979**). ''Adaptation, saturation and physiological masking in single auditory-nerve fibers,'' J. Acoust. Soc. Am. **65**, 166–178.

Smith, R. L., and Brachman, M. L. (**1982**). ''Adaptation in auditory-nerve fibers: a revised model,'' Biol. Cybern. **44**, 107–120.

Tong, Y. C., and Clark, G. M. (**1986**). ''Loudness summation, masking, and temporal interaction for sensations produced by electric stimulation of two sites in the human cochlea,'' J. Acoust. Soc. Am. **79**, 1958–1966.

Zeng, F.-G., Turner, C. W., and Relkin, E. M. (**1991**). ''Recovery from prior stimulation II: Effects upon intensity discrimination,'' Hearing Res. **5592**, 223–230.

Zwislocki, J. J., and Pirodda, E. (**1952**). ''On the adaptation, fatigue and acoustic trauma of the ear,'' Experientia **8**, 279–284.

The masking-level difference in low-noise noise

Joseph W. Hall III and John H. Grose
The Division of Otolaryngology/Head & Neck Surgery, University of North Carolina Medical School, Chapel Hill, North Carolina 27599-7070

William Morris Hartmann
Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824

(Received 16 November 1996; revised 22 August 1997; accepted 15 January 1998)

In experiment 1 NoSo and NoSπ thresholds for a 500-Hz pure tone were obtained in a low-fluctuation masking noise and a high-fluctuation masking noise for six normal-hearing listeners. The noise bandwidth was 10 Hz. In agreement with previous investigations, the NoSo thresholds were lower in low-fluctuation noise than in high-fluctuation noise. For three listeners, NoSπ thresholds were similar for the two types of noise, while for the other three listeners, NoSπ thresholds were *higher* for low-fluctuation noise than for high-fluctuation noise. In experiment 2, the masker was created by amplitude modulating a 500-Hz pure tone by a 0–10-Hz low-pass noise. The degree of masker fluctuation was controlled by adjusting the average modulation depth (100%, 63%, 40%, and 25%). The signal was a 10-Hz-wide noise centered on 500 Hz. Results were similar to those of experiment 1: for the NoSo conditions, signal detection improved with decreasing degree of fluctuation, and for NoSπ conditions, the results were subject dependent. For three listeners, NoSπ thresholds were again similar in the two types of noise, while for the other three listeners, NoSπ thresholds were again higher in low-fluctuation noise than in high-fluctuation noise. The results showed that a high degree of masker fluctuation sometimes facilitates NoSπ detection. It is possible that the binaural detection mechanism utilizes the relatively good signal-to-noise ratios that occur in the low power or ''dip'' regions of fluctuating masker waveforms. [S0001-4966(98)01305-8]

PACS numbers: 43.66.Dc, 43.66.Pn [RHD]

INTRODUCTION

Previous experiments using low-noise (Pumplin, 1985) have investigated the effects of masker amplitude fluctuation on monaural signal detection (Hartmann and Pumplin, 1988). Conditions have contrasted detection in noise having a random phase relation among components (where noise fluctuation is generally high) with detection in noise having components whose phases are selected such that the noise fluctuation is relatively low (low-noise noise). Results (Hartmann and Pumplin, 1988) indicate that monaural signal detection is better for low-noise narrow-band noise than for random-phase narrow-band noise. Similar effects were reported by Margolis and Small (1974) for a low-fluctuation noise produced by frequency modulating a sinewave by a thermal noise. The low-noise noise result is consistent with the notion that random variation in the energy of the masker is detrimental to the detection of the signal (e.g., Bos and de Boer, 1966).

There is reason to believe that studies manipulating noise fluctuation may also be of interest in conditions of binaural unmasking. For example, it is possible that results from low- and high-fluctuation noise conditions may have bearing upon models of the masking-level difference (MLD) (Hirsh, 1948). Specifically, different binaural models may be associated with different predictions concerning the effect of masker fluctuation. In the equalization–cancellation (EC) model of Durlach (1963), the stimulus waveforms at the two ears undergo auditory filtering, level/time equalization, and subtraction. For the sake of simplicity, we will consider only NoSo and NoSπ stimulation. With perfect equalization, the subtraction process would completely eliminate an No masker, and masked signal detection threshold would be similar to that obtained in quiet. However, the model assumes that there is both time and amplitude ''jitter'' in the process, resulting in an No noise being reduced in level rather than eliminated. In contrast, the subtraction process for the Sπ signal essentially results in an in-phase addition of the signal, with a consequent boost in signal level. Thus, for NoSπ detection, the signal-to-noise ratio is effectively increased by the EC process. For an So condition, the subtraction process results in a reduction in signal level. However, because the same jitter parameters hold for both signal and noise, the noise and signal are reduced by similar amounts in the NoSo case. Thus, for NoSo, the signal-to-noise ratio is essentially the same at the EC input and the EC output.

In the EC model, both NoSo and NoSπ detection reduce to the problem of detecting signal energy in noise. If it is supposed that the decision statistic for detection is similar for the NoSo and NoSπ cases, then the effects of low-noise noise should be similar for NoSo and NoSπ detection. Because power fluctuations are relatively minor in low-noise noise, both NoSo and NoSπ thresholds are expected to be relatively lower than in noise with prominent power fluctuations. Because effects of low-noise noise are expected to be similar for NoSo and NoSπ detection, the MLD would then be expected to be similar for low-noise noise and random-phase noise.

In contrast, the MLD is expected to be relatively smaller for low-noise noise conditions from the standpoint of cross-correlation models (Jeffress, 1948; Colburn, 1973; Stern and Colburn, 1978; Stern and Trahiotis, 1992) of the MLD. One

reason for this prediction is related to the fact that, in cross-correlation models, the decision statistic for binaural detection is proposed to be radically different from that associated with monaural, or NoSo, detection. As in the EC model, monaural or NoSo detection is usually hypothesized to be based upon an analysis of stimulus energy. Thus, a relatively low NoSo threshold is predicted for a low-noise noise masker. In contrast, binaural (e.g., NoSπ) detection is assumed to be based upon a cross-correlation process. In the physiological realization of such a process, the decision statistic is presumed to be related to a change in the activity pattern of neurons responding to different degrees of interaural delay. If it is assumed that interaural correlation will be just as high for a random-phase No noise as for a low-noise No noise, there is no reason to expect that the relatively high degree of fluctuation associated with random-phase noise should be deleterious for NoSπ detection. According to this interpretation, the MLD should be smaller in low-noise noise because the NoSo threshold is expected to be lower in low-noise noise than in random-phase noise, but the NoSπ threshold is not expected to be lower in low-noise noise than in random-phase noise.

A second reason that MLDs might be smaller in low-noise noise is directly related to the depth of fluctuation of the masker envelope. Masker dips (where noise power is low for short temporal epochs) are associated with a relatively high signal-to-noise ratio. There is evidence that monaural detection for a signal presented in a band of random noise does not appear to be able to benefit from the good signal-to-noise ratios associated with masker dips (Buus *et al.*, 1996). This may be because it is difficult for the auditory system to determine whether a signal occurring in a masker dip is actually a signal or simply a random fluctuation of the masker. A cross-correlation mechanism offers a potential solution to this problem in that an $S\pi$ signal occurring in an No masker dip will result in interaural correlation quite different from that associated with the masker alone. If such short-duration changes in interaural correlation are effective in cueing detection, $S\pi$ thresholds might actually be better in random-phase noise (where dips are relatively deep) than in low-noise noise (where dips are relatively shallow).

In the course of this investigation, we learned of a similar on-going study of the MLD in low-noise noise maskers, with similar results, by David Eddins and Laura Barber. The manuscripts resulting from his work and the work in our laboratory have therefore been submitted as companion papers.

I. EXPERIMENT 1: MLD IN LOW-NOISE NOISE

A. Method

1. Subjects

Subjects were six listeners with normal hearing, aged between 24 and 44 years. All had previous experience listening in MLD and CMR paradigms.

2. Stimuli

All noise bands were centered on 500 Hz and had a bandwidth of 10 Hz. Each band was composed of 11 equal-amplitude sinusoidal components. For the high-fluctuation noise, the phase relation among the 11 components was random. For the low-noise noise, the phase relations were chosen by the low-fluctuation noise algorithm (Pumplin, 1985). For each type of noise (random-phase or low-noise noise), ten different noise files were generated. One of the ten files was selected randomly from interval to interval in order to reduce possible effects associated with a particular frozen noise sample (Hanna and Robinson, 1985). Each noise file was 4096 samples long, using a sampling rate of 4096 Hz. Noise was delivered via a 12-bit Data Translation D/A converter. The signal was a 400-ms, 500-Hz pure tone, shaped with a 50-ms squared-cosine rise/fall. The masker was gated on (50-ms squared-cosine) 100 ms before the signal was gated on, and was gated off simultaneously with the signal. The masker was always interaurally in phase (No) and the signal was either interaurally in phase (So) or interaurally out of phase (Sπ).

3. Masker statistics

The average kurtosis for the ten low-noise masker waveforms was 1.655 (compared to 3.0 for Gaussian noise). Figure 1 of Hartmann and Pumplin (1988) shows that 1.655 is well outside the expected distribution for kurtosis for random-phase noises. The average crest factor for the ten low-noise noise maskers was 1.710, compared with an expected value of 2.76 for random-phase noises (Hartmann and Pumplin, 1988, Eq. C4). In contrast with the fluctuations in signal power, the low-noise noise maskers showed no special character in their instantaneous phases. Phase variation caused by the center frequency was extracted and the residual phase variation (Hartmann, 1997, Eq. 18.6) was studied. Neither the overall phase variance nor any time-dependent behavior indicated differences between low-noise noise and random-phase maskers. This analysis is relevant because the rate of change in interaural difference cues for an NoSπ stimulus depends upon the changes in the instantaneous phase of the masker. The analysis suggests no obvious differences between low-noise noise and random-noise maskers in this regard.

4. Procedure

Thresholds were determined using a three-alternative forced choice (3AFC) three-down, one-up adaptive procedure, estimating the 79.4% detection threshold (Levitt, 1971). An initial step-size of 8 dB was reduced to 4 dB after two reversals, and further reduced to 2 dB after two more reversals. A threshold run was stopped after 12 reversals, and the average of the last 8 reversals was taken as the threshold for a run. Four threshold runs were averaged to compute the final threshold, unless the range of the runs was greater than 3 dB; in that case, a fifth run was obtained and included in the average. Each trial was preceded by a 300-ms warning light. Each interval was marked by a 400-ms light. The interstimulus interval was 300 ms. Visual feedback was provided after each response. The stimuli were delivered binaurally by means of Sony MDR V6 earphones.

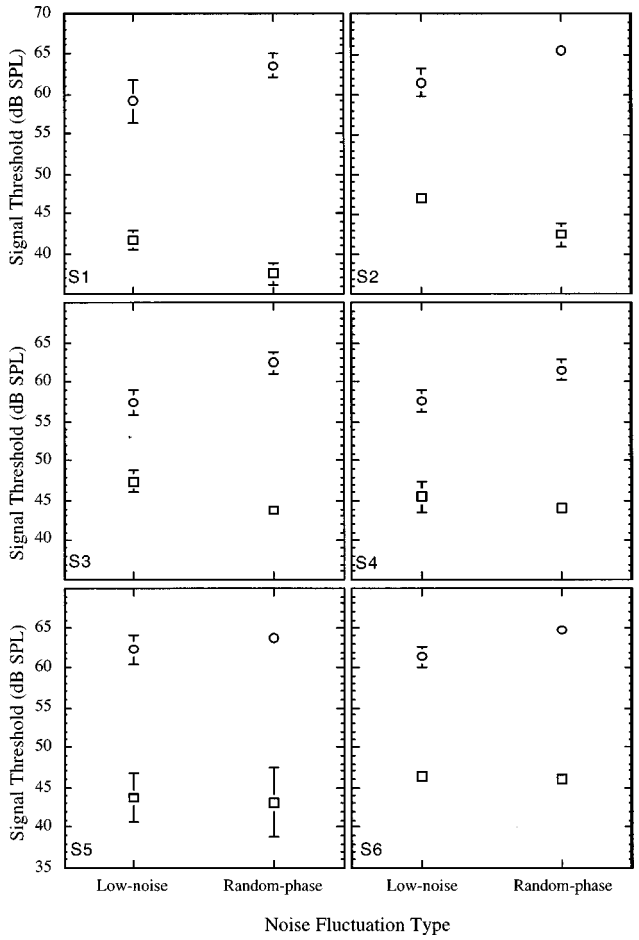


FIG. 1. The NoSo (circles) and NoSπ (squares) thresholds for the six listeners, in low-noise noise (Low) and random noise (High). Error bars show plus and minus one standard deviation. Standard deviations are not shown when they are smaller than the data symbol.

B. Results and discussion

Threshold data for the six subjects are summarized in Fig. 1. Individual rather than average data are shown because there were different patterns of results across subjects. The most consistent finding, of course, was that NoSπ thresholds were lower than NoSo thresholds. A second consistent finding was that, for NoSo conditions, signal thresholds were lower (by about 5 dB) in low-noise noise than in random-phase noise. This result is consistent with that reported by Hartmann and Pumplin (1988). However, results for low-noise noise and random-phase noise were not consistent across subjects for NoSπ detection. In NoSπ conditions, subjects 1, 2, and 3 showed higher thresholds for low-noise noise than for random-phase noise. As can be seen in Table I, these subjects had considerably larger MLDs (by about 9 dB) in random-phase noise than in low-noise noise. For these subjects, MLDs were relatively small in low-noise noise partly because the NoSo thresholds were relatively low in low-noise noise, and partly because the NoSπ thresholds were relatively high in low-noise noise. For subjects 4, 5, and 6, NoSπ thresholds were similar between low-noise noise and random-phase noise (see Fig. 1). For these subjects, MLDs were again smaller in low-noise noise than in random-phase noise, but only by about 2 to 5 dB. In these subjects, the smaller MLD in low-noise noise than in random-phase noise was due primarily to the fact that the NoSo thresholds were lower in low-noise noise than in random-phase noise. Across all subjects, the average MLD for random-phase noise was 20.7 dB, and the average MLD for low-noise noise was 14.5 dB (see Table I).

TABLE I. Individual and mean MLDs (dB) for random phase noise and low-noise masking noise. Data are for experiment 1.

Noise type	S1	S2	S3	S4	S5	S6	Mean
High fluctuation	26.0	23.0	18.7	17.4	20.6	18.7	20.7
Low fluctuation	17.3	14.4	9.8	12.7	18.5	15.1	14.5

The most striking new result of this experiment was the higher NoSπ threshold in low-noise noise than in high-fluctuation noise for three of the six subjects. This result suggests that a cue related to a relatively high degree of masker fluctuation can actually aid binaural signal detection. In order to examine the generality of this result, we performed a second experiment employing a different method for generating maskers with varying degrees of fluctuation.

II. EXPERIMENT 2: MASKING NOISE CREATED BY AMPLITUDE MODULATING A PURE TONE BY A LOW-PASS NOISE

A. Method

1. Subjects

The subjects were the same as those who participated in the first experiment.

2. Stimuli and procedure

The masker was created by multiplying a 500-Hz pure tone by a DC-shifted low-pass noise (0–10 Hz). The masking waveform was defined by

$$A(t)=A_0[1+N(t)]\cos(\Omega_0 t),$$

where t is time, A_0 is the amplitude of the 500-Hz carrier, $N(t)$ is the low-pass noise waveform, and Ω_0 is 2π times the carrier frequency. A similar method was used in a MLD study by Grantham and Robinson (1977). However, in the Grantham and Robinson study, the modulator was a bandpass noise from 43 to 77 Hz, and modulation of the 500-Hz carrier by this bandpass noise resulted in a modulated masker with spectral components extending from approximately 423 to 577 Hz. In the present study, the masker components were restricted to frequencies between approximately 490 and 510 Hz. The degree of fluctuation was adjusted by varying the depth of modulation. Average modulation depths (in terms of percent of modulation) were 100%, 63%, 40%, and 25%, where 100% corresponds to a noise waveform $[N(t)]$ with rms value of approximately 0.707. The level of the masker (carrier plus sidebands) was held constant at approximately 59 dB SPL. The signal was a 10-Hz-wide band of noise centered on 500 Hz. An inverse fast Fourier transform (FFT) incorporating a sampling rate of 11.025 kHz and buffer size of 2^{17} discrete points was used to create the signal. This resulted in a stimulus with approximately 0.08-Hz frequency resolution that, upon cyclical output, had an overall period-

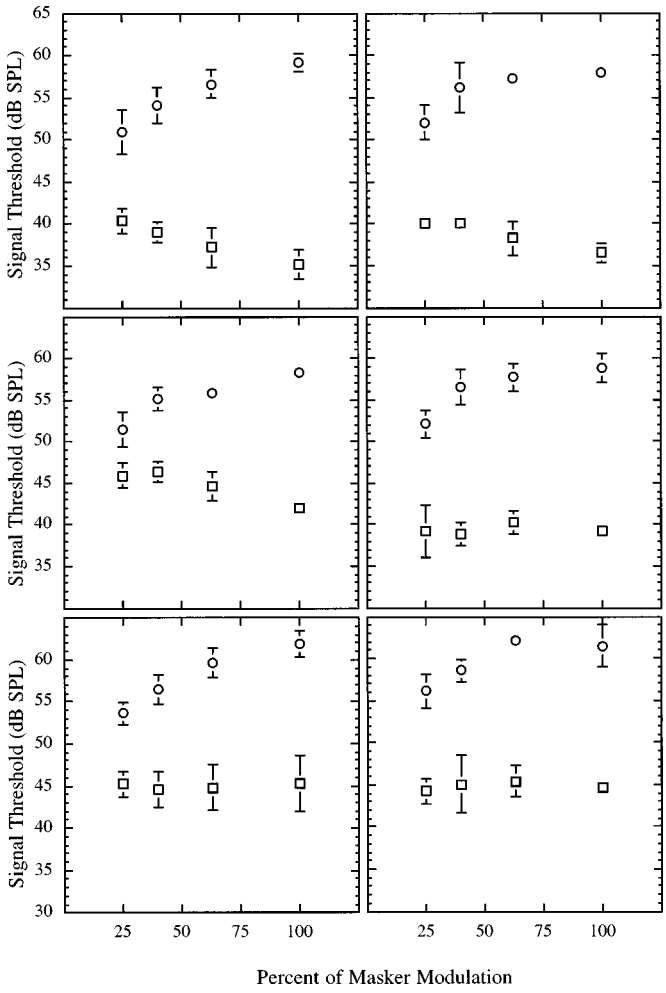


FIG. 2. The NoSo (circles) and NoSπ (squares) thresholds for the six listeners, as a function of the average percent of modulation of the masker. Error bars show plus and minus one standard deviation. Standard deviations are not shown when they are smaller than the data symbol.

icity of approximately 11.89 s. The noise signal was played continuously through a 20-bit digital-to-analog converter and low-pass filtered at 3 kHz, and was gated via a Tucker-Davis SW2 gate. A noiseband was used as a signal, rather than a pure tone, so that the phase between the carrier of the amplitude-modulated masker would be random with respect to the phase of the signal. The signal was 400 ms in duration, and had 50-ms squared-cosine rise/fall. The masker was presented continuously. The masker was always interaurally in phase (No) and the signal was either interaurally in phase (So) or interaurally out of phase (Sπ). The threshold procedure was the same as that used in the first experiment.

B. Results and discussion

Threshold data for the six subjects are summarized in Fig. 2. In several respects, the pattern of results was similar to that obtained in experiment I. As in experiment 1, NoSo thresholds consistently improved as the masker fluctuation decreased. The average NoSo threshold was 59.6 dB for 100% modulation, but improved to 52.7 dB for 25% modulation. As in experiment 1, there were individual differences in the effect of masker fluctuation for NoSπ detection: subjects 1, 2, and 3 again showed higher thresholds with decreases in masker fluctuation, whereas subjects 4, 5, and 6 again showed relatively stable NoSπ thresholds across the different conditions of masker fluctuation. Across all subjects, the average MLDs were 19.1, 16.4, 13.8, and 10.2 dB for 100%, 63%, 40%, and 25% modulation, respectively (see Table II).[1]

TABLE II. Individual and mean MLDs for maskers varying in average percent of modulation. Data are for experiment 2.

Percent modulation	S1	S2	S3	S4	S5	S6	Mean
100	23.9	21.4	16.2	19.7	16.6	16.9	19.1
63	19.4	18.9	11.2	17.4	14.8	16.7	16.4
40	15.0	16.1	8.7	17.7	12.0	13.6	13.8
25	10.5	11.9	5.6	12.9	8.4	11.9	10.2

Again, the most striking finding was that for subjects 1, 2, and 3, NoSπ thresholds increased as a function of decreased masker fluctuation. This result again suggests that a cue related to a relatively high degree of masker fluctuation can aid binaural signal detection. Subjects 4, 5, and 6 again showed little change in the NoSπ threshold with changes in masker modulation depth.

III. GENERAL DISCUSSION

It was pointed out in the Introduction that a straightforward EC model interpretation would predict that NoSo and NoSπ thresholds would depend similarly on masker fluctuation statistics. The findings for the Sπ thresholds did not agree with this prediction. The present findings were in better agreement with a cross-correlation mechanism. For a cross-correlation mechanism, it was predicted that detection for NoSπ would would either not vary as a function of the degree of masker fluctuation, or that Sπ detection might be worse in low-fluctuation noise than in high-fluctuation noise. In three of the listeners tested, Sπ thresholds were similar in low-fluctuation noise and in high-fluctuation noise, and in the three other listeners, NoSπ thresholds were *lower* in high-fluctuation noise than in low-fluctuation noise. The latter result is in direct contrast to the situation for monaural detection, where high noise fluctuation results in relatively poor detection (Bos and de Boer, 1966). It would appear that, at least in some listeners, a high degree of noise fluctuation is favorable for signal detection. It seems likely that this effect is related to low-energy or ''dip'' regions in the masker. In a masker having a relatively low degree of fluctuation, the signal-to-noise ratio is relatively constant over the duration of a signal. However, in a masker having a relatively high degree of fluctuation, the signal-to-noise ratio will be relatively poor in masker peak regions, but will be relatively good in masker dip regions. It is possible that the relatively short but large binaural cues that exist during masker dips are effective in cuing NoSπ detection.

The above interpretation is consistent with Isabelle's (1995) account of NoSπ data in a frozen noise experiment. As in previous experiments by Gilkey and his colleagues (Gilkey *et al.*, 1985; Gilkey and Robinson, 1986), Isabelle attempted to obtain information about the cues accounting for binaural detection by examining performance for particu-

lar digitized samples of noise alone and signal-plus-noise. While Isabelle found that NoSπ performance was not correlated significantly with stimulus energy, modest, albeit significant, correlations were found for some decision statistics based upon variability of the interaural time and/or interaural intensity cues. Interestingly, the highest correlations were found for instantaneous interaural time differences. Isabelle suggested that the auditory system may utilize the infrequent but large interaural time differences occurring near the minima of the masker envelope, noting that ''this strategy may be likened to 'listening in the valleys of the noise.' ''

One feature of binaural analysis that might be seen as in conflict with the notion of relatively fine temporal resolution has been termed ''binaural sluggishness'' (Grantham and Wightman, 1979). This term refers to the fact that listeners are relatively insensitive to dynamic variation in interaural difference cues, in that changes in binaural cues occurring at rates of more than a few Hz are not heard as movement in perceived location, but, instead as a ''blur'' or as diffuseness of location (Blauert, 1972; Grantham and Wightman, 1978). Grantham and Wightman (1979) reported data consistent with an interpretation that binaural sluggishness may be associated with small MLDs. They investigated the detectability of a brief Sπ tone burst presented in a noise masker whose interaural phase varied sinusoidally between 1.0 and -1.0. When the short signal was presented at a time when the masker had interaural correlation near 1.0, a MLD occurred, provided that the sinusoidal modulation of interaural masker phase was very slow (e.g., 0.5 Hz). However, essentially no MLD occurred when the modulation rate was raised to only 4 Hz. This result is consistent with an interpretation that binaural sluggishness prevented the auditory system from taking advantage of the short temporal epochs when the stimulus was in NoSπ configuration. However, it reasonable to assume that the use of binaural signal detection information during short temporal epochs is poor only when both the masker and signal-plus-masker contain dynamically varying interaural cues (as was the case in the Grantham and Wightman study). When the masker has a stable interaural phase, it is likely that the binaural system can take advantage of binaural detection information in short temporal epochs. Indeed, the robust MLDs that occur for short Sπ signals in stable No maskers (Blodgett *et al.*, 1958; Green, 1966; Robinson and Trahiotis, 1972; Grantham and Wightman, 1979) provide strong evidence that the binaural system can make good use of binaural detection information occurring in a short temporal epoch. We therefore do not view binaural sluggishness as incompatible with a detection process involving relatively fast sampling of interaural correlation.

ACKNOWLEDGMENTS

Constantine Trahiotis and Huanping Dai made several helpful suggestions concerning the interpretation of the results. We thank Steve Colburn for helpful suggestions and for bringing Scott Isabelle's dissertation to out attention. Raymond Dye and two anonymous reviewers made helpful suggestions on a previous version of this manuscript. We appreciate technical assistance provided by Saju Joy. This research was supported by grants from NIH (NIDCD R01 DC00397 and NIDCD R01 DC00181) and AFOSR.

[1]It should be mentioned that during the review process for this manuscript Kohlrausch *et al.* (1997) published results comparing MLD data for random-phase and low-fluctuation noise. They used a 1000-Hz center frequency and examined masking bandwidths from 5 to 100 Hz. For the NoSo conditions, their average data (four subjects) indicated lower thresholds for low-fluctuation noise, and for the NoSπ conditions, their average data indicated similar thresholds between random-phase and low-fluctuation noise (similar to our data for subjects 4, 5, and 6).

Blauert, J. (**1972**). ''On the lag of lateralization caused by interaural time and intensity differences,'' Audiology **11**, 265–270.

Blodgett, H. C., Jeffress, L. A., and Taylor, R. W. (**1958**). ''Relation of masked threshold to signal-duration for various interaural phase-combinations,'' Am. J. Psychol. **71**, 283–290.

Bos, C. E., and de Boer, E. (**1966**). ''Masking and discrimination,'' J. Acoust. Soc. Am. **39**, 708–715.

Buus, S., Zhang, L., and Florentine, M. (**1996**). ''Stimulus-driven, time-varying weights for Comodulation Masking Release,'' J. Acoust. Soc. Am. **99**, 2288–2297.

Colburn, H. S. (**1973**). ''Theory of binaural interaction based on auditory nerve data. I. General strategy and preliminary results on interaural discrimination,'' J. Acoust. Soc. Am. **54**, 1458–1470.

Durlach, N. I. (**1963**). ''Equalization and cancellation theory of binaural masking-level differences,'' J. Acoust. Soc. Am. **35**, 1206–1218.

Gilkey, R. H., and Robinson, D. E. (**1986**). ''Models of auditory masking: a molecular psychophysical approach,'' J. Acoust. Soc. Am. **79**, 1499–1510.

Gilkey, R. H., Robinson, D. E., and Hanna, T. E. (**1985**). ''Effects of masker waveform and signal-to-masker phase relation on diotic and dichotic masking by reproducible noise,'' J. Acoust. Soc. Am. **78**, 1207–1219.

Grantham, D. W., and Robinson, D. E. (**1977**). ''Role of dynamic cues in monaural and binaural signal detection,'' J. Acoust. Soc. Am. **61**, 542–551.

Grantham, D. W., and Wightman, F. L. (**1979**). ''Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation,'' J. Acoust. Soc. Am. **65**, 1509–1517.

Grantham, D. W., and Wightman, F. L. (**1978**). ''Detectability of varying interaural temporal differences,'' J. Acoust. Soc. Am. **63**, 511–523.

Green, D. M. (**1966**). ''Interaural phase effects in the masking of signals of different durations,'' J. Acoust. Soc. Am. **39**, 720–724.

Hanna, T. E., and Robinson, D. E. (**1985**). ''Phase effects for a sine wave masked by reproducible noise,'' J. Acoust. Soc. Am. **77**, 1129–1140.

Hartmann, W. M. (**1997**). *Signals, Sound, and Sensation* (AIP, Springer-Verlag, New York).

Hartmann, W. M., and Pumplin, J. (**1988**). ''Noise power fluctuations and the masking sine signals,'' J. Acoust. Soc. Am. **83**, 2277–2289.

Hirsh, I. J. (**1948**). ''Influence of interaural phase on interaural summation and inhibition,'' J. Acoust. Soc. Am. **20**, 536–544.

Isabelle, S. (**1995**). ''Binaural detection performance using reproducible stimuli,'' Ph.D. dissertation, Boston University.

Jeffress, L. A. (**1948**). ''A place theory of sound localization,'' J. Comp. Physiol. Psychol. **41**, 35–39.

Kohlrausch, A., Fassel, R., van der Heijden, M., Kortekaas, S., van der Par, S., and Oxenham, A. J. (**1997**). ''Detection of tones in low-noise noise: Further evidence for the role of envelope fluctuations,'' Acustica **83**, 659–669.

Levitt, H. (**1971**). ''Transformed up–down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Margolis, R. H., and Small, Jr., A. M. (**1974**). ''Masking with narrow-band FM noise,'' J. Acoust. Soc. Am. **56**, 692–694.

Pumplin, J. (**1985**). ''Low-noise noise,'' J. Acoust. Soc. Am. **78**, 100–104.

Robinson, D. E., and Trahiotis, C. (**1972**). ''Effects of signal duration and masker duration on detectability under diotic and dichotic listening conditions,'' Percept. Psychophys. **12**, 333–334.

Stern, R., and Colburn, H. (**1978**). ''Theory of binaural interaction based on auditory-nerve data IV,'' J. Acoust. Soc. Am. **64**, 127–140.

Stern, R. M., and Trahiotis, C. (**1992**). ''The role of consistency of interaural timing over frequency in binaural lateralization,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, K. Horner, and L. Demany (Pergamon, Oxford).

The influence of stimulus envelope and fine structure on the binaural masking level difference[a)]

David A. Eddins and Laura E. Barber
Psychoacoustics Laboratory, Department of Speech and Hearing Sciences, Indiana University, Bloomington, Indiana 47405

(Received 16 November 1996; revised 10 September 1997; accepted 15 January 1998)

A masking level difference (MLD) paradigm was used to investigate the influence of stimulus envelope and stimulus fine-structure characteristics on monaural and binaural hearing. The degree of masker envelope fluctuation was manipulated by selecting narrow-band noises (50 Hz) on a continuum of values of the normalized fourth moment of the envelope. The noises were specified as low-noise noise (LNN), medium-noise noise (MNN), and high-noise noise (HNN). Fine-structure cues were studied by measuring thresholds at 500 and 4000 Hz, regions in which the availability of such cues to the auditory system differ substantially. In addition, thresholds were measured for Gaussian noise maskers (GN) and for maskers having a flat magnitude spectrum, termed equal-magnitude noise (EMN) maskers. The results indicated lower NoSo thresholds for LNN than for the other four masker types. Furthermore, there were no differences in threshold for maskers having moderate and high degrees of envelope fluctuation (MNN and HNN). The NoSπ thresholds were not significantly different across masker type and were characterized by large individual differences among the seven listeners. The results are considered in relation to models of monaural and binaural processing. Consistent with previous reports, the results indicate that binaural detection depends on interaural differences in the stimulus envelope and fine structure at low frequencies and changes in the envelope at high frequencies. [S0001-4966(98)01405-2]

PACS numbers: 43.66.Ba, 43.66.Dc, 43.66.Mk [RHD]

INTRODUCTION

Since the duplex theory of sound localization was first suggested by Lord Rayleigh (1907), it has become abundantly clear that binaural auditory processing relies on interaural time and level differences to differing degrees across stimulus frequency. Consistent with this theory, results obtained in binaural masking level difference (MLD) paradigms using broadband masking noise indicate that the masking release is maximal at low signal frequencies and diminishes with increased frequency. Likewise, the lateralization of sinusoids and clicks based on interaural timing differences is limited to low stimulus frequencies (e.g., Klumpp and Eady, 1956; Yost *et al.*, 1971). These results are often explained in terms of the reduced accuracy in temporal coding at high frequencies due to the degradation in neural phase locking to the temporal fine structure with increasing stimulus frequency (e.g., Durlach, 1964).

It is clear, however, that interaural disparities conveyed by slowly fluctuating temporal envelopes can contribute to binaural processing at high stimulus frequencies in binaural detection and lateralization tasks (e.g., Klumpp and Eady, 1956; McFadden and Pasanen, 1976, 1978; Zurek and Durlach, 1987; Bernstein and Trahiotis, 1992a, 1992b). Thus, while phase locking to stimulus fine structure may provide binaural cues at low frequencies, interaural differences in temporal envelope may provide interaural cues at both low and high stimulus frequencies (Bernstein and Trahiotis, 1985a, 1985b). The present experiment was designed to explore the contributions of stimulus envelope and stimulus fine-structure cues to monaural and binaural detection using a MLD paradigm.

In the classic MLD paradigm, masked signal detection measured under homophasic (NoSo) conditions is compared to detection under antiphasic (NoSπ) conditions. For homophasic conditions, like monaural detection, thresholds are similar at low and high signal frequencies. Thus stimulus fine structure likely plays little role in these conditions. Changes in the masker envelope fluctuation, however, may have a substantial influence on detection in homophasic conditions (e.g., Hartmann and Pumplin, 1988). For antiphasic conditions, masked signal detection may be influenced by both changes in stimulus fine structure and envelope fluctuations, as noted above. One means of assessing the role of stimulus fine structure in antiphasic detection is to make measurements at relatively low (e.g., 500 Hz) and high (e.g., 4000 Hz) frequencies, where the availability of fine-structure information differs widely. To gain insight into the importance of the stimulus envelope in antiphasic detection, as well as homophasic detection, the present experiments employed masker types that differed in degree of envelope fluctuations. The first masker type was narrow-band, Gaussian noise, which is characterized by pronounced envelope fluctuations. The second type of masker was characterized by *minimal* envelope fluctuations.

To minimize envelope fluctuations, low-noise noise maskers were used, the details of which have been described

[a)]Portions of this research were presented at the 129th meeting of the Acoustical Society of America, Indianapolis, Indiana, May 1996 [J. Acoust. Soc. Am. **99** 2471(A) (1996)].

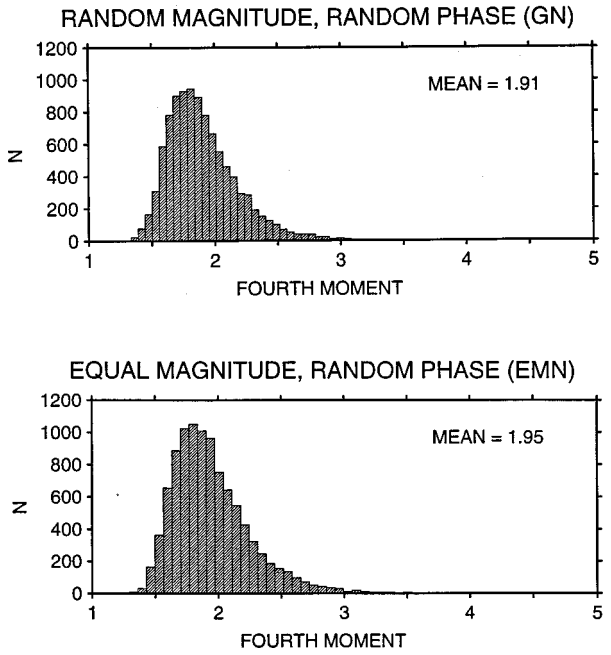


FIG. 1. Distribution of normalized envelope fourth moments obtained in 10 000 random samples of Gaussian noise (upper panel) and equal-magnitude noise (lower panel).

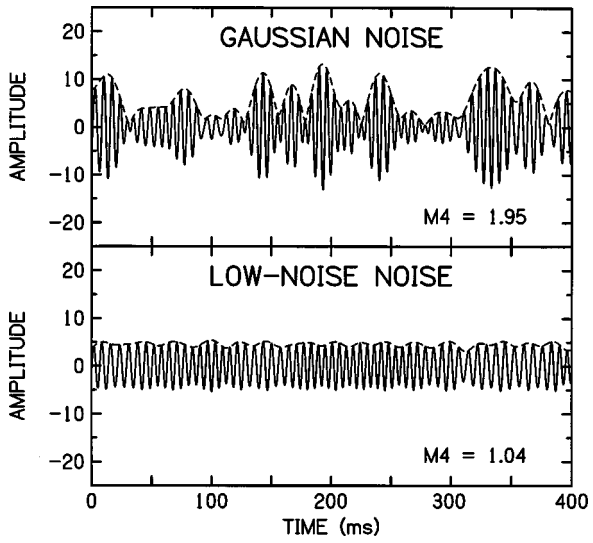


FIG. 2. The upper panel shows a sample of Gaussian noise having a normalized envelope fourth moment approximately equal to the mean of the distribution of fourth moments (see Fig. 1). The lower panel shows the sample of low-noise noise used in the present experiment. The waveforms are shown by the solid curves and the stimulus envelope by the dashed curves. The noise bandwidth is 50 Hz. The fourth moment of the envelope is denoted M4.

by Pumplin and Hartmann in a series of papers (Pumplin, 1985; Hartmann, 1987; Hartmann and Pumplin, 1988, 1991a). Briefly, a sample of low-noise is obtained by minimizing the fluctuations in the power of the waveform. Beginning with a magnitude spectrum, a directed search of the multidimensional space of possible component phases results in the set of phases that produces minimal variance in the power of the waveform. In practice, however, use of the algorithm becomes increasingly computationally intensive with increases in the number of spectral components comprising the stimulus. Fortunately, Hartmann and Pumplin (1991b) published a series of tables providing the phase spectra necessary to generate a number of different low-noise noises. As noted by Hartmann (1987), for psychoacoustic consideration, it may be more appropriate to consider the normalized fourth moment of the *envelope* rather than the fourth moment of the *waveform*, as the power fluctuations of the waveform contain frequencies too high to be perceptually relevant. For narrow-band noise, however, the normalized fourth moment of the waveform, W, and envelope, Y, are related in a simple way, $Y=2/3W$ (Hartmann and Pumplin, 1988). The term fourth moment will be used hereafter to refer to the normalized fourth moment of the envelope, $Y=\langle E\rangle^4/(\langle E^2\rangle)^2$, where E is the stimulus envelope.[1]

To facilitate the comparison of thresholds obtained with Gaussian noise (GN) and low-noise noise (LNN) maskers, it is helpful to understand the range of possible fourth moments under consideration. The upper panel of Fig. 1 illustrates the distribution of fourth moments obtained in 10 000 random samples of narrow-band (50 Hz) GN composed of 21 sinusoidal components with component magnitudes selected from a Rayleigh distribution and phases selected from a uniform distribution (0 to 2π). Note that the distribution of fourth moments is skewed toward the upper edge (range $=1.31$ to 4.35, mean$=1.91$, median$=1.86$, s.d.$=0.29$). A sample of noise having a fourth moment close to the mean of this distribution is shown in the upper panel of Fig. 2 (solid curve). The envelope of this waveform is shown by the dashed curve and has a fourth moment of 1.95.

Because the phase search necessary to generate low-noise noises was prohibitively long for 21-component stimuli, we chose to use a sample of noise taken from the tables of Hartmann and Pumplin (1991b). In this sample, each component of the magnitude spectrum had the same value and component phases were chosen to minimize the fourth moment. As Hartmann (1987) has shown, the distribution of waveform sample values for equal-magnitude noises (EMN) differs from noises composed of Rayleigh-distributed magnitudes for a very small number of components (<10), but is nearly indistinguishable for larger numbers of components. The same is true for the distribution of fourth moments for EMN (Fig. 1, lower panel). Note that this distribution is nearly identical to the distribution for Gaussian noise shown in the upper panel (range$=1.29$ to 4.62; mean$=1.95$; median$=1.89$; s.d.$=0.32$). The LNN waveform (solid curve) and envelope (dashed curve) are shown in the lower panel of Fig. 2. The fourth moment of the LNN sample used here, 1.04, is not likely to occur in random sampling from this distribution.

By manipulating the nature and availability of stimulus envelope and fine-structure cues to the auditory system, a better understanding of the dependence of masked threshold on various monaural and binaural cues can be achieved. Eddins (1994, 1996) has demonstrated the usefulness of LNN in identifying stimulus envelope contributions to complex detection tasks such as profile analysis and comodulation masking release (CMR). In those experiments, the choice of reference threshold was shown to have a strong influence on the estimated masking release. In the present experiment the

homophasic (NoSo) threshold will serve as the reference when computing the MLD. Since lower thresholds are typically observed with LNN than with GN maskers (e.g., Hartmann and Pumplin, 1988), it appears likely that any differences in MLD magnitude obtained with these two masker types will reflect, at least in part, changes in the NoSo (reference) condition. Specifically, with respect to homophasic detection, and by inference, monaural detection, changes in the stimulus fine structure are not expected to yield substantial differences in masked threshold; however, minimizing the degree of envelope fluctuation should result in lower masked thresholds.

It is not entirely clear how NoSπ thresholds might differ with the two maskers. While the addition of a sinusoidal signal to either GN or LNN will produce changes in the envelope relative to the masker-alone stimulus, there are two important differences in these masker conditions. First, the amount of fluctuation in the maskers themselves differs, providing very different standard or reference stimuli. Second, the envelope fluctuations will change in opposite directions with the addition of a sinusoidal signal to the two noise types. The degree of envelope fluctuation will decrease for GN maskers and will increase for LNN maskers. How will such changes, and the resulting interaural differences, influence masked threshold in the NoSπ condition?

To obtain some insight into this question, we considered the model described by Bernstein (1996) and Bernstein and Trahiotis (1996), based simply on a cross correlation of the stimuli at the two ears after processing by a half-wave rectifier and low-pass filter. This model is similar, in many respects, to most models of binaural interaction, which center around a binaural comparator based on the cross correlation, or the coincidence, of input from the two ears. Excellent reviews of such models are provided by Colburn and Durlach (1978), Durlach and Colburn (1978), Hafter and Trahiotis (1994), and Stern and Trahiotis (1995). Bernstein and Trahiotis (1996) have shown that such a model provides a reasonable account of the detection of interaural time differences as well as masked signal detection under antiphasic conditions over a wide range of stimulus frequencies. Furthermore, the change in correlation resulting from the addition of an antiphasic signal at threshold is comparable to direct measures of binaural correlation discrimination, leading to the interpretation that binaural detection relies on a constant change in interaural correlation of the stimuli following peripheral processing. It should be noted that the interaural correlations computed here are *normalized correlations* and not *normalized covariances* (i.e., Pearson product-moment correlation).[2]

If detection thresholds are mediated by changes in the interaural correlation produced by an antiphasic signal, then an examination of the output of such a model should predict the influence of the opposing cues obtained with GN and LNN maskers. In order to account for differences in listener performance across stimulus frequency, Bernstein and Trahiotis (1996) included half-wave rectification and low-pass filtering prior to the correlation computation. The low-pass filter stage, in essence, serves to limit the availability of interaural differences in the stimulus fine structure as stimulus frequency increases, and is based, in theory, on physiological processes in the auditory periphery (cf. Weiss and Rose, 1988). Model predictions were obtained by computing interaural correlations resulting from NoSπ conditions using either GN or LNN maskers following half-wave rectification and low-pass filtering. The model details are presented in the Appendix. The model simulations indicated a *larger* change in interaural correlation when the signal was added to LNN than when added to GN. Thus, assuming that a listener requires a *fixed criterion change in interaural correlation* to detect the presence of a signal, the model predicts that thresholds should be 3 to 5 dB *lower* for the LNN than the GN masker in the NoSπ condition. Combining the above assumption and model predictions with the report by Hartmann and Pumplin (1988) for the detection of a tone in LNN and GN, one may formulate general predictions for the present experiment. For homophasic conditions, masked thresholds should be lower for LNN than for GN and should be similar at 500 and 4000 Hz. For antiphasic conditions, thresholds should also be lower for LNN than GN, and should be higher for both masker types at 4000 Hz than at 500 Hz, the latter due to the loss of fine-structure coding.

The purpose of the present investigation was to explore the contributions of stimulus envelope and fine structure to monaural and binaural processing using a MLD paradigm. Characteristics of the masker envelope were manipulated by using noises differing in the normalized fourth moment of the envelope while the fine structure was manipulated by varying the stimulus frequency. The results will be considered in terms of current models of monaural and binaural processing. The reader should note that the authors became aware of an independent, ongoing investigation of the MLD using LNN maskers by Hall *et al.* (1998) during the course of this study. Discussions between the authors revealed that, although substantial differences in experimental method were employed, the results obtained seemed wholly compatible. Thus, the authors decided to submit the works as companion papers.

I. GENERAL PROCEDURES

A. Conditions

Masked pure-tone thresholds were obtained at 500 and 4000 Hz for homophasic (NoSo) and antiphasic (NoSπ) stimulus configurations using five masker types. The maskers of primary interest were Gaussian noise (GN) and low-noise noise (LNN). New samples of Gaussian noise were generated on each observation interval as the sum of sinusoidal components having Rayleigh-distributed magnitudes and uniformly distributed phases (0 to 2π). The LNN masker was composed of a single set of sinusoidal components having equal magnitudes and phases chosen such that the fourth moment of the waveform was minimized.[3] The same noise sample was presented in every observation interval of a given psychophysical trial and thus can be considered frozen noise in a statistical sense. Any differences in masked thresholds obtained with these two noise types might be attributed to three factors: (1) magnitude spectrum (contoured versus flat), (2) waveform regularity (random versus

frozen), or (3) temporal envelopes (fluctuating versus flat). Additional masking conditions were included to control for the first two possibilities.

To control for differences in the spectral shape or contour, masked thresholds were obtained with random samples of EMN. Any differences between GN and LNN attributable to differences in spectral shape might arise as differences in threshold between the GN and EMN maskers. Based on the statistical attributes of the two noise types, substantial differences in thresholds with these two maskers were not expected. To control for the random versus frozen difference, a sample of EMN having a fourth moment equal to the mean of the distribution of fourth moments (1.95) was chosen. This masker, termed medium-noise noise (MNN), is similar to the LNN masker in spectral shape (equal magnitudes) and regularity (same noise presented in every observation interval). On the basis of envelope fluctuation alone, thresholds obtained with EMN and MNN should be similar, since on average they will have the same fourth moment. However, if the frozen nature of the MNN influences detection, then the obtained thresholds should differ.

It is important to note that, for all frozen-noise maskers used in this study, the envelope starting phase was randomized across intervals in an attempt to reduce the possibility of improved detection resulting from the memorization of various aspects of the frozen maskers through repeated presentation. This was accomplished by the addition of a phase shift to each sinusoidal component proportional to the component frequency. The amount of phase shift was random from interval to interval. This operation resulted in a new starting point for the waveform presented in each interval and had no influence on the fourth moment of the envelope.

Finally, for completeness a high-noise noise (HNN) masker was included. It was reasoned that if thresholds were highly dependent on the magnitude of the fourth moment, then perhaps there would be a systematic change in threshold along the continuum of fourth moment values (e.g., LNN, MNN, and HNN). The HNN was arbitrarily defined as the sample of EMN which was the same distance from the mean of the EMN distribution as the LNN sample (0.91 units). Thus, the HNN had a fourth moment of 2.88.

B. Stimulus generation

All stimuli were generated using a digital array processor (TDT AP2) and two 16-bit D/A converters with a sampling period of 24.4 μs (40 984 Hz). Noise bands were generated in the frequency domain by filling portions of two 16 384-point buffers with the appropriate component magnitudes or phases. The magnitudes were scaled so that each noise sample had the same overall energy and a random, frequency-proportional phase shift was added to each component. A subsequent inverse FFT on the complex buffer pair yielded the desired noise waveform having a bandwidth of 50 Hz and a pressure spectrum level of 50 dB SPL. Pure-tone signals were generated in the time domain and were random with respect to the masker. The masker and signal were shaped by 10-ms $\cos^2$ onset and offset ramps and output simultaneously (400 ms) via separate D/A converters, and routed through two dual-channel adders (TDT SM1). One channel was equipped with a phase-inverter which, when engaged, introduced a phase shift of π radians.

C. Procedures

An adaptive, two-interval, two-alternative forced-choice method with a three-down, one-up tracking algorithm was used to estimate the signal level corresponding to 79.4% correct detection (Levitt, 1971). The step size in the adaptive procedure was 2 dB. Threshold was estimated by averaging the signal level at each of the last even number of reversals, excluding the first three reversals. The final threshold for a condition was taken as the average threshold across six 60-trial blocks. Three additional runs were included if the standard deviation across the six runs exceeded 2.3 dB. Each trial, consisting of two 400-ms observation intervals separated by 400 ms and marked by LEDs, was followed by visual feedback. Stimuli were presented via ER-2 (Etymotic) insert earphones and testing was conducted in a sound-attenuating chamber.

D. Listeners

Seven listeners participated, ranging in age from 22–30 years. All had pure-tone thresholds of 15 dB HL or better at octave frequencies from 250 to 8000 Hz and normal middle-ear function based on a screening tympanogram (Y, 226 Hz). Only one of the listeners had previous experience in psychoacoustic listening tasks. Excluding the two authors, listeners were paid an hourly wage and received a 10% bonus upon completion of the study.

The possible effects of practice on listener performance were examined and found to be minimal. Subjects who initially showed small MLDs at 500 Hz were given additional practice (approximately 1200 trials) prior to data collection with no substantial improvement. Upon the completion of data collection, thresholds for several conditions were re-estimated and no substantial differences (<2 dB) were found for any listener.

II. RESULTS

The mean data are shown in Fig. 3, with the upper two panels displaying masked thresholds (in dB SPL) and the lower two panels displaying the corresponding MLDs (in dB). The five masker conditions are listed on the abscissa. Data for the 500- and 4000-Hz frequencies are shown in the left and right columns, respectively.

A. Homophasic detection

In the homophasic condition (filled squares), the GN, EMN, MNN, and HNN maskers produced roughly the same amount of masking, with thresholds ranging from 60–65 dB SPL (mean$=62.9$, s.e.$=0.5$). The LNN produced about 5.5 dB less masking than the other four masker types, the average threshold being 57.4 dB SPL. This difference is consistent with previous comparisons of masked thresholds in GN and LNN for frozen noise maskers (Hall *et al.*, 1996; Hart-

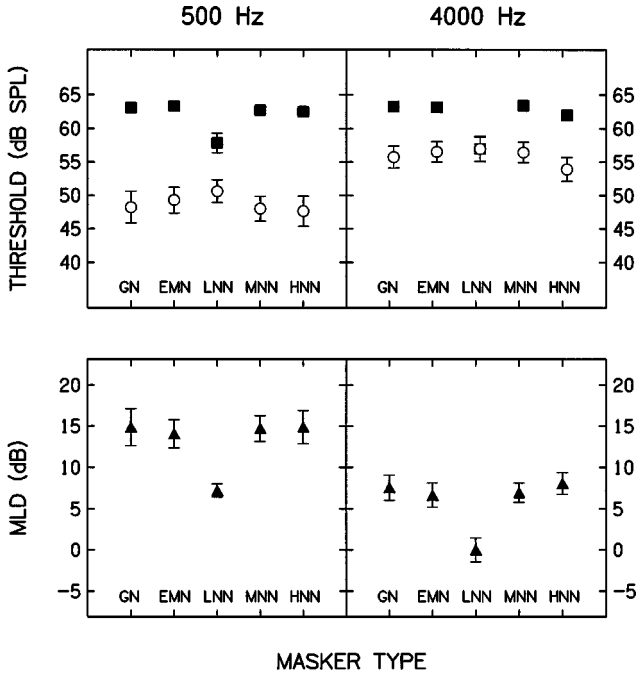


FIG. 3. Mean thresholds in dB SPL (upper panel) and corresponding masking level differences in dB (lower panel). The results are shown for 500 Hz on the left and 4000 Hz on the right. The five masker types are shown on the abscissa. Symbols represent the NoSo (filled circles) and NoSπ (open squares) conditions.

mann and Pumplin, 1988; Kohlrausch, 1996) and for random noise maskers (Kohlrausch, 1996). One of the listeners, however, showed no threshold difference between the LNN and the other masker types at either 500 or 4000 Hz. This listener generally had the highest NoSo and NoSπ thresholds and the smallest MLDs of the seven listeners. Interestingly, one of the six listeners in the Hartmann and Pumplin (1988) investigation showed equivalent homophasic thresholds in their LNN and GN conditions. That listener also had the highest thresholds across conditions among their six listeners.

Differences among thresholds obtained with the GN and LNN maskers can be better understood by first considering the relatively small differences in NoSo thresholds obtained with GN, EMN, MNN, and HNN. The lack of a threshold difference for the GN and the EMN maskers indicates that the shape of the short-term magnitude spectrum had little influence on tone detection in narrow-band noise, as expected. Further, thresholds were not strongly influenced by the use of frozen noise samples, as demonstrated by threshold comparisons between the MNN and EMN maskers. This result is contrary to many previous studies comparing masked thresholds in frozen and random narrow-band noise (e.g., Isabelle and Colburn, 1991; Kidd *et al.*, 1993). Evidence from those studies and from this laboratory (Eddins, 1996) indicate that listeners often are able to ''learn'' the pattern of envelope fluctuations of a frozen noise, facilitating detection and discrimination. The ''frozen'' stimuli used here, however, were randomized with respect to time, and thus were somewhat ''thawed,'' a process that apparently was successful in limiting listener's abilities to learn specific masker characteristics. Finally, although the fourth moments associated with the LNN, MNN, and HNN maskers increased progressively (1.04, 1.95, and 2.88), thresholds in these conditions did not reflect this systematic change. The NoSo thresholds were clearly lowest with the LNN masker, but there was essentially no difference between thresholds in the MNN and HNN conditions. Thresholds do not appear to be highly dependent on the magnitude of the fourth moment over the entire range of possible values. Given these results, it is reasonable to assume that the difference between masked thresholds in the GN and LNN conditions are the result of temporal differences between the two maskers.

B. Antiphasic condition

Thresholds for the antiphasic condition (open circles) were similar across all five masker types. Thus, the advantage gained for NoSo detection in LNN did not extend to NoSπ detection. In fact, for each listener, the NoSπ threshold was slightly worse for LNN than for GN. In addition, there were marked differences across frequency, with thresholds averaging 7.2 dB higher at 4000 Hz than at 500 Hz. A repeated-measures analysis of variance revealed significant main effects of frequency [$F(1,24)=7.20$, $p=0.01$], interaural condition [$F(1,24)=52.53$, $p<0.01$], and masker type [$F(4,96)=8.59$, $p<0.01$] as well as significant interactions of frequency and interaural condition [$F(1,24)=60.00$, $p=0.01$] and interaural condition and masker type [$F(4,96)=3.37$, $p<0.01$]. *Post hoc* analyses (Tukey-HSD) of the interaural condition–frequency interaction confirmed that while NoSo thresholds did not differ across frequency, the NoSπ thresholds differed from each other at 500 and 4000 Hz and from the NoSo thresholds at either frequency. Similar analyses of the interaural condition–masker type interaction revealed three significant differences: (1) lower NoSo thresholds for the LNN masker than the other four maskers, (2) higher NoSo than NoSπ thresholds for all masker types, and (3) higher NoSπ thresholds for the LNN masker than for the HNN masker.

The lower two panels of Fig. 3 show corresponding MLDs at 500 (left) and 4000 Hz (right). The MLDs were similar in magnitude for the GN, EMN, MNN, and HNN maskers, averaging 14.6 dB at 500 Hz and 7.3 dB at 4000 Hz. Most interesting are the very small MLDs obtained with LNN maskers, averaging 7.2 and 0.0 dB at 500 and 4000 Hz, respectively. A repeated-measures analysis of variance on the MLDs revealed significant main effects of masker type [$F(4,48)=61.70$, $p<0.01$] and signal frequency [$F(1,12)=6.51$, $p<0.01$] but no masker-frequency interaction. *Post hoc* analyses (Tukey-HSD) revealed that the MLDs obtained with the GN, EMN, MNN, and HNN maskers were not significantly different from each other but were all significantly higher than the MLDs for the LNN maskers. Thus, MLDs were smaller for LNN maskers than for maskers with higher fourth moment values and were smaller at 4000 than 500 Hz.

C. Individual differences among subjects

Despite having only a 5-dB range of thresholds across listeners in the NoSo conditions, large individual differences in the NoSπ thresholds were obtained. These are reflected in the error bars about the open symbols of Fig. 3. As an ex-

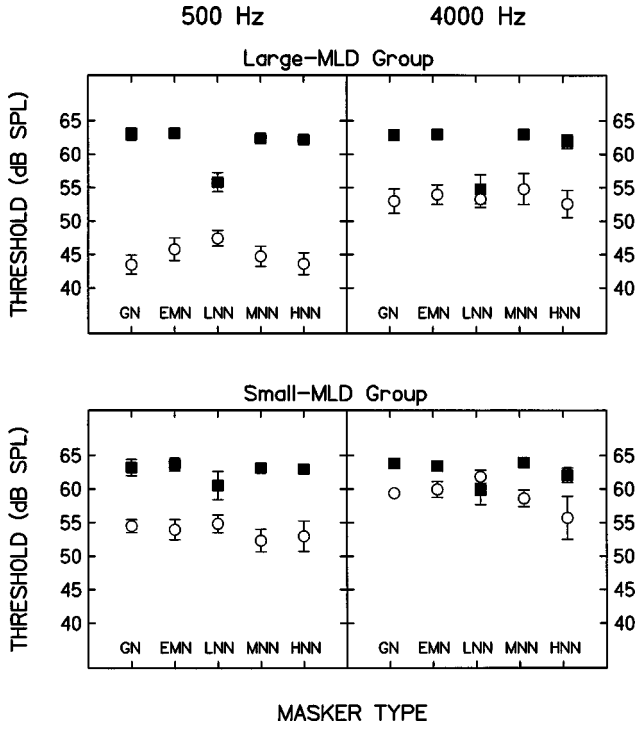


FIG. 4. Thresholds in dB SPL for the large-MLD group (upper panel) and the small-MLD group (lower panel). The results are shown for 500 Hz on the left and 4000 Hz on the right. The five masker types are shown on the abscissa. Symbols represent the NoSo (filled circles) and NoSπ (open squares) conditions.

ample, consider the GN masker at 500 Hz. Four of seven listeners had MLDs exceeding 17 dB, while the three remaining listeners had MLDs less than 10 dB, almost entirely due to differences in the NoSπ conditions. As a result, the mean MLDs for the GN masker, as shown in Fig. 3, are somewhat smaller than those typically reported for very narrow masker bandwidths (Bernstein and Trahiotis, 1992b; Hall and Harvey, 1984; Kohlrausch *et al.*, 1995; McFadden and Pasanan, 1978; Metz *et al.*, 1968; Staffel *et al.*, 1990; Zurek and Durlach, 1987). To better appreciate the rather large individual differences here, as well as the large differences among stimulus conditions obtained for some listeners, the data are considered below for a large-MLD group (four listeners) and a small-MLD group (three listeners), separated on the basis of their MLDs in the 500-Hz, GN masker condition. The results for these two groups are plotted in the upper and lower panels of Fig. 4 as a function of masker type. The data for the 500- and 4000-Hz conditions are shown in the left and right columns, respectively.

Considering first the thresholds for the NoSo conditions (filled squares), it is clear that the GN, EMN, MNN, and HNN maskers produced roughly the same amount of masking for each subgroup. Thus, differences in performance across the groups would not appear to result from a general deficit in masked signal detection. Substantial differences between the two groups were obtained for the LNN conditions. The LNN produced about 8 dB less masking than GN for the large-MLD group at both test frequencies, whereas the difference was only about 3.5 dB for the small-MLD group.

The NoSπ thresholds (open circles) revealed a pattern of results that were similar at 500 and 4000 Hz and generally similar across subject groups. While both subject groups demonstrated higher NoSπ thresholds at 4000 Hz than at 500 Hz, performance for the small-MLD group averaged 7.1 dB worse than the large-MLD group. The average MLDs for the GN, EMN, MNN, and HNN maskers were 18.3 dB at 500 Hz and 9.1 dB at 4000 Hz for the large-MLD group, as compared to 9.8 (500 Hz) and 4.9 dB (4000 Hz) for the small-MLD group. The results for the large-MLD group are consistent with many previous studies with Gaussian, narrow-band noise maskers in terms of the magnitude of the masking release and the difference across frequency (Bernstein and Trahiotis, 1992b; Hall and Harvey, 1984; Kohlrausch *et al.*, 1995; McFadden and Pasanan, 1978; Metz *et al.*, 1968; Staffel *et al.*, 1990; Zurek and Durlach, 1987). For LNN, however, the MLDs were relatively small with substantial between-group differences. The LNN MLDs at 500 and 4000 Hz were 8.4 and 1.5 dB for the large-MLD group and 5.7 and -2.0 dB for the small-MLD group, respectively, which reflected a decrease in NoSo threshold combined with a slight increase in NoSπ threshold. Despite their differences, the groups demonstrated a similar pattern of thresholds and MLDs across masker type and frequency, suggesting that their binaural processing was not fundamentally different. Relative to the small-MLD group, however, the large MLD group consistently showed better detection in all of the antiphasic conditions *and* in the NoSo LNN conditions. These differences may reflect basic differences in encoding, processing, and/or attention to the rather subtle cues (monaural and binaural) carried by the temporal envelope.

III. DISCUSSION

A. Individual differences

The large individual differences reported here for the NoSπ conditions are not uncommon in the MLD literature, although the most substantial individual differences typically occur with narrow-band maskers at high-signal frequencies. At frequencies of 1000 Hz and below, threshold differences among listeners as large as 5 to 9 dB have been reported (Hall and Grose, 1995; Hall and Harvey, 1984; Jain *et al.*, 1991; Staffel *et al.*, 1990). Several factors may account for the very large differences (>16 dB) at 500 Hz among our seven listeners. First, general deficits in binaural processing cannot be ruled out. Unfortunately, no other tests of binaural hearing were conducted with the three listeners in the small-MLD group to either confirm or refute this possibility. Second, large individual differences in masked thresholds, particularly in masking release paradigms, have been associated with ''burst'' masking paradigms, in which the signal and masker are gated on and off synchronously (e.g., Fantini *et al.*, 1993; McFadden and Wright, 1992; Hatch *et al.*, 1995), although this is not a universal finding (e.g., Hall and Grose, 1990; Kohlrausch, 1986). Although a comprehensive evaluation of the burst-continuous continuum was not pursued, an informal test using several masker-fringe conditions was conducted for one listener (who had the highest thresh-

olds in NoSπ and in LNN conditions) with no substantial change in threshold for any condition.

In addition to the observation of rather large individual differences, several important results related to both monaural and binaural hearing emerge from these data. The following discussion addresses first the results obtained in the NoSo conditions, along with relevant models of monaural processing, followed by a discussion of the NoSπ conditions and models of binaural processing.

B. Homophasic detection

Implicit here is that the diotic condition, NoSo, primarily reflects monaural auditory processing as this condition does not introduce interaural differences, *per se*. This assumption is supported, in part, by the common result that thresholds are equivalent in the NmSm and NoSo stimulus configurations (e.g., Blodgett *et al.*, 1958).[4] The three most notable results obtained in NoSo detection are: (1) improved detection in LNN relative to the other maskers, (2) the lack of a systematic change in threshold with changes in the fourth moments of the maskers (LNN, MNN, and HNN), and (3) the similarity among thresholds at 500 and 4000 Hz. Any comprehensive explanation should address these three findings. The results of several recent investigations of monaural perception of tones masked by frozen and random narrow-band noise suggest that the masker envelope may have considerable influence on signal threshold, although the exact way in which listeners use such cues remains unclear (e.g., Kidd *et al.*, 1993; Kohlrausch, 1996; Richards, 1992). Indeed, the 6-dB difference between thresholds obtained with the GN and LNN maskers (8 dB for the large-MLD group; 3 dB for the small-MLD group) likely reflects differences in the masker envelope. The magnitude of this threshold difference is in good agreement with previous reports on the detection of tones in GN and LNN. For example, Hartmann and Pumplin (1988) reported a 5-dB difference in thresholds for frozen maskers using GN and LNN, whereas Kohlrausch (1996) reported differences of 8 dB in a frozen noise condition and 5 dB in a random-noise condition. Though minor, the differences among the three studies might well reflect differences in the stimulus bandwidth (100 Hz for Hartmann and Pumplin and Kohlrausch versus 50 Hz here), individual differences among listener performance (as noted here and by Hartmann and Pumplin), differences in the specific noise samples, and/or differences in the random nature of the noises (e.g., Kohlrausch).

Since the GN and LNN maskers were equal in overall energy, an explanation based on a simple energy detector model appears inconsistent with the large difference in thresholds reported, a conclusion common to several recent investigations using narrow-band maskers (Hartmann and Pumplin, 1988; Kidd *et al.*, 1993; Richards *et al.*, 1991; Richards, 1992).[5] In addition, the fact that thresholds did not differ across frequency suggests that the fine structure of the stimulus was not an important determinant of threshold in the NoSo conditions. Thus it seems that the present results favor an explanation for these differences based on changes in the temporal envelope of the signal-plus-masker stimulus.

In four of the five masker types employed here (GN, EMN, MNN, and HNN), the masker itself had salient fluctuations and the addition of a signal tended to reduce those fluctuations. A reasonable decision rule for signal detection, therefore, would be ''choose the interval containing the stimulus with the 'smoother' temporal envelope.'' For the LNN masker, however, the detection strategy might be quite different. In this case, the noise masker alone has a very smooth temporal envelope, with minimal fluctuations and the addition of a tonal signal *increases* the amount of envelope fluctuation. Thus, an appropriate decision rule would be the inverse of the rule stated above.

A variety of models and/or decision statistics have been suggested to account for the detection of a tonal signal masked by a narrow-band noise. In an attempt to illuminate potential cues related to signal detection in the present experiment, a series of computer simulations was conducted employing several common models of signal processing based upon an analysis of the stimulus envelope (e.g., Green *et al.*, 1992; Forrest and Green, 1987; Hartmann and Pumplin, 1988; Richards, 1992; Viemeister, 1979). As none of the models employed here are new, only a brief description will be provided. The interested reader is referred to the original papers for additional detail.

Due to the narrow bandwidth of the stimuli used in the present experiment, the effect of peripheral filtering on stimulus processing was judged to be relatively minor, and no attempt was made to incorporate such filtering in the simulations. For each of the six models, the stimulus envelope was extracted from the digital signal[1] prior to computing a particular decision statistic. In the first model examined (power spectrum of the envelope), psychometric functions were generated for each masker type, based on 1000 stimuli per level. From these functions, the signal level corresponding to 79.4% correct was estimated and taken as threshold. In the remaining five models, threshold estimates were based upon the same adaptive tracking procedure described in Sec. I C. In each case, 20 threshold estimates were obtained for each of the 10 stimulus conditions (2 frequencies by 5 masker types). Unless otherwise noted, the predictions of each model were essentially the same for the 500- and 4000-Hz stimuli and are only reported for the 500-Hz stimuli. While three of the models yielded predicted thresholds that were markedly different from those obtained behaviorally, the other three provided predictions that closely mimicked the pattern of behavioral thresholds obtained.

The power spectrum of the envelope model (PSEN) is a modification of the quasi-optimum detector of changes in the shape of envelope power spectrum proposed by Green *et al.* (1992). In this model, it is assumed that the auditory system extracts the waveform envelope and bases subsequent decisions upon changes in the shape of the power spectrum of that envelope resulting from the addition of a signal. Green *et al.* showed that (1) the powers of various envelope components are nearly exponentially distributed, and (2) an optimum decision statistic would be the cross correlation of the normalized envelope power spectrum and the weights $(\eta - \tau)$, where η is the exponential parameter representing the masker and τ is the exponential parameter for the signal plus

TABLE I. Behavioral thresholds for the large-MLD group and corresponding predictions for ten models of monaural perception for each of the five masker types listed on the left. Differences between predicted and obtained thresholds are shown in parentheses. Stimulus conditions are identical to those described in the Methods section. See text for descriptions of the various models.

	THD LG MLD	PSEN	EMS ENV	NUM MAX	$M4$	σ/μ ENV	Mx/Mn ENV
GN	63.00	63.00 (0.00)	60.43 (2.57)	62.82 (0.18)	63.85 (−0.85)	64.08 (−1.08)	65.34 (−2.34)
LNN	55.80	53.90 (1.90)	55.56 (0.24)	53.9 (2.71)	28.65 (27.15)	31.39 (24.41)	43.12 (12.68)
EMN	63.10	63.00 (0.10)	60.41 (2.69)	63.04 (0.06)	64.14 (−1.04)	63.75 (−0.65)	64.45 (−1.35)
MNN	62.40	59.40 (3.00)	60.43 (1.97)	60.85 (1.55)	63.55 (−1.15)	62.97 (−0.57)	52.30 (10.10)
HNN	62.20	56.10 (6.10)	58.12 (4.08)	59.61 (2.59)	58.52 (3.68)	54.78 (7.42)	62.22 (−0.02)

masker. Following Green *et al.*, the fixed set of component weights ($\eta_i - \tau_i$) was obtained separately for each masker type by subtracting the normalized envelope power for a signal added to the noise at an E/No of 10 dB from the normalized envelope power for the masker alone. Next, it was assumed that the estimation of optimal weights and the extraction of the stimulus envelope by human observers would be imperfect. Thus, internal noises were introduced at both of these stages, modeled as independent additive Gaussian processes. The model predictions obtained for a 500-Hz signal frequency are shown in column 2 of Table I (labeled PSEN) along with the mean thresholds obtained for the large-MLD group in column 1 (labeled THD).[6] The difference between predicted and obtained thresholds are listed below each model prediction. On average, the model predictions were slightly better than the behavioral thresholds (2.2 dB), however, there was a substantial discrepancy for the HNN masker (6.1 dB). Because the pattern of predicted and obtained results differs somewhat, several additional models are considered below.

In the second model, employed successfully by Richards (1992) and Bernstein and Trahiotis (1994), detection is based on the ratio of the envelope mean slope of the signal-plus-masker and masker-alone waveforms. The envelope slope was estimated as follows: the envelope was extracted, normalized, and the absolute values of the differences between successive points in the digital representation of the envelope were averaged. Again it was assumed that this process would be influenced by internal noise associated with coding the envelope slope. Thus, an internal (Gaussian) noise was added to this process. For GN, EMN, MNN, and HNN, the decision rule was to choose the waveform having the smaller mean difference (''smoother'' envelope) as the signal-plus-noise waveform. For LNN, an opposite rule was used. The thresholds obtained from simulations employing an adaptive threshold tracking procedure are shown in column 3 of Table I (labeled EMS). The predicted thresholds are similar to those obtained behaviorally, and similar in pattern to those obtained with the PSEN model. The greatest discrepancy between predicted and obtained thresholds again occurred for the HNN masker (4.08 dB).

Despite the close agreement between behavioral thresholds and those estimated from the models above, a model based simply on a count of the number of envelope maxima provided the most accurate threshold predictions of all. As before, the model included additive internal noise and the decision rule was to choose as the signal the stimulus containing the fewer envelope maxima. The results of this simulation are shown in column 4 of Table I (NUM MAX). In this case, the predicted thresholds were on average 1.4 dB better than behavioral thresholds, with the greatest discrepancy occurring for the LNN (2.7 dB).

As seen in columns 5–7, the threshold predictions obtained with the next three models were, in general, *higher* than behavioral thresholds for the GN, EMN, and MNN maskers, and substantially *lower* than behavioral thresholds for the HNN and LNN conditions. In order of presentation, these models based decisions on the normalized fourth moment of the envelope (column 5, $M4$), the standard deviation of the normalized envelope (column 6, σ/μ), or the ratio of the maximum to minimum value of the normalized envelope (column 7, Mx/Mn ENV). Of the six models evaluated, the first three models provide a much more accurate description of the data than the latter three. The model which counts the number of envelope maxima is perhaps most appealing, because of the accuracy achieved, its simplicity, and the fact that it is physiologically tractable. It is not difficult to imagine a mechanism that reduces the envelope information to a simple sum of envelope peaks. To be successful over a modest range of stimulus bandwidths, the task would require accurate coding of approximately 5–125 events per second (representing bandwidths of about 10–200 Hz).

C. Antiphasic conditions

Several features of the dichotic results merit further discussion, including the small increase in $\mathrm{NoS}\pi$ thresholds for the LNN masker relative to the remaining maskers and the change in $\mathrm{NoS}\pi$ thresholds across frequency. Although $\mathrm{NoS}\pi$ thresholds were not significantly different across the five masker types, the large MLD group exhibited thresholds that were, on average, 4 dB worse for the LNN than for the

GN condition. It should be noted that Hall *et al.* (1996) reported that about half of their six listeners had poorer NoSπ thresholds with LNN than GN maskers, although the remaining half performed equally as well in the two conditions. Also note that the signal used in that experiment was narrow-band noise rather than a sinusoid.

Clearly the present results are not in accord with the predictions of the interaural envelope correlation model described in the Introduction. Recall that the predicted NoSπ thresholds were some 3–5 dB lower for LNN than GN, whereas the obtained thresholds are in the opposite direction. The work of Bernstein (1996) and Bernstein and Trahiotis (1996) indicated that binaural detection/discrimination could be explained on the basis of a constant-criterion change in the normalized correlation. The model estimates of interaural normalized correlations produced by a signal at threshold ranged from 0.91–0.99 across the five masker types at 500 Hz. This range is consistent with previous direct measures of the jnd for interaural correlation (Bernstein and Trahiotis, 1996; Koehnke *et al.*, 1986) and the range of correlations produced by threshold ITD values measured by McFadden and Pasanen (1976) and Nuetzel and Hafter (1981),[7] although the present interaural correlations tend toward the lower end of performance in those studies. At 4000 Hz, the large range of interaural normalized correlations (0.72–0.91) suggests that performance in the different masker conditions did not depend on a single-valued change in correlation, as predicted by Bernstein and Trahiotis. In a more general analysis, the low-pass filtering introduced by Bernstein and Trahiotis was omitted, and, as shown in the bottom panel of Fig. A1, the interaural normalized correlations are still larger for GN than LNN. Thus, a simple examination of interaural cross correlation does not help to explain the data reported here. Perhaps more detailed modeling, taking into account the cross-correlation space across time and frequency, would be more informative (e.g., Stern *et al.*, 1988; Stern and Shear, 1996).

Several recent models of binaural processing have emphasized, among other parameters, the relative importance of interaural time or phase differences of relatively *small* magnitude (e.g., ITDs$<$200 μs), a concept known as ''centrality'' (e.g., Jeffress, 1972; Stern and Trahiotis, 1992). This weighting of small interaural differences is based on physiological evidence that there are more neurons that are sensitive to small interaural delays than to large interaural delays (Kuwada *et al.*, 1987). An examination of the distributions of interaural differences as a function of signal-to-masker ratios (SMR) for GN waveforms indicates increasing standard deviations of IPDs and IIDs (about a mean of zero), and a shift in the maxima or modes of the probability density functions for IPDs and IIDs away from zero with increasing signal level (Zurek, 1991). Thus the *variability* and *magnitude* of interaural disparities increase with SMR (at least over a 20-dB range about threshold). Similar analyses for the five masker types used here reveal that, for a fixed signal level, the standard deviation of IPDs and IIDs increases sharply with increasing masker fourth moment (i.e., from LNN to HNN). This change is consistent with improved detectability with increasing masker fourth moment, similar to the present results. Based upon these interaural disparities, however, the predicted difference in detection among the five maskers is some 8–10 dB rather than the 3–5 dB observed. Conversely, the maxima or modes of the probability density functions shift away from mean zero as the masker fourth moment *decreases*. Taken together, these distributions indicate that IPDs and IIDs tend to be *larger* (higher modal values) and *more regular* as the masker fourth moment decreases (i.e., from HNN to LNN). These properties favor detection in LNN over GN or HNN maskers, however, the concept of centrality suggests that the *shorter* IPDs (i.e., higher fourth moment maskers) are given relatively greater weight. Thus, one possible explanation of the present results is that the change in the magnitude of interaural disparities with masker type, mediated by centrality, results in the pattern of detection thresholds reported here.

Differences in NoSπ detection at 500 and 4000 Hz are consistent with differences between the distributions of interaural phase and intensity based upon the stimulus waveform and those based upon the stimulus envelope. In particular, as suggested by Bernstein and Trahiotis (1985a, 1985b, 1986), at low signal frequencies, interaural differences may be conveyed by both the stimulus fine structure and stimulus envelope, whereas at high signal frequencies, only the envelope-based cues would be available. While the distribution of envelope-based IPDs and IIDs paralleled those based upon the waveform (see above), the standard deviation and maxima of the distributions were uniformly smaller for envelope-than waveform-based interaural differences. This is consistent with the finding of higher NoSπ thresholds at 4000 than at 500 Hz.

IV. SUMMARY AND CONCLUSIONS

The present study investigated the influence of masker envelope and fine structure on NoSo and NoSπ detection commonly measured in a masking level difference paradigm. The fourth moment of the envelope was the metric used to quantify the masker envelope fluctuations, and was systematically varied from low to high values using a low-noise noise masker as defined by Pumplin (1985) and medium- and high-noise noise maskers defined on the basis of the distribution of fourth moments obtained from a large sample of random noises. The fine structure was not manipulated *per se*, but the availability of fine-structure cues to the auditory system was varied by manipulating the stimulus frequency.

In general, NoSo thresholds were lower in the LNN masker than any of the other four masker types. The threshold values were nearly the same at 500 and 4000 Hz. Similar NoSo thresholds in the GN, EMN, MNN, and HNN maskers revealed a negligible influence of (1) the random nature of the magnitude spectrum of the narrow-band noises (Rayleigh distributed versus equal magnitude), (2) the constraints on the particular noise sample (random versus frozen), and (3) rather substantial variation in the masker envelope fluctuations over the range of moderate to high fourth moment values. Detection thresholds in the NoSo condition were best described by a model that was based on the number of envelope maxima in the masker alone or masker-plus-signal.

The NoSo thresholds were similar across the seven normal-hearing subjects; however, substantial individual differences in NoSπ thresholds were obtained. On the basis of the size of the MLD obtained for GN at 500 Hz, the listeners were separated into two groups, a large-MLD group and a small-MLD group, for the purpose of summary and discussion. Listeners in the large-MLD group had MLDs exceeding 17 dB while listeners in the small-MLD group had MLDs less than 10 dB in this condition.

Mean NoSπ thresholds were about 4 dB worse in the LNN masker than the GN masker at 500 Hz for the large-MLD group. This pattern of results is similar to recent results obtained by Hall *et al.* (1996) and Kohlrausch (1996). Over both frequencies, however, differences in NoSπ thresholds among the five masker types failed to reach significance for either listener group. This pattern of thresholds is inconsistent with the predictions of a simple interaural-cross-correlation model based on normalized correlation subsequent to half-wave rectification and low-pass filtering. Such a model, which has been highly successful in accounting for detection and discrimination results in the NoSπ conditions of previous investigations, predicts *better* performance ($\approx$5 dB) in the LNN masker than the GN masker. The observed pattern of binaural (antiphasic) thresholds may be understood by consideration of the distributions of interaural disparities. In general, the results are consistent with the use of changes in both envelope- and fine-structure-based interaural cues at 500 Hz, but only envelope-based cues at 4000 Hz.

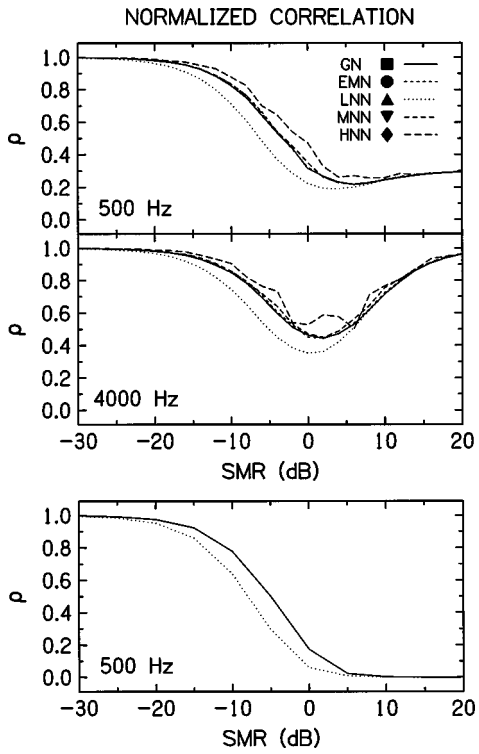


FIG. A1. Interaural normalized correlations at the output of the cross-correlation model described in the text. The curves in each panel represent different masker types. Obtained predictions from the Bernstein and Trahiotis (1996) model for the 500-Hz (upper panel) and 4000-Hz (center panel) signal frequencies and obtained predictions based upon half-wave rectified waveforms for the 500-Hz signal frequency (lower panel).

ACKNOWLEDGMENTS

The authors would like to thank Dr. Ann Clock Eddins, Dr. Larry Humes, and two anonymous reviewers for helpful comments on previous versions of this manuscript. We also thank Dr. Les Bernstein for helpful discussions concerning the cross-correlation model and Dr. Joseph W. Hall III, Dr. John H. Grose, and Dr. Armin Kohlrausch for interesting discussions of masking by LNN.

APPENDIX: NORMALIZED CROSS-CORRELATION ANALYSES

This Appendix provides a description of the interaural normalized cross-correlation model presented in the Introduction and a description of model predictions for the conditions of the present experiment. The model was first described by Bernstein (1996) and Bernstein and Trahiotis (1996) to explain NoSπ detection and is based simply on the normalized cross correlation of the stimuli at the two ears after processing by a half-wave rectifier and low-pass filter. Stimulus generation followed the methods described in Sec. I B. Interaural correlations were based on 100 trials per level and were computed over a 50-dB range of signal-to-masker ratios (SMR) in 2-dB steps.

The results of the computer simulations are shown as the curves in Fig. A1 for the 500-Hz (upper panel) and 4000-Hz (middle panel) frequencies. The lower panel will be described below. The different line types correspond to each of the five masker types of the present experiment as described in Sec. I A [Gaussian noise, GN; equal-magnitude noise, EMN; low-noise noise, LNN; medium-noise noise, MNN; and high-noise noise, HNN]. A comparison of the general form of the functions in the two panels illustrates the differences imposed by low-pass filtering. Note that for a given SMR, smaller interaural correlations correspond to better detection. The work of Bernstein (1996) and Bernstein and Trahiotis (1996) indicated that binaural detection/discrimination could be explained on the basis of a constant-criterion change in the normalized correlation. Under such an assumption, the model predicts that detection would be best for LNN, followed by GN, EMN, HNN, and MNN (see Method section for descriptions of EMN, MNN, and HNN). Over the range of behaviorally relevant correlation values (e.g., >0.90), the curves predict some 3–5 dB better performance for the LNN than the GN condition. Threshold in these conditions corresponded to SMRs from -16 to -19 dB. At 4000 Hz, the simulations yielded the same pattern of results for small SMRs. Here, the SMRs at threshold ranged from -8 to -11 over the five masker types.

The lower panel of Fig. A1 shows interaural correlations for GN and LNN without the low-pass filter stage of the Bernstein and Trahiotis model, and thus represents a more general analysis of interaural normalized correlation. Again note that the correlations are smaller for LNN than GN, implying that the predictions reported above are not specific to the Bernstein and Trahiotis implementation of the interaural cross-correlation model.

[1]The envelope was taken as the root-mean square of the real and imaginary parts of the analytic signal derived via Hilbert transform.

[2]Important distinctions between the *normalized correlation* and the *normalized covariance* (i.e., Pearson product-moment correlation) have been detailed in recent papers by van de Par and Kohlrausch (1995), Kohlrausch *et al.* (1995), and Bernstein and Trahiotis (1996). Briefly, the two measures are equivalent if the two arguments of the correlation have a mean of zero (e.g., stimulus waveforms) but differ if their means are not equal to zero (e.g., stimulus envelopes). These studies have shown that the normalized correlation provides a better index of performance in binaural tasks than does the normalized covariance.

[3]The reader is referred to Hartmann and Pumplin (1991b), table IV B, for a list of the appropriate phases used to generate the low-noise noise.

[4]Langhans and Kohlrausch (1992) reported substantial differences in NoSo and NmSm thresholds using a frozen wideband noise and short duration tone bursts. They also reported lower NoSo thresholds for frozen noise than for running noise. In addition to duration and bandwidth differences between that study and the present experiment, no differences were seen between our running (EMN) and frozen (MNN) conditions. Thus, the randomization of the time axis in the present experiment seems to have eliminated many of the effects of ''frozen'' noise typically reported. Despite the results of Langhans and Kohlrausch, we maintain that the NoSo thresholds here primarily reflect monaural detection.

[5]Hartmann and Pumplin (1988) have suggested that an energy detector combining multiple, brief observations spanning the duration of the signal might account for the reduced masking of LNN. They did not provide a detailed description of or analysis with such a model. Perhaps a decision statistic based on the variance of those multiple observations would be able to account for the present results as well as recent results obtained with of tone-in-noise detection. It is not obvious that such a model would yield predictions that differ from the envelope-based models presented below.

[6]Any suitable model would have to account for the results of all listeners rather than the ''average'' listener or any subset of listeners. Since there were substantial differences in masked threshold across the subjects, we chose the conservative approach of gauging model performance by the thresholds obtained for the large-MLD group. In the case of each model considered, the shortcomings would also hold for the small-MLD group or the average performance across all listeners.

[7]This statement is based on the transformations of Bernstein and Trahiotis (1996) in which the threshold ITDs reported by McFadden and Pasanen (1976), and Nuetzel and Hafter (1981) were converted to corresponding normalized correlations. The reader should note that Koehnke *et al.* (1986) measured the jnd for the normalized covariance (Pearson Product Moment Correlation) rather than the normalized correlation, and for comparison the values reported in that paper should be converted to normalized correlations.

Bernstein, L. R. (**1996**). ''The index of interaural correlation: Accounting for binaural detection across center frequency,'' J. Acoust. Soc. Am. **99**, 2470(A).

Bernstein, L. R., and Trahiotis, C. (**1985a**). ''Lateralization of low-frequency, complex waveforms: The use of envelope-based temporal disparities,'' J. Acoust. Soc. Am. **77**, 1868–1880.

Bernstein, L. R., and Trahiotis, C. (**1985b**). ''Lateralization of sinusoidally amplitude-modulated tones: Effects of spectral locus and temporal variation,'' J. Acoust. Soc. Am. **78**, 514–523.

Bernstein, L. R., and Trahiotis, C. (**1986**). ''Lateralization of bands of noise and sinusoidally amplitude-modulated tones: Effects of spectral locus and bandwidth,'' J. Acoust. Soc. Am. **79**, 1950–1957.

Bernstein, L. R., and Trahiotis, C. (**1992a**). ''Detection of antiphasic sinusoids added to the envelopes of high-frequency bands of noise,'' Hearing Res. **62**, 157–165.

Bernstein, L. R., and Trahiotis, C. (**1992b**). ''Discrimination of interaural envelope correlation and its relation to binaural unmasking at high frequencies,'' J. Acoust. Soc. Am. **91**, 306–316.

Bernstein, L. R., and Trahiotis, C. (**1994**). ''Detection of interaural delay in high-frequency sinusoidally amplitude-modulated tones, two-tone complexes, and bands of noise,'' J. Acoust. Soc. Am. **95**, 3561–3567.

Bernstein, L. R., and Trahiotis, C. (**1996**). ''On the use of normalized correlation as an index of interaural envelope correlation,'' J. Acoust. Soc. Am. **100**, 1754–1763.

Blodgett, H. C., Jeffress, L. A., and Taylor, R. W. (**1958**). ''Relation of masked threshold to signal duration for various interaural phase-combinations,'' Am. J. Psychol. **71**, 283–290.

Colburn, H. S., and Durlach, N. I. (**1978**). ''Models of binaural interaction,'' in *Hearing, Vol. IV of Handbook of Perception*, edited by E. C. Carterette and M. P. Friedman (Academic, New York).

Durlach, N. I. (**1964**). ''Note on binaural masking-level differences at high frequencies,'' J. Acoust. Soc. Am. **36**, 576–581.

Durlach, N. I., and Colburn, H. S. (**1978**). ''Binaural Phenomena,'' in *Hearing, Vol. IV of Handbook of Perception*, edited by E. C. Carterette and M. P. Friedman (Academic, New York).

Eddins, D. A. (**1994**). ''Temporal and spectral processing in increment detection,'' J. Acoust. Soc. Am. **95**, 2963(A).

Eddins, D. A. (**1996**). ''Temporal cues in monaural and binaural detection and discrimination,'' J. Acoust. Soc. Am. **99**, 2471(A).

Fantini, D. A., Moore, B. C. J., and Schooneveldt, G. P. (**1993**). ''Comodulation masking release as a function of type of signal, gated or continuous masking, monaural or dichotic presentation of flanking bands, and center frequency,'' J. Acoust. Soc. Am. **93**, 2106–2115.

Forrest, T. G., and Green, D. M. (**1987**). ''Detection of partially filled gaps in noise and the temporal modulation transfer function,'' J. Acoust. Soc. Am. **82**, 1933–1943.

Green, D. M., Berg, B. G., Dai, H., Eddins, D. A., Onsan, Z., And Nguyen, Q. (**1992**). ''Spectral shape discrimination of narrow-band sounds,'' J. Acoust. Soc. Am. **92**, 2586–2597.

Hafter, E. R., and Trahiotis, C. (**1994**). ''Functions of the binaural system,'' in *Handbook of Acoustics*, edited by M. J. Crocker (Wiley, New York).

Hall, III, J. W., and Grose, J. H. (**1990**). ''Comodulation masking release and auditory grouping,'' J. Acoust. Soc. Am. **88**, 119–125.

Hall, III, J. W., and Grose, J. H. (**1995**). ''Amplitude discrimination in masking release paradigms,'' J. Acoust. Soc. Am. **98**, 847–852.

Hall, III, J. W., and Harvey, A. D. G. (**1984**). ''NoSo and NoSπ thresholds as a function of masker level for narrow-band and wideband masking noise,'' J. Acoust. Soc. Am. **76**, 1699–1703.

Hall, III, J. W., Grose, J. H., and Hartmann, W. M. (**1996**). ''The masking-level difference in low-noise noise,'' J. Acoust. Soc. Am. **99**, 2595(A).

Hall, III, J. W., Grose, J. H., and Hartmann, W. M. (**1998**). ''The masking-level difference in low-noise noise,'' J. Acoust. Soc. Am. **103**, 2573–2577.

Hartmann, W. M. (**1987**). ''Temporal fluctuations and the discrimination of spectrally dense signals by human listeners,'' in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ).

Hartmann, W. M., and Pumplin, J. (**1988**). ''Noise power fluctuations and the masking of sine signal,'' J. Acoust. Soc. Am. **83**, 2277–2289.

Hartmann, W. M., and Pumplin, J. (**1991a**). ''Periodic signals with minimal power fluctuations,'' J. Acoust. Soc. Am. **90**, 1986–1999.

Hartmann, W. M., and Pumplin, J. (**1991b**). See AIP Document No: PAPS-JASMA-90-1986-15 for 15 pages of extended tables for minimum power fluctuations. Order by PAPS number and journal reference from American Institute of Physics, Physics Auxiliary Publication Service, Carolyn Gehlbach, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999. Fax: 516-576-2223, e-mail: paps@aip.org. The price is $1.50 for each microfiche (98 pages) or $5.00 for photocopies of up to 30 pages, and $0.15 for each additional page over 30 pages. Airmail additional. Make checks payable to the American Institute of Physics.

Hatch, D. R., Arne, B. C., and Hall, III, J. W. (**1995**). ''Comodulation masking release (CMR): Effects of gating as a function of number of flanking bands and masker bandwidth,'' J. Acoust. Soc. Am. **97**, 3768–3774.

Isabelle, S. K., and Colburn, H. S. (**1991**). ''Detection of tones in reproducible narrow-band noise,'' J. Acoust. Soc. Am. **89**, 352–359.

Jain, M., Gallagher, D. T., Koehnke, J., and Colburn, S. (**1991**). ''Fringed correlation discrimination and binaural detection,'' J. Acoust. Soc. Am. **90**, 1918–1926.

Jeffress, L. A. (**1972**). ''Binaural signal detection: Vector theory,'' in *Foundations of Modern Auditory Theory, Vol. II*, edited by J. V. Tobias (Academic, New York).

Kidd, Jr., G., Uchanski, R. M., Mason, C. R., and Deliwala, P. S. (**1993**). ''Discriminability of narrow-band sounds in the absence of level cues,'' J. Acoust. Soc. Am. **93**, 1028–1037.

Klumpp, R. G., and Eady, H. R. (**1956**). ''Some measurements of interaural time difference thresholds,'' J. Acoust. Soc. Am. **28**, 859–860.

Koehnke, J., Colburn, H. S., and Durlach, N. I. (**1986**). ''Performance in several binaural-interaction experiments,'' J. Acoust. Soc. Am. **79**, 1558–1562.

Kohlrausch, A. (**1986**). ''The influence of signal duration, signal frequency

and masker duration on binaural masking level differences,'' Hearing Res. **23**, 267–273.
Kohlrausch, A. (**1996**). ''New masking experiments using low-noise noise maskers,'' J. Acoust. Soc. Am. **99**, 2543(A).
Kohlrausch, A., Par, S. v. d., and Houtsma, A. J. M. (**1995**). ''A new approach to study binaural interaction at high frequencies,'' in *Proceedings of the 10th international Symposium on Hearing, Irsee 94*, edited by G. Manley, G. Klump, G. Koppl, H. Fastl, and H. Oeckinghaus (World Scientific, Singapore).
Kuwada, S., Stanford, T. R., and Batra, R. (**1987**). ''Interaural phase-sensitive units in the inferior colliculus of the unanaesthetized rabbit: Effects of changing frequency,'' J. Neurophysiol. **57**, 1338–1360.
Langhans, A., and Kohlrausch, A. (**1992**). ''Differences in auditory performance between monaural and diotic conditions. I: Masked thresholds in frozen noise,'' J. Acoust. Soc. Am. **91**, 3456–3470.
Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.
McFadden, D., and Pasanen, E. G. (**1976**). ''Lateralization at high frequencies based on interaural time differences,'' J. Acoust. Soc. Am. **59**, 634–639.
McFadden, D., and Pasanen, E. G. (**1978**). ''Binaural detection at high frequencies with time-delayed waveforms,'' J. Acoust. Soc. Am. **63**, 1120–1131.
McFadden, D., and Wright, B. A. (**1992**). ''Temporal decline of masking and comodulation masking release,'' J. Acoust. Soc. Am. **92**, 144–156.
Metz, P. J., von Bismark, G., and Durlach, N. I. (**1968**). ''Further results on binaural unmasking and the EC model II: Noise bandwidth and interaural phase,'' J. Acoust. Soc. Am. **43**, 1085–1091.
Nuetzel, J. M., and Hafter, E. R. (**1981**). ''Discrimination of interaural delays in complex waveforms: Spectral effects,'' J. Acoust. Soc. Am. **69**, 1112–1118.
Pumplin, J. (**1985**). ''Low-noise noise,'' J. Acoust. Soc. Am. **78**, 100–104.
Rayleigh, Lord (**1907**). ''On our perception of sound direction,'' Philos. Mag. **13**, 214–232.
Richards, V. M. (**1992**). ''The detectability of a tone added to narrow bands of equal-energy noise,'' J. Acoust. Soc. Am. **91**, 3424–3435.
Richards, V. M., Heller, L. M., and Green, D. M. (**1991**). ''The detection of a tone added to a narrow band of noise: The energy model revisited,'' J. Exp. Psychol. **43**, 481–501.
Staffel, J. G., Hall, III, J. W., Grose, J. H., and Pillsbury, H. C. (**1990**). ''NoSo and NoSπ detection as a function of masker bandwidth in normal-hearing and cochlear-impaired listeners,'' J. Acoust. Soc. Am. **87**, 1720–1727.
Stern. R. M., and Trahiotis, C. (**1992**). ''The role of consistency of interaural timing over frequency in binaural lateralization,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, K. Horner, and L. Demany (Pergamon, Oxford).
Stern, R. M., and Trahiotis, C. (**1995**). ''Models of binaural interaction,'' in *Hearing: Handbook of Perception and Cognition*, edited by B. C. J. Moore (Academic, San Diego).
Stern, R. M., and Shear, D. G. (**1996**). ''Lateralization and detection of low-frequency binaural stimuli: Effects of distribution of interaural delay,'' J. Acoust. Soc. Am. **100**, 2278–2288.
Stern, R. M., Zeiberg, A. S., and Trahiotis, C. (**1988**). ''Lateralization of complex stimuli: A weighted image model,'' J. Acoust. Soc. Am. **84**, 156–165.
van der Par, S., and Kohlrausch, A. (**1995**). ''Analytical expressions for the envelope correlation of certain narrow-band stimuli,'' J. Acoust. Soc. Am. **98**, 3157–3169.
Viemeister, N. F. (**1979**). ''Temporal modulation transfer functions based upon modulation thresholds,'' J. Acoust. Soc. Am. **66**, 1364–1380.
Weiss, T. F., and Rose, C. (**1988**). ''A comparison of synchronization filters in different auditory receptor organs,'' Hearing Res. **33**, 175–180.
Yost, W. A., Wightman, F. L., and Green, D. M. (**1971**). ''Lateralization of filtered clicks,'' J. Acoust. Soc. Am. **50**, 1526–1531.
Zurek, P. M. (**1991**). ''Probability distributions of interaural phase and level differences in binaural detection stimuli,'' J. Acoust. Soc. Am. **90**, 1927–1932.
Zurek, P. M., and Durlach, N. I. (**1987**). ''Masker-bandwidth dependence in homophasic and antiphasic tone detection,'' J. Acoust. Soc. Am. **81**, 459–464.

Masker fluctuation and the masking-level difference

John H. Grose and Joseph W. Hall III
The Division of Otolaryngology Head & Neck Surgery, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7070

(Received 22 December 1996; revised 27 July 1997; accepted 15 January 1998)

Previous work suggests that, for some listeners, masker fluctuation may be advantageous for NoSπ detection. This study tested the hypothesis that the benefit of masker fluctuation to binaural analysis is based on the ability to take advantage of epochs of low masker energy in the fluctuating masker where the cues underlying binaural signal detection are more salient. The hypothesis was evaluated using the time-domain version of COSS analysis derived by Buus *et al.* [J. Acoust. Soc. Am. **99**, 2288–2297 (1996)] which measures the perceptual weight applied by a listener within a relatively brief time window as a function of the masker level during the window. The results indicated a dependency of signal detection on short-term masker level in the NoSπ condition but not in the NoSo condition. This finding supports a new perspective indicating that binaural signal detection depends upon the envelope of a masker in a way that is fundamentally different from that typically associated with monaural detection. © *1998 Acoustical Society of America.*
[S0001-4966(98)01505-7]

PACS numbers: 43.66.Pn, 43.66.Dc [RHD]

INTRODUCTION

In its most common form, the masking-level difference (MLD) represents the difference in threshold levels between a signal presented interaurally in phase (So) and the same signal presented interaurally out of phase (Sπ) when both signals are masked by a noise which is presented interaurally in phase (No) (Hirsh, 1948). The magnitude of the MLD is greater at low frequencies than at high frequencies, consistent with interpretations based on the dominant role of stimulus temporal fine structure (Durlach, 1963; Hirsh, 1948; Rabiner *et al.*, 1966; Webster, 1951). However, the role of envelope cues in the binaural analysis of high-frequency stimuli has received considerable attention (Bernstein and Trahiotis, 1992; McFadden and Pasanen, 1978; Zurek and Durlach, 1987). Several recent studies have also examined the role of masker envelope, or fluctuation, in the binaural analysis of low-frequency signals (Eddins, 1996; Hall *et al.*, 1998).

The study by Hall *et al.* (1998) compared MLD performance in random-noise maskers and ''low-fluctuation noise'' maskers for a low-frequency signal. In the low-fluctuation noise condition, the degree of masker fluctuation was minimized either by selecting component phases which resulted in a low fourth moment of the waveform or by modulating a pure tone with an attenuated dc-shifted low-pass noise. The intent of the study was to test predictions arising from two classes of model of the MLD: cross-correlation models (Colburn, 1973; Jeffress, 1948; Stern and Colburn, 1978; Stern and Trahiotis, 1992) and the Equalization/Cancellation model (Durlach, 1963). For NoSo detection, it was expected that thresholds would be lower for low-fluctuation noise maskers than for random-noise maskers because, by extension from monaural conditions (Hartmann and Pumplin, 1988), tonal signals are more readily extracted from maskers whose energy fluctuations are minimal. For NoSπ detection, the two classes of model differ in their predictions. Because an Equalization/Cancellation model reduces to an energy-based mechanism for detecting signals in noise, it can be argued that such a model would predict a correspondingly lower NoSπ threshold in low-fluctuation noise than in random noise. The net result would be a constancy of MLD magnitude measured in the two masker types. A cross-correlation model on the other hand, being based on the pattern of interaural correlation, might predict no change in NoSπ threshold between the two masker types since the pattern of correlation is not expected to be a function of envelope fluctuation. Such a model would therefore predict a smaller MLD in low-fluctuation noise than in random noise. The results of the Hall *et al.* (1998) study showed partial support for the cross-correlation model in that half of the subjects showed the predicted pattern of reduced NoSo thresholds in low-fluctuation noise relative to random noise with no corresponding change in NoSπ thresholds. However, the unpredicted result was that half of the subjects showed a *decrease* in NoSπ threshold (better performance) in the random noise masker. This result suggested that, for some listeners, there is an advantage of masker fluctuation for NoSπ detection. The results of Eddins (1996) also support this interpretation.

The purpose of this study was to directly test the hypothesis that masker fluctuation is advantageous for NoSπ detection. The hypothesis is based on the premise that cues underlying binaural signal detection are more salient during epochs of low masker energy in a fluctuating masker than during epochs of high masker energy. One interpretation of the heightened cue saliency during masker ''dips'' is that the auditory system attributes greater weight to stimulus epochs where masker energy is low and where the signal-to-masker ratio is correspondingly high. This strategy amounts to ''listening in the dips,'' a strategy proposed to underlie comodulation masking release (CMR) (Buus, 1985; Buus *et al.*, 1996). Buus *et al.* (1996) tested the dip-listening hypothesis

of CMR by modifying a quantitative method known as COSS analysis (Berg, 1989) to apply to the time domain. The time-domain version of COSS analysis put forward by Buus *et al.* (1996) measures the weight attributed to a relatively brief ''listening window'' as a function of the masker level during the window. A dip-listening strategy is suggested by the result that the signal weights applied by the listener are high during epochs where the short-term masker level is low (during the envelope dips), and that weights are low when the short-term masker level is high (during the envelope peaks). In the present context, it is hypothesized that if masker fluctuation is advantageous to NoSπ detection, and if this advantage is due to a dip-listening strategy, then a time-domain COSS analysis should indicate relatively high listener weights during minima in the masker envelope and correspondingly low weights during envelope maxima. For NoSo detection, where no benefit of masker fluctuation is expected, it is predicted that listener weights should be independent of short-term masker level.

I. METHOD

A. Time-domain COSS analysis

In its original form, COSS analysis (Berg, 1989) derives the perceptual weights applied by a listener to the individual spectral components within a complex signal by measuring the sensitivity of the listener to random perturbations placed independently on each component during a signal detection task [specifically, Profile Analysis (Berg and Green, 1990)]. Buus *et al.* (1996) modified the technique to apply to a single-component signal. Rather than perturbing the individual (simultaneous) components of a spectrally complex signal, individual (nonsimultaneous) segments of a single-tone signal were perturbed. This was achieved by employing a train of iso-frequency tone bursts as the signal, where the amplitude of each tone burst was independently varied. By measuring the sensitivity of the listener to perturbations in the amplitude of individual tone bursts as a function of the short-term masker levels in which each tone burst occurred, the ability of the listener to independently weight temporal epochs of the masker could be assessed. The complete quantitative derivation of the time-domain COSS analysis is provided in Buus *et al.* (1996); key elements in the procedure are described at the appropriate junctures below.

B. Subjects

Four listeners with normal hearing participated in the experiment. Each listener had audiometric thresholds better than 20 dB HL across the octave frequencies 250–8000 Hz. All listeners were highly experienced in psychoacoustic tasks.

C. Stimuli

The signal to be detected was a train of 500-Hz tone bursts masked by a narrow band of noise centered at 500 Hz. The masking noise was obtained by digitizing a 50-Hz-wide band of Gaussian noise centered at 500 Hz. This noise was generated by passing a broadband, constant spectrum level noise through a digital filter (IHR Universal). The sampling rate of the filter was 5000 Hz and the output was lowpass filtered at 2000 Hz (VBF8 90 dB/oct). The 50-Hz-wide band of noise was then adjusted in level to optimize the dynamic range of the ADC (TDT DD1). Next, 100 segments of the noise, each segment 600 ms in duration, were digitized at a sampling rate of 5000 Hz. Following digitization, each segment was shaped with a 20-ms cosine-squared rise/fall time and stored on disk. For each observation interval of a two-alternative forced choice (2AFC) trial, one of the 100 noise segments was randomly selected for presentation, the only limitation being that the same segment could not be selected for presentation in both intervals of the same trial.

The signal consisted of a train of six tone bursts, each 50 ms in duration, centered at 500 Hz, and digitally generated at a sampling rate of 5000 Hz. Each tone burst was shaped with a 15-ms rise/fall time, where the duration of the rise/fall refers to the interval between the 10% amplitude and the 90% amplitude points. This rise/fall time was selected as being the maximum duration that would allow the peak amplitude of the tone burst to remain approximately equivalent to the peak amplitude of an unshaped 500-Hz pure tone. The level of the tone burst was then expressed in terms of the sound pressure level of the continuous unshaped pure tone. In the first stage of the experiment, the amplitude of the tone bursts was held constant. In the second stage of the experiment, the amplitude of each individual tone burst was randomly perturbed, with the perturbation being drawn from a normal distribution having a mean of 0 and a standard deviation of 3 dB. (The perturbation was limited to ± 2 standard deviations.) When presented together, the masker onset preceded the signal onset by 200 ms; the masker offset therefore lagged the signal offset by 100 ms. In each interval of a 2AFC trial, the selected masker waveform and (if the target interval) the signal waveform were output through separate DAC channels (TDT DD1) at a rate of 5000 Hz. Following low-pass filtering at 2000 Hz (VBF8 90 dB/oct), the masker waveform was presented diotically at an overall level of 80 dB SPL and the signal was presented either diotically (to yield NoSo) or was inverted at one ear (to yield NoSπ). Stimuli were delivered over Sony MDR-V6 headphones to the listener seated in a double-walled sound-treated booth.

D. Procedure

In the first stage of the experiment, masked thresholds were obtained for NoSo detection and NoSπ detection. As noted above, the amplitudes of the six tone bursts in the tone-burst train were equal for this stage of the experiment. A 2AFC procedure was employed which incorporated a three-down, one-up stepping rule to converge upon the 79.4% detection level. The initial step size was 8 dB and this was reduced to 4 dB after two reversals in level direction; after a further two reversals the step size was reduced to its final level of 2 dB. A run was terminated after 12 reversals and the mean of the final 8 reversal levels was accepted as the threshold estimate for that run, provided that the standard deviation of those 8 reversal levels did not exceed 3 dB in which case the run was rejected and another one undertaken. For each signal phase condition, five threshold estimates

were collected and the mean of these five was taken as the threshold signal level. The MLD was calculated as the difference between level of the NoSo threshold and the level of the NoSπ threshold.

In the second stage of the experiment, a 2AFC procedure continued to be used but now the signal level was not varied *adaptively*. Instead, for each listener, the signal level which elicited 79.4% correct detection for each condition was targeted. However, in each trial of a 50-trial block, the level of each tone burst in the burst train was independently perturbed as described above. Each listener undertook 5000 trials and, since each trial contained six independent epochs of signal level, a total of 30 000 signal+masker combinations were collected for each listener and condition. For each listener and condition, a record was kept of every trial, including the masker waveform presented during the signal interval and the amplitudes of the six tone bursts presented in that interval.

E. Data analysis

In a preliminary step, the masker levels within 20-ms rectangular windows centered on the peaks of the signal tone-burst occurrences were computed for each of the 100 masker waveforms. These 600 levels (6 epochs $\times$100 waveforms) were divided into 10 bins of 60 levels each. Thus the bin at one extreme contained masker ''peaks'' and the bin at the other extreme contained masker ''dips.'' A look-up table was constructed in which, for each masker waveform, the sequence of six bins (i.e., the sequence of six short-term masker levels) contained within that waveform was tabulated.

After the 5000 trials for a listener and condition had been collected, the 30 000 random amplitude perturbations were also divided into 10 bins, with approximately the same number of amplitude values per bin. A 10×10 matrix was then constructed which tabulated the 10 masker level bins against the 10 signal level bins. Within each cell of the matrix, the results of the trial which included that signal level/masker level combination were accumulated. Each cell contained approximately 300 observations. For the roughly 300 observations per cell, the percent correct score was computed and converted to d' units using the tables of Green (cf. Green and Dai, 1991).

II. RESULTS

The first phase of the experiment involved the measurement of a standard MLD for the tone-burst train signal where the level of each tone burst was equal. The mean NoSo threshold level for the four listeners was 83.1 dB SPL (standard deviation=1.4 dB) and the mean NoSπ threshold level was 59.4 dB SPL (standard deviation=3.3 dB). The mean MLD was therefore 23.7 dB, which is in line with other studies using similar masker bandwidths and signal frequencies (Grose *et al.*, 1997, 1995; Zurek and Durlach, 1987).

The second phase of the experiment involved the derivation of perceptual weights for each listener and condition. The key premise in this step is that, for a given short-term masker level, the function relating d' to the energy of a signal segment (i.e., a single tone burst) should be a straight line whose slope is proportional to the perceptual weight applied to that signal segment (cf. Buus *et al.*, 1996, equation 6).[1] Accordingly, for each listener, signal phase condition, and short-term masker level, a plot was generated of d' as a function of signal intensity (i.e., the tone-burst intensity centered in each of the ten signal level bins). Figure 1 shows the results of this analysis for each of the four listeners for the two extremes of short-term masker level: the masker peaks and the masker dips. Each panel in the upper row shows the results for one listener for the NoSo condition; the lower row of panels shows the corresponding NoSπ results. In each panel, the squares indicate data for masker peaks and triangles indicate data for masker dips. Notice that the dependency of d' on signal intensity appears more marked for masker dips than for masker peaks in the NoSπ condition in comparison to the NoSo condition.[2]

In order to extract the perceptual weights, a linear regression was undertaken for each plot and the slope of the line taken as an indicator of weight for that combination of signal phase (So or Sπ) and short-term masker level. These slopes were then normalized for each listener and condition according to the equation

$$\text{normalized weight}_k = \frac{10b_k}{6\sum_{i=1}^{10} b_i}, \tag{1}$$

where the normalized weight applied to the signal during the kth range of short-term masker level is equal to the slope, b, of the regression line for that kth masker level range scaled by a value proportional to the sum of all the slopes for that condition. This normalization procedure results in a weight of 0.167 for the situation where equal weights are applied to the signal across all ranges of short-term masker level. The normalized weights computed for the individual listeners were then averaged for each signal phase condition and the results are shown in Fig. 2. Mean normalized weights are plotted against short-term masker level for both the NoSo condition (squares) and NoSπ condition (triangles). The solid line indicates the result expected if perceptual weights were uniform across short-term masker level. These data were submitted to a regression analysis which indicated that the perceptual weights for the NoSo condition did not depend upon short-term masker level [$F(1,8)=1.37$; $p=0.28$]. However, the weights for the NoSπ condition did depend upon short-term masker level [$F(1,8)=162.58$; $p<0.0001$], with greater weight being given to the signal during low masker levels.

III. DISCUSSION

The finding that listeners placed greater weight on the signal during epochs of low masker energy in the NoSπ condition but not in the NoSo condition supports the hypothesis that masker fluctuation is advantageous to NoSπ detection and that this advantage is associated with the masker minima. This is consistent with the findings of Isabelle (1995) who observed that the largest interaural differences (in terms of both phase and level) brought about by the addition of an antiphasic signal to a diotic masker occurred at

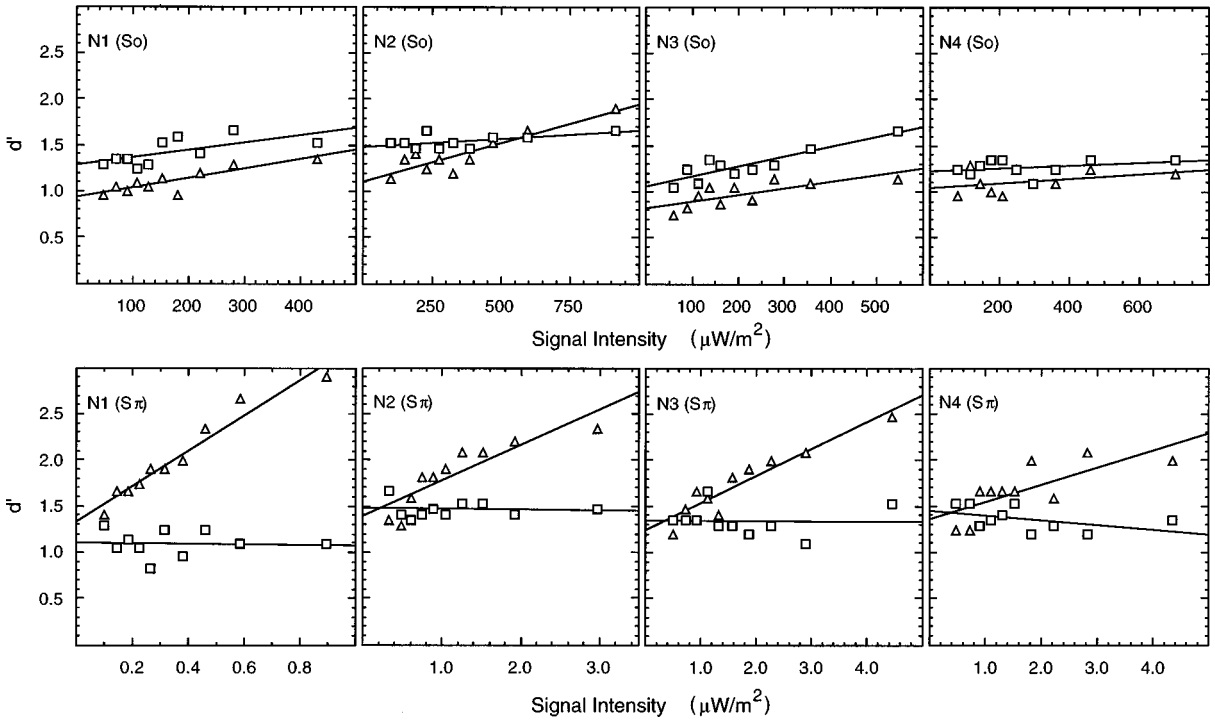


FIG. 1. Individual data for the two extremes of short-term masker level: masker peaks (squares) and masker dips (triangles). Each panel plots d' as a function of signal intensity (the tone-burst intensity centered in the range of each of the ten signal level bins). The upper row of panels show results for the NoSo condition; the lower row of panels shows the corresponding NoSπ results. Each column of panels represents an individual listener. The solid lines are best-fitting linear regressions.

minima in the masker envelope. This resulted in a significant correlation of NoSπ performance with instantaneous interaural time differences in some subjects. However, before interpreting the present finding as supporting a dip-listening strategy, some clarification of terminology is necessary. The strategy of ''listening in the dips,'' or ''dip listening,'' is associated with the phenomenon of CMR and connotes a mechanism whereby the extracted envelope of a masker remote from the signal is used to weight the contribution of stimulus epochs at the target frequency to the signal detection process. In a sense, the degree to which information at the target frequency is attended to is determined by the status of some reference envelope. In the present case of NoSπ detection, it is less parsimonious to invoke the existence of a reference envelope to control the distribution of weights. Rather, a simple monitoring of interaural differences (with uniform weight across time) would yield the same result since the influence of the signal on momentary interaural differences is relatively greater during masker minima. Thus the weights attributed to the signal in the present case of binaural detection should probably be considered a metric of cue effectiveness.

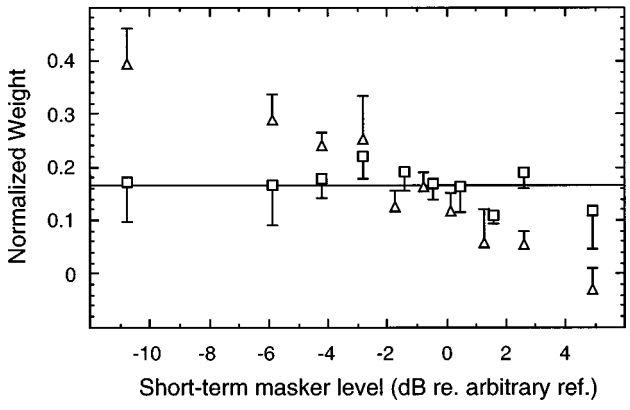


FIG. 2. Mean normalized weights as a function of short-term masker level for both the NoSo condition (squares) and NoSπ condition (triangles). Error bars are one standard deviation. The solid line indicates the result expected if perceptual weights were uniform across short-term masker level.

In conclusion, the results of this study provide new, direct evidence indicating that binaural detection differs fundamentally from monaural detection for signals in random noise in terms of the particular temporal stimulus epochs that determine signal detection. Specifically, information in the masker dips is relatively more effective in NoSπ detection than NoSo detection. It is interesting that a similar dependency of threshold on masker dip information is seen in the special monaural case of CMR. Whether the processes accounting for the masker dip phenomena in monaural CMR experiments and the present MLD experiment are related remains to be determined. The hypothesis that masker fluctuation is advantageous to NoSπ detection because of the implementation of a dip-listening strategy can be further tested. First, the hypothesis predicts that perceptual weights should be more uniform across short-term masker levels for NoSπ detection in low-fluctuation noise maskers. Second, the benefit of masker fluctuation for NoSπ detection should be evident in higher frequency regions. Evaluation of these predictions should shed further light on the role of masker fluctuation in binaural analysis.

ACKNOWLEDGMENTS

The comments of two anonymous reviewers on a previous version of this paper are gratefully acknowledged. This work was supported by the NIDCD (R01-DC01507).

[1]The decision variable in a 2AFC task is normally computed as the difference between the observations in the signal and standard intervals. In the present application, as equations 2–4 of Buus *et al.* (1996) show, this reduces to the weighted sum of the signal energies because (a) the masker is assumed to be independent of the signal in the signal interval and (b) the difference between the sums of the weighted masker energies in the two intervals is expected to be zero.

[2]The NoSo data (upper panels of Fig. 1) provide the impression that a signal of any given intensity is more detectable when presented in the peaks of the masker than in the dips. This counterintuitive trend is also evident in the analogous Single Band data of Buus *et al.* (1996) [see their Fig. 1]. Interpretation of this trend awaits further investigation.

Berg, B. (**1989**). ''Analysis of weights in multiple observation tasks,'' J. Acoust. Soc. Am. **86**, 1743–1746.

Berg, B., and Green, D. M. (**1990**). ''Spectral weights in profile listening,'' J. Acoust. Soc. Am. **88**, 758–766.

Bernstein, L., and Trahiotis, C. (**1992**). ''Detection of antiphasic sinusoids added to the envelopes of high-frequency bands of noise,'' Hearing Res. **62**, 157–165.

Buus, S. (**1985**). ''Release from masking caused by envelope fluctuations,'' J. Acoust. Soc. Am. **78**, 1958–1965.

Buus, S., Zhang, L., and Florentine, M. (**1996**). ''Stimulus-driven, time-varying weights for Comodulation Masking Release,'' J. Acoust. Soc. Am. **99**, 2288–2297.

Colburn, H. S. (**1973**). ''Theory of binaural interaction based on auditory nerve data. I. General strategy and preliminary results on interaural discrimination,'' J. Acoust. Soc. Am. **54**, 1458–1470.

Durlach, N. I. (**1963**). ''Equalization and cancellation theory of binaural masking-level differences,'' J. Acoust. Soc. Am. **35**, 1206–1218.

Eddins, D. A. (**1996**). ''Temporal cues in monaural and binaural detection and discrimination,'' J. Acoust. Soc. Am. **99**, 2471(A).

Green, D., and Dai, H. (**1991**). ''Probability of being correct with 1 of M orthogonal signals,'' Percept. Psychophys. **49**, 100–101.

Grose, J. H., Hall, J. W., and Dev, M. (**1997**). ''The MLD in children: Effects of signal and masker bandwidths,'' J. Speech Hear. Res. **40**, 955–959.

Grose, J. H., Hall, J. W., and Mendoza, L. (**1995**). ''Developmental effects in complex sound processing,'' in *10th International Symposium on Hearing*, edited by G. A. Manley, G. M. Klump, C. Koppl, H. Fastl, and H. Oekinghaus (World Scientific, Singapore).

Hall, J. W., Grose, J. H., and Hartmann, W. M. (**1998**). ''The masking-level difference in low-noise noise,'' J. Acoust. Soc. Am. **103**, 2573–2577.

Hartmann, W. M., and Pumplin, J. (**1988**). ''Noise power fluctuations and the masking sine signals,'' J. Acoust. Soc. Am. **83**, 2277–2289.

Hirsh, I. J. (**1948**). ''Influence of interaural phase on interaural summation and inhibition,'' J. Acoust. Soc. Am. **20**, 536–544.

Isabelle, S. (**1995**). ''Binaural detection performance using reproducible stimuli,'' Ph.D. dissertation, Boston University.

Jeffress, L. A. (**1948**). ''A place theory of sound localization,'' J. Comp. Physiol. Psychol. **41**, 35–39.

McFadden, D. M., and Pasanen, E. G. (**1978**). ''Binaural detection at high frequencies with time-delayed waveforms,'' J. Acoust. Soc. Am. **63**, 1120–1131.

Rabiner, L. R., Laurence, C. L., and Durlach, N. I. (**1966**). ''Further results on binaural masking and the E.C. model,'' J. Acoust. Soc. Am. **40**, 62–70.

Stern, R., and Colburn, H. (**1978**). ''Theory of binaural interaction based on auditory-nerve data IV,'' J. Acoust. Soc. Am. **64**, 127–140.

Stern, R. M., and Trahiotis C. (**1992**). ''The role of consistency of interaural timing over frequency in binaural lateralization,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, K. Horner, and L. Demany (Pergamon, Oxford).

Webster, F. A. (**1951**). ''The influence of interaural phase on masked thresholds: the role of interaural time deviation,'' J. Acoust. Soc. Am. **23**, 452–462.

Zurek, P. M., and Durlach, N. I. (**1987**). ''Masker-bandwith dependence in homophasic and antiphasic tone detection,'' J. Acoust. Soc. Am. **81**, 459–464.

Spontaneous classification of complex tones at high and ultrasonic frequencies in the bat, *Megaderma lyra*

Annemarie Preisler and Sabine Schmidt
Zoologisches Institut der Universität, Luisenstrasse 14, 80333 München, Germany

(Received 13 January 1997; accepted for publication 20 January 1998)

Megaderma lyra, a bat species using harmonically structured calls for echolocation, exploits the spectral content of its echoes for texture discrimination. It is the aim of the present study to test according to which sensory qualities harmonic complex tones are spontaneously classified by this bat. The applied experimental paradigm is especially adapted to the preference of *M. lyra* to use absolute pitch cues. Three animals were trained in a 2-AFC procedure to classify three-component stimuli as low or high, with all their harmonics below or above a pure tone reference of 33 kHz, respectively. Later, the original tones were interspersed with ''incomplete'' test stimuli, with their fundamentals (and lower harmonics) missing. These were ambiguous in that their possible virtual, i.e., collective pitches were below the reference whereas their pure tone pitches were above it. Bat 1 classified 22 of 23 test stimuli with missing fundamentals between 5.3 and 28.3 kHz according to their collective pitches, whereas bat 2 judged all presented ambiguous tones on the basis of their pure tone pitches. Bat 3 failed the pitch control criterion which is why in this case results cannot be interpreted unequivocally. The implications of these findings are discussed with respect to the bats' behavioral context, as well as to psychoacoustical models of the formation of the pitch of complex tones. © *1998 Acoustical Society of America.* [S0001-4966(98)00405-6]

PACS numbers: 43.66.Gf, 43.80.Lb [JWH]

INTRODUCTION

Echolocating bats predominantly perceive their environment by acoustical cues extracted from echoes of their sonar calls (Schnitzler and Henson, 1980). Recent studies on texture perception in bats with a broadband echolocation system have shown that the animals use the spectral characteristics of their echoes to discriminate between targets (Schmidt, 1992; Mogdans *et al.*, 1993). However, the sensory qualities extracted from the broadband spectral information are still unknown.

The present study was designed to test the significance of different sensory qualities, as already established for human perception, in *Megaderma lyra*, a bat species using harmonically structured sonar calls.

Human listeners are able to classify complex tones with respect to the pitches of the most prominent spectral components or to the pitch of the fundamental, and to timbre. Although pitch sensations evoked by single spectral components are likely to reflect the most elementary stage of neural processing (Whitfield, 1980), analytical listening requires training and seems to be the exception rather than the rule (Lichte, 1941; Hartmann, 1988).

In 1843, Seebeck made the unexpected observation that the perceived fundamental pitch is not affected when the first harmonic is removed from a spectrum. Later psychoacoustic experiments revealed that the case of the *missing fundamental* is just a special aspect of complex pitch perception. Terhardt *et al.* (1982) report that ''complete'' complex tones with a first harmonic below 800 Hz elicit fundamental pitch sensations that are more strongly determined by upper harmonics than by the fundamental frequency itself. For tones with fundamentals at higher frequencies, pitch sensations evoked by single spectral components become predominant. In the literature, for the latter type of sensations the term *spectral pitch* has been adopted. But as ''spectral'' refers to the physical domain and ''pitch'' to the subjective perceptual domain, we prefer the expression *pure tone pitch*. Low-frequency pitch sensations, not directly evoked by corresponding spectral components, have been termed *residue pitch* (Schouten, 1940a, b; De Boer, 1956, 1976; Ritsma, 1967; Patterson, 1973), *periodicity pitch* (Licklider, 1954; Walliser, 1969), and *virtual pitch* (Terhardt, 1974). Unfortunately, all these expressions still imply assumptions about the underlying mechanisms and thus are counterproductive for an unbiased discussion. We therefore, when referring to the pitch of complex tones, in the following confine ourselves to speak of *collective pitch* sensations. For the sake of conformity, the expression is also applied to models and data from literature, where originally a different terminology was used.

While collective pitch sensations do not depend on the presence/absence of single partials, i.e., on the exact spectral composition of a complex tone (Patterson, 1973; Terhardt, 1974; Preisler, 1993), timbre depends primarily on the frequency spectrum, although it also depends on the sound pressure and the temporal characteristics of the sound (ANSI S1.1-1994, 12.09). In other words, a complex tones' timbre is characterized by its specific spectral composition irrespective of its collective pitch. Thus, according to this definition, collective pitch and timbre represent complementary higher-order perceptual qualities.

Smoorenburg (1970) developed a two-alternative forced choice (2-AFC) test in order to test the predominance of pure tone versus collective pitch sensations. The method involves sequences of ''incomplete'' tones that are ambiguous in that their missing fundamentals rise whereas their present har-

monics fall in pitch (or vice versa). Smoorenburg (1970) reports that for two-tone stimuli less than half of the tested subjects spontaneously classified most of the sequences according to their collective pitches. However, subsequent experiments revealed that the proportion of ''collective pitch perceivers'' increases with the number of presented harmonics and strongly depends on the energy distribution of the entire spectrum (Preisler, unpublished data). Moreover, a low signal-to-noise ratio increases the percentage of decisions in favor of collective pitch (Hall and Peters, 1981).

For comparative psychoacoustical studies in higher vertebrates, modified versions of the Smoorenburg test, making use of operant conditioning methods, were applied. Results indicate that starlings (Cynx and Shapiro, 1986) and cats (Heffner and Whitfield, 1976) predominantly categorize complex tones according to the pitch of their missing fundamentals. Using a different paradigm, Tomlinson and Schwartz (1988) showed that rhesus monkeys are able to match tones with their fundamentals missing with comparison harmonic complexes having the same pitch.

It has to be emphasized that 2-AFC procedures, allowing a stimulus classification in terms of a dichotomous criterion (e.g., pure tone versus collective pitches), only indicate which characteristic is predominantly perceived by the tested subjects. Obtained preferences, however, do not reveal whether the alternative aspect of information, in principle, is available to these subjects.

Since *M. lyra*, like starlings (Hulse and Cynx, 1985), cebus monkeys, and rats (D'Amato, 1988), has been found to rely mainly on absolute pitch (Schmidt *et al.*, 1994; Preisler and Schmidt, 1995), Sedlmeier (1992a, b) applied a testing procedure especially adapted to this preference. In his experiments, five bats were trained by operant conditioning to classify the collective pitches of complex tones, consisting of the first four harmonics, as higher or lower than a fixed pure tone reference. When the first harmonic was gradually attenuated to a sound-pressure level of 0 dB and masked by bandpass noise in addition, while maintaining positive reinforcement, the animals' classification behavior remained unchanged. Sedlmeier (1992a, b) interpreted this result an evidence for collective pitch sensations in bats. However, in view of the applied experimental paradigm, the possibility of the bats simply having learned the responses by stepwise generalization cannot be ruled out. So, the bats may have memorized each of the tested spectra on the basis of the absolute pitches of their higher harmonics and reacted to them consistently on the basis of positive reinforcement.

The aim of the present study was to find out by which criteria bats spontaneously classify complex tone stimuli and thus to evaluate the significance of different sensory qualities for the animals' natural behavioral context. Therefore, in contrast to Sedlmeier's study, all responses to critical test and control stimuli were rewarded in order to avoid biased operant conditioning.

I. METHODS

A. Experimental setup

The experiments were carried out in a sound-attenuated chamber (2.5 m×3 m×2.2 m). The bats' starting position was placed 20 cm from the middle of the wall opposite to the entrance. The loudspeaker was positioned at a distance of 1.9 m, 40° below the starting position, and directly aligned with the bats' head. Two feeding dishes were mounted symmetrically to the left and right of the loudspeaker. The experimenter, seated behind the speaker, triggered the start of a trial by pressing a button. A schematic top view of the experimental setup is given in Fig. 1.

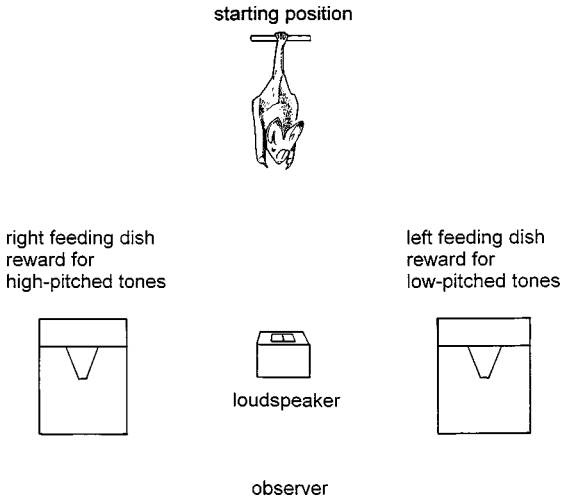


FIG. 1. Experimental chamber (for further explanations, see text).

Stimulus selection, control of external devices, and data collection were performed by a 486 personal computer. Acoustical stimuli were generated using a four-channel-synthesizer (HP 8904). They were fed into a custom-made envelope shaper, followed by an attenuator (Leader LAT-45), a buffer amplifier (chip: Burr Brown, 3553 AM), and a loudspeaker (Technics EAS-10TH800C).

B. Stimuli

1. Time pattern

The time pattern of the stimulus presentation is shown in Fig. 2. The trapezoidally shaped tone pulses had a duration of 400 ms, with rise and decay times of 1 ms. During training, a sequence of at most ten pulse pairs (which in case of reaction was interrupted by the bat's return to the perch) consisting of a reference tone followed by a comparison tone was presented [Fig. 2(a)]. The intertone interval was 400 ms and the interpair interval was 1200 ms. In the testing ses-

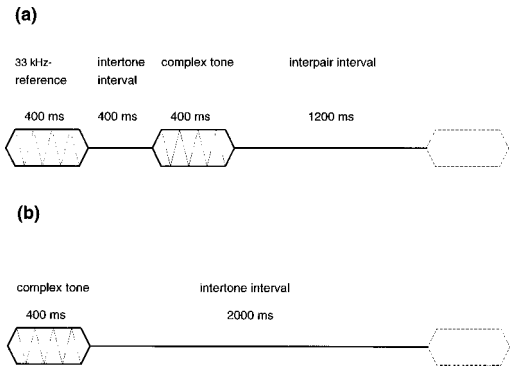


FIG. 2. Time structure of (a) training stimuli and (b) test stimuli.

TABLE I. Spectra and sound-pressure levels of low and high standard stimuli for bats 1–3. The frequencies of the present fundamentals are indicated by f_0.

	Low standard stimuli					High standard stimuli			
No. of task	f_0 (kHz)	Harmonics	dB SPL cond. 1	dB SPL cond. 2	No. of task	f_0 (kHz)	Harmonics	dB SPL cond. 1	dB SPL cond. 2
1	9	1–3	47	65	13	34	1–3	49	67
2	9.131	1–3	47	65	14	34.495	1–3	51	69
3	9.264	1–3	46	64	15	34.996	1–3	49	67
4	9.398	1–3	46	64	16	35.505	1–3	45	63
5	9.535	1–3	49	67	17	36.022	1–3	50	68
6	9.674	1–3	47	65	18	36.546	1–3	46	64
7	9.815	1–3	47	65	19	37.077	1–3	47	65
8	9.957	1–3	49	67	20	37.617	1–3	48	66
9	10.102	1–3	49	67	21	38.164	1–3	49	67
10	10.249	1–3	49	67	22	38.719	1–3	45	63
11	10.398	1–3	48	66	23	39.282	1–3	49	67
12	10.549	1–3	47	65	24	39.853	1–3	48	66

sions, the reference was omitted and a series of at most ten comparison tones was presented with an intertone interval of 2000 ms [Fig. 2(b)].

2. Intensity and spectral composition

The experiments were carried out at two different intensity levels (conditions 1 and 2). Standard stimuli were the same for both conditions, whereas test and control stimuli differed in order to prevent learning transfer.

Overall sound-pressure levels, as measured at the position of the bats' perch, and spectral compositions of all complex tones are listed in Tables I–III. The sound-pressure levels were balanced for standard, test, and control stimuli in order to prevent their discrimination on the basis of level cues.

For all complex tones, the relative intensities of the present harmonics decreased by 6 dB per octave. This decay was chosen as a compromise between the requirements to present test tones having a timbre similar to the bats' own calls (Schmidt, 1992) and to guarantee that all included spectral components lie above the auditory threshold of *M. lyra* (Schmidt *et al.*, 1983).

a. Reference tone. As a reference stimulus, a pure tone pulse with a frequency of 33 kHz was presented (Fig. 3, dashed line).

b. Standard stimuli. We used 24 complex tones, consisting of the first three harmonics, as standard stimuli (Table I). As indicated by the range bars in Fig. 3, the fundamentals (f_0) of the 12 low standards [Fig. 3(a)] varied between 9 and 10.5 kHz, those of the 12 high standards [Fig. 3(b)] between 34 and 39.8 kHz. For both types of standards, the fundamental frequencies were varied in steps of 25 cents ($\frac{1}{8}$ of a whole tone), a distance expected to roughly represent the bat's fre-

TABLE II. Spectra and sound-pressure levels of test and control stimuli for condition 1. The frequencies of the missing fundamentals are indicated by f_0.

No. of task	f_0 (kHz)	Harmonics	dB SPL
Test stimuli (bat 1 only)			
25	5.3	7–10	46
26	6.68	5–8	47
27	8.41	4–7	47
Test stimuli (bats 1–3)			
28	10.6	4–7	44.5
29	13.35	3–6	46.5
30	16.81	2–5	52
31	21.19	2–5	49
32	26.7	2–4	47
High control stimuli (bat 1)			
33	36	2–4	46
34	37	2–4	47
Low control stimuli (bats 2 and 3)			
35	7	2–4	48.5
36	8	2–4	48.5

TABLE III. Spectra and sound-pressure levels of test and control stimuli for condition 2. The frequencies of the missing fundamentals are indicated by f_0.

No. of task	f_0 (kHz)	Harmonics	dB SPL
Test stimuli (bat 1 only)			
37	5.61	6–9	63.5
38	6.3	6–9	65
39	7.07	5–8	64.5
40	7.94	5–8	67
41	8.91	4–7	67
Test stimuli (bats 1–3)			
42	10	4–7	63
43	11.22	3–6	67.5
44	12.6	3–6	66
45	14.14	3–6	64
46	15.87	3–6	62
47	17.82	2–5	68
48	20	2–5	67.5
49	22.45	2–5	69
50	25.2	2–4	65
51	28.28	2–4	64
High control stimuli (bat 1)			
52	34	2–4	64
53	35	2–4	66
Low control stimuli (bats 2 and 3)			
54	9	2,3	65
55	10	2,3	67

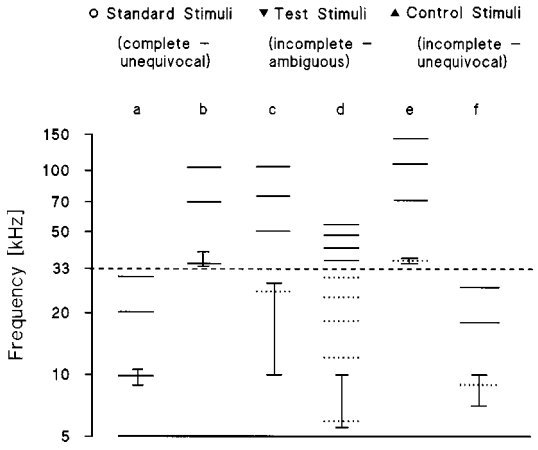


FIG. 3. Examples, indicating the spectral composition of the stimuli [dashed line: pure tone reference at 33 kHz, solid lines: presented harmonics, dotted lines: omitted harmonics, range bars: regions of the present or missing fundamentals (f_0) of all applied complex tones]: (a) low standard stimuli, (b) high standard stimuli, (c) test stimuli for bats 1–3, (d) additional test stimuli for bat 1, (e) control stimuli for bat 1, and (f) control stimuli for bats 2 and 3.

quency resolution for successively presented tones (Sedlmeier, 1992a). Since all single and collective pitches that may be elicited by a given tone are either below or above the reference, these stimuli do not discriminate between both sensory qualities and are unequivocal in this respect. Standard stimuli were presented at a sound-pressure level of 65.8 ± 1.6 dB in condition 1 and of 47.8 ± 1.6 dB in condition 2.

c. Test stimuli. Test stimuli were ''incomplete'' complex tones in which all spectral components below 33 kHz were omitted [Fig. 3(c) and (d)]. These tones are ambiguous in that their spectral components are above the reference, whereas their missing fundamentals are below it.

(i) Bats 1–3. For condition 1, five test tones with missing fundamentals between 10.6 and 26.7 kHz were presented at a sound-pressure level of 47.8 ± 2.8 dB [Table II; Fig. 3(c)]; for condition 2, ten test tones with missing fundamentals between 10 and 28.3 kHz were presented at a sound-pressure level of 65.6 ± 2.4 dB [Table III; Fig. 3(c)].

(ii) Bat 1. In addition, bat 1 was presented with three test tones with very low missing fundamentals between 5.3 and 8.4 kHz (sound-pressure level: 46.7 ± 0.6 dB) for condition 1 [Table II; Fig. 3(d)] and five test tones with very low missing fundamentals between 5.6 and 8.9 kHz (sound-pressure level: 65.4 ± 1.6 dB) for condition 2 [Table III; Fig. 3(d)].

The frequencies of the missing fundamentals were chosen in steps of 400 cents in condition 1 and 200 cents in condition 2.

d. Control stimuli.

(i) High control stimuli [Fig. 3(e)]. For animals classifying the ambiguous test tones as low, four incomplete tones, which are unequivocal in that both their possible collective and pure tone pitches are above the reference, were employed as control stimuli. Their missing fundamentals were set to 36 and 37 kHz (condition 1; Table II) and to 34 and 35 kHz (condition 2; Table III). Since the spectra of both the controls and the highest test tones consist of the second to fourth harmonic, the frequency of the missing fundamentals provides the only cue for correctly discriminating these stimulus types by means of the 33-kHz reference [cf. Fig. 3(c) and (e)]. Thus, an identification of the control stimuli as high indicates collective pitch formation. If test and control stimuli are not discriminated, the bats' behavior has to be accounted for by a characteristic both types have in common, i.e., the specific timbre of complex tones without a fundamental.

(ii) Low control stimuli [Fig. 3(f)]. For animals classifying the ambiguous test tones as high, incomplete tones, which were unequivocal in that both their possible collective and pure tone pitches were below the reference, were chosen as control stimuli. In condition 1, the frequencies of the missing fundamentals were 7 and 8 kHz (Table II); in condition 2, they were 9 and 10 kHz (Table III).

If the controls were categorized as low, the classification of the test tones is to be attributed to pure tone pitch perception. Otherwise the test tones and the controls may have been classified according to some other quality than pitch, e.g., timbre [cf. Fig. 3(c) and (f)].

3. Calibrations

The frequency response of the speaker and the spectral composition of the tone pulses were checked using a B&K measuring amplifier (2610) with a $\frac{1}{4}$-in. microphone (4135), positioned at the bats starting position and connected to a spectrum analyzer (HP3590A). Between 9 and 100 kHz, the sensitivity of the speaker dropped almost linearly by about 10 dB with local deviations of less than ± 1.5 dB.

Inspection of all incomplete tone spectra [Fig. 3(c)–(f)] revealed that nonlinear distortions were at least 60 dB down (see Fig. 4).

C. Experimental animals

Megaderma lyra (Megadermatidae; Geoffroy 1810), a gleaning bat preying on insects and small vertebrates, emits broadband (20–120 kHz) sonar pulses that comprise the first five harmonics. The sound-pressure levels of these calls vary between 75 and 80 dB (determined for hand-held animals at a distance of 10 cm to the microphone), which is relatively low as compared to other bat species that achieve sound-pressure levels of up to 120 dB (Neuweiler, 1990). The relative intensities of the spectral components show a considerable variability in the course of one call and between the individuals. Relative to the frequency region with the most prominent energy, ranging from the second to fourth harmonic, the fundamental is significantly attenuated by 20–35 dB and may have sound-pressure levels between 40 and 60 dB (Schmidt, 1992). In the course of a typical call, having a duration of 0.5 ms, the frequency of the first harmonic sweeps downward from about 26 to 19 kHz. *M. lyra*

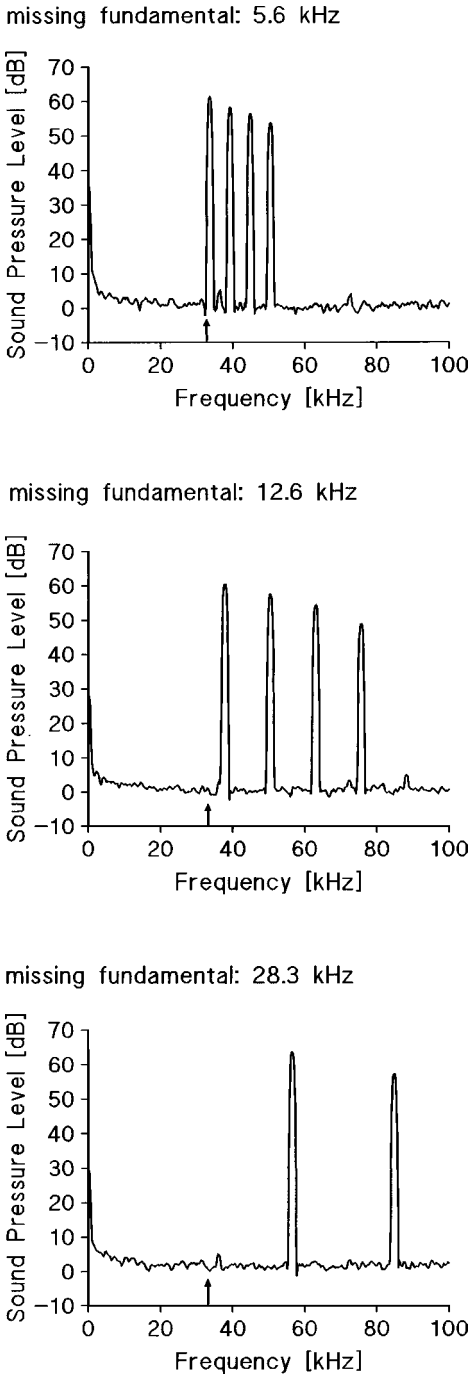


FIG. 4. Spectra of three test tones with missing fundamentals in the low-, middle-, and high-frequency region, measured at a sound-pressure level of 75 dB; S/N ratio: 60 dB (spectrum analyzer HP 3590 A). Harmonics above 100 kHz are outside the analysis range of the analyzer. The arrow at 33 kHz indicates the frequency of the pure tone reference. Please note that no nonlinear distortions are seen below the reference.

has a nearly flat audiogram with low hearing thresholds below a sound-pressure level of 0 dB in the range between 16 and 85 kHz (Schmidt *et al.*, 1983).

Three individuals (bat 1: male, 1 year old; bat 2: female, 4 years old; bat 3: female, unknown age) were used. Bats 1 and 2 were born in captivity as descendants of animals from Tirunelveli (South India). Bats 1 and 3 had already participated in psychoacoustic tests earlier.

The animals were fed mealworms exclusively during the experimental sessions. In addition, a house mouse was supplied every two weeks. The weight of the animals was monitored daily and kept at the normal individual preexperimental weight between 30 and 40 g. The bats were kept in a weakly illuminated flight room at 28 °C with free access to water.

D. Training procedure

The bats were first trained to compare the pitch of the standard stimuli to the fixed pure tone reference with a frequency of 33 kHz. They were rewarded for selecting the left or right feeding dish for standard stimuli below or above the reference tone, respectively. For every trial, a stimulus was selected randomly from a set of 24 standard tones. A correct response was rewarded by opening the chosen feeding dish, allowing the bat to take out some mealworms. Stimulus presentation stopped automatically when the animals returned to the starting position. Trials in which the bats did not select one of the feeding dishes during the ten presentations of the stimulus pair (i.e., within 24 s) were discarded. After four weeks, all three animals had learned the training tasks with a hit rate of more than 90%.

E. Testing procedure and data analysis

Since in previous studies (Sedlmeier, 1992a, b; Schmidt *et al.*, 1994) *M. lyra* has been found to have an excellent memory for absolute pitch, the reference was omitted after the training procedure without affecting the bats' discrimination performance. In the trials of each experimental session, up to 34 different stimuli were presented in random order. The standard tones were interspersed with the ambiguous test stimuli with a probability of about 25%. After the first sessions, in addition, the standard and test stimuli were interspersed with adequate control stimuli [Fig. 3(e) and (f)] in order to test whether the categorization of incomplete tones was due to perceived pitch or to any other cues.

On average, 30 trials per day were obtained from each individual. The data were collected in a period of seven months. Data points were calculated for each individual as the percentage of high categorizations for a given stimulus; each data point represents at least 30 classifications.

To evaluate the significance of the bats' decision behavior, one-dimensional χ^2 tests were performed on the obtained frequencies of left/right decisions for each stimulus. For samples of $N=30$, 75% choices to one side indicate a preference, statistically significant at $p<0.01$. For comparison of two frequency distributions, four-field χ^2 tests were calculated.

The average response latencies for different types of stimuli were compared using a t test for dependent samples. In contrast to simply recording the preferred type of reaction, latency analyses determine subtle behavioral differences that have to be attributed to underlying perceptual differentiation. When two types of stimuli are classified by means of the same predominant criterion (e.g., choice of the right feeding dish because of perceived high pitch), response latencies provide information whether these stimuli might be discerned on the basis of subdominant cues as well. In case the nature

TABLE IV. Results of bat 1 for conditions 1 and 2. The percentage of correct trials refers to their absolute number. The significance of the bat's preference for one side is evaluated by a one-dimensional χ^2 test. The average latencies are given separately for all, correct, and incorrect responses. For the ambiguous test stimuli, latencies refer to trials categorized as low or high, respectively.

Stimulus type	No. of trials	% of trials	χ^2	p	Average latency (ms)		
Standards-low		Correct			Overall	Correct	Incorrect
Condition 1	449	95.6	373	<0.001	1226	1322	3643
Condition 2	700	99.3	680	<0.001	469	446	3601
Standards-high		Correct			Overall	Correct	Incorrect
Condition 1	393	92.6	286	<0.001	4750	4980	1859
Condition 2	664	95	537	<0.001	5532	5404	3964
Test stimuli		Low			Overall	Low	High
Condition 1	259	91.5	179	<0.001	977	710	3855
Condition 2	478	91.4	328	<0.001	1565	1337	3996
Controls (high)		Low			Overall	Low	High
Condition 1	60	1.7	56	<0.001	5797	5827	4001
Condition 2	70	7.1	51	<0.001	3952	4121	1761

of such secondary characteristics is not directly evident from the experimental situation, it has to be subject to further considerations.

II. RESULTS

The classification behavior of the bats for different types of stimuli is summarized in Tables IV–VI. Reactions to each single task are shown in Figs. 5–7.

All bats discerned the standard stimuli as low or high with an error rate of less than 6% (Figs. 5–7; open squares). The test stimuli (Figs. 5–7; filled downward triangles) were classified differently by each bat. Nevertheless, in each individual the decision behavior was highly stable over time.

A. Animal 1

The reaction behavior for both intensity conditions is shown in Table IV and Fig. 5. The animal significantly categorized 22 of the 23 test stimuli as low [condition 1: $\chi^2_{\text{test}}=179$, $N=259$, $p<0.001$; condition 2: $\chi^2_{\text{test}}=328$, $N=478$, $p<0.001$], whereas it classified each of the four control stimuli as high [condition 1: $\chi^2_{\text{control}}=56$, $N=60$, $p<0.001$; condition 2: $\chi^2_{\text{control}}=51$, $N=70$, $p<0.001$]. This bat therefore based its final decisions on the pitches of the missing fundamentals.

Average response latencies (ARLs) for correctly classified high standard [Fig. 3(b)] and control stimuli [Fig. 3(e)], both having present or missing fundamentals between 34 and 37 kHz, differ significantly (condition 1: $\text{ARL}_{\text{control high correct}}$: 5827 ms, $\text{ARL}_{\text{standard high correct}}$: 4980 ms, $t=19.6$, df$=421$, $p<0.001$; condition 2: $\text{ARL}_{\text{control high correct}}$: 4121 ms, $\text{ARL}_{\text{standard high correct}}$: 5404 ms, $t=26$, df$=694$, $p<0.001$). As the overall intensities of both types of stimuli were balanced (cf. Tables I–III), this result cannot be accounted for by perceived level differences. Thus, it must be concluded that, despite the predominance of collective pitch sensations, the spectral composition, i.e., the timbre, of the tones was available to the animal as a subdominant cue, as well.

B. Animal 2

Results are summarized in Table V and displayed in Fig. 6. Bat 2 categorized all test stimuli as high (condition 1: $\chi^2_{\text{test}}=61$, $N=155$, $p<0.001$; condition 2: $\chi^2_{\text{test}}=193$, N

TABLE V. Results of bat 2 for conditions 1 and 2. For further explanations, see Table IV.

Stimulus type	No. of trials	% of trials	χ^2	p	Average latency (ms)		
Standards-low		Correct			Overall	Correct	Incorrect
Condition 1	285	99.6	281	<0.001	1707	1697	4589
Condition 2	526	99.8	522	<0.001	962	964	180
Standards-high		Correct			Overall	Correct	Incorrect
Condition 1	308	93.2	230	<0.001	2906	2860	3537
Condition 2	493	94.1	384	<0.001	2668	2626	3371
Test stimuli		Low			Overall	Low	High
Condition 1	155	18.7	61	<0.001	2892	3300	2798
Condition 2	304	10.2	193	<0.001	1786	2494	1706
Controls (low)		Low			Overall	Low	High
Condition 1	61	98.4	57	<0.001	3443	3434	4001
Condition 2	67	82.1	170	<0.001	8487	8377	8993

TABLE VI. Results of bat 3 for conditions 1 and 2. For further explanations, see Table IV.

Stimulus type	No. of trials	% of trials	χ^2	p	Average latency (ms)		
Standards-low		Correct			Overall	Correct	Incorrect
Condition 1	361	92.8	264.5	<0.001	3091	2953	4871
Condition 2	709	94.5	562	<0.001	1529	1467	2595
Standards-high		Correct			Overall	Correct	
Condition 1	311	92	219	<0.001	4313	4255	4979
Condition 2	576	94.3	543	<0.001	4743	4740	4784
Test stimuli		Low			Overall	Low	High
Condition 1	156	7.7	112	<0.001	4626	5634	4543
Condition 2	312	6.4	237	<0.001	5065	3189.5	5193
Controls (low)		Low			Overall	Low	
Condition 1	74	51.4	0.05	n.s.	4454	4841	4045
Condition 2	72	0	72	<0.001	9264	···	9264

$=304$, $p<0.001$). The animal met the control criterion by correctly classifying all four unequivocal incomplete complex tones [Fig. 3(f)] as having a low pitch (condition 1: $\chi^2_{\text{control}}=57$, $N=61$, $p<0.001$; condition 2: $\chi^2_{\text{control}}=170$, $N=67$, $p<0.001$). This bat thus focused its attention on the pure tone pitches of the present harmonics.

When presented with the ambiguous test stimuli [Fig. 3(c)], the animal decided for the low side to a somewhat higher extent than when presented with the unequivocal high standards [Fig. 3(b)]: condition 1: test tones: 18.7%, high standard tones: 6.8%, $\chi^2=15.1$, $N=463$, $p<0.001$; condition 2: test tones: 10.2%, high standard tones: 5.9%, $\chi^2=5$, $N=797$, $p=0.02$. This result may indicate that the bat based some of its decisions on the collective pitch of the test stimuli. Alternatively, it may reflect an uncertainty of the animal due to the unfamiliar timbre of the test tones.

A comparison of the average response latencies for correctly classified high standard stimuli [Fig. 3(b)] with those for test tones [Fig. 3(c)] categorized as high reveals that the animal left its starting position more quickly when presented with the latter type of stimulus (condition 1: $\text{ARL}_{\text{standard high correct}}$: 2860 ms, $\text{ARL}_{\text{test high}}$: 2798 ms, $t=16.8$, df$=411$, $p<0.001$; condition 2: $\text{ARL}_{\text{standard high correct}}$; 2626 ms, $\text{ARL}_{\text{test high}}$: 1706 ms, $t=23.9$, df$=735$, $p<0.001$). This result reveals that although both types of tones were predominantly classified according to their pure tone pitches, they were discernible on the basis of subdominant cues as well. Again, it cannot be decided from the present data whether collective pitch or timbre sensations are responsible for the observed differences. Moreover, the response latencies for correctly classified low standard [Fig. 3(a)] and low control stimuli [Fig. 3(f)] differ significantly (condition 1: $\text{ARL}_{\text{standard low correct}}$: 1697 ms, $\text{ARL}_{\text{control low correct}}$: 3434 ms; $t=11.6$, df$=342$, $p<0.001$; condition 2: $\text{ARL}_{\text{standard low correct}}$: 962 ms, $\text{ARL}_{\text{control low correct}}$: 8487 ms; $t=11.12$, df$=578$, $p<0.001$). This result can only be explained on the basis of perceived differences in timbre.

C. Animal 3

Results are summarized in Table VI and displayed in Fig. 7. Bat 3 classified all test stimuli as high (condition 1: $\chi^2_{\text{test}}=112$, $N=156$, $p<0.001$; condition 2: $\chi^2_{\text{test}}=237$,

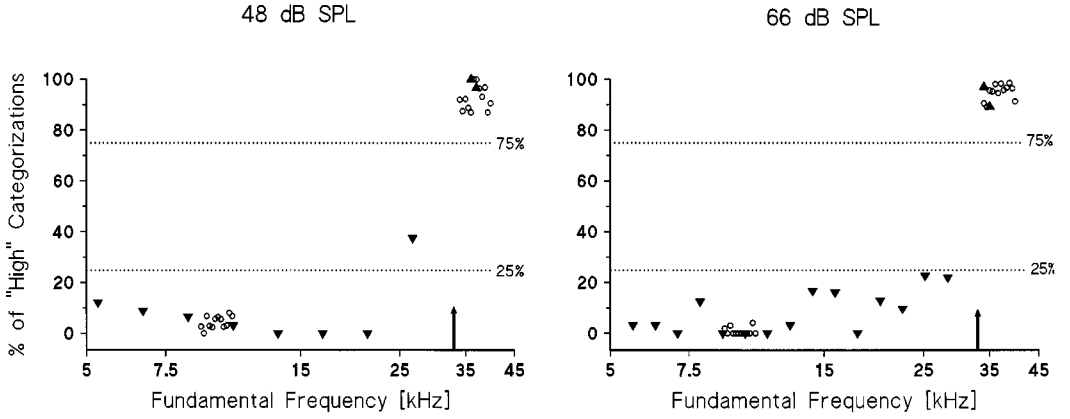


FIG. 5. Pitch classifications by bat 1 for the low-level stimuli (sound-pressure level: ~48 dB; condition 1) and the higher-level stimuli (sound-pressure level: ~66 dB; condition 2). The abscissa represents the frequency of the present or missing fundamentals of the complex tones. The arrow at 33 kHz marks the reference. On the ordinate, the percentage of choices to the side associated with a high pitch is plotted. For all stimuli, at least 30 trials were obtained per data point, corresponding to a significance level of $p<0.01$ for 75% of the choices to one side (dotted lines) for a one-dimensional χ^2-test. Small open circles (○) represent the classification behavior for the standard stimuli, filled downward triangles (▼) for the ambiguous test tones, and filled upward triangles (▲) for the control tasks.

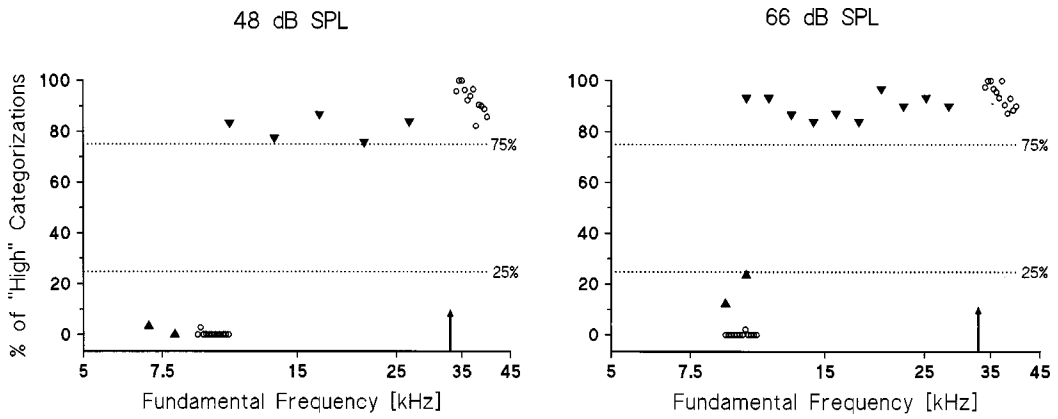


FIG. 6. Pitch classifications by bat 2 for conditions 1 (sound-pressure level: ~48 dB) and 2 (sound-pressure level: ~66 dB). For further details, see Fig. 5.

$N=312$, $p<0.001$). In contrast to bat 2, three of the four control tasks [Fig. 3(f)] were categorized in the same way as the test stimuli, i.e., the bat preferred the feeding dish associated with a high pitch (condition 1: $\chi^2_{\text{control}}=0.05$, $N=74$, n.s.; condition 2: $\chi^2_{\text{control}}=72$, $N=72$, $p<0.001$). This result is incompatible with the assumption that this bat used either collective or pure tone pitches as the decisive cue. We hypothesize that bat 3 preferably paid attention to the similar timbre of all incomplete tones and arbitrarily associated this characteristic with the right feeding dish.

III. DISCUSSION

The present results show that *M. lyra* is not restricted to a single criterion when classifying complex tone stimuli, but rather is able to extract different features from a given spectrum. The predominant classification criterion was found to vary between individuals and to persist over a long period of time in each individual.

This finding is in good agreement with results obtained from human listeners. Although a human subject is normally able to concentrate on each of the three qualities ''pure tone pitch,'' ''collective pitch,'' and ''timbre,'' attention is focused on maximally two of these aspects simultaneously, namely on timbre and on spectral or collective pitch (Smoorenburg, 1970; Houtsma and Goldstein, 1972). For a given ambiguous stimulus, the sensational predominance of one aspect of pitch over the other has been found to be a quite stable individual characteristic (Albrecht, 1972; Deutsch and Steeger, 1978). However, a complex tone's spectral composition (Preisler, unpublished data), its context (Houtgast, 1976), and the actual S/N ratio (Houtgast, 1976) highly determine the probability with which one perceptual mode will predominate the other.

In all three bats, the average response latencies differed significantly for tones including their fundamentals from tones with their fundamentals missing. Hence, it can be concluded that each animal was able to discern the stimuli on the basis of their spectral composition, i.e., their timbre. However, only in bat 3 choices in favor of one feeding dish were determined by that criterion. In contrast, bats 1 and 2 based their decisions solely on pitch criteria, namely on perceived collective and pure tone pitch, respectively. In these animals, the spectral composition of the tones did not directly affect the classification behavior and thus seems to have been of lesser importance.

Subsequently, the following questions will be dealt with:

(a) Can the behavior of bat 1 be explained by any effects other than collective pitch formation, such as a detection of difference tones or focused attention on specific spectral characteristics of the stimulus components?

(b) What are the implications of collective pitch formation at high and ultrasonic frequencies for current models of pitch perception?

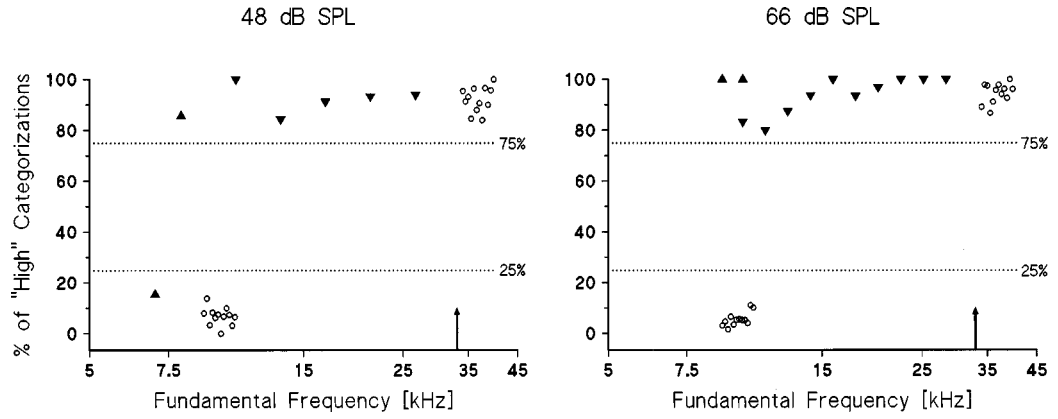


FIG. 7. Pitch classifications by bat 3 for conditions 1 (sound-pressure level: ~48 dB) and 2 (sound-pressure level: ~66 dB). For further details, see Fig. 5.

(c) What is the significance of the observed differences in classification behavior for *M. lyra*?

A. Alternative hypotheses to explain the behavior of bat 1

Collective pitch sensations at ultrasonic frequencies are an interesting phenomenon for two reasons. On the one hand, they bear on theories of how the mammalian auditory system extracts pitch from complex tones. On the other hand, they provide insight of how bats might cope with the problem of the rapidly fluctuating echo spectra in their natural environment.

Therefore, alternative hypotheses to explain why bat 1 consistently classified test tones with missing fundamentals between 5.3 and 28.3 kHz as low have to be tested conscientiously before drawing any further conclusions from this result. In the following, we discuss whether nonlinear distortions originating from the experimental apparatus or the cochlea, or specific characteristics of the incomplete complex tones, may have provided alternative decision cues to this bat.

With our equipment, difference tones from the apparatus were not measurable. As the noise floor in our experiments was 60 dB down (cf. Fig. 4), the sound-pressure levels of possible distortions were at least below -13 dB in condition 1, i.e., below the behavioral auditory threshold of *M. lyra* (Schmidt *et al.*, 1983).

Cubic difference tones (CDTs) are the most prominent cochlear distortions, also arising at low sound-pressure levels (Zwicker, 1990; Kössl, 1992). Subsequently, only the first lower sideband CDT at a frequency of $2f_1-f_2$ will be dealt with. We estimated at what intensities CDTs were to be expected below the 33-kHz reference. For test tones with missing fundamentals between 16.5 and 33 kHz (cf. Table II, tasks 30–32; Table III, tasks 47–51), the frequency ratio of the primary tones f_2/f_1 (3rd/2nd harmonic) is 1.5 and the frequency of a possible CDT matches that of the missing fundamental. The CDTs arising from higher adjacent harmonics (e.g., 4rd/3th, 5th/4th harmonic) can be neglected, since they coincide with the low harmonics of the stimulus. For test tones with a collective pitch below 16.5 kHz (cf. Table II, tasks 25–29; Table III, tasks 37–46), as a consequence of the omission of more than one harmonic, ratios of adjacent primaries are smaller than 1.5. The CDTs arising from the two lowest present harmonics have frequencies below the reference, but do not coincide with the missing fundamentals.

Kössl (1992) measured the intensities of low side CDTs in *M. lyra* as a function of the sound-pressure levels and frequency ratios of the primary tones. As the levels of the CDTs were found to systematically drop for primary ratios exceeding 1.2, larger ratios, like the one of 1.5 which is relevant for the present stimuli, were not especially considered in this study.

In case the lower primary tone had a sound-pressure level between 40 and 50 dB, as relevant for our experiments in condition 1 (average sound-pressure level: 45.4 dB), a CDT having a sound-pressure level between -20 and 0 dB was obtained at the largest tested frequency ratios.

In a later study (Kössl, personal communication), CDTs were measured for different primary tone levels at a constant frequency ratio of 1.5. According to these data, a CDT having a sound-pressure level of less than -13 dB is predicted for a lower primary at a sound-pressure level of 45 dB. For test tones with missing fundamentals above 16.5 kHz, the distortion product thus remains well below the behavioral auditory threshold of *M. lyra* (Schmidt *et al.*, 1983) at least in condition 1.

For test tones with a lower collective pitch, the distances between the remaining adjacent primaries get increasingly smaller (frequency ratios: 1.33–1.17) and higher CDT levels are to be expected (Kössl, 1992). However, CDTs with frequencies below the reference but above the missing fundamentals still remain below the auditory threshold in condition 1.

Moreover, paying attention to local cues, like the prominence of single harmonics or spectral regions, could not have provided useful cues for consistent classifications since 23 different test stimuli of varying overall intensity (± 2 dB) and frequency composition (f_0: 5.3–28.3 kHz) were presented in random order.

B. Implications for current models of the perception of the pitch of complex tones

Since Seebeck's (1843) discovery of the phenomenon of the missing fundamental several attempts were made to explain this ''auditory illusion.'' Different theories attribute collective pitch sensations to an analysis of the acoustical signal either within the time domain, the frequency domain, or both (for review, see De Boer, 1976; Preisler, 1993).

In an early model, Schouten (1938, 1940a, b) regarded collective pitch as an inevitable by-product of the restricted frequency resolution of the basilar membrane. When a cochlear filter is activated simultaneously by more than one harmonic, the period of the superimposed signal is the reciprocal of the highest common divisor of the included frequencies. Hence, in the harmonic case the period of the envelope matches the reciprocal of the frequency of the present or missing fundamental. The model later received physiological support from the observation that neural discharge patterns tend to synchronize to the frequency of single components (Javel, 1986) as well as to the compound signal envelope of spectrally unresolved components (Rose, 1970). To account for the present results by Schouten's theory, one would have to claim phase coupling to modulation frequencies between 5.3 and 28.3 kHz.

In a more recent model, Langner (1981, 1983, 1984, 1988) postulates that information about the missing fundamental is derived from a cross-correlation of the neural spike patterns, driven by the carrier and the modulation frequency of the spectrally unresolved signal. Therefore, to elicit collective pitch sensations at 28 kHz, a carrier at 84 kHz (center frequency of a three-component stimulus, consisting of harmonics 2–4) needs to be coded in the time domain.

Other theories (Moore, 1982; Lazzaro and Mead, 1989; Meddis and Hewitt, 1991; Delgutte and Cariani, 1992) claim that for each cochlear filter channel first an autocorrelation of interspike intervals is performed. At a higher stage of neural

processing, a summation autocorrelogram, including the outputs of all channels, is supposed to account for the final pitch sensation. Again, such models require temporal coding of either the compound signal envelope or of at least the two lowest spectrally resolved components.

In mammals, phase locking has not been reported for frequencies above 5 kHz (Hind *et al.*, 1967; Javel, 1986; Joris and Yin, 1992). This upper frequency limit (3–4 kHz) was confirmed by Fourier analyses of the period histograms, as obtained from the discharge patterns of auditory nerve fibers in cats (Horst *et al.*, 1986), and may thus reflect an intrinsic property of the ascending auditory pathway. It has been suggested that synchronization breaks down at higher frequencies because the ac responses of the hair cells become small in comparison with the dc responses (Russell and Sellick, 1978), the temporal jitter in the initiation of nerve impulses limits synchronous firing (Fay, 1984), and the synapses between hair cells and auditory nerve fibers have low-pass transfer characteristics (Weiss and Rose, 1988).

In bats, neural synchronization to steady-state signals has been found for frequencies up to 3 kHz at the level of the cochlear nucleus (Suga *et al.*, 1975; Vater, 1982), and only up to 350 Hz in the inferior colliculus of the auditory midbrain (Schuller, 1979). Hence, in this respect the bats' temporal resolution is comparable to that of other mammals and does not show specific adaptations for the use of ultrasound.

Thus, neither the original theory of Schouten (1938, 1940a, b), nor later models, regarding phase locking as an essential supposition for collective pitch formation (Langner, 1981, 1983, 1984, 1988; Moore, 1982; Lazzaro and Mead, 1989; Meddis and Hewitt, 1991; Delgutte and Cariani, 1992), are able to account for the present result.

On the other hand, collective pitch sensations at high and ultrasonic frequencies could well be explained on the basis of spectral pattern recognition models (Goldstein, 1973; Wightman, 1973; Terhardt, 1974, 1976; Cohen *et al.*, 1995). The only requirement that has to be met is that, on the basis of a place mechanism, the included partials are sufficiently resolved by the auditory system.

Terhardt (1974, 1976) conjectured that collective pitch formation is a by-product of speech acquisition. The model, which has been critically reviewed and not generally shared by the experts within the last two decades, starts from the assumption that the harmonic structure of voiced sounds of speech (Stoll, 1980) is learned unconsciously in an early period of life as an acoustical reference pattern. Later, stimulation by a single pure tone should be sufficient to produce pitch cues corresponding to its subharmonics. Accordingly, a collective pitch sensation should correspond to the frequency at which most of the subharmonics of a complex tone's spectrally resolved components coincide. However, it is misleading to suppose that pure tones, *per se*, evoke imaginations of their possible subharmonics. Rather, stimulation of a collective pitch by a single pure tone is possible only if the missing information can be inferred from context and can be considered masked by distracting sound (Houtgast, 1976).

If ''speech'' is replaced by the more general term ''acoustical communication,'' Terhardt's concept may be applied to all species whose calls comprise harmonic portions, whose auditory system guarantees spectral resolution, and who possess sufficient learning capabilities. Many birds and mammals meet these requirements.

The widths (Sedlmeier and Schmidt, 1989) and shapes (Schmidt *et al.*, 1994) of auditory filters in *M. lyra* allow the spectral resolution of most of the applied stimuli's single components. A detailed examination of the spectra reveals that this holds for each of the 23 test stimuli.

In view of the available physiological data it is unlikely that collective pitch sensations at high and ultrasonic frequencies originate from an analysis of neural interspike-interval patterns. On the other hand, several psychoacoustical and physiological findings in humans reveal the temporal encoding of pitch information, namely the pure tone pitch sensations evoked by repetitive electrical stimulation of the auditory nerve (Eddington, 1980; Townshend *et al.*, 1987; McKay *et al.*, 1994; Pijl and Schwartz, 1995) and the faint collective pitch sensations elicited by dichotically delayed wideband noise (Bilsen and Goldstein, 1974), sinusoidally modulated or pulsed white noise (Miller and Taylor, 1948; Small, 1955; Harris, 1963; Rosenberg, 1965; Burns and Viemeister, 1976, 1981), and spectrally unresolved frequency components (Houtsma and Smurzynski, 1990). De Boer (1956) suggested that collective pitch sensations are due to two different mechanisms, a phase-sensitive mechanism for harmonics that are too close to be aurally resolved and a phase-insensitive mechanism for widely spaced components. Subsequent studies revealed that perceived collective pitch is predominantly determined by the low resolved components (Ritsma, 1967; Plomp, 1967; Preisler, 1993; Clarkson and Rogers, 1995). Shackleton and Carlyon (1994) and Carlyon and Shackleton (1994) report that the pitch of stimuli consisting of unresolved harmonics increases by an octave when the harmonics are presented in alternating sine–cosine phase, but not in sine phase. This is not the case for complex tones only comprising spectrally resolved components. Moreover, the authors report that when listeners were required to compare the collective pitches of two simultaneously presented complexes, performance was better when the two complexes consisted of the same ''type'' than when one complex consisted of resolved harmonics and one consisted of unresolved harmonics. Based on these results, Carlyon and Shackleton (1994) favor the idea that collective pitch sensations, in contrast to the basic assumptions of autocorrelation models (Moore, 1982; Lazzaro and Mead, 1989; Meddis and Hewitt, 1991; Delgutte and Cariani, 1992), are elicited by two different mechanisms. Meanwhile, this view has been further corroborated by Plack and Carlyon (1995).

In principle, the following alternative conclusions may be drawn from studies supporting a dual hypothesis of collective pitch formation:

(1) There are two distinct mechanisms, both operating on periodicity detection.
(2) For high, unresolved harmonics the periodicity of the composite waveform is coded by neural phase locking, whereas low, resolved harmonics give rise to collective pitch impressions on the basis of spectral pattern recognition, mediated by

(a) a pure place code (Wightman, 1973; Terhardt, 1974, 1976) and
(b) a combination of place and time code (Goldstein, 1973; Srulovicz and Goldstein, 1983; Cohen *et al.*, 1995).

Since harmonics, far extending over the ultrasonic region, are spectrally but not temporally resolvable by the peripheral auditory system of *M. lyra*, our results are only compatible with interpretation (2)(a). We thus infer that spectral pattern recognition, based on a pure place code, is in principle sufficient to elicit collective pitch sensations in mammals. However, in view of the highly reliable results of cochlear implant research (Eddington, 1980; Townshend *et al.*, 1987; McKay *et al.*, 1994; Pijl and Schwartz, 1995; Nelson *et al.*, 1995), interpretation (2)(b) is likely to be valid for species hearing in the normal frequency range.

C. The significance of different modes of complex tone classification for *M. lyra*

According to a stimulus-related definition, the timbre of a steady-state signal is determined by the relative intensities and ratios of the perceptible partials. The absolute pitches of the spectral components play a negligible role, since they affect the sensation only in case the entire spectrum is shifted over a frequency range wide enough to allow an influence of varying auditory thresholds. Collective pitch is an abstract category complementary to timbre, since it depends on the employed spectral frequencies, but does not necessarily vanish with the cancellation of single partials. Thus, collective pitch assignment makes melodic information highly robust in the presence of spectral variations due to selective filtering and masking.

In *M. lyra*, paying attention to single components analytically may increase the animals' ability to detect spectral irregularities within the echoes which provide cues about the position and texture of reflecting objects (Schmidt, 1992; Neuweiler, 1993).

However, the bats may derive such information in a holistic way from variations of timbre, as well. Moreover, like many other higher vertebrates, bats are likely to exploit characteristics of timbre for the recognition and classification of different environmental sounds.

The results of comparative studies in different mammalian species (Heffner and Whitfield, 1976; Tomlinson and Schwartz, 1988) and birds (Cynx, 1986) provide evidence that collective pitch formation is widespread and thus seems to be an essential characteristic of vertebrate auditory systems.

In bats, the horizontal angle between the alignment of the outer ear and the position of the reflecting object determines frequency specific interference patterns within the ear, which are the basis for further neural encoding (Neuweiler, 1993). In species emitting broadband signals, the gain/attenuation of specific frequencies depends on the elevation (vertical angle) between the object and the ear, as well. Hence, the spectral composition of an echo is highly sensitive to spatial changes of the bat, the target, or both. Moreover, echoes originating from structured objects comprise spectral notches and gains which are directly related to the grain size of the texture (Schmidt, 1992). Thus, collective pitch, which is unaffected by single spectral notches and gains, may be employed as an invariant reference within the rapidly fluctuating spectral content of the echoes.

In view of the facts that in *M. lyra* the sound pressure level of the echolocation calls' fundamental frequency may vary between 40 and 60 dB (Schmidt, 1992) and that the hearing thresholds are below a sound-pressure level of 0 dB in the relevant frequency region (Schmidt *et al.*, 1983), these bats should be easily able to perceive the fundamentals of their emitted calls. In case a bat emits a sonar signal at a range of 30 kHz, which is slightly above the fundamental frequency of the echolocation calls of *M. lyra*, towards a small, spherical target at a distance of 3 m, the echo is attenuated by about 65 dB because of geometrical spreading loss and atmospheric attenuation. If the reflecting object is a large, planar surface, the 30-kHz component may be maximally perceived up to a target distance of 12 m due to an attenuation of 65 dB (Lawrence and Simmons, 1980). To account for the energy scattered elsewhere or absorbed by the target's shape during echo formation, further attenuation must be added. Moreover, masking by environmental noise may deteriorate the S/N ratio and so further reduce the distance of perceptibility. Thus, it has to be inferred that the first harmonic is only occasionally perceptible within the reflected echo spectra. It might hence be advantageous for *M. lyra* to replace the faint or missing fundamentals of its echoes by corresponding collective pitch sensations.

In human listeners it has been shown that low signal-to-noise ratios induce a shift of attention from the analytical to the more reliable comprehensive aspect of pitch (Houtgast, 1976; Hall and Peters, 1981). We hypothesize that bats benefit from collective pitch sensations, when listening to their own echoes or to other complex signals in noisy surroundings, in a comparable way.

Since the extensive exploitation of different sound characteristics bears numerous advantages, it has to be assumed that during evolution of the vertebrate auditory system there was a selective pressure to utilize all available cues for signal identification and interpretation. From this point of view, the different classification criteria, as employed by the tested bats, are an interesting but not unexpected finding.

ACKNOWLEDGMENTS

This study was supported by Humboldt-Foundation and SFB 204 ''Gehör,'' Munich. The authors would like to thank Professor G. Neuweiler for his support and D. I. S. Kiesslich for providing the software.

Albrecht, E. M. (**1972**). ''Das akustische Residuum,'' Ph.D. dissertation, University of Vienna, Austria.
ANSI (**1994**). ANSI 1.1-1994, ''American National Standard Acoustical Terminology'' (Acoustical Society of America, New York).
Bilsen, F. A., and Goldstein, J. L. (**1974**). ''Pitch of dichotically delayed noise and its possible central basis,'' J. Acoust. Soc. Am. **55**, 292–296.
Burns, E. M., and Viemeister, N. F. (**1976**). ''Nonspectral pitch,'' J. Acoust. Soc. Am. **60**, 863–869.
Burns, E. M., and Viemeister, N. F. (**1981**). ''Played-again SAM: Further observations on the pitch of amplitude-modulated noise,'' J. Acoust. Soc. Am. **70**, 1655–1660.

Carlyon, R. P., and Shackleton, T. M. (**1994**). ''Comparing the fundamental frequencies of resolved and unresolved harmonics: Evidence for two pitch mechanisms?,'' J. Acoust. Soc. Am. **95**, 3541–3554.
Clarkson, M. G., and Rogers, E. C. (**1995**). ''Infants require low-frequency energy to hear the pitch of the missing fundamental,'' J. Acoust. Soc. Am. **98**, 148–154.
Cohen, M. A., Grossberg, S., and Wyse, L. L. (**1995**). ''A spectral network model of pitch perception,'' J. Acoust. Soc. Am. **98**, 862–879.
Cynx, J., and Shapiro, M. (**1986**). ''Perception of the missing fundamental by a species of songbird (*Sturnus vulgaris*),'' J. Comp. Psychol. **100**, 356–360.
D'Amato, M. R. (**1988**). ''A search for tonal pattern perception in cebus monkeys: Why monkeys can't hum a tune,'' Music Percept. **4**, 453–480.
De Boer, E. (**1956**). ''On the residue in Hearing,'' Ph.D. dissertation, University of Amsterdam, Amsterdam, The Netherlands.
De Boer, E. (**1976**). ''On the 'Residue' and Auditory Pitch Perception,'' in *Auditory System: Clinical and Special Topics*, edited by W. E. Keidel and W. D. Neff (Springer-Verlag, Berlin), pp. 479–583.
Delgutte, B., and Cariani, P. (**1992**). ''Coding of the pitch of harmonic and inharmonic complex tones in the interspike intervals of auditory nerve fibers,'' in *The Processing of Speech*, edited by M. E. H. Schouten (Mouton-DeGruyer, Berlin), pp. 37–45.
Deutsch, W. A., and Steeger, R. (**1978**). ''Binaurales Residuum versus Lokalisation,'' Mitteilungen der Kommission für Schallforschung (Österreichische Akademie der Wissenschaften) **5**, 41–53.
Eddington, D. K. (**1980**). ''Speech discrimination in deaf subjects with cochlear implants,'' J. Acoust. Soc. Am. **68**, 885–891.
Fay, R. (**1984**). ''Temporal processing by the auditory system of fishes,'' in *Time Resolution in Auditory Systems*, edited by A. Michelsen (Springer-Verlag, Berlin), pp. 28–57.
Goldstein, J. L. (**1973**). ''An optimum processor theory for the central formation of the pitch of complex tones,'' J. Acoust. Soc. Am. **54**, 1496–1516.
Hall, J. W., and Peters, R. W. (**1981**). ''Pitch for nonsimultaneous successive harmonics in quiet and noise,'' J. Acoust. Soc. Am. **69**, 509–513.
Harris, G. G. (**1963**). ''Periodicity perception by using gated noise,'' J. Acoust. Soc. Am. **35**, 1229–1233.
Hartmann, W. M. (**1988**). ''Pitch perception and the segregation and integration of auditory entities,'' in *Auditory Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 623–645.
Heffner, H., and Whitfield, I. C. (**1976**). ''Perception of the missing fundamental by cats,'' J. Acoust. Soc. Am. **59**, 915–919.
Hind, J. E., Anderson, D. J., Brugge, J. F., and Rose, J. E. (**1967**). ''Coding of information pertaining to paired low-frequency tones in single auditory nerve fibers of the squirrel monkey,'' J. Neurophysiol. **30**, 794–816.
Horst, J. W., Javel, E., and Farley, G. R. (**1986**). ''Coding of spectral fine structure in the auditory nerve. I. Fourier analysis of period and interspike interval histograms,'' J. Acoust. Soc. Am. **79**, 398–416.
Houtgast, T. (**1976**). ''Subharmonic pitches of a pure tone at low S/N ratio,'' J. Acoust. Soc. Am. **60**, 405–409.
Houtsma, A. J. M., and Goldstein, J. L. (**1972**). ''The central origin of the pitch of complex tones: Evidence from musical interval recognition,'' J. Acoust. Soc. Am. **51**, 520–529.
Houtsma, A. J. M., and Smurzynski, J. (**1990**). ''Pitch identification and discrimination for complex tones with many harmonics,'' J. Acoust. Soc. Am. **87**, 304–310.
Hulse, S. H., and Cynx, J. (**1985**). ''Relative pitch perception is constrained by absolute pitch in songbirds (*Mimus*, *Molothrus*, and *Sturnus*),'' J. Comp. Psychol. **99**, 176–196.
Javel, E. (**1986**). ''Basic response properties of auditory nerve fibers,'' in *Neurobiology of Hearing*, edited by R. A. Altschuler, D. W. Hoffmann, and R. P. Bobbin (Raven, New York), pp. 213–245.
Joris, P. X., and Yin, T. C. T. (**1992**). ''Responses to amplitude-modulated tones in the auditory nerve of the cat,'' J. Acoust. Soc. Am. **91**, 215–232.
Kössl, M. (**1992**). ''High frequency distortion products from the ear of two bat species, *Megaderma lyra* and *Carollia perspicillata*,'' Hearing Res. **60**, 156–164.
Langner, G. (**1981**). ''Neuronal mechanisms for pitch analysis in the time domain,'' Exp. Brain Res. **44**, 450–454.
Langner, G. (**1983**). ''Evidence for neural periodicity detection in the auditory system of the guinea fowl: Implications for pich analysis in the time domain,'' Exp. Brain Res. **52**, 333–355.
Langner, G. (**1984**). ''Time coding and periodicity pitch,'' in *Time Resolution in Auditory Systems*, edited by A. Michelsen (Springer-Verlag, Berlin), pp. 108–121.
Langner, G. (**1988**). ''Physiological properties of units in the cochlear nucleus are adequate for a model of periodicity analysis in the auditory midbrain,'' in *Auditory Pathway*, published by J. Syka and R. B. Masterton (Plenum, New York).
Lawrence, B. D., and Simmons, J. A. (**1980**). ''Measurements of atmospheric attenuation at ultrasonic frequencies and the significance for echolocation by bats,'' J. Acoust. Soc. Am. **71**, 585–590.
Lazzaro, J., and Mead, C. (**1989**). ''Silicon modeling of pitch perception,'' Neurobiology **86**, 9597–9601.
Lichte, W. H. (**1941**). ''Attributes of complex tones,'' J. Exp. Psychol. **28**, 455–480.
Licklider, J. C. R. (**1954**). ''Periodicity pitch and place pitch,'' J. Acoust. Soc. Am. **26**, 945(A).
McKay, C. M., McDermott, H. J., and Clark, G. M. (**1994**). ''Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees,'' J. Acoust. Soc. Am. **96**, 2664–2673.
Meddis, R., and Hewitt, J. (**1991**). ''Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I. Pitch identification,'' J. Acoust. Soc. Am. **89**, 2866–2882.
Miller, G. A., and Taylor, W. G. (**1948**). ''The perception of repeated bursts of noise,'' J. Acoust. Soc. Am. **20**, 171–180.
Mogdans, J., Schnitzler, H. U., and Ostwald, J. (**1993**). ''Discrimination of two-wavefront echoes by the big brown bat, *Eptesicus fuscus*: Behavioral Experiments and Receiver Simulations,'' J. Comp. Physiol. A **172**, 309–323.
Moore, B. C. J. (**1982**). *An Introduction to the Physiology of Hearing* (Academic, London).
Nelson, D. A., Van Tasell, D. J., Schroder, A. C., and Soli, S. (**1995**). ''Electrode ranking of 'place pitch' and speech recognition in electrical hearing,'' J. Acoust. Soc. Am. **98**, 1987–1999.
Neuweiler, G. (**1990**). ''Auditory adaptations for prey capture in echolocating bats,'' Physiol. Rev. **70(3)**, 615–641.
Neuweiler, G. (**1993**). *Biologie der Fledermäuse* (Thieme, Stuttgart).
Patterson, R. D. (**1973**). ''The effects of relative phase and the number of components on residue pitch,'' J. Acoust. Soc. Am. **53**, 1565–1572.
Pijl, S., and Schwarz, D. W. F. (**1995**). ''Melody recognition and musical interval perception by deaf subjects stimulated with electrical pulse trains through single cochlear implant electrodes,'' J. Acoust. Soc. Am. **98**, 886–895.
Plack, C. J., and Carlyon, R. P. (**1995**). ''Differences in frequency modulation detection and fundamental frequency discrimination between complex tones consisting of resolved and unresolved harmonics,'' J. Acoust. Soc. Am. **98**, 1355–1364.
Plomp, R. (**1967**). ''Pitch of complex tones,'' J. Acoust. Soc. Am. **41**, 1526–1533.
Preisler, A. (**1993**). ''The influence of spectral composition of complex tones and of musical experience on the perceptibility of virtual pitch,'' Percept. Psychophys. **54**, 589–603.
Preisler, A., and Schmidt, S. (**1995**). ''Relative versus absolute pitch perception in the bat, *Megaderma lyra*,'' in *Proceedings of the 23rd Göttingen Neurobiology Conference 1995*, edited by N. Elsner and R. Menzel (Georg Thieme Verlag, Stuttgart), p. 309.
Ritsma, R. J. (**1967**). ''Frequencies dominant in the perception of the pitch of complex sounds,'' J. Acoust. Soc. Am. **42**, 191.
Rose, J. E. (**1970**). ''Discharges of single fibers in the mammalian auditory nerve,'' in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Leiden), pp. 176–192.
Rosenberg, A. E. (**1965**). ''Effects of masking on the pitch of periodic pulses,'' J. Acoust. Soc. Am. **38**, 747–758.
Russell, I. J., and Sellick, P. M. (**1978**). ''Intracellular studies of hair cells in the mammalian cochlea,'' J. Physiol. (London) **284**, 261–290.
Schmidt, S. (**1992**). ''Perception of structured phantom targets in the echolocating bat, *Megaderma lyra*,'' J. Acoust. Soc. Am. **91**, 2203–2223.
Schmidt, S., Preisler, A., Sedlmeier, H. (**1994**). ''Aspects of pitch perception in the ultrasonic range,'' in *Advances in Hearing Research*, edited by G. A. Manley, G. Klump, C. Köppl, H. Fastl, and H. Oeckinghaus (World Scientific, Singapore), pp. 374–382.
Schmidt, S., Türke, B., and Vogler, B. (**1983**). ''Behavioural audiogram from the bat, *Megaderma lyra*,'' Myotis **21/22**, 62–66.
Schnitzler, H. U., and Henson, O. W. (**1980**). ''Performance of airborne animal sonar systems. I. Microchiroptera,'' in *Animal Sonar Systems*, edited by R. G. Bunsel and J. F. Fish (Plenum, New York), pp. 109–181.

Schouten, J. F. (**1938**). ''The perception of subjective tones,'' Proc. K. Ned. Akad. Wet. **41**, 1086–1094.
Schouten, J. F. (**1940a**). ''The residue, a new component in subjective sound analysis,'' Proc. K. Ned. Akad. Wet. **43**, 356–365.
Schouten, J. F. (**1940b**). ''The residue and the mechanism of hearing,'' Proc. K. Ned. Akad. Wet. **43**, 991–999.
Schuller, G. (**1979**). ''Coding of small sinusoidal frequency and amplitude modulations in the inferior colliculus of the 'CF-FM' bat *Rhinolophus ferrumequinum*,'' Exp. Brain Res. **34**, 117–132.
Sedlmeier, H. (**1992a**). ''Tonhöhenwahrnehmung beim falschen Vampir Megaderma lyra,'' unpublished doctoral dissertation, Dept. Zool., University of Munich, Germany.
Sedlmeier, H. (**1992b**). ''Pitch sensation in the Indian false vampire bat (*Megaderma lyra*),'' in *Proceedings of the 20th Göttingen Neurobiology Conference*, edited by N. Elsner and W. Richter (Thieme Verlag, Stuttgart), p. 233.
Sedlmeier, H., and Schmidt, S. (**1989**). ''Masked auditory thresholds from the Indian false vampire bat (*Megaderma lyra*),'' in *Proceedings of the 17th Göttingen Neurobiology Conference*, edited by N. Elsner and W. Singer (Georg Thieme Verlag, Stuttgart), p. 292.
Seebeck, A. (**1843**). ''Über die Sirene,'' Ann. Phys. Chem. **60**, 449–481.
Shackleton, T. M., and Carlyon, R. P. (**1994**). ''The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination,'' J. Acoust. Soc. Am. **95**, 3529–3540.
Small, A. M. (**1955**). ''Some parameters influencing the pitch of amplitude modulated signals,'' J. Acoust. Soc. Am. **27**, 751–760.
Smoorenburg, G. F. (**1970**). ''Pitch perception of two-frequency stimuli,'' J. Acoust. Soc. Am. **4**, 924–942.
Srulovicz, P., and Goldstein, J. L. (**1983**). ''A central spectrum model: A synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum,'' J. Acoust. Soc. Am. **73**, 1266–1276.
Stoll, G. (**1980**). ''Psychoakustische Messungen der Spektraltonhöhenmuster von Vokalen,'' in *Proceedings of the DAGA* (VDE Verlag, Berlin), pp. 631–638.
Suga, N., Simmons, J. A., and Jen, P. H.-S. (**1975**). ''Peripheral specialization for fine analysis of doppler-shifted echoes in the auditory system of the 'CF-FM' bat *Pteronotus parnelii*,'' J. Exp. Biol. **63**, 161–192.
Terhardt, E. (**1974**). ''Pitch, consonance, and harmony,'' J. Acoust. Soc. Am. **55**, 1061–1069.
Terhardt, E. (**1976**). ''Ein psychoakustisch begründets Konzept der musikalischen Konsonanz,'' Acustica **36**, 121–137.
Terhardt, E., Stoll, G., and Seewann, M. (**1982**). ''Pitch of complex signals according to virtual pitch theory: Tests, examples and predictions,'' J. Acoust. Soc. Am. **71**, 671–678.
Tomlinson, R. W. W., and Schwarz, D. W. F. (**1988**). ''Perception of the missing fundamental in nonhuman primates,'' J. Acoust. Soc. Am. **84**, 560–565.
Townshend, B., Cotter, N., Van Compernolle, D., and White, R. L. (**1987**). ''Pitch perception by cochlear implant subjects,'' J. Acoust. Soc. Am. **82**, 106–115.
Vater, M. (**1982**). ''Single unit responses in cochlear nucleus of horseshoe bats to sinusoidal frequency and amplitude modulated signals,'' J. Comp. Physiol. **149**, 369–388.
Walliser, W. (**1969**). ''Zusammenhänge zwischen dem Schallreiz und der Periodentonhöhe,'' Acustica **21**, 319–329.
Weiss, T. F., and Rose, C. (**1988**). ''Stages of degradation of timing information in the cochlea: A comparison of hair-cell and nerve-fiber responses in the alligator lizard,'' Hearing Res. **33**, 167–174.
Whitfield, I. C. (**1980**). ''Auditory cortex and the pitch of complex tones,'' J. Acoust. Soc. Am. **67**, 644–731.
Wightman, F. L. (**1973**). ''The pattern-transformation model of pitch,'' J. Acoust. Soc. Am. **54**, 407–416.
Zwicker, E., and Fastl, H. (**1990**). *Psychoacoustics—Facts and Models* (Springer-Verlag, New York), Chap. 14, pp. 247–256.

The pitch of a mistuned harmonic: Evidence for a template model

Jian-Yu Lin and William M. Hartmann
Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824

(Received 18 July 1996; accepted for publication 15 January 1998)

A harmonic of a periodic complex tone can be heard out as a separate entity if the harmonic is slightly mistuned from its correct frequency. Pitch matching experiments show that the pitch of such a mistuned harmonic differs systematically from its frequency. The shift in pitch is found to be an exaggeration of the frequency mistuning. This article considers two classes of model for the pitch shift. In the first class are tonotopically local interaction models which attribute the pitch shift to interactions between the mistuned harmonic and neighboring harmonics, where the neighborhood is established by peripheral filtering. The second class of model attributes the pitch shift to a contrast between the mistuned harmonic and a broadband harmonic template. This article describes six pitch matching experiments using complex tones having spectral gaps, strategically chosen to compare local interaction and template models. The results show that when a competition is set up between local interactions and a template, the template proves to be dominant. A parallel between the pitch shifts of mistuned harmonics and periodicity pitch, also attributed to a harmonic template, is seen as the frequency range of the mistuned harmonic is changed. Tonotopically local influences are evident in several experiments, but they are of secondary importance. © *1998 Acoustical Society of America.* [S0001-4966(98)00305-1]

PACS numbers: 43.66.Hg, 43.66.Ba [JWH]

INTRODUCTION

A complex periodic tone with many harmonics is normally perceived as a single entity with a pitch approximately equal to the fundamental frequency of the tone. The individual harmonics are not heard separately. However, if a harmonic is mistuned from its correct frequency, it may be detected as a separate entity (Moore *et al.*, 1986; Hartmann, 1988). The mistuned harmonic sounds like a sine tone against the background of the low-frequency complex tone. The background, with a complex timbre, has a pitch close to the fundamental, though this pitch can be shifted as a result of the mistuning (Moore *et al.*, 1985; Darwin *et al.*, 1994). The mistuned harmonic appears as a pure tone, and its pitch is also shifted. Matching it with a (nonsimultaneous) sine tone gives a matching frequency that systematically disagrees with the mistuned harmonic frequency (Hartmann *et al.*, 1990).

The pitch shift of the mistuned harmonic can be described succinctly as an exaggeration of the frequency mistuning. For example, the third harmonic of a 200-Hz tone is normally at 600 Hz. If the harmonic is mistuned by 8%, its frequency becomes 648 Hz. Typically its pitch would increase to 655 Hz, corresponding to a pitch shift of 1% $[(655-648)/648=0.01]$. If the harmonic is mistuned by -8%, its frequency decreases to 552 Hz, but its pitch would typically decrease to 546 Hz, corresponding to a pitch shift of -1% $[(546-552)/552=-0.01]$. The reader will notice that pitch shifts are computed with respect to the mistuned harmonic frequency. The present article is a study of the pitch shifts of mistuned harmonics.

To find a model for the pitches of mistuned harmonics, it is natural to begin with previous work, which has been on the pitches of correctly tuned harmonics. In 1979, Terhardt presented a semi-empirical algorithm to calculate the pitches of the harmonic components of complex tones. These then became the *spectral pitches* that were combined to form a virtual pitch for the tone as a whole (see also Terhardt *et al.*, 1982b). According to the algorithm, the pitches of individual harmonics are shifted due to partial masking caused by the excitation patterns of neighboring harmonics.

The algorithm is based on experimental data showing that when a sine tone is partially masked by a sound having a lower frequency than the sine itself, the pitch of the sine is shifted upward. Similarly, masking by a sound having a higher frequency than the sine tone tends to shift the sine tone pitch downward.[1] Terhardt extended these sine tone masking ideas to the pitches of components in a complex tone. A harmonic, somewhere in the middle of the spectrum, tends to be shifted upward by neighboring harmonics having lower frequencies, and it tends to be shifted downward by neighboring harmonics having higher frequencies. Because of the asymmetry of auditory filters there is an upward spread of masking, and the usual net result is that the pitch of a harmonic is shifted upward. An exception occurs for the fundamental component, where there is no lower frequency excitation. Then the model predicts that the pitch shift is downward.

Because Terhardt's model is a place-based excitation pattern model, it can also be used to predict pitch shifts for inharmonic components. For example, Terhardt *et al.* (1982a) used the model to predict the pitches of church-bell tones. However, when the model was applied to mistuned harmonics it predicted positive pitch shifts no matter whether the mistuning was positive or negative. Although it correctly predicted the pitches found experimentally for positive frequency mistuning, it failed to predict the negative shifts for negative mistuning (Hartmann *et al.*, 1990; Hartmann and Doty, 1996).

To try to explain the observed pitch shifts for mistuned harmonics, Hartmann and Doty developed an alternative model based on neural timing. The model began with the idea that pitch should be determined by the peaks of the interspike interval (ISI) histogram, as suggested by Goldstein and Srulovicz (1977). According to the model, the pitch of a component is shifted when excitation from neighboring harmonics appears in the same auditory filter, leading to a complicated temporal pattern. In general, the effect of such interaction on the ISI histogram is to cause a harmonic to be attracted in pitch to its nearest neighbors. Therefore the higher neighbor tends to make a positive pitch shift, and the lower neighbor tends to make a negative pitch shift, an ''attraction'' effect which is opposite to Terhardt's ''repulsion'' effect. Decreasing the spacing in frequency between the harmonic and a neighbor by mistuning the harmonic results in an increased attraction.

The timing model was capable of producing the positive and negative pitch shifts that occur for positive and negative mistunings. With appropriate model parameters, calculated pitch shifts could be brought into agreement with experimental results for mistuned harmonics 2, 3, 4, 5, 7, 9, and 11 [Doty (Smith), 1989; Hartmann and Doty, 1996]. However, there was one glaring exception, namely the mistuned fundamental. The model said that the pitch shifts should be unusually small and opposite in sign to the mistuning. In contrast, experiment showed unambiguously that the mistuned fundamental behaves just like any other mistuned harmonic—pitch shifts are substantial and in the same direction as the frequency shift.

This difficulty led us to consider an alternative paradigm. Terhardt's excitation pattern model and the timing model of Hartmann and Doty are local interaction models is the sense that the pitch shifts depend mainly on neighboring harmonics. The two nearest neighbors are especially important. A limitation of such local interaction models is that they give no recognition to the role of the entire spectrum in integrating a correctly tuned harmonic into a complex tone, or to the role that a broadband harmonic template might play when the harmonic is segregated by mistuning. A harmonic template is an internalized pattern for the components of a periodic tone. Current models of complex tone pitch perception (Goldstein, 1973; Terhardt, 1974) say that the pitch is the fundamental frequency, or harmonic spacing, of the best fitting template. Further, the ''pitch meter'' studied by Duifhuis *et al.* (1982) and Scheffers (1983) uses a self-consistent harmonic template to determine whether each resolved spectral component of the physical tone should be integrated into a complex tone percept or should be segregated. The experiments of this article were designed to search for a role for a broadband template and to establish a contrast between template models and local interaction models in the pitch of mistuned harmonics.

I. EXPERIMENTS

A. Method

Experiments were performed to measure a quantity defined as the *pitch shift gradient*. The gradient concept originated in the observation that pitch shifts for mistuned harmonics are in the same direction as the frequency mistuning. Therefore, if one takes the difference between the pitch shifts observed for positive and negative mistunings, the result should be larger and more robust than either pitch shift alone. An example is shown in Fig. 1; part (a) plots pitch, and part (b) plots pitch shift. The pitch shift observed for a mistuning of $+8\%$ is $\Delta p_+ = 1.1\%$. The pitch shift observed for a mistuning of -8% is $\Delta p_- = -1.3\%$. The pitch shift gradient is $\Delta p_+ - \Delta p_- = 2.4\%$, and there is an average pitch shift of $(\Delta p_+ + \Delta p_-)/2 = -0.1\%$.

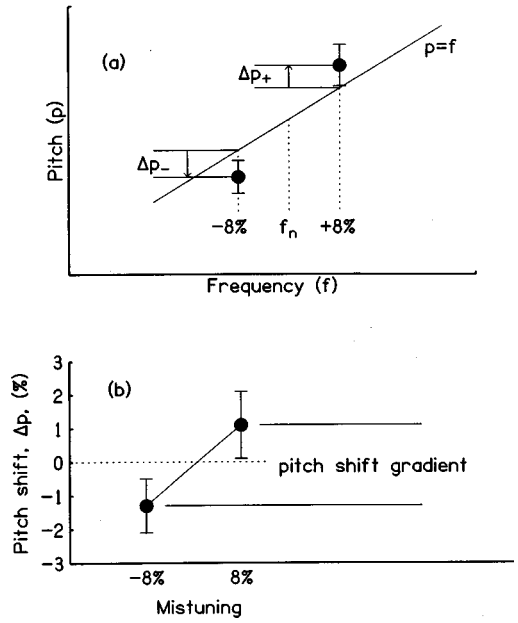


FIG. 1. Pitch shift gradient. A positive pitch shift gradient occurs when pitch changes more rapidly than frequency. Part (a) shows the mean and standard deviation of pitch matching data for a positively mistuned harmonic ($+8\%$) and a negatively mistuned harmonic (-8%). In the absence of a gradient, the pitch shifts Δp_+ and Δp_- would be the same. In fact, however, these shifts even have different signs. Part (b) plots the values of Δp_+ and Δp_- to show the pitch shift gradient. This gradient would be called *large* because the error bars for matching positive and negative mistunings do not overlap. The error bars are two standard deviations in overall length. All the data presented in this article are represented by graphs like part (b). To measure a pitch shift gradient does not necessarily require positive and negative mistuning of a harmonic. A gradient can be measured using any two frequencies that are not greatly different.

A pitch shift gradient is a perturbation in the smooth functional relationship between pitch and frequency. It occurs when, over a small range in frequency, the pitch changes differently from the frequency itself. Therefore, a pitch shift gradient can be measured by measuring the percentage pitch shifts for any two frequencies that are rather close together and taking the difference. If the measured pitch shifts turn out to be a constant percentage of the frequencies, then the pitch shift gradient is zero. The experiments below all used mistunings of plus and minus 8% to measure the gradients. Doty's data indicate that pitch shifts grow with increasing mistuning up to mistunings of about $\pm 4\%$ and remain flat, or decrease slightly out to $\pm 8\%$. Therefore, $\pm 8\%$ is a reasonable standard. The results would likely have been the same had the mistunings been as small as $\pm 4\%$.[2]

1. General procedure

The listener was seated in a sound-treated enclosure, holding a response box that controlled the events of an experimental trial. The listener pressed a yellow button to hear a complex tone with one of its harmonics mistuned. The listener pressed an orange button to hear a matching sine tone with a frequency that the listener could adjust by means of a ten-turn potentiometer on the box. Each tone was preceded by a 300-ms silent interval to minimize interactions between intervals. The listener could call up the complex tone or the matching tone in any sequence, with no time limit. When satisfied with the match, the listener pressed the green button to finish the trial. The stimulus and matching frequencies were recorded, and then the next trial began, with a different complex tone. There was no feedback to the listener.

The experiments used a wide variety of tones with as few as three and as many as six different complex spectra in a run. Each experimental run included at least one presentation of each spectrum with a mistuned harmonic mistuned by $+8\%$ and -8%. Therefore, a run included from 6–12 trials. It took from 10–20 min for a listener to finish a run. After a run was completed, the listener could come out to rest. Runs continued until each listener had done at least 12 runs in each experiment. Every data point shown in the figures that follow was based on the final 12 matches.

2. Stimuli

The complex tones with mistuned harmonics contained from 8–16 partials of equal amplitude. The fundamental frequencies varied from 150–800 Hz. Complex tones were generated by sound files, which were specific as to the fundamental frequency, the harmonic content, the mistuned harmonic number, and the percentage of mistuning—either plus or minus 8%. For a given trial, the appropriate sound file was loaded into a digital buffer 16k (16 384) samples long and was converted by a 16-bit DAC at a nominal sample rate of 16k/s. To prevent the listener from using pitch memory in the task, the sample rate was actually different on every trial. It was randomized over a range of $+10\%$ to -10%, with a rectangular distribution. Because of this randomization, all the numerical values of frequencies given in this article are nominal values.

The output of the DAC was low-pass filtered at 7 kHz, -115 dB/oct, and then shaped by a computer-controlled amplifier to give it an envelope with a 10-ms raised-cosine onset and offset, and a full-on duration of 1 s. The signal was presented to the listener via Sennheiser HD480 headphones such that each component had a level of 58 dB SPL. Therefore, a 16-component complex tone had a level of 70 dB SPL.

The matching tone was generated by repeatedly cycling a buffer using fractional-addressing technology. One cycle of a sine wave was loaded into a 16k buffer and sampled at the rate of 32k/s. The fractional address increment was determined by the potentiometer on the control box, as read by a 12-bit ADC. An exponential frequency control law was applied in software, leading to a frequency resolution of 0.06%.

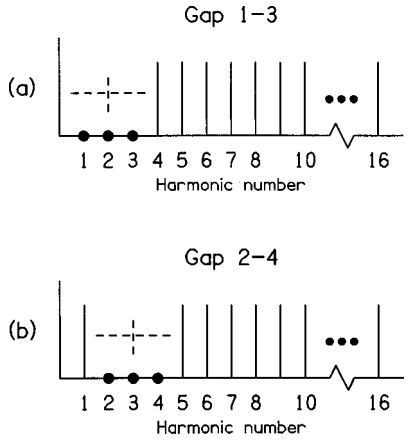


FIG. 2. Spectra for experiment 1. (a) Gap 1–3. A single mistuned harmonic (number 1, 2, or 3) added to harmonics 4–16. Vertical solid lines represent perfectly tuned harmonics. A vertical dashed line represents a mistuned harmonic. A dot is put on the horizontal axis where a harmonic is omitted. (b) Gap 2–4.

The matching tone was filtered and enveloped like the complex tone. The matching tone level was 55 dB SPL, 3 dB lower than a component of the complex tone, causing the matching tone to be approximately as loud as the mistuned harmonic. It was expected that equal loudness would make the pitch matching task easier.

3. Listeners

There were four listeners, three males, B, J, T, and one female, C. Their ages ranged from 19–55. All the listeners had negative otological histories and had some training as performers of musical instruments. They could perform accurate sine–sine pitch matching in the range 150–1000 Hz. Listeners J and B were the authors.

B. Experiment 1

1. Spectra of the complex tones

The spectra of the two complex tones in experiment 1 are shown in Fig. 2. The stimulus in Fig. 2(a) was made by starting with a periodic complex tone with a nominal fundamental frequency of 200 Hz and harmonics 4–16. Then one (and only one) of the omitted harmonics (1, 2, or 3) was reinserted as a mistuned harmonic, mistuned by either $+8\%$ or -8%. There were, therefore, six different stimuli in this experiment. This experiment is called ''gap 1–3.''

The stimulus in Fig. 2(b) was made in the same way except that harmonics 2, 3, and 4 were initially omitted. Then one of those harmonics was reinserted as a mistuned harmonic, mistuned by either $+8\%$ or -8%. This experiment also had six different stimuli, and it is called ''gap 2–4.''

2. Results—The zigzag effect

Figure 3 shows the results for gap 1–3 and gap 2–4 experiments, one panel for each subject. There are several observations to be made: First, there is a large pitch shift gradient for each plus–minus mistuned pair. Second, there are average shifts for some of the pairs, though these differ in size from subject to subject. Finally, although the magni-

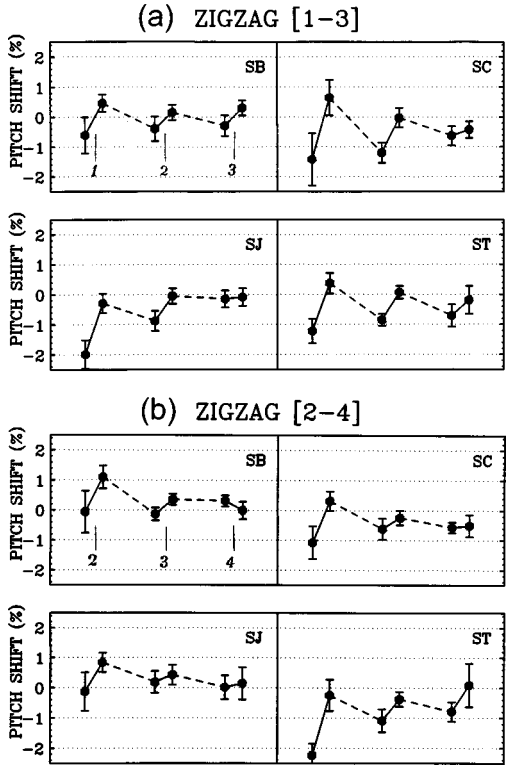


FIG. 3. Results of experiment 1. There are four panels (a) for the gap 1–3 experiment, and four panels (b) for the gap 2–4 experiment. In each panel there are three pairs of data points representing the ±8% mistuned pairs for three mistuned harmonics. Members of a pair are connected with a solid line so that one can compare the gradients of different pairs. A dashed line connects pairs to illustrate the ragged dependence of the pitch shift on the mistuned harmonic frequency, indicating the *zigzag effect*. The horizontal axis is monotonic (but not linear) with mistuned harmonic frequency.

tudes of the gradients are different for different listeners, there is a common tendency for low mistuned harmonics to have larger gradients. This low-frequency tendency will be seen in other experiments too.

Experiment 1 shows that the pitch shifts do not vary smoothly with mistuned harmonic frequency. The patterns shown in Fig. 3(a) and (b) are zigzags. These zigzag functions are evidence against an important class of models in which pitch shifts are mediated by local interaction. The gap 1–3 experiment is critical. The stimuli for gap 1–3 can be regarded as six target sine tones, increasing in frequency from harmonic 1 mistuned by −8% to harmonic 3 mistuned by +8%. As its frequency increases, the target tone gets closer to the rest of the spectrum (harmonics 4–16), which is responsible for the pitch shifts. Local interaction models based on the interaction of excitation patterns logically require that a shift varies smoothly with the separation between the target and the rest of the spectrum because the interaction decreases monotonically with increasing separation in local models. This prediction is contrary to the zigzag pattern observed experimentally. Local interaction models based on timing, like Hartmann–Doty, do not necessarily predict a smooth behavior for the gap 1–3 experiment, but the peripheral filtering in these models causes them to predict that gradients should increase in magnitude as the mistuned harmonic number increases from 1 to 3. In fact, the zigzag pattern does the reverse. It is largest for mistuned harmonic

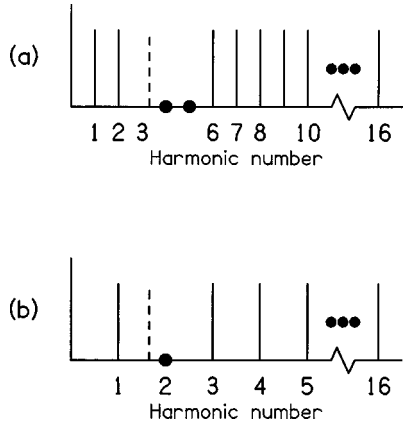


FIG. 4. Spectra for experiment 2. Part (a) shows mistuned harmonic number 3. Part (b) shows mistuned harmonic number 1.5. A dot appears on the horizontal axis where a harmonic is omitted.

number 1 and smallest for number 3. The zigzag pattern seen for gap 1–3 occurs also for the gap 2–4, except that the gradients are somewhat smaller in gap 2–4.[3]

The results of experiment 1 suggest that local interaction models are not appropriate models for the pitches of mistuned harmonics. Instead, it suggests that there is something special about the integer harmonic numbers, even when they are in a gap. A large pitch shift gradient occurs when the pair straddles a special frequency where a harmonic is supposed to be. This, in turn, is evidence that the complex tone forms a harmonic template within the auditory system, signalling the location of harmonics.

C. Experiment 2

1. Spectra of the complex tones

The purpose of experiment 2 was to make a direct comparison between two stimuli: (a) in which the target tone was a mistuned harmonic, and (b) in which it was not. If a harmonic template is the dominant influence on the pitch shifts of mistuned harmonics, as suggested by experiment 1, then shifts should be evident for the mistuned harmonic in (a) but not for the target in (b). The key to experiment 2 was that the local spectral structure around the target was similar for both stimuli. Therefore, local interaction models predict similar pitch shifts for both (a) and (b).

The spectra for the tones of experiment 2 are shown in Fig. 4(a) and (b). For stimulus (a), the fundamental frequency was 200 Hz. Harmonics 4 and 5 were omitted, and the third harmonic was mistuned, either by +8% or −8%, i.e., it was nominally either 648 or 552 Hz. For stimulus (b), the fundamental frequency was 400 Hz, and the second harmonic was omitted to make a gap. A target, having the same frequency as the target in stimulus (a), was inserted in the gap. Figure 4(b) shows that the target, with frequency of 648 or 552 Hz, looks like a mistuned component with harmonic number 1.5. Therefore, stimulus (b) makes an adequate test: the harmonics neighboring the target have the same frequencies as in stimulus (a) [the spectral differences between (a) and (b) are six (Cambridge) critical bandwidths away from the target region], but 1.5 is not a legitimate harmonic number.

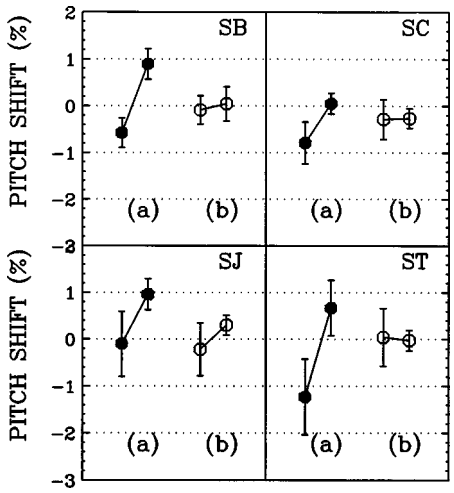


FIG. 5. Results of experiment 2. There are four subjects. Letters (a) and (b) refer to spectra in Fig. 4. The gradient is seen to be smaller in case (b), mistuned harmonic number 1.5, indicating the *integer effect*.

2. Results—The integer effect

The results of experiment 2 are shown in Fig. 5, one panel per subject. For all the subjects there are substantial shift gradients for stimulus (a), where the frequencies of the pair straddle harmonic 3, but there are no similar gradients for stimulus (b), where the frequencies of the pair are not close to an integer harmonic. The effect is significant $[F(1,3)=16, p<0.03]$. The result of experiment 2 is called the ''integer effect,'' meaning that large pitch shift gradients occur when the frequencies of the pair straddle integer harmonic frequencies. The integer effect suggests that a harmonic template plays an important role in the shifts.

For local interaction models, one expects pitch shifts to be mainly determined by the closest frequency components. Therefore, local interaction models tend to predict approximately the same pitch shifts for stimuli (a) and (b). Specific calculations with the tonotopic model of Terhardt *et al.* (1982a, 1982b) and the timing model of Hartmann and Doty both predict that (a) and (b) should have the same pitch shifts to five significant figures. These predictions are contrary to the large difference between (a) and (b) shown in Fig. 5.

D. Experiment 3

1. Spectra of the complex tones

The main purpose of experiment 3 was to see whether spectrally distant harmonics contribute to the pitch shift of a mistuned harmonic. The experiment used five complex tones, all with a period of 1/200 s. All had a mistuned fundamental. The spectra are shown in Fig. 6: Tone (a) was a 16-component complex tone. Tones (b)–(d) were the same as tone (a) but with increasing number of omitted harmonics. Tone (e) had only harmonics 1, 2, and 3.

2. Results—The proximity effect

The results of experiment 3 are shown in Fig. 7. Because listener SB did not complete this experiment, data are shown for only three listeners; the results are essentially the same for all three. As the neighboring harmonics 2, 3, and 4 were

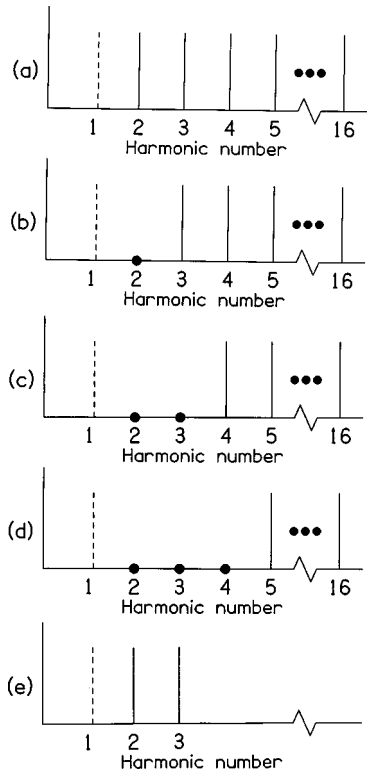


FIG. 6. Spectra for experiment 3. Tones (a)–(e) have a mistuned fundamental. Harmonic spectral blocks run from proximate (a) to remote (d). Tone (e) had only three harmonics. Dots on the horizontal axis mark the positions of omitted harmonics.

omitted one by one, the pitch shift gradient became smaller, but some gradient was always present. In only one case, tone (d) for SJ, was the gradient less than the error bars.

The results of the experiment show that distant harmonics are capable of generating a pitch shift gradient. Because the distant harmonics were well outside the critical band containing the target, the pitch shifts cannot be understood from a model that treats only a single peripheral channel. Specific calculations using Terhardt's model show negligible pitch shift ($<0.1\%$) for a mistuned fundamental when harmonic 2

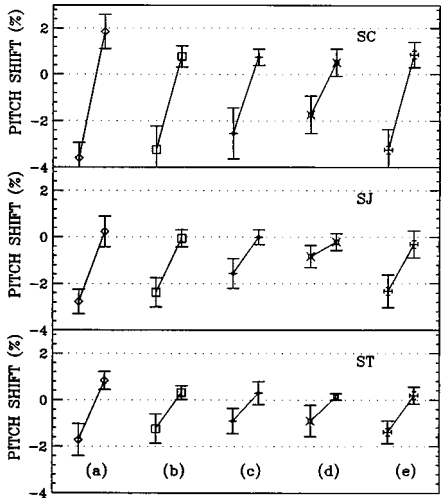


FIG. 7. Results of experiment 3. There are three subjects. Letters (a)–(e) refer to spectra in Fig. 6. Pitch shift gradients decrease in order (a) through (d), indicating the *proximity effect*.

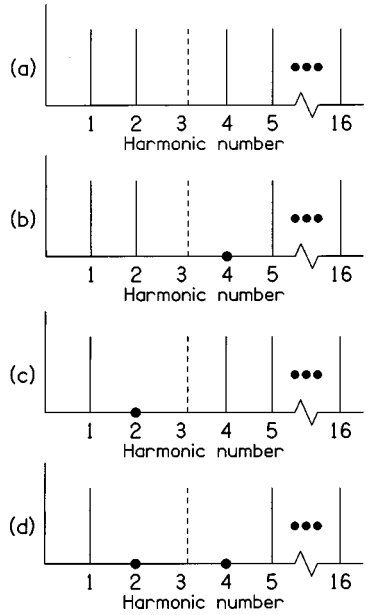


FIG. 8. Spectra for experiment 4. A mistuned third harmonic in symmetrical [(a) and (d)] and asymmetrical [(b) and (c)] harmonic environments. Dots on the horizontal line indicate omitted harmonics.

is omitted and even smaller shifts for other omitted-harmonic stimuli. The Hartmann–Doty model underpredicts the pitch shift for the mistuned fundamental, as always.

The fact that appreciable pitch shift gradients were caused by distant harmonics suggests that pitch shifts originate at a higher level, where the outputs of tuned channels are combined. This result is consistent with a template model for the shifts.

A second result of experiment 3 is that the shift gradients were smaller when the target was farther from the harmonics of the complex tone. This result is a ''proximity effect,'' which might be regarded as evidence for the importance of a local interaction such as partial masking. However, the large size of the shifts themselves in Fig. 7 seems out of proportion to the small weight that local interaction plays in other experiments, such as experiment 2. It seems more likely that the proximity effect represents increasing influence for a harmonic template when the components that establish the template become closer to the target.

E. Experiment 4

1. Spectra of the complex tones

Experiment 4 created spectral asymmetries for the mistuned harmonic target by selectively eliminating nearest-neighboring harmonics. The motivation for this experiment was that both the Terhardt place model and the Hartmann–Doty timing model are sensitive to local spectral structure. The former is particularly sensitive to the removal of a lower harmonic; the latter is sensitive to both lower and upper neighbors of the mistuned harmonic. Experiment 4 looked for the expected large changes in average shifts.

The target was a mistuned third harmonic, mistuned by $+8\%$ or -8%. The fundamental was 200 Hz. The spectral contexts are shown in Fig. 8: In tone (a) all harmonics were present. In tone (b) the fourth harmonic was omitted. In tone (c) the second harmonic was omitted. In tone (d) both the second and the fourth harmonics were omitted. The Terhardt model predicts that eliminating the second harmonic changes the average shift from 1.9% to 0.5%. The Hartmann–Doty model predicts that all shifts, regardless of mistuning, are positive for an omitted second harmonic, and all shifts are negative for an omitted fourth harmonic.

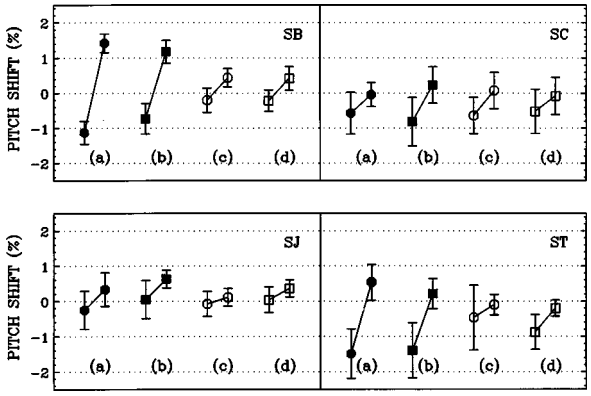


FIG. 9. Results of experiment 4. There are four subjects. Letters (a)–(d) refer to spectra in Fig. 8. Larger gradients occur for tones (a) and (b) where the second harmonic is present, indicating a *local asymmetry effect*.

2. Results—The local asymmetry effect

The data from experiment 4 are shown in Fig. 9, one panel for each subject. The first observation to be made is that the large changes in average shifts (average over positive and negative mistunings) predicted by local interaction models for asymmetrical spectra (b) and (c) did not materialize. The changes in average shifts are small, and the directions of the changes are different for different listeners. This result is evidence against local interaction models. The largest effect seen in this experiment was a reduction in gradient caused by removing the second harmonic (spectra c and d), whether or not the fourth was removed. For SB and ST the change was dramatic. Overall, the reduction in pitch shift upon removal of the second harmonic was significant $[F(1,7)=11.3,\ p<0.02]$. The special importance of the second harmonic in producing the gradient, in comparison with the fourth, is a *local asymmetry*. Generalized from this experience with a mistuned third harmonic, a statement of the local asymmetry effect says that the harmonics below the target are more important than the harmonics above the target in making a pitch shift gradient. Like the proximity effect, the local asymmetry effect has some of the character of a local interaction model. The upward spread of excitation might be expected to result in just this kind of asymmetry. However, local interaction models have the problem that an interaction large enough to generate a large change in gradients (e.g., SB and ST) also tends to generate a large shift in average pitch, which is not seen experimentally.

F. Experiment 5

1. Spectra of the complex tones

Experiment 5 was done to see the how pitch shift gradients for the mistuned fundamental depend on the period of the complex tone. There were three complex tones with eight

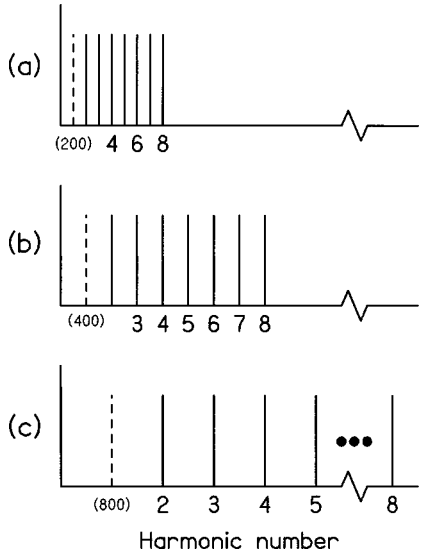


FIG. 10. Spectra for experiment 5. The three complex tones are scaled so that nominal fundamental frequencies are 200, 400, and 800 Hz. The fundamentals are mistuned.

components and mistuned fundamentals. The only difference among the tones was the nominal fundamental frequency: (a) 200 Hz, (b) 400 Hz, and (c) 800 Hz, as shown in Fig. 10.

2. Results—The frequency effect

The results of Experiment 5 are shown in Fig. 11. The results clearly show that the pitch shift gradient decreases with increasing fundamental frequency. Whereas the gradients at 200 Hz are large, greater than 2% for all listeners, the gradients at 800 Hz are almost zero. This is the frequency effect. The gradient is larger for 200 Hz than for 400 Hz [$F(1,3)=53$, $p<0.01$], and it is larger for 400 Hz than for 800 Hz [$F(1,3)=24$, $p<0.05$].

The Terhardt model predicts the frequency effect, largely because of the increased sensitivity to SPL excess as the frequency decreases. It also predicts the dominance of negative pitch shifts seen experimentally. However, the predicted pitch shift gradient has the wrong sign, and, therefore, the model fails. The Hartmann–Doty model fails for a mistuned fundamental, as usual, because it predicts that the gradient is too small to be observed, whatever the fundamental frequency.

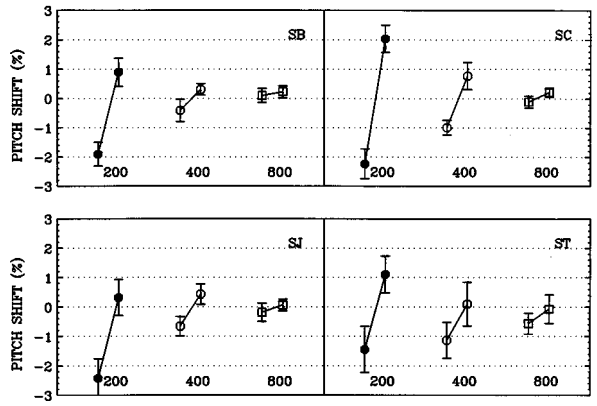


FIG. 11. Results of experiment 5. There are four subjects and three fundamental frequencies with spectra given in Fig. 10. The gradients are seen to be smaller for higher fundamental frequencies, indicating a *frequency effect*.

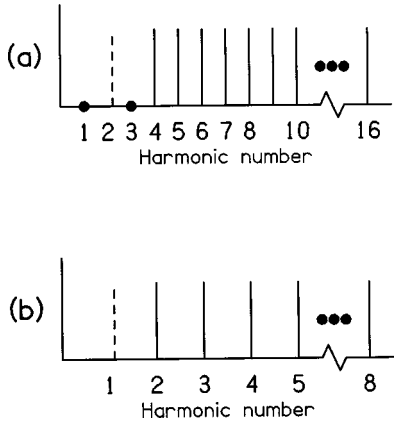


FIG. 12. Spectra for experiment 6. A mistuned ''400-Hz'' component serves as a mistuned second harmonic (a), or as a mistuned fundamental (b).

The small gradients seen at 800 Hz might have been predicted from a harmonic template model because periodicity pitch is weak at 800 Hz. Modern pitch perception models suggest that periodicity pitch, as seen in tones with missing fundamentals, is the result of a central harmonic template. The DWS pitch meter is an example (Scheffers, 1983). According to Ritsma's 1962 amplitude modulation experiments, 800 Hz is the upper limit for the existence of periodicity pitch. It seems natural to attribute this limit to a weakness in template influence for fundamental frequencies as high as 800 Hz. An ineffective template, in turn, would produce small pitch shift gradients, as observed in experiment 5. A possible further interpretation for the results of experiment 5 is that templates with low fundamental frequencies or long periods (e.g., 1/200 s are more effective than templates with short periods (e.g., 1/800 s). This interpretation was tested in experiment 6.

G. Experiment 6

1. Spectra of the complex tones

Experiment 6 had two goals. The first was to test the template period conjecture from experiment 5, the second was to explore the dependence of pitch shift gradients on harmonic number. The idea was to create a similar local spectral environment for mistuned harmonics, subject to two different overall periodicities. The spectra are shown in Fig. 12. In tone (a) the second harmonic of the complex tone was mistuned and its neighbors (the fundamental and the third harmonic) were omitted. The fundamental frequency was 200 Hz. In tone (b) the stimulus was a 400-Hz complex tone with its fundamental mistuned. The local spectral structures for the mistuned harmonic were similar for both (a) and (b), but the harmonic number in (a) was 1 (the fundamental), and the harmonic number in (b) was 2 (the second harmonic). Both stimuli had the same power. According to local interaction models both should have the same pitch shift (less than 0.2% in Terhardt's model).

2. Results

The results of experiment 6 are shown in Fig. 13. For all the subjects, the lower harmonic number (the fundamental)

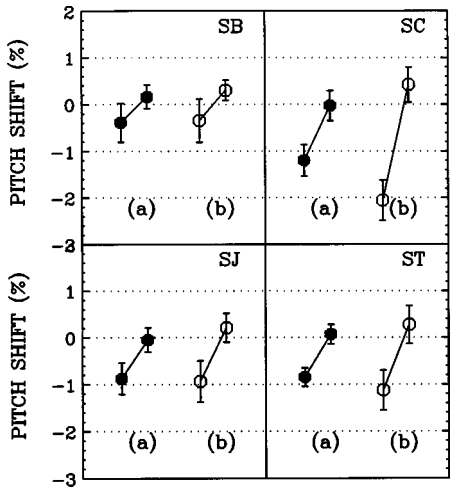


FIG. 13. Results of experiment 6. There are four subjects. Letters (a) and (b) refer to spectra in Fig. 12. The gradients for (b) tones are slightly larger, indicating a *harmonic number effect*.

suffered the larger gradient. This is a harmonic number effect. The effect is not significant according to our usual test $[F(1,3)=4.4]$ because the size of the difference varies among subjects leading to a large variance. It is important, however, that the reverse effect did not occur.

The results of experiment 6 affect our interpretation of experiment 5. If long-period templates are more effective than short, as conjectured in the discussion of experiment 5, then experiment 6 should find a larger pitch shift gradient in case (a) (200-Hz complex tone) than in case (b) (400-Hz complex tone). In fact, this did not occur. Instead, the larger gradient occurred for the 400-Hz complex tone. Therefore, our interpretation of the frequency effect in experiment 5 is that templates are less effective in shifting the pitches of higher-frequency fundamentals.

II. DISCUSSION AND CONCLUSION

A. Outcomes and interpretations

The pitch matching experiments of this article confirmed that the pitch of a mistuned harmonic differs from its frequency and that the shift in pitch exaggerates the frequency difference caused by mistuning. The experiments used computer-generated complex tones with specially tailored spectra. The principal goal was to distinguish between two classes of models for the pitch shifts, one based on tonotopically local interactions, the other based on a broadband template.

In a local interaction model, either place based or timing based, the pitch shifts are attributed to the effect of neighboring harmonics. The effects of distant harmonics are imagined to be attenuated by peripheral auditory filters. It is reasonable to try to apply local interaction models to mistuned harmonic pitch because these models have provided a qualitative account of many other pitch shift effects that occur for sine tones when noise or other tones are added to the signal (Webster and Muerdter, 1965; Terhardt, 1971; Rakowski and Hirsh, 1980; Burns and Osterle, 1980; Houtsma, 1981; Hartmann, 1996, 1997).

In a template model, by contrast, it is hypothesized that pitch shifts result from the influence of a broadband framework established by all the harmonics of the complex tone. Because the template model is more central than local interaction models, it is less clear how one should proceed to make it quantitative. The fact that pitch shifts are almost always found to be exaggerations of the frequency mistuning means that the template acts in such a way as to enhance contrasts. The enhancement favors the segregation of the mistuned harmonic as an independent entity.

Important evidence in favor of the template model came from the *zigzag* effect in experiment 1, where the presence of upper harmonics led to a ragged pitch shift pattern for mistuned lower harmonics. We found it helpful to think about this experiment in the following physical terms: In experiment 1 (gap 1–3), there is an empty spectral region where harmonics 1–3 have been omitted. Above this region there is spectral strength in harmonics 4–16. Experiment 1 consisted of placing a sine tone probe somewhere in the empty spectral region and observing ''forces'' on it caused by harmonics 4–16. The imagined forces are the influences that shift the pitch of the sine tone. The zigzag effect indicated an apparent force whereby pitches are ''repelled'' from sites where harmonics ought to be. The gap 2–4 variation of experiment 1 could be analyzed similarly.

The *integer* effect, found in experiment 2, gave additional evidence for the influence of a harmonic template because no significant pitch-shifting forces were found for sine tones half way between harmonic sites. By contrast, a large pitch shift gradient appeared, for the same local spectrum, when the sine tone was near a harmonic site.

Experiments 3 and 4 led to the *proximity* and *local asymmetry* effects, which indicated a role for local interaction. These effects are strong enough that any complete model for pitch shifts must take them into account. These effects essentially showed that spectral components near the mistuned harmonic target are more important in shifting the pitch of the target than are remote components. This result suggests a role for peripheral filtering of the kind that is always observed in masking experiments, but it is not clear that the effects actually resemble masking. Masking models, or other local interaction models consistent with the proximity effect, also tend to predict large average pitch shifts for asymmetrical spectra, and these were not observed. Instead, the proximity and local asymmetry effects may indicate a restricted ''range of interaction'' for a template.

Experiment 5 studied mistuned fundamentals for three tones with spectra that were identical except for octave scaling factors. This experiment found that a low-frequency mistuned fundamental was shifted considerably more than a high-frequency mistuned fundamental. We interpreted this observation as a *frequency* effect, which says that a template is more effective at a low frequency, such as 200 Hz, than at a higher frequency, such as 800 Hz. Such a frequency effect is well known in experiments that study the strength of periodicity pitch, which has often been attributed to template matching, and it was gratifying to find it in our experiment. An alternative interpretation for experiment 5 would be that templates with long periods are more effective than tem-

plates with short periods. However, evidence contrary to the latter interpretation was found in experiment 6, which systematically changed the period of the template. In general, experiments 3–5 exhibited larger pitch shift gradients for lower frequencies.

B. Caveats and conjectures

The experiments reported in this article measured pitch shifts, defined as the difference between the frequency of a mistuned harmonic and the frequency of a matching sine tone. A measured shift is, therefore, the change in the pitch of a sine tone when it is embedded in an otherwise periodic complex, with the period chosen in such a way that the sine tone becomes a mistuned harmonic. To avoid misinterpretation, it is important to note that our experiments did not measure the change in the pitch of a harmonic as the harmonic frequency changes from correctly tuned to mistuned. Therefore, our measured shifts can be interpreted as the pitch shifts caused by mistuning a harmonic only if it is assumed that the pitch of a correctly tuned harmonic is equal to the frequency of the harmonic. The preponderance of evidence suggests that this this assumption may be valid. In 1971 and 1979 Terhardt reported that the pitches of harmonics of a complex tone differed from their frequencies. However, extensive experiments by Peters *et al.* (1983) failed to confirm that result. Further, when the zero-frequency-shift data were extracted from the pitch measurements made by Doty (1989), it was found that harmonic pitches were statistically the same as harmonic frequencies, in agreement with Peters *et al*. Therefore, it is likely that our measurements actually do reflect the change in pitch as a spectral component becomes inharmonic, even though that is not what we measured. Our focus on pitch shift gradient made it unnecessary to measure the pitch shift caused by mistuning *per se*.

The focus on pitch shift gradient has another advantage, especially when comparing data to theory. There are pitch shifts with signal level, even for the sine-wave matching tone, and these effects are included in models such as Terhardt's. However, these level effects cancel in the measured gradient. Therefore, they do not confound the spectral context effects of interest here.

As noted in the Introduction, mistuning a harmonic in a complex tone leads to a shift in the low pitch (virtual pitch) of the tone overall as well as a shift in the pitch of the harmonic. Possibly these two shifts are related. Specifically, it might be imagined that pitch of the mistuned harmonic is shifted in order to agree with a harmonic of the shifted virtual pitch. This conjecture, however, does not square with the facts. According to the measurements of Moore *et al.* (1985), changing the frequency of a harmonic by a certain percentage leads to a change in virtual pitch that is always a considerably smaller percentage. By contrast, changing the frequency of a harmonic leads to a *larger* percentage change in the pitch of that harmonic.

It seems evident that the pitch shift observed for a component that has been perceptually segregated from a complex tone is likely to be closely associated with the process that segregates the component in the first place. It is the pitch, after all, that identifies a perceptually segregated component in a mistuned harmonic experiment. Recently, de Cheveigné (1997) has proposed a statistical decision theory model that successfully predicts the pitch shift of a mistuned harmonic as a function of the mistuning as measured by Doty. The model begins by assuming that a component of a complex tone can be represented by a distribution of pitch cues on an internal tonotopic coordinate. Pitch cues that are close to a harmonic place are perceptually integrated with the rest of the complex tone. Pitch cues that are more remote lead to a segregated tone. Precisely because they are more remote, these pitch cues lead to a pitch shift for the segregated tone. Although the model in its present version cannot account for both the pitch shifts and the high hit rates for mistuned harmonic identification found by Hartmann and Doty, this model does appear to be consistent with the pitch shift effects reported in the present article.

III. CONCLUSION

When a harmonic of a complex tone is mistuned, the harmonic can be heard out as a separate entity, distinct from the complex tone background. Pitch matching experiments show that the pitch of the mistuned harmonic does not agree with its frequency. The shift in pitch is, very reliably, found to be an exaggeration of the physical mistuning.

This article considered two possible kinds of model to account for the pitch shift. A tonotopically local model attributes the shift to the influence of neighboring harmonics. The influence may be the result of partial masking of firing rate excitation patterns (Terhardt) or it may stem from alterations in neural spike timing, as appear in the interspike interval histogram (Hartmann–Doty). In either case the effects are restricted by peripheral filtering. A second class of model attributes the pitch shift to a contrast between the mistuned harmonic and a spectral pattern in the mind of the listener that serves as a template against which tones are compared.

To try to distinguish between the two models, pitch matching experiments were done using special spectra. By changing the spectrum near the mistuned harmonic while maintaining the best fitting template (experiment 1) or by changing the template while maintaining the local spectrum (experiment 2) we were able to show a dominant role for a template. Scaling the frequency range of the matching experiment showed that pitch shifts tend to vanish for mistuned fundamentals near 800 Hz, similar to the limit for virtual pitch formation (experiment 5). This agreement is expected if both the pitch shift and virtual pitch are derived from a template.

Further experiments that maintained the best-fitting template while changing the spectrum near the mistuned harmonic (experiments 3 and 4) indicated a role for the local spectrum, which altered the details of the pitch shifts but did not upset the basic pattern established by the template comparison process. In the end, the dominance of the template is overwhelming.

ACKNOWLEDGMENTS

This work was supported by the National Institute on Deafness and Other Communicative Disorders, Grant No.

DC00181. We are grateful to Brad Rakerd, Tim Vander Velde, Alain de Cheveigné, and Ray Meddis for useful discussions.

[1]Terhardt's semiempirical model is unspecific concerning a physiological mechanism for the pitch shift, but in the low-frequency range the model is consistent in spirit with the idea that pitch is determined by a moment (e.g., centroid) of a neural excitation pattern along the tonotopic axis, and that the effect of masking is to erode the excitation pattern caused by the target component, thereby reducing the statistical weight of target excitation closest to the place of the masker. All the mistuned harmonics in the present paper are in this low-frequency range.

[2]Mathematically, a ''gradient'' is a slope, calculated as the ratio of rise-to-run on a graph. Our measure of the ''pitch shift gradient'' would have been more consistent with this definition had we divided our data by the 16% change in mistuned harmonic frequency $[+8\%-(-8\%)=16\%]$. Doing so, however, would have been misleading because the pitch shift saturates as a function of frequency shift. Doty's 1989 measurements show that pitch shifts measured at $\pm 8\%$ hardly differ from pitch shifts measured at $\pm 4\%$. Nevertheless, we use the term ''pitch shift gradient'' because the frequency change (the putative denominator) was always constant in our experiments, and because this term conveys the idea of a wrinkle in the pitch continuum that causes the pitch shift not to be a constant.

[3]Specific calculations using Terhardt's model show negligible pitch shift gradients for gap 1–3 and pitch shifts that are always positive and decreasing with increasing target frequency for gap 2–4. That occurs because this model gives great weight to low frequency partials. Predicted pitch shift gradients for gaps 2–4 have the wrong sign as usual.

Burns, E. M., and Osterle, E. L. (**1980**). ''Pitch shifts associated with partial masking,'' J. Acoust. Soc. Am. Suppl. 1 **67**, S20.

Darwin, C. J., Ciocca, V., and Sandell, G. F. (**1994**). ''Effects of frequency and amplitude modulation on the pitch of a complex tone with a mistuned harmonic,'' J. Acoust. Soc. Am. **95**, 2631–2636.

de Cheveigné, A. (**1997**). ''Harmonic fusion and pitch shifts of mistuned partials,'' J. Acoust. Soc. Am. **102**, 1083–1087.

Doty (Smith), S. L. (**1989**). ''Pitch perception of mistuned harmonics in a complex tone,'' M.S. thesis, Michigan State University, unpublished.

Duifhuis, H., Willems, L. F., and Sluyter, R. J. (**1982**). ''Measurement of pitch in speech: An implementation of Goldstein's theory of pitch perception,'' J. Acoust. Soc. Am. **71**, 1568–1580.

Goldstein, J. L. (**1973**). ''An optimal processor theory for the central formation of the pitch of complex tones,'' J. Acoust. Soc. Am. **54**, 1496–1516.

Goldstein, J. L., and Srulovicz, P. (**1977**). ''Auditory-nerve spike intervals as an adequate basis for aural spectrum analysis,'' in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, New York), pp. 337–345.

Hartmann, W. M. (**1988**). ''Pitch perception and the segregation and integration of auditory entities,'' in *Auditory Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 623–645.

Hartmann, W. M. (**1996**). ''Pitch periodicity, and auditory organization,'' J. Acoust. Soc. Am. **100**, 3491–3502.

Hartmann, W. M. (**1997**). *Signals, Sound, and Sensation* (AIP Press, Springer-Verlag, New York), Chap. 12.

Hartmann, W. M., and Doty, S. L. (**1996**). ''On the pitches of the components of a complex tone,'' J. Acoust. Soc. Am. **99**, 567–578.

Hartmann, W. M., McAdams, S., and Smith, B. K. (**1990**). ''Matching the pitch of mistuned harmonic in an otherwise periodic complex tone,'' J. Acoust. Soc. Am. **88**, 1712–1724.

Houtsma, A. J. M. (**1981**). ''Noise-induced shifts in the pitch of pure and complex tones,'' J. Acoust. Soc. Am. **70**, 1661–1668.

Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (**1985**). ''Relative dominance of individual partials in determining the pitch of complex tones,'' J. Acoust. Soc. Am. **77**, 1853–1860.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1986**). ''Thresholds for hearing mistuned partials as separate tones in harmonic complexes,'' J. Acoust. Soc. Am. **80**, 479–483.

Peters, R. W., Moore, B. C. J., and Glasberg, B. R. (**1983**). ''Pitch of components of a complex tone,'' J. Acoust. Soc. Am. **73**, 924–929.

Rakowski, A., and Hirsh, I. J. (**1980**). ''Post-stimulatory pitch shifts for pure tones,'' J. Acoust. Soc. Am. **68**, 467–474.

Ritsma, R. J. (**1962**). ''Existence region of tonal residue I,'' J. Acoust. Soc. Am. **34**, 1224–1229.

Scheffers, M. T. M. (**1983**). ''Simulation of auditory analysis of pitch: An elaboration on the DWS pitch meter,'' J. Acoust. Soc. Am. **74**, 1716–1725.

Terhardt, E. (**1971**). ''Pitch shifts of harmonics, an explanation of the octave enlargement phenomenon,'' *Proc. 7th ICA*, Budapest **3**, 621–624.

Terhardt, E. (**1974**). ''Pitch, consonance and harmony,'' J. Acoust. Soc. Am. **55**, 1061–1069.

Terhardt, E. (**1979**). ''Calculating virtual pitch,'' Hearing Res. **1**, 155–182.

Terhardt, E., Stoll, G., and Seewann, M. (**1982a**). ''Pitch of complex signals according to virtual-pitch theory: Test, examples, and predictions,'' J. Acoust. Soc. Am. **71**, 671–678.

Terhardt, E., Stoll, G., and Seewann, M. (**1982b**). ''Algorithm for extraction of pitch and pitch salience from complex tonal signals,'' J. Acoust. Soc. Am. **71**, 679–688.

Webster, J. C., and Muerdter, D. R. (**1965**). ''Pitch shifts due to low-pass and high-pass noise bands,'' J. Acoust. Soc. Am. **37**, 382–383.

Monaural phase discrimination by macaque monkeys: Use of multiple cues

David B. Moody, Colleen G. Le Prell, and Andrew J. Niemiec[a)]
Kresge Hearing Research Institute, University of Michigan Medical School, Ann Arbor, Michigan 48109-0506

(Received 14 May 1996; revised 27 May 1997; accepted 15 January 1998)

Research examining the discrimination of monaural phase change has suggested that temporal envelope shape, which varies with phase, may be an important cue. Much of that research employed stimuli consisting of three components, a center frequency (F_c), which is varied in phase, and an upper and lower sideband separated from the carrier by some frequency (ΔF). As the phase of the center component is varied, both temporal envelope and temporal fine structure change. The present research explored the salience of both envelope and fine structure as cues in a phase discrimination task. Monkeys were trained to report detection of a change from a three-tone complex with 90° starting phase for the center component to one in which the starting phase was smaller. In general, for the values of F_c tested, thresholds for phase change decreased as ΔF increased. When tested with comparison stimuli that had a temporal envelope closely matched to that of the standard, but 0° starting phase, subjects had difficulty discriminating these stimuli from the standard for smaller ΔF, but readily discriminated them at larger ΔF values. These findings suggest that temporal envelope is a critical cue in discrimination of three-tone complexes on the basis of the starting phase of the center component at small values of ΔF, but that other cues are used at larger ΔF values.
 [S0001-4966(98)01005-4]

PACS numbers: 43.66.Nm, 43.66.Gf, 43.80.Lb [RHD]

INTRODUCTION

Under appropriate conditions, changes in the phase of a monaurally presented auditory stimulus can be detected. The principal conditions required for such detection to occur are that the stimulus consists of two or more components, and that the component varied in phase is close enough in frequency to at least one of the other components so that it cannot be separately resolved by the auditory system. The resulting perceptual differences are described as changes in ''roughness'' or ''timbre'' (Mathes and Miller, 1947; Goldstein, 1967; Patterson, 1988) or in pitch (Buunen *et al.*, 1974; Moore, 1977; Moore and Glasberg, 1988). In general, the results of these studies suggest that the shape of the temporal envelope, which varies with change in component phase, produces the resulting perceptual differences.

The stimulus configuration most commonly employed in these experiments consists of three sinusoidal components: a carrier (F_c) and two half-amplitude sidebands separated from the carrier by some frequency (ΔF). The stimulus complex thus consists of $F_c-\Delta F$, F_c, and $F_c+\Delta F$. When the starting phase of the carrier and sidebands are all 0°, these components produce a sinusoidally amplitude-modulated (SAM) tone at F_c with a modulation frequency equal to ΔF. When the starting phase of the carrier is 90° relative to the sidebands, the stimulus is referred to as a quasi-frequency-modulated (QFM) condition. Although the long-term power spectra of these two signals are identical, the shape of the temporal envelope and the temporal fine structure differ significantly. In the SAM signal, the envelope of the waveform reaches zero amplitude at periods of $T=1/\Delta F$ and the waveform crosses zero at periods of $T=1/F_c$. In the QFM signal, the envelope does not reach zero, with the waveforms being much less modulated and the envelope reaching minima at periods of $T=1/2\Delta F$. In addition, the times between zero crossings do not occur at the same period but vary during the modulation cycle, thus resembling a frequency-modulated stimulus. As phase is varied between 0° and 90°, there is a gradual transition in the features of the stimulus between being SAM and QFM. Examples are given in Sec. I.

Mathes and Miller (1947) were among the first to demonstrate the role of limited auditory frequency resolution in the perception of monaural phase effects. von Helmholtz, a century earlier (as reprinted in 1954), had suggested on theoretical bases that phase relations of closely spaced tones had important implications for auditory theory. However, his empirical observations led to what is known as Ohm's acoustic law or Helmholtz's phase rule, which essentially states that the ear is phase deaf. Using both SAM and QFM stimuli, Mathes and Miller (1947), and later Goldstein (1967), clearly demonstrated that the two types of stimuli are discriminable only when the components are within a limited frequency region. Both studies showed that the size of that frequency region was level dependent, with estimates ranging from 15% of F_c at 20 dB SL to 40% of F_c at 60 dB SL (Mathes and Miller, 1947). These results indicate that the critical bandwidth for phase effects differs from the critical bandwidth observed with other auditory phenomena in that it is wider as well as being more level dependent.

Nelson (1994) defined the critical bandwidth for phase

[a)]Current address: Department of Psychology, Kenyon College, Gambier, OH 43022-9623.

(CB_{phs}) as the ΔF at which subjects could no longer distinguish QFM from SAM waveforms. He showed that the level-dependent behavior of CB_{phs} could be inferred from the filter slopes derived from forward-masking psychophysical tuning curves. Since the ability to discriminate phase changes is tied to the failure to resolve the frequency components of the complex stimulus, it might be assumed that poor frequency resolution, such as might occur in certain types of hearing impairment, would yield good phase discrimination. Indeed, Rosen (1984, 1986, 1987) has shown wider CB_{phs} values for hearing-impaired listeners; that is, they could detect the difference between SAM and QFM signals over a wider range of frequencies at equivalent sound pressure levels. Using a task similar to that of the present study, Rosen and Smith (1988) showed greater sensitivity to phase change for impaired listeners relative to that of normal-hearing subjects, a result consistent with wider CB_{phs}. Results indicating wider CB_{phs} were also shown by Moore and Glasberg (1988, 1989) using different phase discrimination tasks. Nelson and Schroder (1995), on the other hand, showed narrower or equivalent CB_{phs} values in hearing-impaired listeners at equivalent SPLs.

When the individual components of a signal cannot be separately resolved by the auditory system, it is assumed that discrimination is based on temporal changes in the stimulus, most notably on the temporal envelope which is strongly affected by phase changes. In other words, the output of the auditory filter that is stimulated presents a time-varying output to the nervous system resulting from the interaction of the stimulus components contained within its passband. Based on the demonstrations of de Boer (1961) that only phase changes which affect the envelope of a signal give rise to subjective differences, Buunen (1975) concluded that the only signal parameter needed to describe time-structure detection was the envelope of the signal. However, changes in temporal fine structure might also be detected. Buunen (1975) and Nelson (1994) considered cues that might be provided by an ''internal spectrum'' produced by the phase-dependent interaction of nonlinear distortion products with each other and with the acoustic components. Both authors suggest that these two explanations are not mutually exclusive, and may simultaneously contribute to phase discrimination.

We know of no studies of monaural phase discrimination in nonhuman subjects. The present experiment extends the study of phase discrimination to the macaque monkey, which shows frequency selectivity (Serafin *et al.*, 1982) and critical band functions (Gourevitch, 1970) similar to those from humans, but somewhat poorer frequency discrimination (Prosen *et al.*, 1990). The intent of the study was not to determine the range of ΔF within which phase could be discriminated as has been done in previous research with human listeners, but rather to explore the conditions under which temporal envelope appears to be the cue that is used to make such a discrimination. We believe such data will help to identify those features of complex auditory signals which may be most salient to these animals in the discrimination of complex signals such as are used in communication.

I. METHOD

A. Subjects

Four male macaque monkeys served as subjects in this experiment. One of the subjects, M141, was a *Macaca nemestrina*; the remaining three subjects, M161, M172, and M174, were *Macaca fuscata*. At the start of testing, the ages of the subjects were as follows: M141, >13 yr; M161, 13 yr; M172, 8 yr; and M174, 8 yr. All subjects had previous experience in psychoacoustic experiments, having served in studies of pure-tone frequency discrimination (M141, M161, and M174), discrimination of frequency modulation (M141 and M172) and discrimination of amplitude modulation (M141 and M161).

The subjects were individually housed and were maintained on a restricted food regimen in which they received a daily allotment of whole-diet, banana-flavored reward pellets (Bio-Serve, 190 mg) and Purina high-protein monkey chow. The amount of food given was sufficient to maintain steady weight gains throughout the experiment. Fresh fruit was provided three times per week. Water was freely available in the home cage at all times. Environmental enrichment in the form of toys and foraging puzzles was provided in the home cages. The animal facilities were AAALAC accredited and animal husbandry met or exceeded applicable standards, including the NIH Guide for the Use and Care of Laboratory Animals.

B. Apparatus

Prior to each testing session, the animals were placed in primate restraining chairs (Moody *et al.*, 1970) and transported from the animal quarters to the laboratory. Once in the laboratory, they were placed in double-walled sound rooms (Industrial Acoustics). Their heads were restrained, earphones (TDH-39P), mounted in universal joints, were placed over each ear, and a food-delivery chute was mounted on the chair. When required by the experimental procedure, the banana-flavored reward pellets could be automatically delivered down the chute such that the monkey could pick them up with his tongue. Mounted in front of the chair at the monkey's eye level was a cue light surrounded by a metal cylinder. Contact with the cylinder served as the behavioral response, and was sensed by a circuit that detected a minute electrical current that passed from the restraint chair through the monkey.

An Intel-based computer system controlled all aspects of the experiment, including generation and presentation of stimuli, implementation of experimental contingencies, and recording of data.

C. Stimuli

The stimuli were digitally synthesized using additive synthesis techniques and a 20-kHz clock rate. Each of the three sinusoidal components was separately generated and all three were then added together. The starting phase of the center component was varied as required to produce the test stimulus sets; sidebands both started at 0° and had amplitudes half that of the center component. Thus, the power

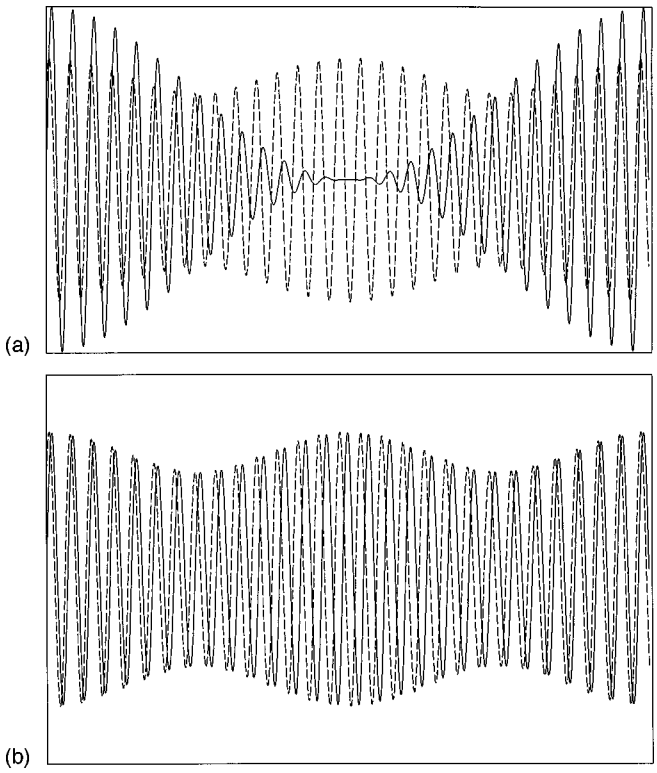


FIG. 1. (a) Examples of three-component stimuli having identical power spectra but with 0° (solid line) or 90° (dashed line) starting phases for the center component. The 0° condition results in a sinusoidally amplitude-modulated (SAM) signal at the frequency of the center component. The 90° condition results in a quasi-frequency-modulated (QFM) signal. (b) Examples of three-component stimuli having closely matched temporal envelopes but different starting phase, sideband spacing, and sideband envelopes. The dashed line is identical to that in panel (a); the solid line is a signal with 0° starting phase for the center component, but with sidebands spaced at twice the frequency of those of the SAM signal in panel (a) and with amplitudes of 0.079 that of the center component.

spectra of all stimuli for a particular combination of F_c and ΔF were identical. Schematic examples of stimuli with 0° and 90° center components are shown in Fig. 1(a).

The test stimuli were 250 ms in duration and were generated with 20-ms rise/fall times using a cumulative normal distribution function as the gating waveform. The digitized stimuli were presented through a 12-bit digital–analog converter and were low-pass filtered with an 8-kHz cutoff. All stimuli were presented at a level of 74 dB SPL. The response of the earphones was flat (± 1 dB) within the range of frequencies in the test stimuli.

Three values of F_c (0.5, 1, and 2 kHz) and four values of ΔF (25, 50, 100, and 200 Hz) were employed in the experiment, although not all values of ΔF were tested with each F_c. In addition, special stimuli were generated that had temporal envelopes that were closely matched to each of the 90°-phase (QFM) conditions, but with different sideband structures and with 0° starting phase for the center component. The sidebands for these stimuli were located at $2\Delta F$ of the corresponding QFM stimulus, but their amplitudes were set to 0.079 of the amplitude of the center frequency resulting in a SAM stimulus with a modulation index of 0.158. The overall gain of this stimulus was also adjusted to compensate for the reduction in sideband level. The QFM and envelope-matched stimuli generated using these parameters are depicted in Fig. 1(b). Although these stimuli did not, in fact, have mathematically identical envelopes,[1] we will use the term ''equal envelope'' to refer to the SAM stimuli having envelopes modeled after the QFM stimuli.

D. Procedure

The behavioral procedure was based on one that was originally devised for determining frequency and intensity difference thresholds (Moody *et al.*, 1976). Briefly, the subject was required to wait until the onset of a flashing cue light, and then to make contact with the response cylinder. Once contact was made, and for as long as it was maintained, the cue light remained steady and an auditory stimulus (the standard) was repeatedly presented through the right earphone. Stimuli were on for 250 ms and off for 250 ms. From 1 to 9 s after onset of the first standard stimulus, a test trial was initiated during which a comparison stimulus alternated three times with the standard. If the subject responded to the comparison stimulus presentation during the 2.8-s trial by releasing contact with the cylinder (a hit), a food pellet reward was delivered and the cue light and auditory stimuli were terminated. The flashing cue light was presented again as soon as 3 s passed without a response. Also included in the procedure were catch trials during which no comparison stimulus was presented, but cylinder releases were recorded to provide an estimate of the frequency with which the subject was responding in the absence of an appropriate stimulus (false alarms). Data from sessions during which responses occurred on more than 20% of the catch trials were discarded. Any responses that occurred at any other time than during test trials, or failure to respond to a test trial, resulted in a brief time-out from the experiment. During this 5-s interval, the cue light and stimulus presentation were discontinued and the subject was required to refrain from touching the response cylinder. These contingencies were effective in reducing the frequency of inappropriate responses.

For these experiments, the standard stimulus was the 90°-phase condition (QFM), and the comparison stimuli consisted of three-tone complexes in which the starting phase of the center component was less than 90°. During a given session, eight different comparison stimuli were presented according to the method of constant stimuli, and a threshold was determined as that starting phase that was detected on 50% of the test trials. The eight comparison stimuli were selected such that at least one, but no more than two, of the stimuli were detected on fewer than 50% of the test trials.

Testing was continued on each combination of F_c and ΔF until stability was achieved, as defined by the criterion that the thresholds measured on four of five consecutive days had to be within $\pm 10\%$ or $\pm 1°$ (whichever was greater) of the median threshold for those five days. The median was then taken as the estimate of threshold.

Following determination of each threshold estimate, an additional condition was run in which the most difficult stimulus in the test set (the one closest in phase to the 90° standard) was replaced with one, as described earlier in Sec. I C, having a temporal envelope closely matched to that of the QFM standard. All other experimental conditions re-

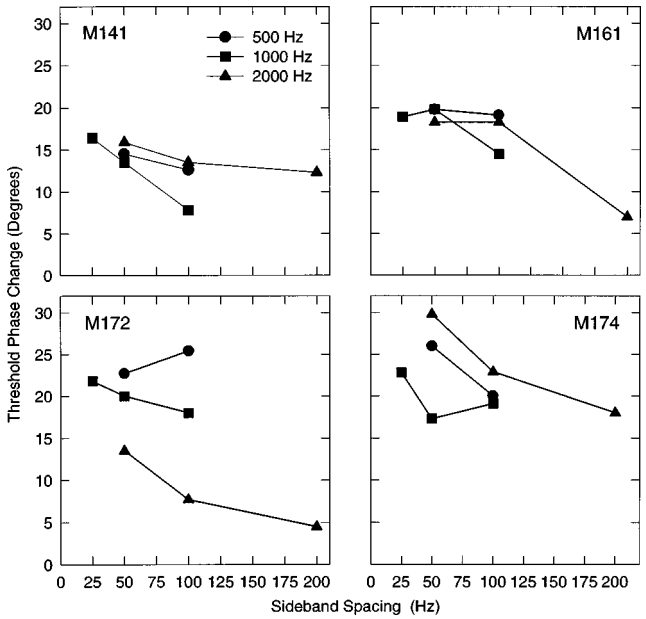


FIG. 2. Thresholds from each of the four subjects for detecting a change in phase of the center component of a three-component complex. The standard stimulus had a 90° starting phase. The ordinate values represent the decrement in starting phase required for the subject to detect the phase difference on 50% of the presentations. The symbols represent different values of the center frequency (F_c).

mained unchanged. Because this stimulus had a temporal envelope like that of the QFM standard stimulus, but was actually a SAM stimulus having lower-amplitude sidebands spaced at twice the frequency of the standard, it provided an indication of the extent to which the phase discrimination was based on the temporal envelope of the waveforms rather than on difference in temporal fine structure or sideband configuration. In order to reduce the likelihood that subjects would learn to ignore the temporal envelope which, in these stimuli, was no longer a reliable cue for discriminating comparisons from the standard, generally no more than three sessions were run before switching to another phase-discrimination condition. For one subject, M174, up to nine sessions with equal-envelope stimuli were carried out to determine if any progressive improvement occurred. During the equal-envelope testing sessions, percent hits to the envelope-matched stimulus were taken as indicators of the salience of the temporal envelope cue. If subjects failed to discriminate the equal-envelope stimulus from the standard, then envelope was assumed to be the cue being used in the phase discrimination. If subjects readily discriminated the equal-envelope from the standard, then it was assumed that cues other than envelope were available.

II. RESULTS AND DISCUSSION

Average threshold values required for each of the four subjects to discriminate a change in phase from the 90° standard are shown in Fig. 2. All subjects were able to discriminate the change in the starting phase of the center component of the three-component complex, in some cases showing remarkable sensitivity. For most subjects, thresholds were in the 10°–20° range, similar to those reported for normal and impaired humans by Rosen and Smith (1988) on a similar task but with quite different stimulus values. Discrimination of all stimulus combinations was expected in this study since all values of ΔF were much less than 40% of F_c, which, based on human data (Mathes and Miller, 1947), would be expected to be within the CB_{phs}. The values of ΔF were also less than the values of the conventionally defined critical band for macaque monkeys (Gourevitch, 1970). Since phase change could be readily discriminated within the range tested, the objective was to determine which features of the stimulus were used by the subjects in making the discrimination.

Although there are a few exceptions, the general tendency is for phase change to become more discriminable as ΔF increases. Since a decrease in center-component phase results in an increase in modulation depth, comparison of these data with those derived from temporal modulation transfer functions (TMTFs) is relevant. Based on TMTFs determined for macaque monkeys (including M141 and M161 from the present experiment; Moody, 1994), the expectation would be that, if anything, temporal envelope shape would become less discriminable at ΔF values of 100 and 200 Hz. Because the modulation rate of the envelope of the QFM standard stimulus is $2\Delta F$, the present data from ΔF values of 100 and 200 Hz are most appropriately compared with TMTF results obtained at modulation frequencies near 200 and 400 Hz, respectively. In the Moody (1994) data, sensitivity to modulation drops dramatically above 160 Hz. These TMTF observations call into question the notion that it is change in the temporal envelope of the waveform that provides the most salient discriminative cue to the subject, at least for the higher values of ΔF.

Comparisons with TMTF results must be carefully qualified, however. The Moody (1994) TMTF data were obtained with sinusoidally amplitude-modulated noise carriers, not pure tones as in the present study. The TMTF data are not normally obtained using sinusoidal carriers because it would be difficult to differentiate between the detection of modulation and the detection of spectral changes in sideband amplitude as modulation is varied. Viemeister (1979) presents TMTF data in which restricted spectral regions are stimulated by bandpass filtering broadband modulated noise. Those data suggest reduced sensitivity at lower values of F_c, but show the same decrease in sensitivity with increases in modulation frequency (ΔF) seen in the broadband data. Thus, those data provide additional evidence that if envelope is the cue being used to make the discriminations in the present study, sensitivity should decrease rather than increase as ΔF increases.

A second important difference is that the TMTF animals in the Moody (1994) study were discriminating an amplitude-modulated stimulus from an unmodulated standard, rather than a modulated stimulus from another modulated stimulus. It seems unlikely, however, that this difference could result in greater sensitivity to envelope change at higher modulation frequencies.

If temporal envelope is not the relevant discriminative cue, then an alternative cue is detection of changes in the temporal fine structure of the QFM standard as the phase of the center component is reduced toward the SAM condition.

Since the amount of change in the temporal fine structure is related to the value of ΔF, there would be greater changes at larger ΔF values, which would be expected to be more easily discriminable. However, certain aspects of the data are not compatible with that interpretation. If one assumes that the ability to discriminate changes in temporal fine structure resulting from changes in phase is related to frequency discrimination ability, then one would conclude that phase discrimination acuity should be related in an orderly way to F_c for a given value of ΔF. That is, the fine-structure changes produced by a given phase change for a given ΔF should be most easily discriminated with the 500-Hz F_c, and least easily discriminated when the F_c is 2000 Hz. The data do not provide strong evidence of such a difference, and for one subject, M172, the reverse ordering of thresholds as a function of F_c was observed.

Thus, two observations suggest that multiple mechanisms might be contributing to the observed results. The first observation is the inference from the TMTF data that discrimination based on envelope changes seems unlikely at higher values of ΔF. The second observation is that discrimination based on fine structure change also seems unable to explain all aspects of the data. The suggestion of multiple mechanisms was evaluated by presenting subjects with stimuli that had envelope structures closely matched to those of the standard QFM stimuli, but that had different sideband spacing, sideband amplitudes, and 0° carrier phase. The extent to which subjects failed to respond to such stimuli during initial exposures was used as an indication of the extent to which envelope was the dominant cue for the phase discrimination.

The results of the equal-envelope tests are presented in Fig. 3, which shows the percent of time the subject detected the equal-envelope stimulus for each of the conditions. The lines connecting the solid symbols indicate the average results from all sessions. The smaller open symbols indicate the hit rate for the first session in which a particular equal-envelope stimulus was presented.

The low hit rate for the smaller values of ΔF clearly indicates that the subjects had difficulty discriminating a QFM stimulus from a stimulus that had a closely matched envelope but that differed in spectral content and phase. This observation strongly suggests that envelope was the cue being used to make the discrimination at low ΔF values. Additionally, for most of the lower values of ΔF, there was no indication of improvement with additional testing, even though responses to the equal-envelope stimuli were reinforced. Even in M172, the animal that was tested nine times with the $F_c = 2$ kHz and $\Delta F = 50$ Hz, no indication of improvement was seen. This lack of improvement with explicit training suggests that, although other cues were present in the stimuli, they were not readily available to the subjects.

At higher values of ΔF, subjects more readily discriminated the equal-envelope stimuli from the QFM standards, even on the first equal-envelope test session. This observation suggests that the subjects might be using nonenvelope cues, possibly temporal fine-structure changes, to discriminate phase change. The general relationship seen in Fig. 3 between F_c and discriminability supports use of a fine-structure cue. At low values of F_c, nonenvelope cues become salient at lower values of ΔF, suggesting a Weber's law type of discriminability. Those suggestions must be further qualified, however, because the discrimination could also have been based on spectral resolution of the wider sidebands in the equal-envelope stimuli. Since this cue was unavailable in the phase-discrimination condition, the discrimination of fine structure, i.e., changes in instantaneous frequency, at larger ΔF values is more parsimonious in the sense of explaining both sets of data. One remaining possibility is that the envelopes were not equated after peripheral filtering, and therefore that discrimination of postfiltering envelope differences may have occurred. Although we believe this explanation to be less likely, we have no direct tests to rule out the possibility. It is likely that different mechanisms are used for discrimination at high versus low values of ΔF.

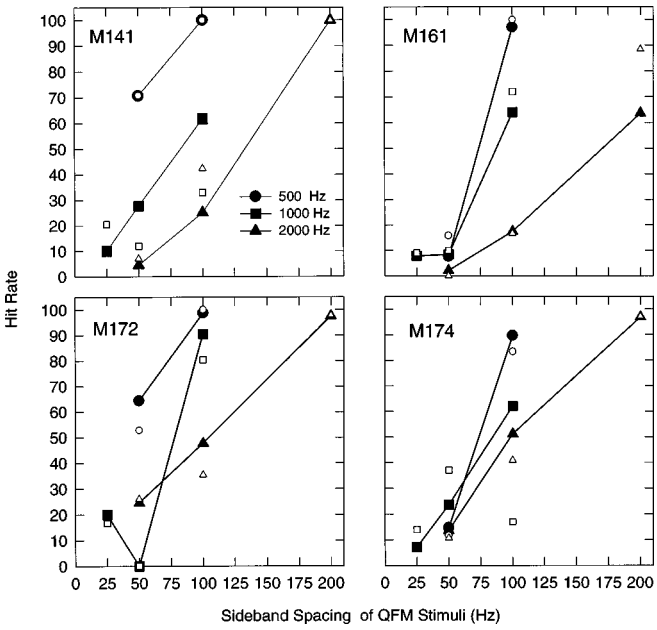


FIG. 3. Percent of time each of the four subjects responded to presentations of a stimulus having a temporal envelope closely matched to that of the 90° standard, but that differed in starting phase, sideband spacing, and sideband amplitude. The filled symbols represent the average of all presentations; the smaller open symbols represent data obtained on the first session of each condition. Low hit rates at narrow sideband spacings on this task suggest that temporal envelope was the most probable cue being used in the discrimination task to detect phase differences. In almost every case, hit rates and false alarm rates were similar when the hit rate was below 20%, suggesting that detection of the equal-envelope stimuli was truly at chance levels.

III. CONCLUSIONS

Within the range of F_c and ΔF values that were tested in this study, macaque monkeys were able, in a psychophysical task, to detect when the starting phase of the center component of a three-component complex was decreased from 90°. In general, phase discrimination thresholds decreased as sideband spacing was increased. Based on TMTF data, this observation suggested that a cue other than temporal envelope was being used to make the discrimination, at least at higher sideband spacing. When tested with a stimulus that had a temporal envelope closely matched to that of the QFM

standard, but that differed in center component phase, sideband spacing, and sideband amplitude, the subjects had difficulty in making the discrimination for narrower sideband spacing, but readily discriminated stimuli with wider spacing. Together these findings suggested that temporal envelope was the cue being used for narrower sideband spacing, but that some other cue was used as sideband spacing increased. It is likely that the cue was temporal fine structure, but the present data do not eliminate other possibilities. Thus, monaural phase discrimination can be based not just on temporal envelope changes, but also on other differences that are available as component phase is varied.

ACKNOWLEDGMENTS

This research was supported by Program Project Grant No. 2 P01 DC00078 from the National Institute on Deafness and Other Communication Disorders. The authors wish to express their appreciation and thanks to Catherine Thompson who had the primary responsibility for the daily testing of the subjects in this experiment, and also to thank the two reviewers who provided many useful suggestions on improving the manuscript and insights into the mathematical relationships involved in describing the envelopes of SAM and QFM stimuli.

[1]For these experiments, the sideband level required to produce closely matching temporal envelopes was determined empirically by adjusting sideband amplitude and overall gain until the two waveforms appeared to match, determined by visual inspection, as shown in Fig. 1(b). Subsequent mathematical determinations, using the equations provided by Rosen and Smith (1988), have indicated that the sideband amplitude required to produce a modulation depth equivalent to that of the QFM stimulus should have been 0.086 of the carrier (SAM index of modulation=0.172). In addition, even when the modulation depth is matched, the envelope of the QFM stimulus slightly departs from being strictly sinusoidal. Neither the mismatch of modulation depth nor the mismatch of envelope shape would be expected to modify the conclusions of the present study.

Buunen, T. J. F. (**1975**). ''Two hypotheses on monaural phase effects,'' Acustica **34**, 98–105.

Buunen, T. J. F., Festen, J. M., Bilsen, F. S., and van den Brink, G. (**1974**). ''Phase effects in a three-component signal,'' J. Acoust. Soc. Am. **55**, 297–303.

de Boer, E. (**1961**). ''A note on phase distortion in hearing,'' Acustica **11**, 182–184.

Goldstein, J. L. (**1967**). ''Auditory spectral filtering and monaural phase perception,'' J. Acoust. Soc. Am. **41**, 458–479.

Gourevitch, G. (**1970**). ''Detectability of Tones in Quiet and in Noise by Rats and Monkeys,'' in *Animal Psychophysics*, edited by W. C. Stebbins (Appleton-Century-Crofts, New York), pp. 67–97.

Mathes, R. C., and Miller, R. L. (**1947**). ''Phase effects in monaural perception,'' J. Acoust. Soc. Am. **19**, 780–797.

Moody, D. B. (**1994**). ''Detection and discrimination of amplitude-modulated signals by macaque monkeys,'' J. Acoust. Soc. Am. **95**, 3499–3510.

Moody, D. B., Stebbins, W. C., and Miller, J. M. (**1970**). ''A primate restraint and handling system for auditory research,'' Behav. Res. Methods Instrum. **2**, 180–182.

Moody, D. B., Beecher, M. D., and Stebbins, W. C. (**1976**). ''Behavioral Methods in Auditory Research,'' in *Handbook of Auditory and Vestibular Research Methods*, edited by C. Smith and J. Vernon (Thomas, Springfield, IL), pp. 439–497.

Moore, B. C. J. (**1977**). ''Effects of Relative Phase of the Components on the Pitch of Three-Component Complex Tones,'' in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 349–362.

Moore, B. C. J., and Glasberg, B. R. (**1988**). ''Effects of the Relative Phase of the Components on the Pitch Discrimination of Complex Tones by Subjects with Unilateral Cochlear Impairments,'' in *Basic Issues in Hearing*, edited by H. Duifhuis, J. W. Horst, and H. P. Wit (Academic, London), pp. 421–430.

Moore, B. C. J., and Glasberg, B. R. (**1989**). ''Difference limens for phase in normal and hearing impaired subjects,'' J. Acoust. Soc. Am. **86**, 1351–1365.

Nelson, D. A. (**1994**). ''Level-dependent critical bandwidth for phase discrimination,'' J. Acoust. Soc. Am. **95**, 1514–1524.

Nelson, D. A., and Schroder, A. C. (**1995**). ''Critical bandwidth for phase discrimination in hearing impaired listeners,'' J. Acoust. Soc. Am. **98**, 1969–1976.

Patterson, R. D. (**1988**). ''Timbre Cues in Monaural Phase Perception: Distinguishing Within-Channel Cues and Between-Channel Cues,'' in *Basic Issues in Hearing*, edited by H. Duifhuis, J. W. Horst, and H. P. Wit (Academic, London), pp. 351–358.

Prosen, C. A., Moody, D. B., Sommers, M. S., and Stebbins, W. C. (**1990**). ''Frequency discrimination in the monkey,'' J. Acoust. Soc. Am. **88**, 2152–2158.

Rosen, S. (**1984**). ''Hyperacute monaural phase sensitivity in the hearing impaired,'' Br. J. Audiol. **18**, 257–258 (abstract).

Rosen, S. (**1986**). ''Monaural Phase Sensitivity: Frequency Selectivity and Temporal Processes,'' in *Auditory Frequency Selectivity*, edited by B. C. J. Moore and R. D. Patterson (Plenum, London), pp. 419–427.

Rosen, S. (**1987**). ''Phase and the Hearing-Impaired,'' in *The Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Martinus Nijhoft, Dortrecht), pp. 481–488.

Rosen, S., and Smith, D. A. (**1988**). ''Auditory Temporal Information in the Profoundly Hearing Impaired,'' in *Basic Issues in Hearing*, edited by H. Duifhuis, J. W. Horst, and H. P. Wit (Academic, New York), pp. 431–439.

Serafin, J. V., Moody, D. B., and Stebbins, W. C. (**1982**). ''Frequency selectivity in the monkey's auditory system: Psychophysical tuning curves,'' J. Acoust. Soc. Am. **71**, 1513–1518.

Viemeister, N. F. (**1979**). ''Temporal modulation transfer functions based on modulation thresholds,'' J. Acoust. Soc. Am. **66**, 1364–1380.

von Helmholtz, H. L. F. (**1954**). *On the Sensations of Tone* (Dover, New York) (reprint of the AJ Ellis English translation of the 1877 German edition).

Lateralization and binaural discrimination of patients with pontine lesions[a)]

Vered Aharonson and Miriam Furst[b)]
Department of Electrical Engineering-Systems, Faculty of Engineering, Tel Aviv University, Tel Aviv 69978, Israel

Robert A. Levine
Eaton Peabody Laboratory, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts 02115

Michael Chaigrecht
Department of Neurology, Sourasky Medical Center, Tel Aviv, Israel

Amos D. Korczyn
Department of Neurology, Sourasky Medical Center, Tel Aviv, Israel and Department of Neurology, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

(Received 24 January 1996; revised 10 October 1997; accepted 5 December 1997)

Lateralization and just-noticeable difference (jnd) measurements relative to the center were tested in a large group of patients with pontine lesions caused either by stroke or multiple sclerosis. Stimuli included binaural clicks, and low- and high-frequency narrow-band noise bursts. Two major types of abnormalities were revealed in the lateralization performances: perception of all stimuli, regardless of interaural differences (time and/or level) in the *center* of head (center-oriented), or lateralization of all stimuli to one side or the other of the head (side-oriented). The only significant correlation between jnd and lateralization performances was that an elevated jnd was always manifested in abnormal lateralization, while abnormality in lateralization did not necessarily indicate an abnormal jnd. Center-oriented lateralization was observed either for both interaural differences or only for one of them, and was found in both MS and stroke patients. All side-oriented stroke patients were similarly unable to center binaural stimuli for both time and level cues, whereas only one MS patient had this abnormality for interaural time differences, while his level performance was normal. More abnormalities were detected in the narrow band stimuli tests, although in some cases performance was more degraded for click stimuli. Lateralization tasks with high-frequency stimuli were more sensitive detectors of abnormality than jnd for any kind of stimulus, or lateralization tasks with low-frequency stimuli or clicks. [S0001-4966(98)03704-7]

PACS numbers: 43.66.Pn, 43.66.Sr [RHD]

INTRODUCTION

The perception of auditory stimuli presented simultaneously to both ears has been investigated for several decades, yet the processes and structures responsible for binaural interaction are far from being fully understood. Considerable evidence indicates that this analysis is performed in the brainstem and that the cues used to perceive the sources of stimuli in auditory space consist of the differences between the stimuli reaching the two ears—either interaural time differences (ITD) or interaural level differences (ILD). Psychophysical experiments, concentrating in general upon people with normal hearing, have been among the most commonly employed for testing interaural faculties in humans (Tobias and Zerlin, 1959; Domnitz, 1973; Domnitz and Colburn, 1977; Durlach and Colburn, 1978; Wightman *et al.*, 1987; Algom *et al.*, 1988). Since the 1980's, interest has increased in the performance of multiple sclerosis (MS) patients in response to binaural stimuli (Hausler and Levine, 1980; Jenkins and Masterton, 1982; Hausler *et al.*, 1983; Hendler *et al.*, 1990; Furst *et al.*, 1990, 1992, 1995; Levine *et al.*, 1993a, b). These experiments were either discrimination tests, measuring just-noticeable differences (jnd) for ITD and ILD (Hausler and Levine, 1980; Levine *et al.*, 1993a); masking level difference (MLD) tests (Hendler *et al.*, 1990; Levine *et al.*, 1994), or lateralization tests using a click stimulus (Furst *et al.*, 1990, 1992, 1995; Furst and Algom, 1995).

It is difficult to arrive at firm conclusions from these studies, as they employed different stimuli and different experimental methods. Yet the following points may be suggested: (a) Abnormal performance in jnd tests is more likely to occur when high-frequency stimuli are presented, while low-frequency and broadband stimuli may still result in good discrimination ability. (b) Inability to perform interaural time discrimination can coexist with normal interaural level discrimination. (c) Some patients cannot discriminate small interaural differences. Of such patients some lateralize all click stimuli to one or the other side of their heads. Other patients never lateralize the clicks but refer the clicks only to the middle of their heads, for all interaural differences of time or level.

[a)]This study has been carried out in accordance with the principles of the Helsinki Declaration.
[b)]Electronic mail: mira@eng.tau.ac.il

These observations do not enable us to answer questions about the relationship between discrimination and lateralization, or the relationship between perception of transients and longer stimuli.

The purpose of the present study was to examine these relationships. We examined both stroke and MS patients with pontine lesions, and compared the performances obtained with clicks, high-frequency narrow-band noise, and low-frequency narrow-band noise. Each patient was tested in both discrimination and lateralization tasks.

I. EXPERIMENTAL METHODS

Eleven young healthy subjects and 22 patients went through the experimental procedure. The healthy subjects were students ranging in age from 20 to 30 years. All had normal audiograms and served as a control group. The patients who participated in the study were 9 multiple sclerosis (MS) patients and 13 stroke patients with pontine lesions, as indicated by clinical symptoms and routine MRI analysis. The MS patients were 27 to 43 years old and the stroke patients ranged in age from 57 to 82. All the patients had normal audiograms according to the standard of their age group.

Preliminary results from a complementary investigation of healthy elderly subjects indicate that there is no significant difference between young and elderly controls with regard to the psychophysical tests of this study (Babkoff *et al.*, 1996).

The test procedures included measurements of lateralization and discrimination ability. All the tests were performed during the same session for each subject. Each test was repeated for three stimulus types: clicks, and two $\frac{1}{3}$-octave noise bursts. Stimuli duration were 0.1 ms for the click and 250 ms for the noise bursts. The high-frequency narrow-band noise (HNB) was centered around 4 kHz and the low-frequency narrow-band noise (LNB) around 0.5 kHz. All stimuli were computer generated by an Ariel board 16+, based on a TI TMS320-C25 DSP chip, and using a 44-kHz sampling rate. The clicks were computer generated rarefaction clicks 40 s long. The narrow-band (NB) noises were generated by MATLAB, as a sum of sinusoidal components whose amplitudes and phases were uniformly distributed. To smooth the generated waveform, a Hanning filter was used.

All the tests took place in a sound-proof chamber, and the stimuli were delivered through calibrated Yamaha HP-2 earphones. Each subject's threshold for pure tones at frequencies from 250 Hz to 8 kHz and for the test stimuli was measured prior to the binaural tests. The stimuli were presented at a level 50 dB above the subject's threshold.

The procedure for the lateralization tests was similar to that described by Furst *et al.* (1995). Interaural time and intensity differences of stimuli were presented to a subject through earphones. In each trial, two successive short bursts were introduced, separated by an interval of 500 s. The first presented diotic stimuli to the two ears, whereas in the second, either the arrival time of the stimulus to one of the earphones was delayed (ITD trials), or the level was increased at one ear and decreased at the other ear by the same amount (ILD trials). Interaural time differences were varied from -900 to 900 s in 225-s steps (negative values indicate delays to the right ear). Interaural level differences ranged between -10 and 10 dB, in steps of 2.5 dB, where negative values indicate a higher intensity to the left ear. Thus both types of trials had a nine-level scale. After listening to both intervals, the subject's task was to indicate the position of the auditory image of the second burst by pushing one of nine equally spaced buttons on a rectangular keyboard. Button Number 1 on the subject's left signified sound to the left side; Button Number 9 signified the right side. Button 5 was considered analogous to the middle of the head with four numbers on each side of the center.

Prior to the lateralization testing each subject underwent a training session, to ensure that the instructions were fully understood. The training lasted a few minutes. The subject pushed the nine different buttons and listened to the ITD and ILD corresponding to each button. Each subject then decided when he/she felt ready to start the lateralization tests. At that point, the experimenter instructed the subject to close his/her eyes. As the experimenter pushed one of the buttons randomly, the subject was asked (1) to point to the position on his/her scalp which most closely corresponded to where the sound was perceived and (2) to relate this position to one of the nine button numbers.

Immediately following this training session the lateralization experiment was begun. Interaural level or time differences were presented randomly for one of the three types of stimuli (click, LNB, or HNB). The run was completed when each of the 17 different interaural differences had been presented at least 4 times for this stimulus. The same procedure was then repeated for the other two types of stimuli.

For all three types of stimuli, just-noticeable differences (jnd) of interaural time and level were measured with respect to the stimulus presented diotically. An adaptive 2-interval 2-alternative forced choice procedure was used (Levitt, 1971). The stimuli and interaural differences were similar to the ones used in the lateralization tests; here, however, time and level jnds were measured separately. The subject listened to successive intervals (300 ms each) with a 300-ms pause between them. His/her task was to press one of two buttons to indicate whether the signal in the second interval was perceived to the right (right button pressed) or left (left button pressed) of the signal in the first interval. Each of four successive presentations had the same interaural difference; the order of the two intervals was randomized for the four presentations. Performance on the four presentations determined the next interaural difference, as follows. All four correct resulted in a smaller interaural difference with the next four presentations. Three correct resulted in a repetition of the same interaural difference with the next four presentations. When only were one or two were correct, or when two successive sets of four presentations yielded one error in each, a larger interaural difference was introduced. The test ended when six reversals had occurred. The jnd was defined as the mean of the six reversals.

II. LATERALIZATION DATA REPRESENTATION

Typical lateralization performance results are shown in the histograms in Fig. 1, obtained from a subject in the con-

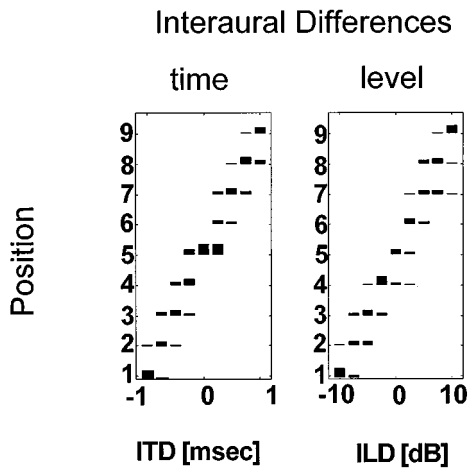


FIG. 1. Estimated joint probability functions, pdf(POS, ITD) and pdf(POS,ILD), represented by histograms, provided by the performance of a normal subject in the lateralization experiment for click stimuli. The histograms were obtained by counting the number of times the subject reported perceiving a certain position in response to dichotic stimuli with different ITDs and ILDs.

trol group. The left panel represents the subject's performance for clicks stimuli with ITDs and the right panel, the response to click stimuli with ILD. In both panels, the y-axis represents the nine positions defined by the test procedure, and the x-axis the interaural difference. Each bin in the histograms in Fig. 1 is defined as $h(\text{pos,ITD})$ or $h(\text{pos,ILD})$, which is equal to the relative number of times the subject indicated a certain position (pos) for a given interaural difference, ITD or ILD. The height of each bin represents the value of $h(\text{pos,ITD})$ or $h(\text{pos,ILD})$. Thus

$$\sum_{\text{pos}}\sum_{\text{ITD}} h(\text{pos,ITD})=1, \quad \text{and} \quad \sum_{\text{pos}}\sum_{\text{ILD}} h(\text{pos,ILD})=1. \tag{1}$$

There is a high correlation between the indicated positions and the interaural differences, as shown in Fig. 1. It is possible to calculate this correlation by fitting linear lines (regression lines):

$$\text{POS}=a^t\cdot \text{ITD}_{\text{pos}}+b^t \tag{2}$$

$$\text{POS}=a^l\cdot \text{ILD}_{\text{pos}}+b^l, \tag{3}$$

when POS$=1,2,...,9$ are the positions perceived, $\text{ITD}_{\text{POS}}=1,2,...,9$ are the introduced positions for ITD$=-900$, -675, -450, -225, 0, 225, 450, 675, 900 μs, respectively, and $\text{ILD}_{\text{POS}}=1,2,...,9$ for ILD$=-10$, -7.5, -5, -2.5, 0, 2.5, 5, 7.5, 10 dB, respectively. The coefficients a^t, b^t, a^l, and b^l in Eqs. (2) and (3) are derived by means of minimum-squared error fitting to the data. For the normal subject whose performance with click stimuli is presented in Fig. 1, the coefficients derived were: $a^t=a^l=0.95$, $b^t=0.27$, and $b^l=0.19$. For all the subjects in the normal control group and for all the stimulus types (click, LNB, and HNB), the regression line fit yielded the following coefficients: $0.84<a^t$, $a^l<1$ and $b^t,b^l<0.7$. The normal means and standard deviations are presented in Table I.

Representative patients' performances in the lateralization experiment are presented in Fig. 2. The performance of three stroke patients for clicks is presented in a histogram similar to that in Fig. 1. Comparison of the performance of the three patients to normal performance (presented in Fig. 1) indicates that Pt35's performance (top panels in Fig. 2) is the closest to normal. The values from his regression line fit [Eqs. (2) and (3)] are $a^t=0.93$, $b^t=0.26$, $a^l=0.87$, and b^l

TABLE I. Mean and standard deviation (s.d.) of the six lateralization indices, the four regression line parameters and the jnds, for ITD and ILD, obtained by the control group for clicks.

	Click		LNB		HNB	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
Θ_c^t	0.35	0.05	0.32	0.04	0.35	0.04
Θ_c^l	0.32	0.03	0.33	0.02	0.32	0.03
Θ_s^t	0.37	0.04	0.33	0.05	0.34	0.04
Θ_s^l	0.31	0.02	0.33	0.02	0.32	0.03
a^t	0.93	0.04	0.94	0.04	0.94	0.04
b^t	0.35	0.20	0.28	0.27	0.23	0.22
a^l	0.93	0.04	0.96	0.03	0.91	0.03
b^l	0.30	0.29	0.16	0.18	0.42	0.17
δ^t	0.03	0.02	0.03	0.02	0.04	0.03
δ^l	0.97	0.25	0.96	0.28	1.04	0.40

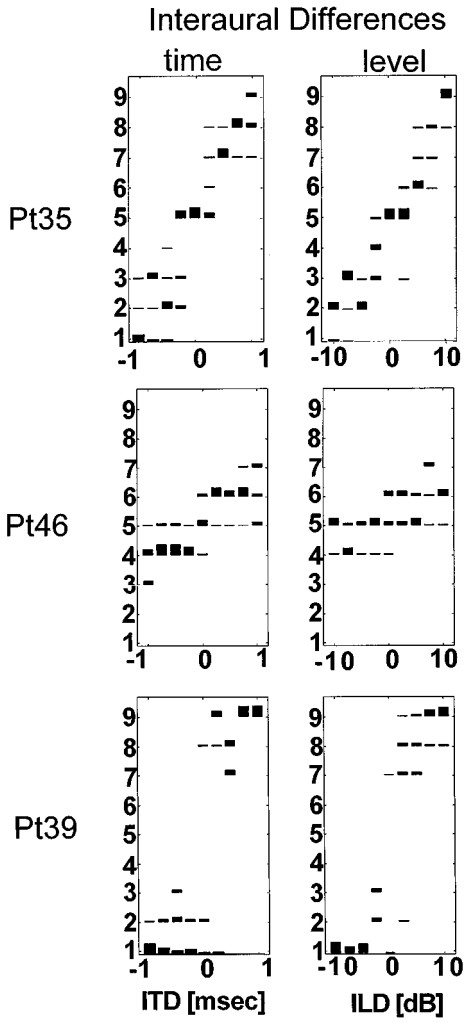


FIG. 2. Histograms of the estimated joint probability functions, pdf(POS, ITD) and pdf(POS, ILD), for clicks for patients: Pt35, Pt46, and Pt39.

TABLE II. Mean and standard deviation (s.d.) of P_{ITD} (pos) obtained by the control group for clicks.

Position:	1	2	3	4	5	6	7	8	9
Mean	0.113	0.105	0.110	0.099	0.122	0.103	0.109	0.112	0.116
s.d.	0.082	0.072	0.044	0.036	0.055	0.039	0.047	0.076	0.072

$=0.42$. Those values are all in the normal range, thus we can consider this patient's performance as normal for click lateralization.

Pt46 (Fig. 2, middle row) perceived most binaural stimuli near the center (positions 4, 5, and 6), resulting in the following regression line coefficients: $a^t=0.44$, $b^t=2.78$, $a^l=0.23$, and $b^l=4.11$ are all significantly different from those of the control group. On the other hand, the performance of Pt39 (Fig. 2, bottom row) is abnormal but in a different way than Pt46. This patient perceived the binaural stimuli only at the sides (positions 1, 2, 3 or 7, 8, 9) and never in the center (positions 4, 5, 6). The regression line slopes were $a^t=1.15$, $a^l=1.21$, and y intercepts ($b^t=-0.97$, $b^l=-1.21$) were significantly abnormal (more than 3 standard deviations from the mean). Although these slopes are well outside the normal range (Table I) from inspection of Pt39's data, it is clear that the regression line statistics do not capture the full external of the abnormality. For this reason another measure was developed to complement the regression line. In order to provide a measure of a listener's tendency to report sounds at some locations and not others, the number of times that each perceived position was selected during an experiment was counted. We then defined the probability of perceiving a given position as

$$P_{\mathrm{ITD}}(\mathrm{pos})=\sum_{\mathrm{ITD}} h(\mathrm{pos},\mathrm{ITD}), \tag{4}$$

and

$$P_{\mathrm{ILD}}(\mathrm{pos})=\sum_{\mathrm{ILD}} h(\mathrm{pos},\mathrm{ILD}), \tag{5}$$

when pos$=1,2...,9$. The performance of all subjects in the control group yields $P_{\mathrm{ILD}}(\mathrm{pos})\approx P_{\mathrm{ITD}}(\mathrm{pos})\approx 0.11$ for every stimulus and every pos$=1,...,9$. In Table II are the mean and standard deviation (s.d.) of $P_{\mathrm{ITD}}(\mathrm{pos})$ obtained for clicks from the normal control group.

We have chosen to further collapse our data for two reasons. First, with 9 positions and 2 types of interaural differences, 18 measures for each stimulus type become unwieldly. Second, if the mean ± 2.5 s.d. were used to define the normal range, then note (Table II) that for almost every position the lower boundary of the normal range is negative. Other stimuli and ILD experiments yielded similar values. Thus if a subject ignores any position [i.e., $P_{\mathrm{ITD}}(\mathrm{pos})=0$ or $P_{\mathrm{ILD}}(\mathrm{pos})=0$] the performance would still be considered normal, and therefore this measure is meaningless. Therefore, we have chosen to characterize lateralization performance by using only three principal positions: left, center, and right. In terms of the lateralization experiment, positions 1, 2, and 3 constitute ''left;'' positions 4, 5, and 6 ''center;'' and positions 7, 8, and 9 ''right.'' Thus instead of considering the probability for perceiving a given position, we now consider the probability for perceiving left, right or center. From reviewing all our data we conclude that the principle features of subject performance are preserved using this collapsing method.

Formally, we then define a lateralization index, Θ_α^β, where the superscript β stands for the interaural difference presented (t for time, l for level) and the subscript α stands for the principal position (C for the center, L for the left, and R for right). Thus

$$\Theta_L^t=\sum_{\mathrm{pos}=1}^{3} P_{\mathrm{ITD}}(\mathrm{pos}), \tag{6}$$

$$\Theta_C^t=\sum_{\mathrm{pos}=4}^{6} P_{\mathrm{ITD}}(\mathrm{pos}), \tag{7}$$

$$\Theta_R^t=\sum_{\mathrm{pos}=7}^{9} P_{\mathrm{ITD}}(\mathrm{pos}). \tag{8}$$

The control subjects displayed little or no difference between the left and right lateralization indices, none of the right–left differences reached a level of statistical significance according to Student's t test ($p<0.01$). Thus side lateralization indices, Θ_S^t and Θ_S^l, were defined the algebraic means of left and right indices for the respective interaural differences. The mean normal values and their standard deviations are given in Table I. As can be seen from this table, with this procedure for collapsing the data the mean $-2.5\cdot$s.d. is now meaningful (>0).

Now if we examine the lateralization indices for the patients of Fig. 2, we find that Pt35, whose performance we consider as normal, also had normal indices ($\Theta_L^t=0.41$, $\Theta_C^t=0.29$ and $\Theta_R^t=0.30$). For Pt39, who ignored the center positions, $\Theta_L^t=0.30$, $\Theta_C^t=0.0$, and $\Theta_R^t=0.70$, only Θ_L^t is in the normal range, while Θ_C^t is much smaller than the normal range and Θ_R^t is much higher. The low value of Θ_C^t indicates that the center positions are ignored. For Pt46, who ignored the side positions $\Theta_L^t=0.05$, $\Theta_C^t=0.89$, and $\Theta_R^t=0.06$. His Θ_C^t is much higher than normal, and his side indices were much lower than the normal range.

If we label the normal range of each lateralization index as $\underline{\Theta}_\alpha^\beta$, we can define side-oriented performance as having $\Theta_C^\beta<\underline{\Theta}_C^\beta$ and $\max\{\Theta_R^\beta,\Theta_L^\beta\}>\overline{\Theta}_S^\beta$, and analogously we can define center-oriented performance when $\Theta_C^\beta>\overline{\Theta}_C^\beta$ and $\min\{\Theta_R^\beta,\Theta_L^\beta\}<\underline{\Theta}_S^\beta$. With these definitions, Pt39's performance is side-oriented for click lateralization, and Pt46 is center oriented.

Both measures, the regression line fit and the lateralization indices, will be used to characterize lateralization performance. The regression line reveals to what extent left and right can be distinguished, whereas the lateralization indices reveal to what extent regions are ignored. Not surprisingly, the values of the regression line parameters and the lateral-

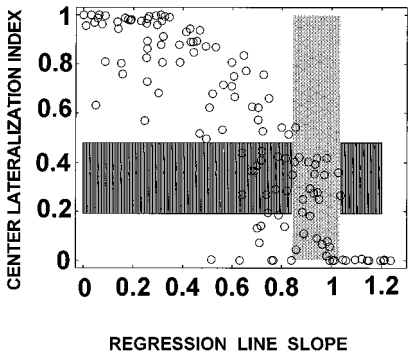


FIG. 3. Regression line slope as a function of the center lateralization index derived by all patients in the different lateralization experiment. The shadowed areas are the normal range of performance.

ization indices are highly correlated. In Fig. 3 the center lateralization index [Eq. (7)] is plotted against the corresponding regression line slope [a^t or a^l, in Eqs. (2) and (3)] for all six lateralization experiments of every patient. As can be seen in Fig. 3, there is a high correlation between the two measures ($r=0.93$). Similar correlation was found between the center lateralization index and the y intercepts of the regression line.

With this approach, a subject's performance for each stimulus can be characterized by the values of 12 parameters. Namely, the six lateralization indices (Θ^t_L Θ^t_C Θ^t_R Θ^l_L Θ^l_C and Θ^l_R) the regression line parameters (a^t, a^l, b^t, and b^l) and the jnd for ITD and ILD (δ^t and δ^l).

III. RESULTS

For the control group, the difference in the means and s.d.'s of Θ^t_C, Θ^l_C, a^t, and a^l as obtained for the different stimuli (click, LNB, and HNB) was not statistically significant ($p<0.01$ according to Student's t test). Thus the normal ranges plotted in Fig. 3 (shadowed areas) reveal the mean $\pm 2.5\cdot$s.d. obtained by the control group for all stimulus types.

Performance characterized as center oriented always had regression line slopes below the normal range. For those experiments in which all stimuli appeared to be heard very close to the center (Θ^t_C, $\Theta^l_C>0.9$), regression line slopes varied widely, ranging between 0 and about 0.4. In other words, some patients (a^t, $a^l=0$) were unable to differentiate at all the position of any of the stimuli on the basis of their interaural differences, while other patients were able to do so (a^t, $a^l>0$) appropriately to some extent. No subjects differentiated position inappropriately (a^t, $a^l<0$).

Performances characterized as side oriented had regression lines whose slopes ranged from beyond the normal range to about 0.5. In some cases the regression line slope was within the normal range. In other words (as typified by Pt39 of Fig. 2), all subjects with side-oriented performances, even if unable to hear stimuli as coming from the center, lateralized the stimuli with interaural differences to the correct side.

For the three stimulus types, means and standard deviations (s.d.) were calculated for each psychophysical parameter from the control group's data, and they are presented in Table I. The jnd values were similar to those reported previously (Levine, 1994; Furst and Algom, 1995).

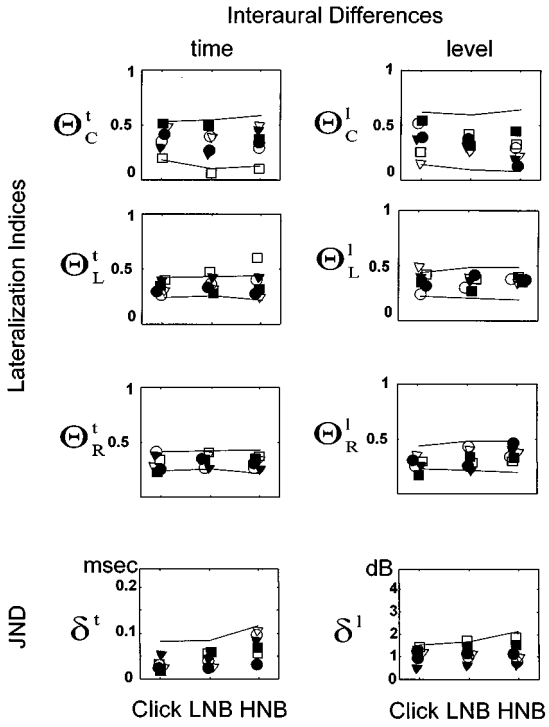


FIG. 4. Lateralization and jnd performances for patients in group A. Stroke patients: Pt23—open circle; Pt25—open triangle; Pt36—open square. MS patients: Pt35—filled triangle; Pt50—filled circle; Pt52—filled square.

Patients were grouped according to their click ITD lateralization. Each group performance is described in a separate figure (see Figs. 4, 6, 8, and 9). For each patient his/her lateralization indices and jnds can be found corresponding to a unique symbol (open=MS, filled=stroke) in one of these four figures. Each figure consists of eight panels, distributed in four rows and two columns. The lateralization indices for center, right, and left are in the top three rows, respectively, and the jnd's, δ^t and δ^l are in the bottom row. ITDs are in the left column and ILDs are in the right column. The three stimulus types, Click, LNB, and HNB form the horizontal axis of each panel. The normal range (mean $\pm 2.5\cdot$s.d. of the control group), is marked by solid lines on each panel. Only the upper limits are indicated for the jnd's. Because the regression line data and the lateralization indices are highly correlated (Fig. 3), the regression line coefficients are not included in these figures, but will be discussed.

Patients were always able to perform the lateralization procedure, but some were not always able to perform the jnd task for some stimuli (for the largest interaural differences our system allows). Thus if a symbol is missing from a panel, it means that the patient was unable to perform that particular jnd.

A. Patients with normal click lateralization

Three patients with MS and three with stroke had normal click lateralization indices and are shown in Fig. 4. Five of these six patients had normal jnd and lateralization indices

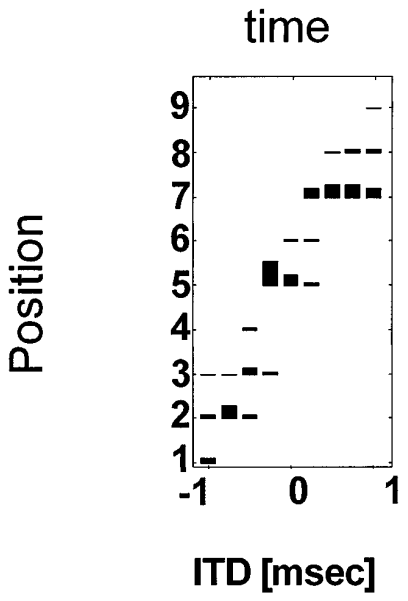


FIG. 5. Histogram of the estimated joint probability function pdf(POS, ILD) for LNB for patient Pt25.

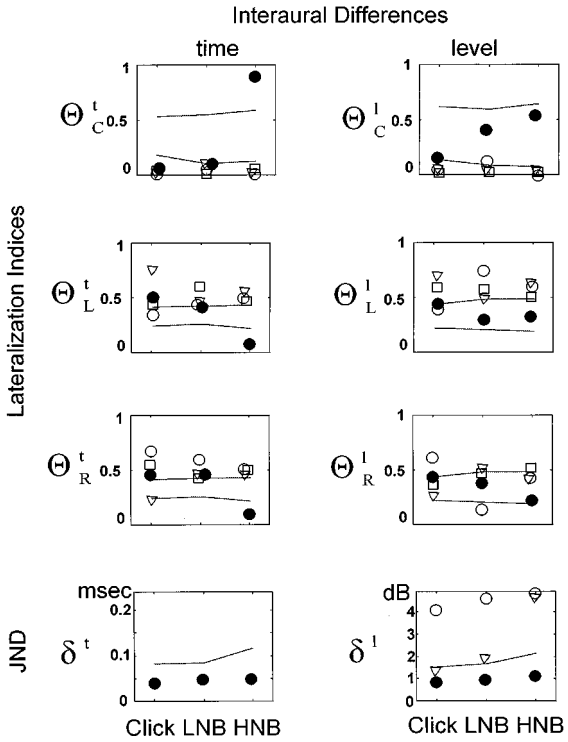


FIG. 6. Lateralization and jnd performances for patients in group B. Stroke patients: Pt39—open circle; Pt43—open triangle; Pt44—open square. An MS patient: Pt40—filled circle.

for all other stimuli as well. Only stroke patient Pt36 (open square) had any abnormalities. His ITD lateralization was left oriented for both LNB and HNB (but was normal for ILDs). The center ITD indices were very low for both types of narrow-band noise ($\Theta_C^t = 0.02$ for LNB, and 0.07 for HNB). The left indices were abnormally high and the right indices were within the normal range. All his regression lines were normal.

The regression lines were not uniformly normal for these subjects. In about a third of the conditions (11 out of 36) the regression-line slopes were too shallow, i.e., below the normal range (between 0.6 and 0.8). The abnormality was never because the slope was too steep. Figure 5 is an example of lateralization performance with an abnormally shallow regression line. Shown is Pt25's lateralization for LNB with ILDs. The regression line slope is 0.72, which reflects in part the fact that the most extreme positions (1 and 9) were infrequently selected.

B. Patients with side-oriented click lateralization

Four patients, three with strokes and one with MS, had side-oriented performance for clicks (Fig. 6). All three stroke patients lateralized symmetrically for click, ITD, and ILD. They all had a very low Θ_C^t and Θ_C^l and at least one abnormally high side index for both time and level differences. The side indices clearly show asymmetry between left and right for these three patients. Each one of them had a very high left or right index for time and level, while the other side's two indices were almost normal. Insight into this asymmetry can be obtained by referring back to Fig. 2 and examining the actual distribution of responses for Pt39 (open circle in Fig. 6). It is obvious that the patients did not confuse left stimuli with right, but perceived most of the center stimuli to the left and a few right stimuli to the left.

In this group an MS patient, Pt40, indicated by filled circles in Fig. 6, presented a different pattern of behavior. The left and right side indices of this patient were symmetrical for both time and level tests.

Like the data for clicks, the performance for NB noise of all three stroke patients was side oriented for ITD and ILD. However, Pt43 (open triangle in Fig. 6) had similar left and right indices for LNB stimuli with both ITD and ILD, while his HNB lateralization was biased to the left as it was for clicks. Pt44 (open square in Fig. 6) displayed symmetry between left and right for HNB and LNB with both ITD and ILD. Pt39 (open circle) had a remarkable preference for the right side for clicks (for both ITD and ILD), but a preference for the left for both LNB and HNB ILDs. On the other hand, for both LNB and HNB this subject's ITD performance was abnormal, but similar to left and right sides.

For the one MS patient, Pt40, the NB tests were normal for ILD, but yielded highly irregular performance for ITD; for LNB, performance had a tendency toward side orientation while for HNB stimuli performance was strongly center oriented: high Θ_C^t and low Θ_L^t and Θ_R^t indices. With both narrow-band stimuli as with clicks, Pt40 perceived stimuli at the left and right positions symmetrically, for both interaural differences.

As shown in Fig. 3, most side-oriented subjects had abnormal high regression line slopes (see also Pt39 of Fig. 2). However, in some of the performances categorized as side oriented, the slope of the regression line fit was either within or even below normal limits. The lateralization performance of Pt43 for click ITDs is such an example (Fig. 7). The regression line yielded a slope of 0.76 (below the normal

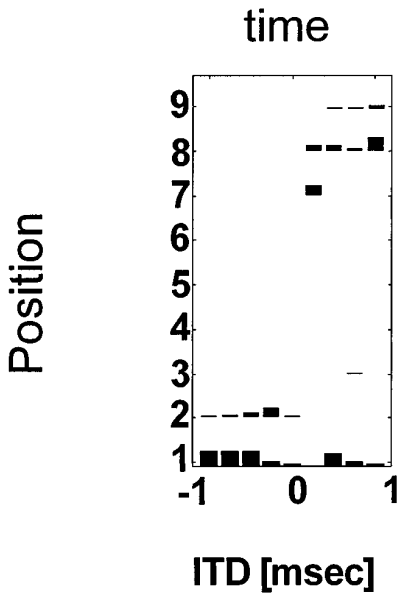


FIG. 7. Histogram of the estimated joint probability function pdf(POS, ITD) for clicks for patient Pt43.

range). This low value was due to the fact that the patient selected left of center positions for some of the right ear leading ITDs. Nonetheless, it is clear that generally Pt43 made correct right–left distinctions.

The single MS patient in this group, Pt40, had normal jnd's for all stimuli. But all the three stroke patients were unable to perform time jnd, and Pt44 (open square) could not perform level jnd as well. Another stroke patient (Pt39, open circle) had abnormally high δ^l values for all stimuli. Pt43 (open triangle) had normal level jnd for click and LNB, and abnormally high jnd for HNB.

C. Patients with center-oriented click lateralization

Center-oriented performance for click lateralization occurred in the other 12 patients. Seven were center oriented for both ITD and ILD click lateralization, three were center oriented for ITD click lateralization only, one was clearly center oriented for ILD clicks but borderline for ITD clicks, and one was borderline for both ITD and ILD.

In Fig. 8 are the data from the seven patients who had similar abnormalities in ITD and ILD click lateralization. Five were stroke patients and two MS patients. One stroke patient (Pt32, open hexagon) never perceived the stimulus as coming from the right ($\Theta^t_R = 0.0$) but did have a normal Θ^t_L. All other patients had abnormally small side lateralization indices, but similar ones on both sides.

All but one of the patients of Fig. 8 had similar results for NB lateralization as for click lateralization. The one exception was a stroke patient Pt29 (open triangle), whose HNB lateralization was center oriented, but his LNB lateralization was not. For LNB, his Θ^t_C was normal, Θ^t_L was low and Θ^t_R was high. For ILD, all his LNB lateralization indices were normal.

Four of the stroke patients and one MS patient whose performances are shown in Fig. 8 had normal or near normal jnd's. Two of the patients presented in Fig. 8, one stroke

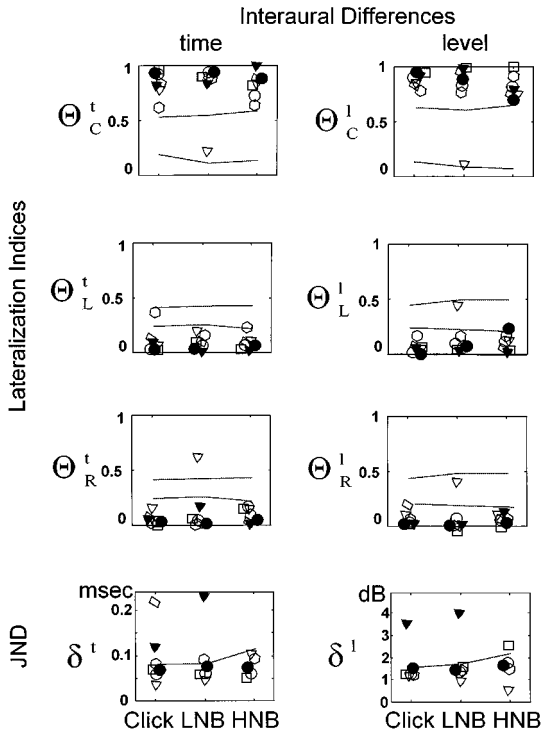


FIG. 8. Lateralization and jnd performances for patients in group C. Stroke patients: Pt29—open triangle; Pt30—open rectangle; Pt32—open hexagon; Pt37—open circle; Pt47—open square. MS patients: Pt46—filled triangle; Pt49—filled circle.

(open rectangle) and one MS (solid triangle), had very high jnd's for both ITD and ILD. In fact, the MS patient could not perform any of the HNB jnd's and the stroke patient was unable to perform any of the jnd's, with the exception of ITD clicks (for which his jnd was highly abnormal).

In Fig. 9 are the data from another five patients whose performance is also classified as center oriented for ITD clicks. All had normal or near normal time and level jnd's, and all were center oriented for HNB ITD. Otherwise they were a heterogeneous group. Three (two stroke patients and one MS patient) performed abnormally for click time lateralization, but normally for click level lateralization. Two of these three were also center oriented for LNB ITDs.

Another two MS patients had near normal performance for click ITD lateralization, but had center-oriented performance for click ILD lateralization. MS patient Pt38 (filled square) was at most slightly center oriented for LNB time lateralization. This patient's ILD lateralization revealed a center orientation in LNB stimuli, but normal HNB performance. The lateralization of MS patient Pt42 (filled circle) was actually side oriented for LNB with ITD, normal for LNB with ILD and center oriented for both HNBs.

Of the three who had normal lateralization for ILD clicks, the lateralization of stroke patient Pt54 (open circle) was highly center oriented for both LNB and HNB with ILD as well as ITD. The MS patient, Pt48 (filled triangle), had high center indices for all ITD but normal center indices for all ILD. The only right–left asymmetry in his lateralization indices occurred for LNB with ITD. For this one stimulus condition his left index was normal, while his right index

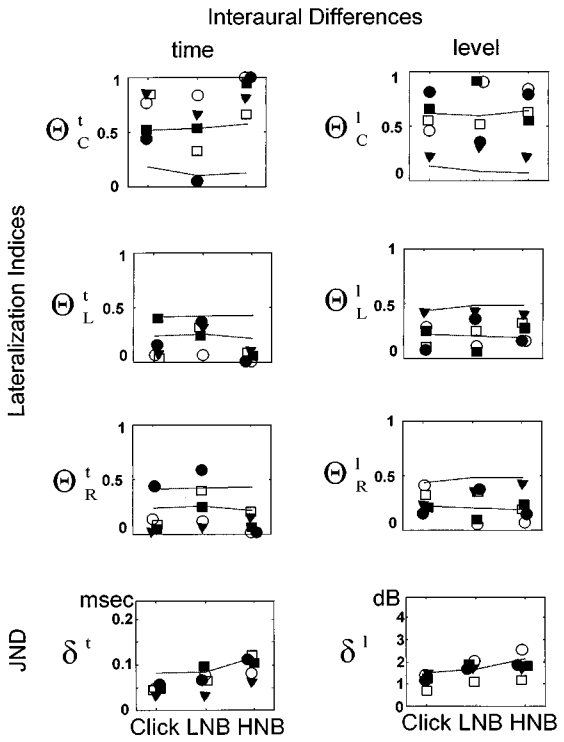


FIG. 9. Lateralization and jnd performances for patients in group C. Stroke patients: Pt31—open square; Pt54—open circle; MS patients: Pt38—filled square; Pt42—filled circle; Pt48—filled triangle.

was very low. Aside from click and HNB ITD, stroke patient Pt31 (open square) had normal lateralization.

IV. DISCUSSION

Many new features were observed in the present study as compared to previous reports on the psychophysical performance of patients with pontine lesions (Furst *et al.*, 1992; Levine *et al.*, 1993a, b; Furst *et al.*, 1995). While previous studies have dealt only with MS patients, the present study includes stroke patients whose ischemic lesions, unlike MS, are pathologically similar to the destructive lesions that have been studied in experimental animals (Masterton *et al.*, 1992).

In the present report we have also studied nine MS patients, which is comparable to the seven MS patients of our previous study (Furst *et al.*, 1995). In Table III patients from both studies are combined together for a total of 29 patients. Subjects are grouped together according to their underlying disease (MS or stroke) and their lateralization performance for clicks. It is clear from Table III that similar center- and

TABLE III. Number of MS and stroke patients according to their lateralization performance. The MS patients group includes also those who were tested in our previous study (Furst *et al.*, 1995). The number of patients in the previous study is indicated in parenthesis.

Lateralization performance:	Normal	Center oriented	Side oriented
MS	5 (2)	7 (2)	4 (3)
Stroke	3	7	3

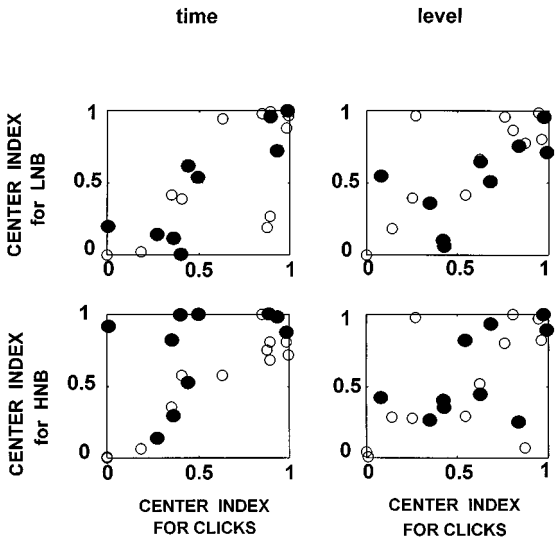


FIG. 10. Center lateralization index for narrow-band noise bursts (LNB and HNB) as a function of the center lateralization index for clicks, of all patients. MS patients are indicated by filled circles and stroke patients by open circles.

side-oriented behaviors were detected for both stroke and MS patients, despite their very different pathologies.

If the lesions responsible for these results were shown to be in the same location then this would imply that abnormal lateralization performance is independent of pathology. A second possible scenario is that there is an interaction between lesion type and lesion location; the type of abnormal lateralization may depend upon both factors. The reason only two types of abnormal lateralization has emerged from our studies is because a disturbance of the pontine binaural processing system can only manifest itself (for our lateralization paradigm) in these two basic abnormal patterns. Whatever the case, these two patterns of abnormal lateralization performance are revealing something fundamental about the neurology of the binaural system. Future studies will further clarify the nature of this processing.

As an elaboration upon previous lateralization procedures (Furst *et al.*, 1992, 1995), where only click stimuli were used, this study also included narrow-band noises, at high and low frequencies. Comparison with the results for click lateralization tests indicated the same abnormal performance types reported by Furst *et al.* (1995). When performance was abnormal, interaural differences of either type (ITD or ILD) were for the most part perceived either to the sides or only at the center of head. Of the parameters we analyzed the center lateralization index was the parameter that best characterized the lateralization abnormality. Center-oriented performance yields a high value of center lateralization index (approaching 1), while side oriented yielded a very low value (close to 0).

Comparison between patients' performance in the lateralization experiment for clicks and their performance for narrow-band noise bursts is shown in Fig. 10. The center lateralization index obtained by each patient for HNB and LNB is plotted as a function of his/her center lateralization index for clicks. For both ITD and ILD stimuli there is a very

high correlation between the performance for clicks and both NB stimuli. For HNB $r=0.70$ for ITD and 0.77 for ILD. Similarly, the correlation between the response to clicks and LNB yielded $r=0.86$ for ITD, and $r=0.70$ for ILD.

Not surprisingly, this correlation is not perfect. For example, as can be seen in Fig. 10, of the subjects that were highly center oriented (very high center indices) for ITD HNB, three had normal ITD click lateralization and one was highly side oriented for ITD click lateralization (very low center index). Similar examples of ''out-liers'' can be found for the three other plots of Fig. 10. The fact that lateralization performance can be so different in some subjects for these two different stimuli implies that lateralization not only depends upon the location of lesions, but also on the characteristics of the stimulus. Because a click is a broadband stimulus, more spectral cues are available for judging lateralization of a click than for either of the narrow-band noise bursts. On the other hand temporally the noise bursts are of much longer duration than the click. Since in general psychophysical performance degrades with shortening of stimulus duration (Tobias and Zerlin, 1959), all the other things being equal, psychophysical performance for the noise bursts should have an advantage over psychophysical performance for clicks.

Previous studies of MS subjects (e.g., Levine *et al.*, 1993a) measured jnds for high- and low-frequency noise and found that tasks with high-frequency stimuli were more sensitive detectors of abnormality than were tasks with low-frequency stimuli. Our results are consistent with this observation, since from review of Figs. 4, 6, 8, and 9 it is clear that the HNB jnd for ITD were always abnormal whenever any other jnd was abnormal. On the other hand, our results show that jnd can be normal while lateralization is distinctly abnormal (Fig. 8). Furthermore our results also indicate that, if any jnd or lateralization task is abnormal, then lateralization for HNB with ITD will always be abnormal. We can extend then the original observation of Levine *et al.* and state that lateralization tasks with high-frequency stimuli are more sensitive detectors of abnormality than (1) jnd for any kind of stimulus and (2) lateralization tasks with low-frequency stimuli or clicks.

The similarity between ITD and ILD abnormalities in lateralization, previously reported by Furst *et al.* (1995) and Furst and Algom (1995) for MS patients, was further explored in the present study. In most cases, for both MS and stroke patients, similar ITD and ILD behaviors were found for all types of performance, for all stimulus types, and for both types of patients, stroke and MS. These results make it clear that binaural processing in the brainstem involves both interaural time and level analysis, perhaps using the same mechanisms and depending on the same anatomical structures, and ITD tasks are more sensitive to disruption.

Furst and Algom (1995) tested both lateralization and discrimination tasks during the same session, and showed that in most cases abnormal lateralization coincided with abnormal jnds with but two exceptions. In the present study, we have found many subjects with normal jnd but markedly abnormal lateralization. In Fig. 11, time jnds are plotted against center lateralization index for the same stimulus (for all patients and all stimuli). The normal ranges for both jnd and the center lateralization index are shaded. For patients who could not perform the jnd task, their jnd was arbitrarily assigned the value 1. It is clear from this data that patients who performed normally in jnd tests varied in their lateralization performance, presenting normal, center-oriented, or side-oriented click lateralization. On the other hand, patients who had abnormal jnds for ITD and/or ILD had extremely abnormal lateralization performance; their center lateralization index could be zero (side oriented), or close to one (center oriented). As in Levine *et al.* (1993a), in most patients abnormality in the jnd task was for high frequencies.

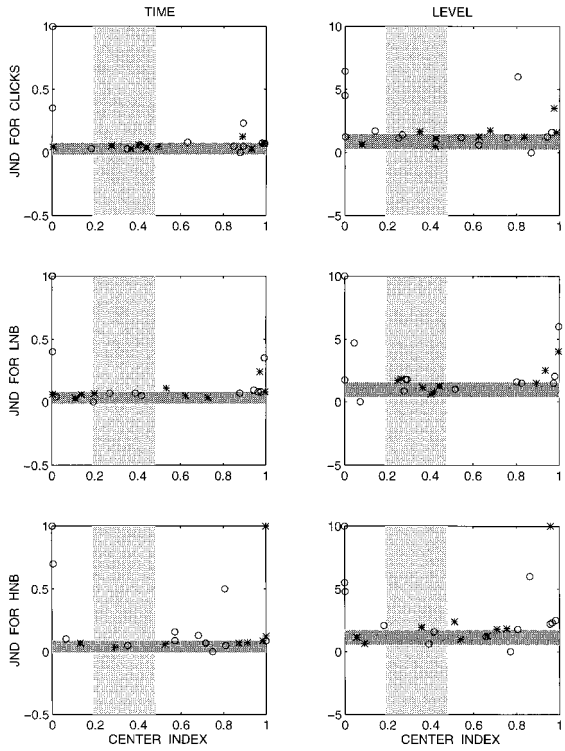


FIG. 11. Time and level jnds as a function of the center lateralization index, for clicks, LNB and HNB stimuli, of all patients. MS patients are indicated by crosses and stroke patients by open circles. The shadowed areas indicate the normal range.

The fact that some patients could have normal jnd and poor lateralization but not vice versa suggests that while performing jnd tasks, subjects use many different cues. The lateralized position, as we have measured it, is only one of the possible cues. It would at first glance appear to be the most natural one and it might be sufficient; however the results of the present study make it clear that lateralized position is not a necessary cue for discrimination (jnd).

For side-oriented lateralization, a subject generally can make right–left distinctions correctly (regression line slope >0.5, see Fig. 3). Hence this right–left cue is probably available for the jnd task, and could explain normal jnd performance when lateralization is side oriented. However, this clue is not available for many subjects with center-oriented lateralization since their regression line slope is zero. Thus the fact that such subjects with regression line slopes can have normal jnd makes it clear that lateralized position is not a necessary cue for interaural discrimination (jnd).

One possible explanation is that the brainstem binaural processing analyzes right–left position in two ways; (a) on a ''gross'' basis that is responsible for our large scale right–left analysis, such as involved in our lateralization task; (b) on a ''micro'' basis that is responsible for detecting tiny right–left changes such as in a jnd task. Furthermore it could be that the ''gross'' right–left detector operates well throughout auditory space, whereas, the ''micro'' right–left detector operates optimally near center space where jnds are smallest.

The results seem to hint that several independent mechanisms participate in the process of binaural time and level perception. The brainstem auditory pathway, where most of this processing probably takes place, contains several levels at which an estimate for either ITD or ILD could occur. Future work will examine the relationship between (a) lesion location (from MRI scan data) and (b) both abnormalities in binaural psychophysical performance and abnormalities in electrophysiology (from binaural auditory evoked potentials). With such an approach it might be possible to develop a qualitative and *quantitative* model for the structure and function of the nuclei and tracts of the brainstem that take part in binaural processing.

ACKNOWLEDGMENTS

We thank all the subjects who participated in this study, which was supported by US–Israel Binational Science Foundation Grant No. 93-00324 and by the Adams Super-Center for Brain Studies, Tel Aviv University, Israel.

Algom, D., Adam, R., and Cohen-Raz, L. (**1988**). ''Binaural summation and lateralization of transients: a combined analysis,'' J. Acoust. Soc. Am. **84**, 1302–1315.

Babkoff, H., Ben-David, N., and Muchnik, C. (**1996**). ''Change in the lateralization of clicks due to aging,'' J. Basic Clin. Physiol. Pharmacol. **7**, 295.

Domnitz, R. (**1973**). ''The interaural time jnd as a simultaneous function of interaural time and interaural amplitude,'' J. Acoust. Soc. Am. **53**, 1549–1552.

Domnitz, R., and Colburn, H. (**1977**). ''Lateral position and interaural discrimination,'' J. Acoust. Soc. Am. **61**, 1586–1598.

Durlach, N. I., and Colburn, H. S. (**1992**). ''Binaural phenomena,'' in *Handbook of Perception: Hearing*, Vol. IV, edited by E. C. Carterette and M. P. Friedman (Academic, New York), pp. 365–466.

Furst, M., and Algom, D. (**1995**). ''Lateralization and discrimination of dichotic clicks: evidence from patients with brainstem lesions and normal cohorts,'' J. Basic Clin. Physiol. Pharmacol. **6**, 149–171.

Furst, M., Eyal, S., and Korczyn, A. D. (**1990**). ''Prediction of binaural click lateralization by brainstem auditory evoked potentials,'' Hearing Res. **49**, 347–360.

Furst, M., Algom, D., Mohaban, L., and Korczyn, A. D. (**1992**). ''Binaural click lateralization: Evidence from psychophysics and brainstem auditory evoked Potentials in normal subjects and multiple sclerosis patients,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, New York), pp. 635–642.

Furst, M., Levine, R. A., Korczyn, A. D., Fullerton, B. C., Tadmor, R., and Algom, D. (**1995**). ''Brainstem lesions and click lateralization in patients with multiple sclerosis,'' Hearing Res. **82**, 109–124.

Hausler, R., and Levine, R. A. (**1980**). ''Brainstem auditory evoked potentials are related to interaural time discrimination in patients with multiple sclerosis,'' Brain Res. **191**, 589–594.

Hausler, R., Colburn, S., and Marr, E. (**1983**). ''Sound localization in subjects with impaired hearing,'' Acta Oto-Laryngol. Suppl. **400**, 1–62.

Hendler, T. (**1985**). ''Auditory evoked potentials in multiple sclerosis: audiological and psychophysical correlates,'' Ph.D. thesis, State University of New York at Stony Brook.

Hendler, T., Squires, N. K. *et al.* (**1990**). ''Psychophysical measures of central auditory dysfunction in multiple sclerosis, neurophysiological and neuroanatomical correlates,'' Ear Hear. **11**, 403–416.

Jenkins, W. A., and Masterton, R. B. (**1982**). ''Sound localization: Effects of unilateral lesions in central auditory system,'' J. Neurophysiol. **47**, 987–1060.

Levine, R. A., Gardner, J. C., Fullerton, B. C., Stufflebeam, S. M., Furst, M., and Rosen, B. R. (**1994**). ''Multiple sclerosis lesions of the auditory pons are not silent,'' Brain **117**, 1127–1141.

Levine, R. A., Gardner, J. C., Stufflebeam, S. M., Fullerton, B. C., Carlisle, W. E., Furst, M., Rosen, R. B., and Kiang, N. Y. S. (**1993a**). ''Binaural auditory processing in multiple sclerosis subjects,'' Hearing Res. **68**, 59–72.

Levine, R. A., Gardner, J. C., Stufflebeam, S. M., Fullerton, B. C., Carlisle, W. E., Furst, M., Rosen, R. B., and Kiang, N. Y. S. (**1993b**). ''Effects of multiple sclerosis brainstem lesions on sound lateralization and brainstem auditory evoked potentials,'' Hearing Res. **68**, 73–88.

Levitt, H. (**1971**). ''Transformed up–down method in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Masterton, R. B., Granger, E. M., and Glendenning, K. K. (**1992**). ''Psychoacoustical contribution of each lateral lemniscus,'' Hearing Res. **63**, 57–70.

Tobias, J. V., and Zerlin, S. (**1959**). ''Lateralization threshold as a function of stimulus duration,'' J. Acoust. Soc. Am. **31**, 1591–1594.

Wightman, F. L., Kistler, D. J., and Perkins, M. E. (**19xx**). ''A new approach to the study of human sound localization,'' in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer-Verlag, New York), pp. 26–48.

Spectro-temporal factors in two-dimensional human sound localization

Paul M. Hofman[a] and A. John Van Opstal
University of Nijmegen, Department of Medical Physics and Biophysics, Geert Grooteplein 21, 6525 EZ Nijmegen, The Netherlands

(Received 6 June 1997; revised 5 December 1997; accepted 29 January 1998)

This paper describes the effect of spectro-temporal factors on human sound localization performance in two dimensions (2D). Subjects responded with saccadic eye movements to acoustic stimuli presented in the frontal hemisphere. Both the horizontal (azimuth) and vertical (elevation) stimulus location were varied randomly. Three types of stimuli were used, having different spectro-temporal patterns, but identically shaped broadband averaged power spectra: noise bursts, frequency-modulated tones, and trains of short noise bursts. In all subjects, the elevation components of the saccadic responses varied systematically with the different temporal parameters, whereas the azimuth response components remained equally accurate for all stimulus conditions. The data show that the auditory system does not calculate a final elevation estimate from a long-term (order 100 ms) integration of sensory input. Instead, the results suggest that the auditory system may apply a ''multiple-look'' strategy in which the final estimate is calculated from consecutive short-term (order few ms) estimates. These findings are incorporated in a conceptual model that accounts for the data and proposes a scheme for the temporal processing of spectral sensory information into a dynamic estimate of sound elevation. © *1998 Acoustical Society of America.* [S0001-4966(98)03005-7]

PACS numbers: 43.66.Qp, 43.66.Ba, 43.66.Mk [RHD]

INTRODUCTION

Human auditory localization depends on implicit binaural and monaural acoustic cues. Binaural cues arise through interaural differences in sound level (ILD) and timing (ITD) that relate in a simple way to the horizontal component of sound direction relative to the head. The auditory system, however, cannot distinguish, from these cues, between all source positions with the same horizontal component that lie on the so-called ''cone of confusion.'' Due to the front–back symmetry of the ITD and ILD cues, a sound azimuth (α) estimate provided by these binaural cues is therefore ambiguous (e.g., Blauert, 1996; Wightman and Kistler, 1992).

Monaural cues consist of direction-dependent linear spectral filtering caused by the torso, head, and pinnae. Incident waveforms are reflected and diffracted in a complex and direction-dependent way, which typically gives rise to strong enhancement and attenuation at particular frequency bands (Hebrank and Wright, 1974; Shaw, 1974; Mehrgardt and Mellert, 1977; Middlebrooks *et al.*, 1989; Lopez-Poveda and Meddis, 1996). These cues are essential for elevation (ϵ) localization and in resolving front–back ambiguities (e.g., Batteau, 1967; Musicant and Butler, 1984; Blauert, 1996).

The relation between the pinna geometry and the direction-dependent filtering has been shown in both experimental and theoretical studies (Batteau, 1967; Teranishi and Shaw, 1968; Han, 1994; Lopez-Poveda and Meddis, 1996). The actual importance of the filtering in two-dimensional localization has been underlined by various behavioral experiments. In some of these studies, the geometrical structure of the pinna was altered artificially, resulting in a degraded localization performance (Gardner and Gardner, 1973; Oldfield and Parker, 1984b). In other studies, manipulation of the sound spectra and the resulting effects on elevation localization could be related to spectral features in the pinna filters (e.g., Middlebrooks, 1992; Butler and Musicant, 1993).

The features in the spectral pinna filter, or head-related transfer function (HRTF), are believed to carry the information about sound location (Kistler and Wightman, 1992). Yet, a fundamental problem arises in the processing of these features into estimates of sound location (Zakarouskas and Cynader, 1993; Hofman and Van Opstal, 1997). At the level of the eardrums, the available spectrum, $y(\omega;\mathbf{r}_S)$, associated with the source position, $\mathbf{r}_S=(\alpha_S,\epsilon_S)$, results from the source spectrum, $x(\omega)$, filtered by the particular direction-dependent HRTF, $h(\omega;\mathbf{r}_S)$:

$$y(\omega;\mathbf{r}_S)=h(\omega;\mathbf{r}_S)\cdot x(\omega) \tag{1}$$

with ω the frequency in octaves. In principle, the auditory system has no knowledge about the relative contributions of neither $h(\omega;\mathbf{r}_S)$ nor $x(\omega)$. Yet, it is necessary to minimize the source influences in order to recognize $h(\omega;\mathbf{r}_S)$ and thus determine the position, $\mathbf{r}_S$ (mainly elevation and front–back angle, as left–right angle is determined predominantly from binaural cues).

So far, only a few computational models have been proposed, which suggest how this could be done. A possible solution to this problem would be to make a priori assumptions about the sound source spectrum. It would enable the auditory system to directly compare the incoming spectrum to the (stored) HRTFs (e.g., Neti *et al.*, 1992; Zakarouskas

[a]Electronic mail: paul@mbfys.kun.nl

and Cynader, 1993; Middlebrooks, 1992). Neti *et al.* (1992) showed that a feed-forward neural network could be trained to accurately extract the location of broadband noise on the basis of the spectral filter properties of the cat's pinna. Since the model was only trained with white noise, however, it is not clear how it withstands spectral variations.

Zakarouskas and Cynader (1993) proposed a model in which the comparison between the cochlear spectrum and the HRTFs is based on spectral derivatives of first and second order (i.e., $dy/d\omega$, $d^2y/d\omega^2$). For a sound source spectrum that is locally constant, or has a locally constant slope, the model was shown to recognize essential spectral features of the underlying HRTF (specifically relevant peaks and notches), and thus correctly extract the direction of the sound.

In an alternative model, proposed by Middlebrooks (1992), the sensory spectrum is compared with the HRTFs by computing a spectral correlation coefficient. This scheme suggested that localization is accurate if the source spectrum is broadband and sufficiently flat, such that the sensory input maintains maximal correlation with the underlying HRTF of the associated source position.

In the latter two models, accurate localization relies on specific spectral constraints on the source spectrum. This assumption is supported by recent experimental results. On one hand, the auditory system appears to tolerate random variations within the broadband sound spectrum (e.g., Wightman and Kistler, 1989). On the other hand, if relative variations in the source spectrum become too large, localization can be disturbed dramatically (e.g., Middlebrooks, 1992). It therefore remains unclear what the actual spectral constraints are.

So far, the majority of localization studies have applied stimuli with stationary spectral properties. Yet, natural sounds possess a high degree of nonstationarity. Although it is commonly accepted that spectral shape cues play an essential role in elevation detection, it is as yet unknown how the auditory system applies the spectral analysis to nonstationary sensory information. In addition, due to the fundamental relation between the temporal and spectral domains, a sufficiently high spectral resolution requires a minimal time window over which the spectral estimation is integrated: $\Delta f \cdot \Delta T = \text{constant}$. The temporal and spectral resolutions needed for adequate sound localization, however, are not well known.

One possibility is that the sensory information is integrated over a time scale of order, say, 100 ms to obtain an average spectrum on which a spectral (shape) analysis can be applied. If true, sounds with the same average power spectrum on that time scale would be localized equally well: It would allow a considerable amount of freedom for the phase spectrum on that time scale.

An alternative possibility could be that the auditory system applies a ''multiple-look'' strategy, in which elevation estimation is based on multiple, consecutive short-term (say, of order a few ms) spectral analyses of the ongoing sensory information. This latter scheme would imply that the spectral-temporal behavior of the stimulus on that short time scale is also important.

In the present study, we specifically focused on these spectro-temporal aspects for sound localization in two dimensions. To our knowledge, such data are not available in the current literature. In two experiments, the sensitivity of the localization process to short-term spectro-temporal variations was of interest. Localization to frequency-modulated tones (''FM sweeps'') and trains of short-duration broadband bursts was measured. These stimuli had similar broadband average power spectra, but fundamentally different spectro-temporal behaviors. In a third experiment, we estimated the minimal time needed by the localization process to complete elevation estimation. To that means, localization to broadband noise of various durations (3–80 ms) was measured. In a recent pilot study by Frens and Van Opstal (1995), it was shown that localization performance systematically deteriorates in elevation, but not in azimuth, when stimulus duration of broadband noise bursts is shorter than 10 ms.

Saccadic eye movements were used to quantify the response accuracy. It enabled accurate measurements (within 1 deg) of a very early spatial percept (below 200 ms) for stimuli presented within the oculomotor range (35-deg eccentricity range in all directions). Subjects were tested under entirely open-loop conditions, i.e., neither acoustic, nor visual or verbal feedback was provided.

The results of our experiments show a consistent and systematic influence of spectro-temporal factors of the stimulus on the elevation component of the localization response in all subjects tested. These findings will be discussed in terms of a conceptual spectro-temporal model of human auditory localization in two dimensions.

I. METHODS

A. Subjects

Seven male subjects participated in the localization experiments. Their ages ranged from 23 to 39 years. Subjects were employees and students of the department. Three of the subjects (PH, JO, and JG) were experienced in sound localization experiments, whereas the other subjects had no such previous experience and were kept naive as to the purpose of this investigation. Inexperienced subjects were given one or two practice sessions to get acquainted with the setup and localization paradigm and to gain stable performance. All subjects reported to have no hearing problems of any kind.

B. Apparatus

Experiments were conducted in a completely dark and sound-attenuated room with dimensions $3\times3\times3$ m. The walls, floor, and ceiling were covered with acoustic foam, that effectively absorbed reflections above 500 Hz. The room had an A-weighted ambient background noise level of 30 dB SPL.

The orientation of the subject's right eye was measured with the scleral search coil technique (Collewijn *et al.*, 1975; Frens and Van Opstal, 1995). The oscillating magnetic fields that are needed by this method were generated by two orthogonal pairs of square 3×3 m coils, attached to the room's edges: one pair of coils on the left and right walls generated a horizontal magnetic field (30 kHz) and a second pair on the ceiling and floor created a vertical magnetic field (40 kHz).

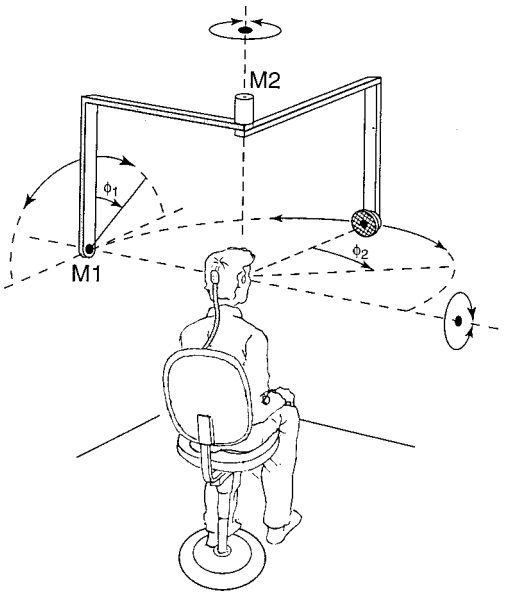


FIG. 1. Experimental setup for delivering acoustic stimuli at various spatial locations. Two stepping motors, M1 and M2, independently control the rotation angles ϕ_1 and ϕ_2, respectively. This construction ensures a fixed distance from the speaker to the center of the subject's eyes (0.9 m) for any stimulus direction (ϕ_1, ϕ_2).

An acoustically transparent frontal hemisphere (consisting of a thin wire framework, covered with black silk cloth) with 85 red light-emitting diodes (LEDs) was used for calibration of the eye-coil measurements and for providing a fixation light at the start of each localization trial. LED coordinates are defined in a two-dimensional polar coordinate system with the origin at the straight-ahead gaze direction. Target eccentricity, $R \in [0,35]$ deg, is measured as the gaze angle with respect to the straight-ahead fixation position, whereas target direction, $\phi \in [0,360]$ deg, is measured in relation to the horizontal meridian. For example, $R=0$ (for any ϕ) corresponds to ''straight ahead,'' and $\phi=0$, 90, 180, and 270 deg (for $R>0$) correspond to ''right,'' ''up,'' ''left,'' and ''down'' positions, respectively. LEDs were mounted at a distance of 85 cm from the subject's eye, at directions $\phi=0,30,60,...,330$ deg and at eccentricities $R=0,2,5,9,14,20,27,35$ deg.

Sound stimuli were delivered through a broad-range lightweight speaker (Philips AD-44725) mounted on a two-link robot (see Fig. 1). The robot consisted of a base with two nested L-shaped arms, each arm driven by a separate stepping motor (Berger Lahr VRDM5) with an angular resolution of 0.4 deg. It enabled rapid (within 2 s) and accurate positioning of the speaker at practically any point on a frontal hemisphere with a radius of 90 cm, the center of which was aligned with the LED hemisphere's center. To prevent spurious echoes, the robot was entirely coated with acoustic foam.

The robot's stepping engines produced some sound while moving which, at first sight, might be suspected to provide additional cues with regard to the speaker position. However, the first engine (M1) always remained in place, at the left of the subject. The sound of the second engine (M2), that moved in the midsaggital plane above the subject, appeared to provide no localizable stimulus cues. This was experimentally verified with two subjects in a previous study (Frens and van Opstal, 1995).

A second potential source for response biases was the speaker displacement between consecutive trials. Especially if a new stimulus position is close to the position of the previous trial, the subject might conclude a small displacement of the speaker from the short duration of motor movements. This potential problem was effectively resolved by incorporating a random movement with a minimal displacement of 20° for each engine and prior to each trial.

Two PCs controlled the experiment: a PC-386 and PC-486 that acted as a master and slave computer, respectively. The master PC was equipped with hardware for data acquisition, stimulus timing and control of the LED hemisphere: eye position signals were sampled with an AD-board (Metrabyte DAS16) at a sampling rate of 500 Hz, stimulus timing was controlled by a digital IO board (Data Translation DT2817) and the LEDs were controlled through a second digital IO board (Philips I2C).

The slave PC controlled the robot and generated the auditory stimuli. It received commands from the master PC through its parallel port. Stimulus generation was done by storing a stimulus in the slave's RAM before a trial and, after receiving a trigger from the timing board in the master PC, passing it through a DA converter (Data Translation DT2821) at a sampling output rate of 50 kHz. The output of the board was fed into a band pass filter (Krohn–Hite 3343) with a pass band between 0.2 kHz and 20 kHz, amplified (Luxman A-331), and passed to the speaker.

C. Sound stimuli

Gaussian white noise (GWN), recorded from a function generator (Hewlett–Packard HO1-3722A) and passed through a band pass filter (Krohn–Hite KH 3343 with pass band 0.2–20 kHz flat within 1 dB) was used as a basis for the noise stimuli. The speaker characteristic was flat within 12 dB between 2 and 15 kHz and was not corrected for. In all experiments (also in each session), the same broadband noise burst with a duration of 500 ms was included as the control stimulus. All stimuli had 1-ms sine-squared onset–offset ramps.

In experiment I (subjects BB, JO, PH), the test stimulus set consisted of broadband noise of various durations, $D=3$, 5, 10, 20, 40, and 80 ms [see Fig. 2(a)]. Each stimulus that was presented within one session was drawn randomly from the recorded noise.

In experiment II (subjects JG, JR, PH), the test stimuli were trains of 3-ms bursts with various duty cycles, $\Delta T=3$, 10, 20, 40, 80 ms [see Fig. 2(b)]. A duty cycle of ΔT ms means that the onsets of consecutive bursts were ΔT ms apart. Each 3-ms burst of the train had been drawn randomly from the recorded noise. A single 3-ms burst stimulus (i.e., same as in experiment I with $D=3$ ms) was included as the limit case $\Delta T=\infty$. The total duration of each burst train was about 500 ms.

In experiment III (subjects KH, PH, VC), the test stimulus set consisted of sweeps of various periods [see Fig. 2(c)]. Inverse Fourier transforms ($N=64,128,...,2048$ points) were

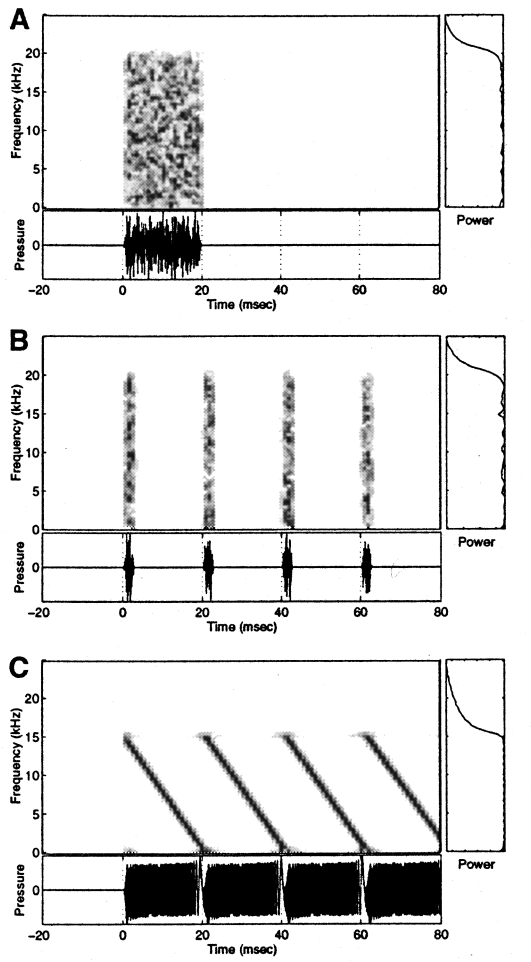


FIG. 2. Examples of the three stimulus types used in the localization experiments: a noise burst with $D=20$ ms (a), a pulse train with $\Delta T=20$ ms (b) and an FM sweep with $T=20$ ms (c). The graphs show the stimuli from 20 ms before stimulus onset until 80 ms after stimulus onset. The large panels show the sonograms, which describe the spectro-temporal behavior of the power spectrum; spectral power is coded by the gray scale, where bright corresponds to low power, dark to high power. The panels on the right contain the time-integrated power spectra (all on the same scale). Finally, the lower panels show the stimulus waveforms.

used to transform an amplitude spectrum, that was flat up to 16 kHz with a corresponding phase spectrum calculated according to Schröder's algorithm, in to the time domain (Schröder, 1970; Wightman and Kistler, 1989). This procedure yields FM sweeps with fixed periods $T=1.28$, 2.56, 5.12, 10.24, 20.48, or 40.96 ms. The instantaneous characteristics of this stimulus are narrow band with a center frequency that (repetitively) traverses the entire frequency range downwards with a constant velocity of $16/T$ kHz/s. The total stimulus duration was 500 ms.

Examples of the spectograms of synthesized stimuli that were used in the experiments (before being passed through the speaker) are shown in Fig. 2: a noise burst with duration $D=20$ ms, a burst train with duty cycle $\Delta T=20$ ms, and a sweep with period $T=20$ ms. As can be seen, the noise burst contains spectral power over the entire frequency range and over the entire stimulus duration. In a burst train, the total stimulus duration is long and the stimulus is broadband, but power is only present during short 3-ms intervals. Like the noise burst, the sweep contains no silence periods throughout the whole stimulus duration. The sweep can be considered ''broadband'' if one regards a whole period, yet narrow band on smaller time scales. Thus all stimuli have flat, broadband time-averaged power spectra, whereas the spectro-temporal behavior is fundamentally different for the three stimulus types.

The control stimulus and all test stimuli had equal rms values. For the burst trains this rms value refers to the non-silence periods. The A-weighted sound level at which the stimuli were delivered was 70 dB, measured at the subject's head position (measuring amplifier Brüel & Kjær BK2610 and microphone Brüel & Kjær BK4144).

D. Stimulus positions

In this paper, the coordinates of both the oculomotor response and the sound source position are described in a double-pole coordinate system, in which the origin coincides with the center of the head. The horizontal component, azimuth α, is defined by the stimulus direction relative to the vertical median plane, whereas the vertical component, elevation ϵ, is defined by the stimulus direction relative to the horizontal plane.

The stimulus positions were confined to 25 ''boxes'' centered at azimuths $\alpha=0,\pm 13,\pm 26$ deg, and elevations $\epsilon=0,\pm 13,\pm 26$ deg [see, e.g., Fig. 5(a)]. The dimensions of each box were 8 deg$\times$8 deg, limiting the total stimulus range in both azimuth and elevation to $[-30,30]$ deg. Sets of 25 stimulus positions were composed by randomly selecting a position within each box. Already for one set, this selection procedure ensured a high degree of uncertainty in the stimulus position while maintaining a homogeneous distribution over the oculomotor range. Importantly, the number of stimulus positions could thus be limited for each stimulus condition, which was highly desirable as each experiment had seven different conditions.

E. Paradigm

The eye position in a head-fixed reference frame was used as an indicator of the perceived sound location (see also Frens and Van Opstal, 1995). In order to calibrate the eye coil, each session started with a run in which all 84 peripheral LEDs on the hemisphere were presented in a random order. Subjects were instructed to generate an accurate saccade from the central fixation LED at 0-deg eccentricity to the peripheral target, and to maintain fixation as long as the target was visible. After calibration, the eye position was known with an absolute accuracy of 3% or better over the full oculomotor range.

In the subsequent runs, sound stimuli were presented. A trial always started with the central visual fixation stimulus. Then, after a random period of 0.4–0.8 s, the LED was switched off and the sound was presented at some peripheral location. The subject's task was to direct the eyes as fast and as accurately as possible toward the apparent sound location without moving the head. A firm head rest enabled the subject to stabilize his head position throughout the session.

A typical run with sound stimuli consisted of the control stimulus and six other sound stimulus conditions. During the 175 consecutive trials, each stimulus was presented once at a set of 25 pseudo-randomly drawn positions (see above). The order of stimulus conditions and positions throughout a session was randomized.

Each subject participated in four sessions on four different days. Hence, each subject traversed a total number of 700 localization trials. Each of the seven temporally defined stimuli was presented 100 times in total four times within each of the 25 stimulus boxes. Subject PH participated in all three experiments (2100 trials).

F. Data analysis

Eye positions were calibrated on the basis of responses to 85 visual stimuli in the first run of the session. From this run, sets of raw eye position signals (AD values of the horizontal and vertical position channel) and the corresponding LED positions (in azimuth and elevation) were obtained. LED azimuth, α, and elevation, ϵ, were calculated from the polar coordinates, (R, ϕ), of the LEDs by:

$$\alpha = \arcsin(\sin R \cos \phi),$$
$$\epsilon = \arcsin(\sin R \sin \phi). \tag{2}$$

These data were used to train a three-layer backpropagation neural network that mapped the raw data signals to calibrated eye position signals. In addition, the network also corrected for small inhomogeneities of the magnetic fields and a slight crosstalk between the horizontal and vertical channels.

A custom-made PC program was applied to identify saccades in the calibrated eye-position signals on the basis of preset velocity criteria for saccade onset and offset, respectively. The program enabled interactive correction of the detection markings. The endpoint of the first saccade after stimulus onset was defined as the response position (see also Sec. II). If saccade latency re. stimulus onset was less than 80 ms or exceeded 500 ms, the saccade was discarded from further analysis. Earlier or later stimulus-evoked responses are highly unlikely for a normal, attentive subject, although the precise values of the boundaries are somewhat arbitrary. Earlier responses are generally assumed to be predictive and are very inaccurate, even for visually evoked saccades. Later responses are considered to be caused by inattention (see also Sec. II).

Response positions versus stimulus positions were fitted for the respective components with a linear fit procedure that minimizes the summed absolute deviation (Press *et al.*, 1992). This method is less sensitive to outliers than the more common least-squares method. For the same reason, the correlation between response- and stimulus positions was quantified by the nonparametric rank correlation coefficient, rather than by Pearson's linear correlation coefficient.

Confidence levels for both the linear fit parameters and the correlation coefficients were obtained through the bootstrap method (Press *et al.*, 1992), since explicit expressions for the confidence levels in the methods described above are not available. In the bootstrap method, one creates N synthetic data sets by randomly selecting, with return, data points from the original set (N typically about 100). A synthetic set has the same size as the original set, so that a given data point from the original set can occur more than once in the synthetic set. The parameter of interest is then computed for each synthetic data set and the variance in the resulting N parameter values is taken as the confidence level.

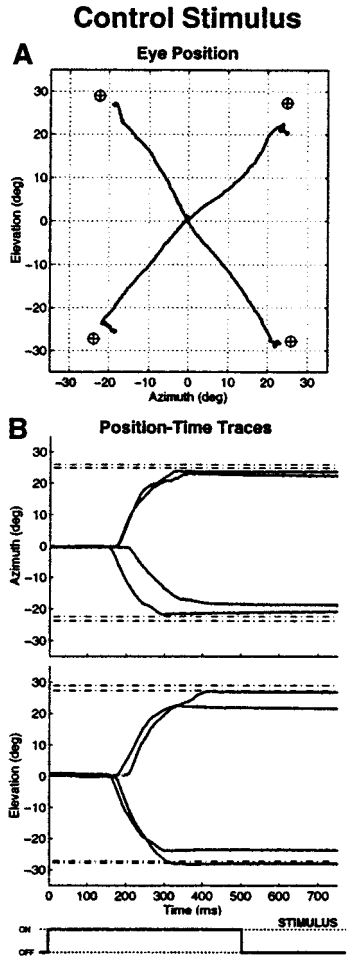


FIG. 3. Typical saccadic eye movement responses towards the control stimulus (GWN, $D = 500$ ms) at four different eccentric positions. (a): Eye-position trajectories during the full duration of the trial. (b): Corresponding position-time traces for both azimuth and elevation during the first 750 ms after stimulus onset. Stimulus positions are indicated by the visor symbols (a) and by the dot-dashed lines (b). Stimulus timing is indicated at the bottom. Data from naive subject BB.

II. RESULTS

A. Control condition

Characteristics of typical responses to four eccentric stimulus positions in the control condition are plotted in Fig. 3. It shows the eye-position trajectory in space and the separate eye-position components (i.e., azimuth and elevation) versus time. One can see that the offset position of the eye (i.e., the response position) corresponds closely to the stimulus position. The response is accurate for both azimuth and elevation components.

It can be seen from Fig. 3(b) that the responses follow shortly after stimulus onset. In these examples, primary sac-

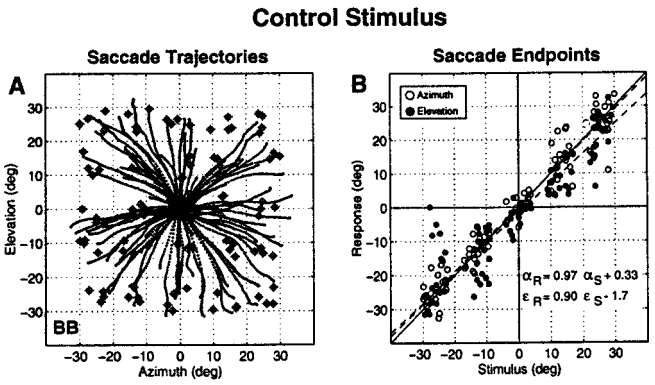


FIG. 4. Saccadic responses to the control stimulus. (a) Trajectories of primary saccades (thick dotted lines, sampled at 2-ms intervals) to control stimuli presented throughout the entire stimulus range. Stimulus positions are indicated by the visor symbols. (b) Endpoint positions of the same primary saccades as in (a) versus stimulus positions for both azimuth (○) and elevation (●). Also the linear-fit results for response positions versus stimulus position are provided (azimuth: α_S vs α_R; elevation: ϵ_S vs ϵ_R). Note large slopes (0.97 and 0.90, respectively; $N=98$) and small offsets (within 2 deg). Subject BB.

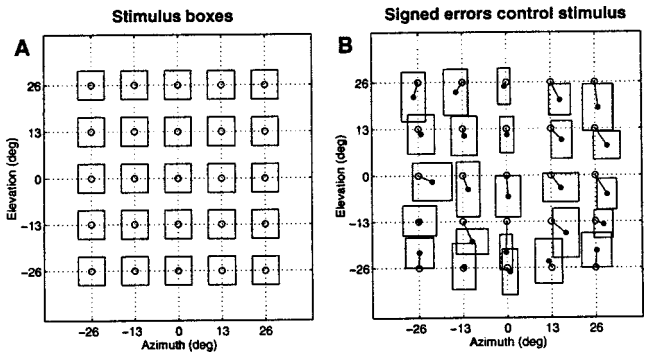


FIG. 5. (a) Stimulus positions were presented within the square boxes, with dimensions 8×8 deg. Sets of 25 stimulus positions were composed by selecting a position at random within each box. (b) The lines from the open symbols to the closed symbols correspond to local mean signed errors. For each stimulus box, the mean signed error is computed by averaging the individual mean signed errors of all seven subjects. The width and height of a box correspond to twice the standard deviation in the signed errors for azimuth and elevation, respectively. The center of each stimulus box is indicated by the open circle [in both (a) and (b)]. Note the clear separation of almost all response boxes, and the larger response scatter in the elevation components as compared to the azimuth components.

cades are initiated with a latency of approximately 200 ms, and are completed within 400 ms after stimulus onset (i.e., before stimulus offset time).

All subjects in this study made accurate localization responses to the control stimulus (500-ms GWN) in all directions (azimuth and elevation). Figure 4 shows all saccade trajectories [Fig. 4(a)] and saccade endpoints [Fig. 4(b)] for all four sessions of one of the subjects (BB). Note how both the stimulus positions and saccade trajectories are distributed over the entire stimulus range. Note also, that the accuracy of response elevation is both quantitatively and qualitatively similar to azimuth localization. A summary of the results for all subjects is listed in Table I.

Figure 5(b) gives an overall impression of the local localization accuracy by showing the averaged signed errors of the saccadic responses for each stimulus box. For each subject and each stimulus box, the mean signed error was computed for all responses to stimuli presented within that box (typically 4–8 responses for each box and each subject). Then, a final mean signed error for each stimulus box was obtained by averaging the subject-mean signed errors over all 7 subjects.

Note that, for the majority of the error boxes in Fig. 5(b), the height is larger than the width. Thus the scatter in the responses is somewhat larger for elevation than for azimuth. These scatter properties underline the fact that azimuth and elevation localization are dissociated processes (see the Introduction). They clearly contrast with the scatter properties of visually evoked saccadic responses, that betray the polar organization of the visuomotor system (e.g., Van Opstal and Van Gisbergen, 1989).

One may observe that the highest accuracy for azimuth localization is reached near the median plane, whereas elevation accuracy is approximately homogeneous over the entire oculomotor field. The largest signed errors were found for

TABLE I. Parameters of the azimuth (α) and elevation (ϵ) response components to the control stimulus (500-ms white noise) for each subject. Columns 2–3: rank correlation coefficient ρ between response position and stimulus position. Note that correlation coefficients $\rho_\epsilon > 0.90$ for elevation and $\rho_\alpha > 0.95$ for azimuth. Columns 4–5: slope, or gain, G, of a straight-line fit for response versus stimulus position. Columns 6–7: standard deviation Δ_{FIT} of the difference between the actual response and the response predicted by the fit (in degrees). Columns 8–9: average absolute localization error Δ_{RESP} (in degrees). Column 10: median response latency (in ms). The two bottom rows present, for each column, the mean and standard deviation, respectively, pooled for all subjects.

	Corr ρ		Gain G		Δ_{FIT}		Δ_{RESP}		
Subject	α	ϵ	α	ϵ	α	ϵ	α	ϵ	Lat.
JO	0.97	0.91	1.1	0.67	4.3	4.9	5.1	6.3	166
BB	0.97	0.94	0.97	0.90	4.5	6.2	3.3	4.9	198
VC	0.96	0.91	0.96	0.81	5.0	6.8	3.8	8.3	226
KH	0.96	0.92	1.0	0.88	5.7	5.5	5.2	4.6	244
JR	0.97	0.95	1.1	1.1	5.0	6.9	4.3	5.9	148
JG	0.97	0.93	1.3	1.0	4.6	6.4	5.4	7.8	156
PH	0.98	0.95	1.1	1.1	3.9	5.9	4.4	6.4	166
mean	0.97	0.93	1.06	0.91	4.6	6.0	4.5	6.3	186
s.d.	0.01	0.02	0.11	0.15	0.7	0.8	0.9	1.5	38

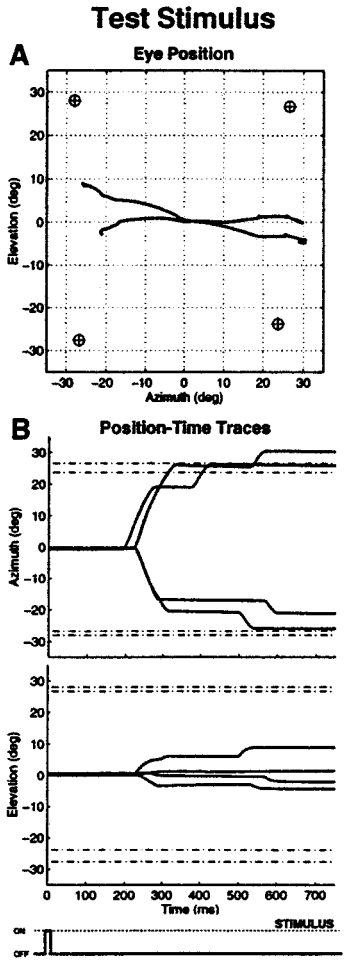


FIG. 6. Typical responses toward short-duration noise bursts ($D=3$ ms) at four different eccentric positions. Note marked undershoot of the elevation components, whereas the azimuth components remain accurate (cf. Fig. 3). Note further that also small secondary saccades are made after approximately 400 ms. See legend of Fig. 3 for further details.

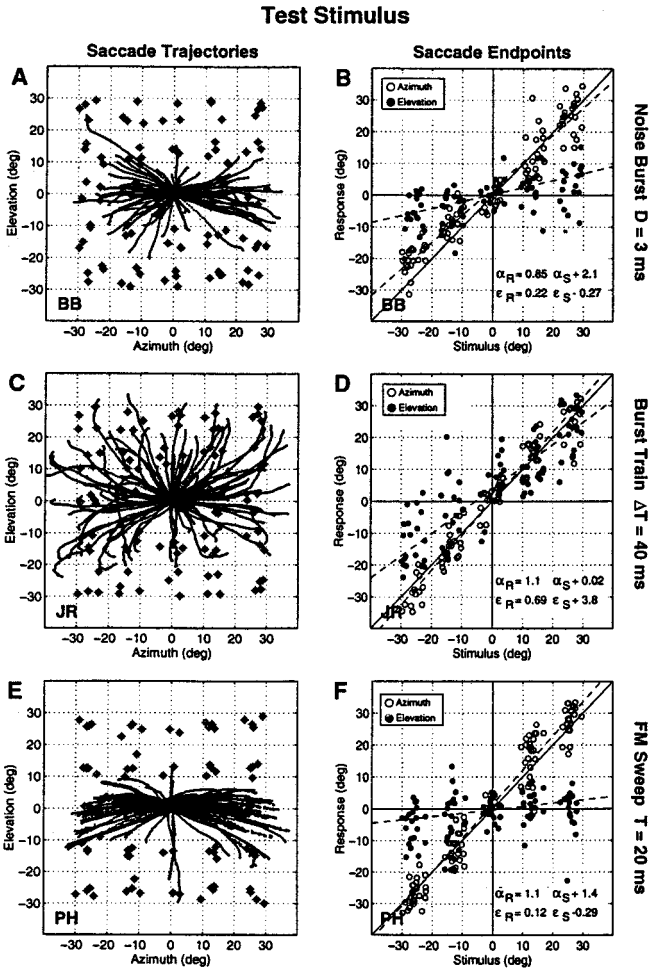


FIG. 7. Saccadic responses to test stimuli. (a) and (b) responses of subject BB to noise bursts of duration $D=3$ ms (correlations for azimuth and elevation: $\rho_\alpha=0.97$, $\rho_\epsilon=0.69$; $N=95$). (c) and (d) responses of subject JR to pulse trains with duty cycle $\Delta T=40$ ms ($\rho_\alpha=0.97$, $\rho_\epsilon=0.86$; $N=97$). (e) and (f) responses of subject PH to FM sweeps with period $T=20$ ms ($\rho_\alpha=0.96$, $\rho_\epsilon=0.38$; $N=143$). Correlations for response position versus stimulus position are listed in Tables II, III, and IV. See legend of Fig. 4 for further details.

stimulus positions within the box at $(\alpha,\epsilon)=(0,-13)$ deg, where response elevations were, on average, about 8 ± 5 deg lower than the actual stimulus elevations.

B. Test conditions

Typical responses to four eccentric stimulus positions in one of the test conditions (noise burst, duration $D=3$ ms) are plotted in Fig. 6. It is clear that azimuth localization is accurate. However, in contrast to the control condition, where elevation detection was also accurate, the responses now exhibit large undershoots in elevation.

Typical responses for the three different stimulus types employed in this study are shown in Fig. 7 (data from subjects BB, JR, and PH, respectively). The most obvious feature of these data, obtained for all three test conditions, is that the saccade trajectories cover only part of the vertical stimulus range [compare Fig. 7(a) with Fig. 4(a)]. Compared to the control condition (see Fig. 4), localization accuracy has clearly deteriorated for elevation, whereas it has remained the same for azimuth [Fig. 7(b), open symbols]. However, although elevation accuracy has deteriorated in these test conditions, the correlation between stimulus elevation and response elevation is still highly significant. This can also be seen in Tables II, III, and IV, which summarize the regression results for all subjects and the test conditions.

In summary, we found for all spectro-temporal stimuli that, although response elevation could be incorrect, it was quite consistent and correlated highly with the real stimulus elevation: $\rho_\epsilon \geqslant 0.4$ for all conditions. This is also expressed by the difference, Δ_{FIT}, between the responses and the straight line fits (dashed lines in Fig. 4 and Fig. 7): Δ_{FIT} remained within 4–8 deg for elevation and within 3–6 deg for azimuth for all stimulus conditions and subjects. Hence, the variance in the responses did not increase substantially with respect to the control condition (compare with Table I).

C. Response gain

Next, we compared the responses for each subject for the different temporal parameters D, ΔT, and T with the results from the control stimulus ($D=500$ ms). The straight

TABLE II. Parameters of the responses to noise-burst stimuli with various durations D (in ms) for three subjects. Note that elevation gain increases systematically with increasing D, but that the variability in the data (Δ_{FIT}) is independent of D. Azimuth gain, however, is not affected by D. Note also, that the median latencies tend to decrease as D increases. See Table I for further explanation of the columns.

		Corr ρ		Gain G		Δ_{FIT}		Δ_{RESP}		
Subject	D	α	ϵ	α	ϵ	α	ϵ	α	ϵ	Lat.
BB	3	0.97	0.69	0.85	0.22	4.9	5.9	4.9	13	210
	5	0.96	0.71	0.81	0.33	5.5	6.4	5.1	11	213
	10	0.96	0.79	0.86	0.30	5.1	4.8	4.7	11	207
	20	0.97	0.89	0.87	0.50	5.4	5.6	4.4	9.1	191
	40	0.97	0.94	0.91	0.63	4.6	5.3	3.7	6.7	188
	80	0.95	0.94	0.90	0.76	6.4	5.8	5.0	5.7	190
JO	3	0.98	0.80	0.99	0.28	3.4	4.1	4.7	12	182
	5	0.98	0.76	1.0	0.35	3.4	5.1	4.1	11	180
	10	0.97	0.80	1.0	0.43	3.8	5.0	4.4	10	173
	20	0.97	0.88	1.0	0.42	4.1	4.6	4.3	9.8	164
	40	0.96	0.92	1.0	0.56	4.3	4.4	4.8	7.2	164
	80	0.97	0.92	1.1	0.61	4.0	4.8	4.7	7.0	163
PH	3	0.97	0.85	1.1	0.67	4.4	8.0	4.4	10	188
	5	0.97	0.87	1.1	0.70	4.2	7.1	4.4	9.4	180
	10	0.97	0.87	1.1	0.82	4.6	7.5	4.7	10	172
	20	0.96	0.89	1.1	0.81	4.4	7.1	4.9	8.3	164
	40	0.98	0.93	1.1	0.99	3.3	6.2	3.7	6.7	164
	80	0.97	0.94	1.1	1.0	4.1	6.0	4.2	6.1	162

line appeared to yield a reasonable description of the relation between response versus stimulus position, and the slope of the line turned out to be a characteristic parameter for the responses in a given condition. One can immediately see from Fig. 8 that the response gains for elevation varied systematically with all three temporal parameters (bottom panels). The azimuth component of the saccades was unaffected by the stimulus parameters (top panels).

For the noise burst, the same systematic variation with D is observed for all three subjects: the response elevation gain increases gradually with stimulus duration D, from $D=3$ ms up to $D=80$ ms. Although there is some intersubject variability as to the absolute values of these gains, all subjects followed a similar trend. Note that similar quantitative inter-subject differences were also obtained for the control stimuli. Furthermore, as the control stimulus is a noise burst with $D=500$ ms, the results suggest that for broadband noise bursts responses stabilize at roughly $D=80$ ms. Note that the

TABLE III. Parameters of the responses to the burst-train stimuli with various duty cycles ΔT (in ms). Elevation gain, but not azimuth gain, depends systematically on ΔT. $\Delta T=\infty$ refers to the single 3-ms burst stimulus. Median latencies tend to decrease with ΔT for all three subjects. See Table I for explanation of the columns.

		Corr ρ		Gain G		Δ_{FIT}		Δ_{RESP}		
Subject	ΔT	α	ϵ	α	ϵ	α	ϵ	α	ϵ	Lat.
JR	3	0.98	0.94	1.1	1.1	4.4	6.9	3.8	6.4	148
	10	0.98	0.93	1.1	1.0	4.4	7.3	3.9	5.7	154
	20	0.98	0.90	1.2	0.86	4.2	7.7	4.0	7.0	160
	40	0.97	0.86	1.1	0.69	5.0	7.7	4.2	8.6	166
	80	0.97	0.85	1.1	0.61	4.9	6.7	4.0	8.9	168
	∞	0.96	0.74	1.1	0.49	5.1	7.8	4.2	10	166
JG	3	0.98	0.94	1.2	1.0	4.4	5.9	5.2	6.6	156
	10	0.98	0.87	1.2	0.82	4.1	7.7	4.9	8.0	158
	20	0.96	0.93	1.2	0.73	5.5	4.7	5.6	7.1	166
	40	0.97	0.92	1.2	0.77	4.8	5.8	5.5	7.1	170
	80	0.97	0.92	1.3	0.77	4.9	6.2	5.7	6.1	182
	∞	0.97	0.83	1.3	0.55	5.7	7.9	6.8	12	176
PH	3	0.98	0.92	1.1	1.0	4.4	7.5	4.2	6.1	164
	10	0.97	0.96	1.1	0.97	4.4	5.1	4.1	5.8	178
	20	0.97	0.91	1.1	0.83	4.3	6.5	4.6	9.0	170
	40	0.97	0.91	1.1	0.71	4.3	6.0	4.7	10	182
	80	0.97	0.82	1.1	0.59	4.0	6.9	4.7	11	190
	∞	0.97	0.84	1.1	0.76	4.4	8.1	4.5	11	182

TABLE IV. Parameters of the responses to sweep stimuli with various repetition periods T (in ms). Elevation gain, but not azimuth gain, depends systematically on T. Median latencies tend to increase with T for subject PH, but not for subjects VC and KH. See Table I for explanation of the columns.

		Corr ρ		Gain G		Δ_{FIT}		Δ_{RESP}		
Subject	T	α	ϵ	α	ϵ	α	ϵ	α	ϵ	Lat.
VC	1.3	0.95	0.79	0.92	0.58	5.4	8.5	4.4	13	244
	2.6	0.96	0.87	0.84	0.63	4.2	7.2	4.5	9.8	232
	5.1	0.96	0.70	0.86	0.27	5.0	6.9	5.2	12	236
	10	0.95	0.55	0.87	0.15	5.4	6.8	4.8	13	237
	20	0.96	0.43	0.94	0.14	5.4	5.8	4.6	14	234
	41	0.95	0.47	0.97	0.15	5.6	7.2	4.7	14	240
PH	1.3	0.97	0.91	1.1	0.91	4.3	7.2	4.0	10	164
	2.6	0.97	0.91	1.1	0.81	5.6	7.6	4.7	6.5	163
	5.1	0.96	0.86	1.0	0.64	4.8	6.9	4.0	7.9	170
	10	0.97	0.43	1.2	0.15	5.3	8.1	5.5	14	168
	20	0.96	0.38	1.1	0.12	4.8	6.5	4.3	14	188
	41	0.96	0.63	1.2	0.20	6.2	4.1	7.1	14	184
HK	1.3	0.94	0.84	0.95	0.55	5.7	7.1	4.9	8.3	251
	2.6	0.95	0.84	0.81	0.66	5.3	7.4	5.0	8.4	246
	5.1	0.96	0.85	0.94	0.43	4.7	5.0	4.4	12	269
	10	0.95	0.57	0.87	0.29	4.8	7.5	4.3	15	243
	20	0.93	0.41	0.79	0.13	4.7	5.9	5.3	20	258
	41	0.97	0.60	0.93	0.19	4.6	5.3	3.9	18	239

short bursts already account for relatively large gains: for example, the gains at $D=10$ ms are already about 40%–80% of the final gains obtained for $D=80$ ms.

For the burst-train stimulus, response gain decreases monotonously with the duty cycle ΔT for subjects PH and JR. The response gain of subject JG, although displaying a similar overall trend, does not vary significantly for the intermediate values of $\Delta T=10,20,40,80$ ms. In this experiment, the intersubject variability for the gains is small for both the burst trains and the control condition. For the shortest duty cycles applied ($\Delta T=3$ ms), the gain is similar as for the control condition (indicated by C). In addition, for subject PH who also participated in experiment I, the gain for $\Delta T=80$ ms is very similar to the gain observed for the single noise burst at $D=3$ ms (condition $D3$ in Fig. 8). This result could be expected: the latency of this subject's response lies around 150 ms (see Table III), so that only the first burst (i.e., at $t=0$ ms) and maybe the second one (at $t=80$ ms) could actually have been processed by the auditory system for generating the first saccade.

For the sweeps, the elevation gain decreases when the period T increases. In contrast to the noise burst and the burst train, the change with the temporal parameter is more sudden. Approximately $T=5$ ms seems to be a critical value, where the response elevation changes most rapidly with T. From $T=10$ ms (subject KH) or $T=20$ ms (subjects PH and VC) there is little change in the gain which lies between 0.1 and 0.2. Although the gains for $T\geqslant 10$ ms are relatively low, correlations for the sweep data were still 0.38 or higher.

D. Response latency

The latencies of primary saccades are typically well below 300 ms. Figure 9(a) shows a latency distribution for saccades to the control stimulus (subject PH). One can see that latencies peak near 170 ms and remain within the [100,300]-ms interval. This interval was typical for all subjects, which can be further appreciated from the cumulative latency distributions in Fig. 9(b). Data from three different subjects for the control condition are presented: latencies that were relatively short (subject JR), long (VC), and intermediate (PH). Note that, also for subject VC, more than 90% of the latencies remained below 300 ms.

It may be observed that in Fig. 9(b) the curves are nearly linear and roughly parallel to each other (see Sec. III). This is most obvious for the distributions of subjects PH and VC, which are well-defined, since they consist of a relatively high number of saccades. For other subjects, fewer saccades were available for each condition, but the distributions exhibited roughly similar characteristics [e.g., subject JR, Fig. 9(b)].

For subject PH, also cumulative distributions are shown that resulted from the several short-noise burst experiments. One can see that latencies systematically increase as the duration of the noise burst decreases. The same trend was observed for the other subjects (BB, JO) who participated in this experiment (see Table II). Also in the other experiments, a systematic shift of the (reciprocal) latency distribution as function of the temporal stimulus parameter was observed as well (except for subjects VC and KH in response to FM sweeps). Median latencies also show this trend (see Tables II, III, and IV). For different stimulus conditions, the reciprocal latency distributions differed in their offset, but retained their shape.

E. Primary and secondary saccades

The endpoint of the primary saccade was accepted as a valid response when its onset latency fell in the interval [80,500] ms. For the average subject, a valid response was measured in $97\%\pm 2\%$ of the trials. In the same time inter-

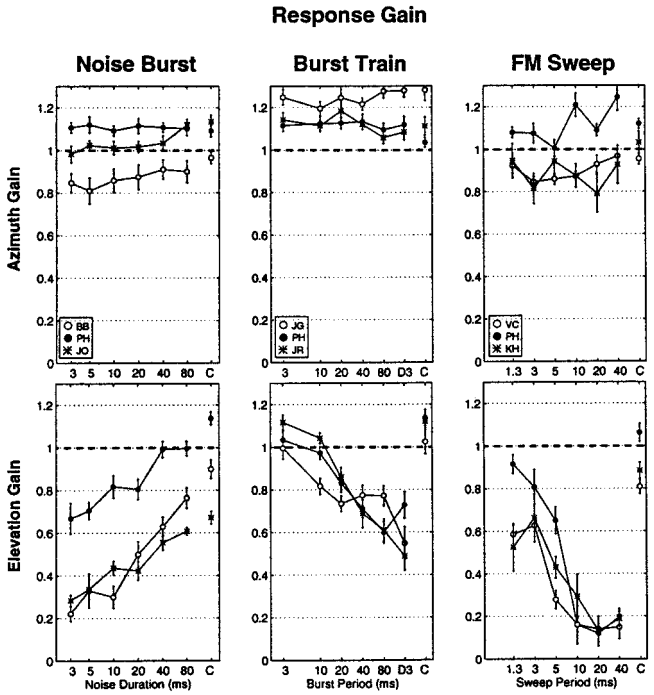


FIG. 8. Response gains for all subjects and all conditions. The gain is defined as the slope of the straight line fitted through response component versus stimulus component (see also Fig. 4 and Fig. 7). Top panels: azimuth gains. Lower panels: elevation gains. Gains are plotted as function of the temporal stimulus parameters D, ΔT, and T. The control condition is indicated by C. The single-burst condition in the burst-train experiments is labeled by D3. Optimal agreement of response and stimulus position is associated with gain 1.0 (dashed lines). Lowest correlations ρ_ϵ obtained for noise bursts, burst trains and sweeps were 0.58, 0.83, and 0.38, respectively, but still highly significant (see Tables II, III, and IV).

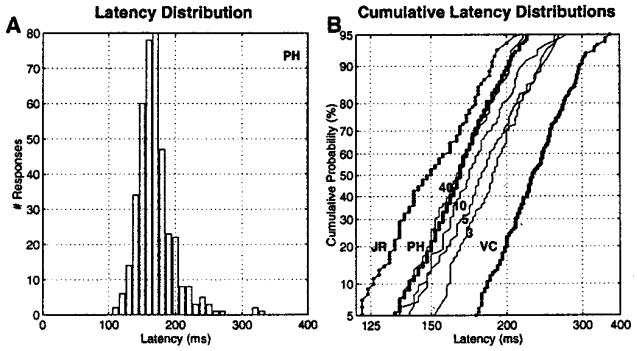


FIG. 9. (a) Response latency distribution for the control condition. Subject PH. (b) Cumulative response latency distributions. The latency axis has a reciprocal scale, whereas the abscissa is on probit scale. In this format, a Gaussian distribution of reciprocal latency results in a straight line. The thick dotted lines refer to the control conditions for subject JR, VC, and PH; the distributions consist of 99, 219, and 398 saccades, respectively. The thin lines show the cumulative distributions of responses to short noise bursts with $D=3$, 5, 10, 40 ms (subject PH). For each condition D, about 116 saccades are included.

val, a secondary saccade was observed in $23\% \pm 9\%$ of the trials, and a third saccade in less than 4% of the trials.

To test whether the secondary saccade was corrective (which is known to be the case for visually evoked saccades), the unsigned errors after the primary saccade, and after the secondary saccade were compared for each subject. This was done for saccade azimuth α, elevation ϵ, amplitude R and direction ϕ. The analyses revealed that incorporating the second saccade did not significantly change (i.e. neither improve nor deteriorate) response accuracy (data not shown).

To further check for a possible relation of the secondary saccade with the stimulus position, an additional analysis was performed by comparing the directions of the primary saccade, the secondary saccade and the stimulus. First, the difference in direction, $\Delta\phi_{12}$ between the primary and the secondary saccade was computed. For all subjects pooled, it was found that $\Delta\phi_{12}=4\pm48$ deg. The individual results for each subject were similar. Thus the secondary saccade generally proceeds in the same direction as the primary saccade. Next, $\Delta\phi_{12}$ was compared to the difference in direction, $\Delta\phi_{01}$, between the primary saccade and the stimulus position (i.e., the direction error). No significant correlation between $\Delta\phi_{12}$ and $\Delta\phi_{01}$ was found. Thus the secondary saccade does not correct for a residual direction error after the primary saccade either.

III. DISCUSSION

A. General findings

In the present study, human localization experiments were performed using a wide range of acoustic stimuli. The time-averaged power spectra of the stimuli were always broadband and identical in shape, but the spectro-temporal behavior on a millisecond time scale was fundamentally different. Localization performance varied systematically with the experimental parameters T, ΔT, and D. Whereas elevation detection appeared to be very sensitive to the spectro-temporal stimulus behavior, azimuth localization remained unaffected and was equally accurate for all conditions.

Our findings provide new insights into the spectro-temporal processing of acoustic sensory information. The data suggest specific temporal constraints for accurate acoustic localization of stimulus elevation. Moreover, these results underline the presence of separate dynamical processes underlying the analysis of the different acoustic cues for the detection of azimuth (ITD, IID) and elevation (spectral shape cues).

B. Saccadic eye movements

1. Orienting to sounds through saccadic eye movements

Saccadic eye movements were used to measure the perceived sound direction. The results show that this method yields a highly reproducible and accurate measure of the acoustic localization percept for stimuli presented within the oculomotor range (approximately 35 deg in all directions). Correlations for both the horizontal and vertical components of responses to control stimuli exceeded 0.9 in experienced as well as in naive subjects. The oculomotor response forms an important part of the natural repertoire of stimulus-evoked orienting (including head and body movements) and does not require any specific training of the subjects. Moreover, the response is fast (latencies remain well below 300 ms; see also Frens and Van Opstal, 1995).

In previous localization studies with human subjects, different response methods have also been used to quantify the localization percept: arm pointing (e.g., Oldfield and Parker, 1984a), the naming of learned coordinates (e.g., Wightman and Kistler, 1989), and head pointing (e.g., Makous and Middlebrooks, 1990). The first two methods are substantially slower than the eye-movement method and may thus be assumed to measure a later acoustic percept, possibly also incorporating cognitive aspects. The head-pointing method is potentially faster. Although latencies of head movements can be similar to eye movement latencies, head-movements dynamically alter the acoustic input for long-duration stimuli, which was deemed to be an undesirable factor for the purpose of this study.

2. Potential artefacts

We are confident that our main results cannot be ascribed to peculiar properties of the oculomotor system. First, the sound stimuli were presented well within the oculomotor range and stimulus types and stimulus positions were always presented in a randomized order. Therefore, the results cannot be due to a (conscious or subconscious) strategy adopted by the subjects.

A second argument for believing that the results reflect properties of the auditory system, rather than of the visuomotor system, is that quite different behaviors were obtained for the azimuth components of the responses, than for the elevation components. Such a behavior is never encountered in visually evoked eye movements toward stimuli at similar locations.

We also believe that the auditory stimuli were always presented well above the detection threshold, since similar temporal-dependent response behaviors were obtained for all three acoustic stimulus types. The finding that the azimuth components of the responses remained unaffected by the stimulus parameters, indicates that the stimuli were well-perceived by the auditory system, despite the fact that the overall acoustic energy of the stimuli varied greatly. This is also supported by the fact that the standard deviations of the latency distributions were hardly affected [see Fig. 9(b), and below].

Finally, also the result that response variability did not change systematically with the temporal stimulus parameters (Tables II, IV, and III) argues against the possibility that the auditory stimuli may have been close to the detection threshold for the shortest stimulus durations D or longest duty cycles ΔT.

3. Latency characteristics

There are some interesting aspects regarding the auditory-evoked saccadic latency distributions. First, the cumulative latency probabilities in Fig. 9(b) are plotted on a so-called probit scale, which is the inverse of the Error function. If the cumulative distribution on probit scale yields a straight line, the variable follows a Gaussian distribution. This linear feature of reciprocal latency distributions has been shown to be characteristic for visually evoked saccades (Carpenter, 1995). In the present study, straight lines were also obtained for auditory saccades, despite the different origin and encoding format of acoustic sensory information.

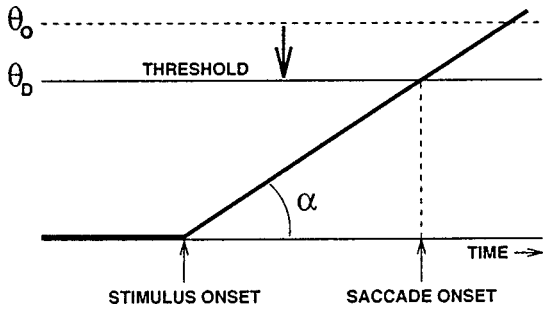


FIG. 10. A simple model that accounts for the observed characteristics of the latency distribution (see also Carpenter, 1995). At stimulus onset, a decision signal rises with a constant rate, α, and upon exceeding a fixed threshold, at θ_D, a saccade is initiated. In each trial, the increase rate, α, is drawn from a gaussian distribution. In order to incorporate the condition-dependent modulations on the latency in this study, it is proposed that the threshold is not constant, but instead decreases during stimulus presentation, from the initial value θ_0 to the final value θ_D.

Carpenter (1995) proposed a simple model for visual saccade initiation that accounts for this characteristic property (Fig. 10). The model assumes that a decision signal starts to increase after stimulus onset and, subsequently, initiates the saccade upon exceeding a fixed threshold (θ_0). Consequently, the time-to-threshold (i.e., the saccade latency) is inversely proportional to the increase rate of the decision signal. The increase rate (slope α) is assumed to vary from trial to trial, but remains constant during a trial. If this rate is distributed as a gaussian over trials, then the reciprocal latency is necessarily distributed likewise. Our data suggest that a similar mechanism could underlie the initiation of auditory-evoked saccades.

Yet, the question remains how stimulus-related shifts in the distributions could arise. For example, Fig. 9(b) shows that latencies become larger for shorter stimulus durations. The simple model described above has only a few parameters that can cause a shift in the distribution: the initial level of the decision signal, its mean increase rate, and the threshold level. The first two parameters have preset values and are assumed to depend on expectation about the stimulus. Since the different stimuli were presented in random order, expectation is not likely to play a role in the present situation. The shift in the median latencies is qualitatively explained, however, if it is assumed that the threshold decreases as more sensory information comes in (θ_D). For short burst durations, the threshold would then remain systematically higher (yielding longer latencies) than for longer stimulus durations (see also Fig. 10).

4. Secondary saccades

The typical response pattern of an auditory orienting trial usually consists of a large primary saccade followed by a smaller secondary (and occasionally a tertiary) saccade. A similar pattern is also typical for visually elicited responses. The primary saccade carries the eye over roughly 90% of the required trajectory (undershoot), whereas the secondary saccade corrects for the remaining retinal error. It has been proposed (Harris, 1995) that such a motor strategy minimizes

(in a statistical sense) the time needed to fixate the target, by taking into account both the effects of a longer duration needed to complete larger saccades, and the additional time needed to program a corrective saccade in the opposite direction (due to an overshoot). Here, the question was what the underlying strategy for secondary saccades could be in the case of a sound stimulus.

Analysis of the auditory responses indicated that the secondary saccades were not corrective. Note that, in contrast to visual stimulus conditions, the oculomotor system is not provided with any feedback concerning its performance after completing the primary auditory saccade. Notwithstanding, the occurrence of a small secondary saccade in roughly the same direction as the primary saccade was quite typical in all subjects tested. We propose that this phenomenon reflects a property of the programming mechanisms underlying the oculomotor response, rather than auditory processing. The data suggest that the secondary saccade is a pre-programmed movement that in this case, however, does not (cannot) correct for a residual error.

C. Temporal aspects of sound localization

1. Effect of stimulus duration

It should be realized, that the programming of an auditory-evoked saccade consists of at least two main stages: a target localization stage, in which the acoustic information is transformed into an estimate of target location, and a response initiation stage in which the estimated target coordinates are transformed into the appropriate motor commands. We believe that the data obtained from the noise-burst experiments may be interpreted in the light of these two, partly separate, stages.

First, it was found that the response elevation gain increases systematically as function of stimulus duration, and that it reaches a plateau for stimulus durations exceeding 80 ms (Fig. 8). From these data we infer that the auditory localization system needs roughly 80 ms of Gaussian broadband input to reach a stable estimate of target elevation.

Second, the latency data [Fig. 9(b) and Table II] may provide further insights into the processing time of the response initiation stage. Although it takes less time for the short-duration bursts to complete, the associated latencies were about 20 ms longer than for the longest stimuli (see Table II). This suggests that the movement initiation stage takes, at least, about 20 ms longer to initiate a saccade toward the shortest burst. We have no simple explanation for this apparent, yet consistent, discrepancy.

In contrast to sound elevation, sound azimuth can apparently be determined accurately on the basis of only a few milliseconds of sensory information. Even for the shortest stimuli ($D=3$ ms), the accuracy was about the same as for the control condition ($D=500$ ms). In this sense, azimuth localization can be considered as a much ''faster'' process than elevation localization. From our data it is not possible to conclude whether this property is due to mechanisms processing either the interaural phase or intensity differences, since the broadband bursts provided both cues simultaneously.

2. Effect of nonstationarity of the short-term spectrum

The results obtained from the FM sweeps provide some further interesting suggestions regarding the temporal processing of sensory information in sound localization. In contrast to the noise bursts, the duration of the sweeps was kept constant (at 500 ms), but the spectro-temporal behavior was varied by means of the cycle variable T. Response characteristics were similar as in the case of the noise bursts. Although elevation localization accuracy varied with T, responses remained consistent: high correlation coefficients of response versus stimulus position were obtained for all conditions. Because elevation localization performance was most accurate for $T<5$ ms and relatively inaccurate for $T>5$ ms, it is suggested that the auditory localization system discriminates spectro-temporal patterns at a temporal resolution down to about 5 ms.

A possible explanation for this ability is that the auditory system applies a so-called *multiple-look* strategy (e.g., Viemeister and Wakefield, 1991). In such a mechanism, the input spectrum is measured over consecutive short time windows, each lasting only a few milliseconds. Each short-term spectrum is processed into a position estimate, which, at a higher level, is integrated with the earlier estimates into a final estimate (see below, Fig. 12). This explanation is also in line with the interpretation of the burst-duration experiments (see above).

As an example, consider the computation of a short-term spectrum over a 5-ms time window somewhere in the cycle of an FM sweep with a long period, e.g., $T=20$ ms. The outcome will vary substantially for different window positions within the cycle. Elevation updates computed from consecutive (but very different) short-term spectra would then be inconsistent throughout the cycle and prevent the dynamic elevation estimate to stabilize at the actual target elevation. Note, that such a 5-ms time-averaged spectrum in a fast sweep ($T<5$ ms) would be broadband, whereas it would be narrow band for very slow sweeps ($T\gg 5$ ms).

3. Effect of silence gaps in the stimulus

The data obtained from the burst-train stimuli suggest further constraints on the dynamics of the spectral analysis in the localization process. For the shortest burst intervals, ΔT, near-optimal localization performance was obtained, although each burst was only 3 ms long.

Interestingly, the elevation gain was observed to decrease significantly when the individual bursts of the train were all chosen to be identical, rather than randomly drawn as in the experiments presented here, even for $\Delta T=3$ ms (data from subject PH only, not shown). Thus acoustic input presented at later times indeed contributes to the improvement of the final elevation estimate.

The low gains at long intervals are not explained by the fact that, within the latency period of roughly 180 ms, only two to four 3-ms bursts may have contributed to the programming of the first saccade, because the final estimate of the oculomotor responses (even after 500 ms) was not systematically better than the estimate recorded after the primary saccade. Rather, the decrease of elevation gain with

increasing ΔT indicates that the subsequent elevation estimates, based on each 3-ms sound burst, are not kept in acoustic memory forever. Possibly, the integrative mechanism that combines subsequent elevation estimates (see above) is leaky. If so, the data suggest a time constant for this integrator in the order of a few tens of ms.

D. Toward a spectro-temporal model of sound localization

In this section we study the properties of a monaural localization model that relies on spectral correlations in order to estimate sound elevation, ϵ_S (see also Middlebrooks, 1992). In addition, a biologically plausible mechanism for the dynamic implementation of these correlations, based on our data, will be briefly described.

1. Spectral correlations

In the proposed spectral correlation approach, it is assumed that the auditory system bases its comparison between the sensory signal and the HRTF on the log power of these spectral functions. In addition, a logarithmic scaling of the frequency domain, reminiscent to the tonotopic neural representation of sound frequency throughout the auditory system, is applied. By applying the logarithm to the power of the sensory signal measured at the eardrum, Eq. (1) can be rewritten as:

$$Y(\omega;\epsilon_S)=H(\omega;\epsilon_S)+X(\omega), \tag{3}$$

where ω is in octaves, and ϵ_S is the source elevation. The capitals indicate the logarithmic power spectra [e.g., $X(\omega) \equiv \log|x(\omega)|$]. Note, that elevation, ϵ_S, rather than position, $\mathbf{r}_S$, is used in Eq. (3), as it is assumed that azimuth, α_S, is already extracted from binaural cues.

A quantitative scalar measure of similarity for two spectral functions can be given by the spectral correlation coefficient. Therefore, the mean, $\overline{F}$, and variance, σ_F^2, for an arbitrary spectral function, $F(\omega)$, are first introduced:

$$\begin{aligned}\overline{F}&=\langle F(\omega)\rangle\equiv\int_0^\infty d\omega\ p(\omega)F(\omega),\\ \sigma_F^2&=\langle(F(\omega)-\overline{F})^2\rangle,\end{aligned} \tag{4}$$

with $p(\omega)$ a normalized weighting function that is nonzero in the (broad) frequency band of interest. Here $p(\omega)$ was chosen to be uniform in the [2, 16] kHz range, and zero elsewhere. The spectral correlation, $\mathsf{C}(F(\omega),G(\omega))$, between two functions, $F(\omega)$, and, $G(\omega)$, is then defined as

$$\mathsf{C}(F(\omega),G(\omega))\equiv\left\langle\left(\frac{F(\omega)-\overline{F}}{\sigma_F}\right)\left(\frac{G(\omega)-\overline{G}}{\sigma_G}\right)\right\rangle. \tag{5}$$

The outcome varies in $[-1,1]$, where $\mathsf{C}(\cdot)=1$ corresponds to ''maximal similarity,'' and $\mathsf{C}(\cdot)\leqslant 0$ means ''no similarity.''

In the present discussion, the spectral correlation compares the sensory signal, $Y(\omega;\epsilon_S)$, with (neurally stored) HRTFs, $H(\omega;\epsilon)$, for all sound elevations ϵ. By using Eq. (3) and by taking the mean (where $\overline{Y}=\overline{H_S}+\overline{X}$), this comparison reads:

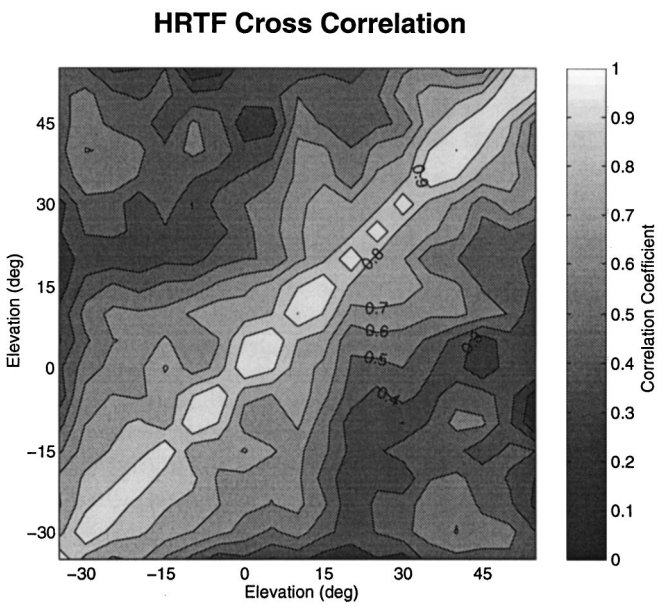


FIG. 11. Spectral correlations of human HRTFs, measured in the midsaggital plane at elevations $\{-35,-30,...,+50,+55\}$ deg. Subject PH. The strength of the correlation is coded by brightness. Markers label the contours at different values. Note that HRTFs correlate strongest with themselves and with HRTFs at neighboring elevations. Correlations were computed in the frequency range [2,16] kHz and HRTFs were sampled at 1/20 octave intervals.

$$\begin{aligned}C_Y(\epsilon;\epsilon_S)&\equiv\mathsf{C}(Y(\omega;\epsilon_S),H(\omega;\epsilon))\\ &=\left(\frac{\sigma_{H_S}}{\sigma_Y}\right)\mathsf{C}(H(\omega;\epsilon_S),H(\omega;\epsilon))\\ &\quad+\left(\frac{\sigma_X}{\sigma_Y}\right)\mathsf{C}(X(\omega),H(\omega;\epsilon))\\ &\equiv\left(\frac{\sigma_{H_S}}{\sigma_Y}\right)C_H(\epsilon;\epsilon_S)+\left(\frac{\sigma_X}{\sigma_Y}\right)C_X(\epsilon),\end{aligned} \tag{6}$$

with σ_Y, σ_{H_S}, σ_X the standard deviations of the respective quantities in Eq. (3). Thus Eq. (6) quantifies the spectral correlation of each stored HRTF (associated with sound elevation, ϵ) with the measured spectrum at the eardrum (resulting from a sound source at elevation, ϵ_S). It proposed that the perception of sound elevation essentially consists of selecting the unique position $\epsilon=\epsilon_P$ at which the global maximum, $C_Y(\epsilon_P;\epsilon_S)$, is attained.

In its decomposed form, Eq. (6) can be readily interpreted. The first term on the right-hand side, $C_H(\epsilon;\epsilon_S)$, is the spectral correlation between the HRTFs $H(\omega;\epsilon)$ and $H(\omega;\epsilon_S)$. Under the assumption that HRTFs are unique with respect to position, $C_H(\epsilon;\epsilon_S)$ reaches a maximum in the neighborhood of ϵ_S, i.e., at $\epsilon\approx\epsilon_S$.

Figure 11 shows that this is indeed a reasonable assumption. In this figure, measured HRTF functions in the median plane of one of our subjects have been correlated with each other. Note that only along the principal diagonal a high correlation is obtained, indicating that, indeed, HRTFs only resemble themselves. Therefore, these functions contain unique information about target elevation, when all frequency bands are allowed to contribute to the computational

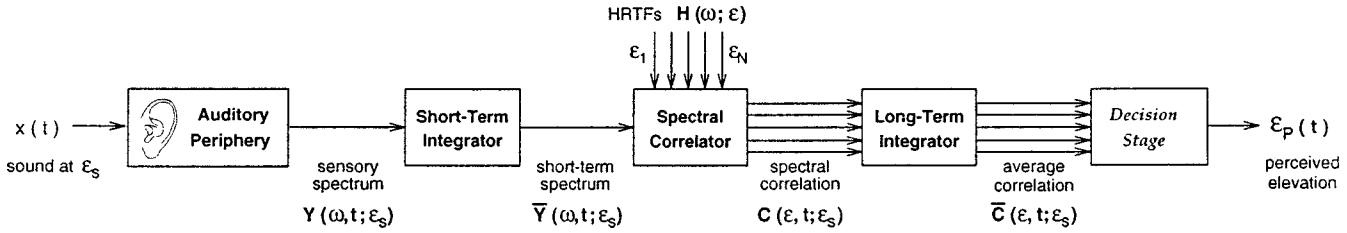


FIG. 12. Conceptual spectro-temporal correlation model of human sound localization. The first stage represents the auditory periphery, where a sound, $x(t)$, at elevation, ϵ_S, is processed from the external ear up to the cochlea. In the second stage, a short-term (order a few ms) averaged spectrum, $\bar{Y}(\omega,t;\epsilon_S)$, is determined around the current time, t, from the ongoing sensory spectrum, $Y(\omega,t;\epsilon_S)$. Then, $\bar{Y}(\omega,t;\epsilon_S)$ is correlated with the neurally stored HRTF representations, $H(\omega;\epsilon)$, for all possible elevations ϵ, i.e., $\epsilon_1,\ldots,\epsilon_N$, yielding short-term correlations $C(\epsilon,t;\epsilon_S)$. Next, integration is done over a longer time span (order several tens of ms), resulting in a long-term correlation up to current time t, $\bar{C}(\epsilon,t;\epsilon_S)$. This integration process may be leaky. In the decision stage, the average correlation at time t yields a perceived elevation, $\epsilon_P(t)$. It is proposed that this decision depends on the maximum of the average correlation, but also on the consistency of that average and on the initial elevation percept.

process. We have verified that this conclusion also holds for the entire elevation domain ($\epsilon \in [-60,90]$ deg, front and back), although the region of maximum correlation broadens appreciably close to zenith positions ($\epsilon \approx +90$ deg; data not shown). It is therefore expected, that stimuli presented within that range are not well localized and discriminated.

The second term in Eq. (6), $C_X(\epsilon)$, expresses the resemblance of the source spectrum $X(\omega)$ with each particular HRTF $H(\omega;\epsilon)$. If the source spectrum does not correlate significantly with any of the stored HRTFs, then $C_X(\epsilon) \approx 0$ for all ϵ [for example, this would be true if $X(\omega)$ is flat and broadband]. In that case, the maximum correlation $C_Y(\epsilon_P;\epsilon_S)$ will be reached at $\epsilon_P = \epsilon_S$, thus the sound would be accurately localized. It is expected that there will be a broad range of naturally occurring acoustic stimuli for which there is little or no resemblance with the stored HRTFs, i.e., $C_X(\epsilon) \approx 0$.

However, the second term does come into play, if the source spectrum $X(\omega)$ does correlate well with a given HRTF, say $H(\omega;\epsilon_*)$. This occurs if $X(\omega)$ contains prominent features that are characteristic for $H(\omega;\epsilon_*)$. Then, $C_X(\epsilon_*) \gg 0$, meaning that $C_Y(\epsilon;\epsilon_S)$ contains a second local maximum at $\epsilon = \epsilon_*$. This could even be the global maximum in which case the perceived position, ϵ_P, would be, ϵ_*, rather than the actual position, ϵ_S, thus $\epsilon_P = \epsilon_* \neq \epsilon_S$. This might occur, for example, when $X(\omega)$ is narrow band (i.e., peaked) with its center frequency at a characteristic peak of $H(\omega;\epsilon_*)$. In the study of Middlebrooks (1992), consistent mislocalizations were attributed to such spurious correlations between the narrow-band source spectrum and one of the HRTFs.

2. Conceptual spectro-temporal model

The data presented in this study clearly indicate that the spectral estimation performed by the auditory localization system is also a temporal process. Therefore, if the auditory system bases its spatial estimation on spectral correlations, the underlying computational mechanisms have to be incorporated within a temporal scheme. Figure 12 provides a conceptual model of the successive stages in this spectral-temporal process (see legend for specific details). The model accounts for the different aspects that have been derived from the data for each of the spectro-temporal stimulus patterns (see above, Sec. III C).

(1) The sweep data suggest that the elevation localization system first measures spectra on a short time scale of about 5 ms (''multiple looks''). Accurate localization requires a broadband, short-term spectrum (fast sweeps, noise bursts). Inaccurate localization results if the short-term spectrum is narrow band (slow sweep).
(2) The noise-burst data suggest that acoustic information needs to be delivered over a (longer) time scale of roughly 80 ms. If short-term estimates are consistent over time (fast sweeps, noise) averaging will enhance the final estimate. Yet, if the estimates vary strongly (e.g., for slow sweeps) the estimates may cancel each other on average.
(3) The burst-train data suggest a power-dependent gating mechanism and leakiness of the long-term integration process. Estimates at different time windows (e.g., in burst trains with a long duty cycle) are not heavily suppressed by the intervening silence periods, thus gating is plausible. Leakiness is inferred from the finding that the elevation gain drops as the silence period increases.

In this scheme, the current estimate, $\epsilon_P(t)$, of target position, ϵ_S, smoothly develops over time as more acoustic information enters the system.

The experimental data further suggest that, in the absence of sufficient spectral processing the auditory localization system stays close to its default initial estimate of elevation, typically near the horizontal plane. An interesting question that remains to be investigated, is whether the system's default elevation estimate depends on the initial gaze direction, or on head orientation. In the present study, initial eye position, the horizontal plane of the head, and the earth-fixed horizon always coincided. This problem could be studied by systematically changing these different frames of reference with respect to each other under similar stimulus conditions as applied in this study.

ACKNOWLEDGMENTS

The authors are indebted to Jeroen Goossens and Jos Van Riswick for their participation in the experiments. We thank Hans Kleijnen and Ton Van Dreumel for valuable technical assistance. BB, KH, and VC are acknowledged for volunteering as experimental subjects. The authors also thank both referees for their useful and constructive com-

ments. This research was supported by the Netherlands Foundation of the Life Sciences (SLW; PH) and the University of Nijmegen (AJVO).

Batteau, D. W. (**1967**). ''The role of pinna in human localization,'' Proc. R. Soc. London, Ser. B **168**, 158–180.

Blauert, J. (**1996**). *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT, Cambridge, MA).

Butler, R. A., and Musicant, A. D. (**1993**). ''Binaural localization: Influence of stimulus frequency and the linkage to covert peak areas,'' Hearing Res. **67**, 220–229.

Carpenter, R. H. S., and Williams, M. L. L. (**1995**). ''Neural computation of log likelihood in control of saccadic eye movements,'' Nature (London) **377**, 59–62.

Collewijn, H., Van der Mark, F., and Jansen, T. C. (**1975**). ''Precise recording of human eye movements,'' Vision Res. **15**, 447–450.

Frens, M. A., and Van Opstal, A. J. (**1995**). ''A quantitative study of auditory-evoked saccadic eye movements in two dimensions,'' Exp. Brain Res. **107**, 103–117.

Gardner, M. B., and Gardner, R. S. (**1973**). ''Problem of localization in the median plane: Effect of pinnae occlusion,'' J. Acoust. Soc. Am. **53**, 400–408.

Han, H. L. (**1994**). ''Measuring a dummy head in search of pinna cues,'' J. Audio Eng. Soc. **42**, 15–37.

Harris, C. M. (**1995**). ''Does saccadic undershoot minimize saccadic flight-time? A Monte-Carlo study,'' Vision Res. **35**, 691–701.

Hebrank, J., and Wright, D. (**1974**). ''Spectral cues used in the localization of sound sources on the median plane,'' J. Acoust. Soc. Am. **56**, 1829–1234.

Hofman, P. M., and Van Opstal, A. J. (**1997**). ''Identification of spectral features as sound localization cues in the external ear acoustics,'' in *Biological and Artificial Computation: From Neuroscience to Technology*, edited by J. Mira, R. Moreno-Díaz, and J. Cabestany (Springer-Verlag, Berlin).

Kistler, D. J., and Wightman, F. L. (**1992**). ''A model of head-related transfer functions based on principal component analysis and minimum-phase reconstruction,'' J. Acoust. Soc. Am. **91**, 1637–1647.

Lopez-Poveda, E. A., and Meddis, R. (**1996**). ''A physical model of sound diffraction and reflections in the human concha,'' J. Acoust. Soc. Am. **100**, 3248–3259.

Makous, J. C., and Middlebrooks, J. C. (**1990**). ''Two-dimensional sound localization by human listeners,'' J. Acoust. Soc. Am. **87**, 2188–2200.

Mehrgardt, S., and Mellert, V. (**1977**). ''Transformation characteristics of the external ear,'' J. Acoust. Soc. Am. **61**, 1567–1576.

Middlebrooks, J. C. (**1992**). ''Narrow-band sound localization related to external ear acoustics,'' J. Acoust. Soc. Am. **92**, 2607–2624.

Middlebrooks, J. C., Makous, J. C., and Green, D. M. (**1989**). ''Directional sensitivity of sound-pressure levels in the human ear canal,'' J. Acoust. Soc. Am. **86**, 89–108.

Musicant, A. D., and Butler, R. A. (**1984**). ''The influence of pinnae-based spectral cues on sound localization,'' J. Acoust. Soc. Am. **75**, 1195–1200.

Neti, C., Young, E. D., and Schneider, M. H. (**1992**). ''Neural network models of sound localization based on directional filtering by the pinnae,'' J. Acoust. Soc. Am. **92**, 3141–3155.

Oldfield, S. R., and Parker, S. P. (**1984a**). ''Acuity of sound localisation: a topography of auditory space. I. Normal hearing conditions,'' Perception **13**, 581–600.

Oldfield, S. R., and Parker, S. P. (**1984b**). ''Acuity of sound localisation: a topography of auditory space. II. Pinna cues absent,'' Perception **13**, 601–617.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (**1992**). *Numerical Recipes in C* (Cambridge U.P., Cambridge, MA), 2nd ed.

Schröder, M. R. (**1970**). ''Synthesis of low-peak-factor signals and binary sequences with low autocorrelation,'' IEEE Trans. Inf. Theory **16**, 85–89.

Shaw, E. A. G. (**1974**). ''Transformation of sound pressure level from free field to eardrum in the horizontal plane,'' J. Acoust. Soc. Am. **56**, 1848–1861.

Teranishi, R., and Shaw, E. A. G. (**1968**). ''External-ear acoustic models with simple geometry,'' J. Acoust. Soc. Am. **44**, 257–263.

Van Opstal, A. J., and Van Gisbergen, J. A. M. (**1989**). ''Scatter in the metrics of saccades and properties of the collicular motor map,'' Vision Res. **29**, 1183–1196.

Viemeister, N. F., and Wakefield, G. H. (**1991**). ''Temporal integration and multiple looks,'' J. Acoust. Soc. Am. **90**, 858–865.

Wightman, F. L., and Kistler, D. J. (**1989**). ''Headphone simulation of free field listening I: stimulus synthesis,'' J. Acoust. Soc. Am. **85**, 858–867.

Wightman, F. L., and Kistler, D. J. (**1992**). ''The dominant role of low-frequency interaural time differences in sound localization,'' J. Acoust. Soc. Am. **91**, 1648–1661.

Zakarouskas, P., and Cynader, M. S. (**1993**). ''A computational theory of spectral cue localization,'' J. Acoust. Soc. Am. **94**, 1323–1331.

Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers

Michael Blomgren and Yang Chen
Department of Communication Sciences, University of Connecticut, Storrs, Connecticut 06269-1085

Manwa L. Ng
Department of Speech Pathology and Audiology, Illinois State University, Normal, Illinois 61790

Harvey R. Gilbert
Department of Communication Sciences, University of Connecticut, Storrs, Connecticut 06269-1085

(Received 15 September 1997; accepted for publication 3 February 1998)

The purpose of the study was to examine the acoustic, aerodynamic, physiologic, and perceptual characteristics of modal and vocal fry production. Twenty normal speakers (10 males, 10 females) participated in the study. Speech material included four sustained vowels (/i/, /ɑ/, /æ/, /u/), and syllable strings of /pi/ repetitions produced in both modal and vocal fry registers. Acoustic data (fundamental frequency, jitter, shimmer, and signal-to-noise ratio), aerodynamic data (airflow and air pressure), and electroglottographic (EGG) data were obtained simultaneously. Results demonstrated considerable differences across voice parameters for the modal and vocal fry registers. Fundamental frequency was significantly lower in vocal fry than in modal register for both males and females, however, significant gender differences existed only in modal register. For both males and females, measurements of jitter and shimmer were significantly higher and signal to noise ratio was significantly lower in vocal fry. In addition, airflow rate in modal register was almost three times as high as the airflow rate in vocal fry register during sustained vowel production. During syllable string production, subglottal air pressure values in modal register were approximately 1.5 times higher than that in the vocal fry register. In general, these data emphasize that the aeromechanical mechanisms of vocal fold vibratory behavior are substantially different between modal and vocal fry registers. A model of vocal fry phonation is presented to account for the present results. [S0001-4966(98)04005-3]

PACS numbers: 43.70.Aj, 43.70.Gr, 43.75.Rs [AL]

INTRODUCTION

According to Hollien (1974) there are three vocal registers. Although variously labeled over the past 40 years, common designations have included (a) falsetto or loft register for the highest of the human vocal frequencies, (b) modal or chest register for the ''normal'' speaking or singing voice, and (c) vocal fry, glottal fry, creak, or pulse register for the phonation produced at the lowest end of the human vocal frequency range. Hollien stated that a vocal register is a series or range of consecutively phonated frequencies that can be produced with nearly identical vocal quality and with little or no overlap in fundamental frequencies (F_0) between adjacent registers. He suggested that a voice register can be operationally defined according to the following parameters: (1) perceptual; (2) acoustic; (3) physiologic; and (4) aerodynamic. In addition, Daniloff *et al.* (1980) reported that ''the complex physiological structure of the larynx can vibrate in three (or more) relatively different ways, giving rise to airflow modulation that yields acoustically and perceptually distinct vocal quality'' (p. 209). The present study explored the acoustic, aerodynamic, physiologic, and perceptual characteristics of two of the three normal voice registers; specifically, modal and vocal fry registers.

Modal register refers to the range of F_0 used in the ''normal mode'' of speaking and singing (Hollien, 1974). Vocal fry is a voice register approximately one octave below the frequencies noted for the normal male modal register. Vocal fry consists of a train of discrete, low-frequency, glottal pulses produced by the larynx (Hollien and Wendahl, 1968). McGlone (1967) indicated that vocal fry exists as a physiologically normal mode of laryngeal vibration which results in a distinctive acoustic signal.

A. Acoustic properties of vocal fry

The F_0 ranges in modal register are characterized by values of approximately 100–140 Hz for males (Hollien and Jackson, 1973; Krook, 1988), and 175–240 for females (Stoicheff, 1981; Krook, 1988). In contrast, vocal fry is associated with a very low F_0. The F_0 of vocal fry samples obtained by Hollien and Wendahl (1968) ranged from 31.6 to 69.1 Hz. McGlone (1967) found values ranging from 10.9 to 52.1 Hz. Murry (1971) obtained values of 22–92 Hz for vocal fry, and 73–98 Hz for the low modal register range. McGlone and Shipp (1971) obtained values of 18–65 Hz for vocal fry. Average frequencies of vocal fry vibration, extrapolated from the literature, range from approximately 20 to 70 Hz with a mean of approximately 50 Hz.

Women usually have a broader F_0 range than men in modal register. In vocal fry register, a negligible difference in F_0 range between males and females has been reported by McGlone (1967). McGlone found that the F_0 range in vocal fry was from 10.9 to 51.0 Hz for males, and from 15.5 to

48.8 Hz for females. Hollien and Wendahl (1968) obtained similar findings for males (24–52 Hz), and females (18–46 Hz). Acoustic parameters of pitch perturbation, shimmer, and signal-to-noise ratio do not appear to have been thoroughly investigated for vocal fry phonation.

B. Aerodynamic properties of vocal fry

1. Airflow

According to McGlone (1967), vocal fry is produced with an average airflow ranging from 2.0 to 71.9 ml/s. McGlone also found that, unlike modal register, airflow rate in vocal fry did not increase as vocal fry frequencies increased. Murry (1971) reported that when vocal fry is sustained, the airflow rate is significantly reduced compared to modal register; about 10–100 ml/s in vocal fry, and 70–180 ml/s in modal register.

2. Subglottal air pressure

Murry (1971) reported subglottal air pressure (Ps) data during vocal fry of 8.0–11.0 cm H_2O. The values were significantly greater than those produced in the modal register (5.0–6.0 cm H_2O). He concluded that Ps increased during the long closing phase to force apart the large vocal fold mass. However, McGlone and Shipp (1971) found that the Ps ranged from 2.5 to 7.6 cm H_2O for vocal fry, and from 2.8 to 7.0 cm H_2O for modal register. Differences between the two vocal registers were not significant. It is clear that the findings of Murry (1971) and McGlone and Shipp (1971) differ and additional research is warranted to further evaluate the vocal fold driving pressures in modal and vocal fry registers.

C. Physiological properties of vocal fry

Vocal fry is characterized by a glottal area function that has sharp, short pulses followed by a long closed glottal interval (Childers and Lee, 1991). There is a very short open period and a very long period where the vocal folds are completely adducted. Due to the short open period in fry phonation, the open quotient (OQ) of vocal fry is less than other registers. Childers and Lee (1991) examined four voice types using electroglottography (EGG): modal, vocal fry, falsetto, and breathy. They found that the glottal pulse widths were smaller (25%–45%) in vocal fry phonation compared to modal register, due to the shorter open phase. Murry (1971) suggested that the low airflow rate in vocal fry was related to the long closed phase of the vocal fold vibratory cycle.

Moore and von Leden (1958), using high-speed laryngeal cinematography, found that during vocal fry the vocal folds opened and closed twice in rapid succession. However, Hollien *et al.* (1977) argued that Moore and von Leden (1958) used a rough type of phonation, and consequently obtained a ''mixed'' phonation which was not indicative of true vocal fry. Later, Whitehead *et al.* (1984) using high-speed laryngeal photography found that the vocal fold vibratory patterns in vocal fry can be associated with single, double, or triple opening and closing gestures followed by a lengthy closed period. The findings by Whitehead *et al.* (1984) appear to support the conclusions of Moore and von Leden (1958), however further research seems warranted to verify the extent and characteristics of multiple opening and closing gestures in vocal fry.

D. Perceptual properties of vocal fry

Vocal fry has been described as sounding similar to the ''popping of corn,'' the imitated sound of a ''motor boat engine,'' or a ''creaky voice.'' In essence, the vocal folds tend to vibrate so slowly that individual vibrations can be perceived (Colton and Casper, 1996). Hollien and Wendahl (1968) reported that the F_0 of their vocal fry samples was very low, and the perceptual regularity of vocal fry provided evidence against classifying it as a voice pathology. Vocal fry may also be regarded as a lower extension of the normal pitch range as brief vocal fry production is common at the very end of sentences when pitch and vocal intensity are beginning to decay (Zemlin, 1988, p. 167).

E. Summary

As vocal fry register is a portion of the full capability of the human vocal range, it appeared that further study should enhance our knowledge of the complex process of phonation. Attempts have been made to understand the phonational properties of vocal fry using a variety of measurement techniques (Childers and Lee, 1991; Hollien and Michel, 1968; McGlone and Shipp, 1971; Murry and Brown, 1971a; Whitehead *et al.*, 1984). Results from these studies vary considerably. To date no studies have measured simultaneously the acoustic, aerodynamic, physiologic, and perceptual characteristics of vocal fry. The purpose of this research is to integrate simultaneous measurement of vocal fry and modal phonational registers to better understand the aeromechanical mechanisms involved with vocal fry production.

I. METHOD

A. Subjects

Twenty normal nonsmoking adult speakers (10 males and 10 females) with ages ranging from 22 to 65 years participated in the present study. Potential subjects were screened for their ability to produce and sustain vocal fry phonation. Each subject's ability to produce vocal fry was judged by two of the experimenters. In order to be considered for inclusion in the present study, speakers were required to produce approximately four seconds of vocal fry in two vowels (/ɑ/ and /i/). Approximately 35 speakers were screened. A total of 24 subjects were selected for recording based on the screening task. Four subjects were subsequently excluded from the study due to their inability to maintain vocal fry in all of the experimental tasks. Subjects were all native speakers of American English, and did not report or exhibit any voice, speech, language, or hearing problems.

B. Speech material

Four vowels (/i/, /ɑ/, /æ/, /u/) were produced by the 20 speakers in both modal and vocal fry registers. The vowels were randomized for each subject. For each register, the vowels were sustained for approximately 6 s and were pro-

duced three times at a normal conversational loudness level. Each speaker also produced a string of seven contiguous /pi/ syllables (/pipi...pi/) three times at normal loudness level in both the modal and vocal fry registers. The speakers were instructed to take a full breath and produce the seven consecutive /pi/ syllables on one breath.

C. Recording procedure

Voice recordings were conducted in a single-wall, sound-isolated booth using an 8-channel PCM data recorder (TEAC, RD-111T). All recordings were digitized, thus avoiding low-frequency distortion inherent in traditional audiotape. Airflow, air pressure, EGG, and acoustic signals were recorded simultaneously on high quality digital tapes. The acoustic signal was acquired via a dynamic microphone (Shure, 515SP) and recorded on one channel of the digital data recorder. A mouth-to-microphone distance of 20 cm was maintained. The EGG signal was obtained using a laryngograph (Kay Laryngograph Ltd.) and recorded on a second channel of the digital data recorder. The EGG signal was sensed by two electrodes, which were placed directly over each subject's thyroid lamina. The electrodes were fixed on a Velcro strip that was then placed around each speaker's neck. The airflow signal was obtained via a pneumotachograph mask that was held securely against each subject's face by one of the researchers to avoid air leakage during the experiment. Air pressure was measured via a 6.0-cm-long polyethylene tube with an internal diameter of 2.5 mm. The tube was passed through the pneumotachograph mask with one end placed in the speaker's mouth, and the other end attached to an air pressure transducer. The airflow and air pressure signals were transduced by using a standard aerodynamic system (Rothenberg, 1973) and were amplified (Glottal Enterprise, MS100-A2) before being recorded. The amplified airflow and air pressure signals were recorded on two separate channels of the data recorder.

The airflow and air pressure signals were calibrated after each recording. During the calibration, two different levels of airflow (0.5 l/s, 1 l/s) and three levels of air pressure (0-cm H_2O, 5-cm H_2O, 10-cm H_2O) were generated by using the Mask Calibration Unit (Glottal Enterprise, MCU-3). All calibration signals were recorded on the PCM data recorder.

D. Analysis

The four signals (airflow, air pressure, EGG, and acoustic) were fed from the data recorder into a computer (80586, CPU) and analyzed by using an integrated computer hardware/software package (Kay Elemetrics, CSL-4300B). The four signals were simultaneously displayed on four separate windows on the computer monitor. For the sustained vowel productions, only airflow and the acoustic signals were analyzed.

The sustained vowel airflow values were obtained by superimposing two markers on the digital flow waveform to delineate a 400-ms segment to be analyzed. The demarcated segment of the flow waveform was then expanded to clearly display the individual peaks corresponding to the individual glottal pulses. The peak airflow of five glottal pulses was measured for each vowel repetition. Thus averaged airflow measures from each subject (in each vowel and in each register) were calculated from 15 peak airflow measures (5 measures per repetition$\times 3$ repetitions). For the /pi/ syllable productions, peak airflow was calculated on the middle five of the seven /pi/ repetitions. Peak airflow in the /pi/ syllable repetition task was measured from the baseline to the maximum point of excursion on the flow waveform following the release of the stop phase of the /p/.

The acoustic analysis was performed on a 1500-ms steady-state segment from the center portion of each sustained vowel production. Vowel samples were digitized using CSL at 20 000 Hz and converted to a waveform file. Waveform analysis (F_0, jitter, shimmer, and signal-to-noise ratio) was conducted with PC-AUDED software (Boston University). Since the vocal fry acoustic waveform was characterized by relatively high variability, the automatic pitch period detection method was not used. Instead, pitch period markers were inserted manually at positive-going zero crossings. The PC-AUDED program allows for a straightforward method of inserting pitch period markers. PC-AUDED scans the signal and automatically imposes the pitch period marker at the exact zero crossing point closest to the manually placed impulse marker. Once the pitch period segments were marked, PC-AUDED computed F_0, jitter, shimmer, and signal-to-noise ratio. Jitter was calculated as a percentage of frequency perturbation (*jitter factor*), which is the mean of the absolute difference between sequential vocal F_0s (in Hz) divided by the mean F_0. This proportion is then multiplied by 100 to attain the percentage score. Shimmer was calculated as the *mean shimmer in decibels* which is the sum of the absolute values of the differences between the current and the previous sample divided by the number of periods minus one. Signal-to-noise ratio was calculated as the difference in dB of the average signal amplitude and the noise signal amplitude. Average signal amplitude and noise signal amplitude were first calculated as rms (root mean square). Then, the average signal amplitude (rms) and the noise signal amplitude (rms) were converted to dB using the following formula: $\mathrm{dB}=20\log_{10}(\mathrm{rms})$.

Subglottal pressure was estimated for the vowel /i/ by measuring and interpolating the peak intraoral pressure during the stop phase of five /pi/ syllable repetitions (Lofqvist *et al.*, 1982; Rothenberg, 1982; Smitheran and Hixon, 1981). The technique is based on the assumption that the vocal folds are fully abducted before the release of the burst during the production of /p/. At that instant, the intraoral pressure is equivalent to the subglottal pressure. For the /pi/ string, only the middle five syllable productions in the vocal fry and modal registers were analyzed. Peak airflow and peak air pressure values were recorded for the middle five productions of /pi/. All airflow and air pressure values were obtained by using a linear regression function calculated from the calibration signals.

The EGG signals were displayed using CSL. Each signal produced in vocal fry register was compared to the EGG signal in modal register. Qualitative visual comparisons were made and deviations in the EGG waveform between the two

TABLE I. Average fundamental frequency (Hz) and standard deviations in semitones (ST), jitter factor (%), mean shimmer (dB), and S/N ratio (dB) for the modal and vocal fry registers. Standard deviations are in parentheses.

	Males		Females	
	Modal register	Vocal fry	Modal register	Vocal fry
F_0 (ST)	117.5 (2.3)	49.14 (3.7)	211.0 (2.0)	48.1 (3.5)
Jitter (%)	1.23 (.83)	14.9 (7.1)	1.79 (.53)	8.8 (4.4)
Shimmer (dB)	0.40 (.21)	1.41 (.66)	0.38 (.20)	1.38 (.82)
S/N ratio (dB)	13.27 (3.87)	0.67 (2.55)	12.84 (3.75)	1.29 (2.74)

registers were noted. Specifically, the number of peaks per cycle were identified. In the present study, an EGG pattern with one opening/closing pattern per vibratory cycle was identified as a *singlet*. The EGG waveforms that contained patterns of more than one peak-per-cycle were identified as *complex repetitive*. Complex-repetitive patterns could include both double peaks (*doublet* cycles) and triple peaks (*triplet* cycles).

E. Perceptual identification of glottal fry

A perceptual task was conducted to determine if listeners were able to differentiate between the modal and glottal fry registers. Forty recordings of the sustained vowel /ɑ/ (one modal and one glottal fry production from each of the 20 subjects) were randomized and presented to a group of 55 listeners. Samples were presented in a sound field via eight overhead speakers. Listener-to-speaker distance ranged from 2 m to 5 m. Several samples of modal and vocal fry production were presented to the listeners for practice before the experimental task. An additional 10 samples were randomly selected from the original samples to serve as an intra-judge reliability test. The listeners were undergraduate and graduate students in the Communication Disorders program at the University of Connecticut. The undergraduate students had no course-work in the area of voice, while the graduate students had a course in voice and some experience listening to voice samples. In a forced-choice format, each listener was asked to rate the 50 samples as either modal or vocal fry register.

II. RESULTS

A. Acoustic data

The F_0 data for the male and female speakers in the modal and vocal fry registers are presented in Table I. A three-way analysis of variance (ANOVA), with gender, vowel type, and register as independent variables, and F_0 as the dependent variable, was performed to determine if significant differences existed among the three independent variables. The analysis indicated no interaction between register and vowel type, or between gender and vowel type with respect to F_0. There was a significant interaction between register and gender [$F(1,159)=372.5$, $p<0.01$]. A one-way ANOVA was performed within in each gender group to examine whether significant differences existed between vocal fry and modal register. The findings indicated that a significant difference existed between vocal fry and modal register for both males and females with respect to F_0 [male: $F(1,79)=341.6$, $p<0.01$; female: $F(1,79)=1355.1$, $p<0.01$]. For the female speakers, the average modal register F_0 was 211 Hz with a range of 175–266 Hz, and the average vocal fry register F_0 was 48 Hz with a range of 24–77 Hz. For the male speakers, the average modal F_0 was 117 Hz with a range of 86–170 Hz, and the vocal fry F_0 was 49 Hz with a range of 24–72 Hz. There was a significant difference between female and male speakers in the modal register [$F(1,79)=342.4$, $p<0.01$] but not in the fry register. The female speakers had a higher mean F_0 than the male speakers only in the modal register.

Separate three-way ANOVAs were performed for the jitter, shimmer, and signal-to-noise ratio (S/N) ratio values, to examine whether an interaction existed among register and gender, register and vowel type, and gender and vowel type. The results indicated no interaction among these three independent variables (register, gender, and vowel type) with respect to the dependent variables of jitter, shimmer, or S/N ratio. The register main effect for jitter, shimmer and S/N ratio values were all significant [jitter: $F(1,159)=117.7$, $p<0.01$; shimmer: $F(1,159)=104.4$, $p<0.01$; S/N ratio: $F(1,159)=474.7$, $p<0.01$]. Specifically, jitter and shimmer values were significantly greater in vocal fry than in the modal register for both males and females (see Table I). S/N ratio values were significantly greater in modal register than in vocal fry register for both males and females (see Table I). The main effect for vowel type was also significant for S/N ratio [$F(3,76)=14.3$, $p<0.01$]. A Newman–Keuls *post hoc* test indicated that the S/N ratio, in modal register only, was significantly higher for /u/ than /æ/ and /ɑ/, and significantly higher for /i/ than /æ/. A gender main effect was identified with jitter [$F(1,159)=68.0$, $p<0.01$]. Specifically, jitter was significantly higher in the male speakers compared to the female speakers only in the vocal fry productions.

B. Aerodynamic data

1. Airflow

Rate of airflow was measured for the four sustained vowels (/i/, /æ/, /u/, /ɑ/) and for the stop consonant in the syllable /pi/. Four one-way ANOVAs were conducted to evaluate whether differences existed among the four sustained vowels with respect to airflow. Vowel differences were evaluated for each register in both gender groups. No significant differences in flow rate values were noted among the four vowels in either modal or vocal fry productions. Consequently, airflow values were collapsed across the four

TABLE II. Average peak airflow rate values (ml/s) and air pressure values (cm H_2O) for modal and vocal fry registers. Airflow in the sustained condition was averaged across the four vowels. Standard deviations are in parentheses.

	Females		Males	
	Modal register	Vocal fry	Modal register	Vocal fry
Airflow on Sustained vowels	154.6 (38.7)	58.2 (28.2)	213.7 (77.4)	69.7 (35.0)
Airflow on /pi/ Syllables	438.0 (124.2)	175.1 (77.5)	648.9 (123.9)	258.0 (87.8)
Air pressure on /pi/ Syllables	7.56 (1.93)	5.25 (1.56)	7.45 (1.88)	5.51 (3.50)

vowels for each gender and register group. The average airflow rates for the modal and vocal fry registers are reported in Table II.

A two-way ANOVA was performed for the flow rate values collapsed across the sustained vowels to examine whether differences existed between gender and register. The results revealed a significant interaction between gender and register [$F(1,158)=13.6$, $p<0.001$]. A Newman–Keuls *post hoc* test indicated that airflow rates for both males and females were higher in modal register than in the vocal fry register, however males had significantly higher airflow rates than females only in the modal register.

A two-way ANOVA was also performed on the flow rate values for /pi/ to determine if differences existed between gender and register. No interaction was found between gender and register. The gender main effect was significant indicating that males exhibited significantly greater airflow rates for /pi/ than females [$F(1,26)=90.7$, $p<0.0001$]. For both males and females, the register main effect was also significant indicating that airflow rates in modal register production were significantly greater than in vocal fry register [$F(1,26)=18.3$, $p<0.0001$].

2. Air pressure

The average air pressure values for the intervocalic stop in the /pi/ syllable are presented in Table II. A two-way ANOVA was performed which revealed no interaction between the two independent variables of gender and register. Only the main effect for register was significant. For both males and females, air pressure values in the modal productions were significantly higher than those in the vocal fry productions [$F(1,19)=11.8$, $p=0.003$].

C. EGG data

The EGG waveforms were evaluated for qualitative differences between modal register and vocal fry. A visual comparison of EGG waveforms in modal register and singlet vocal fry production is provided in Fig. 1. The EGG waveforms that contained patterns of more than one peak-per-cycle were identified as complex repetitive. Both doublet and triplet cycle patterns were observed in the complex-repetitive waveforms (see Fig. 2). The triplet cycles were always observed in association with doublet cycles, while the doublet cycles did not necessarily accompany triplet cycles.

For the male speakers' vocal fry productions, 99 of the 120 EGG waveforms that were analyzed contained complex-repetitive patterns (83%). For the female speakers' vocal fry productions, 45 of the 120 EGG waveforms contained complex-repetitive patterns (38%). A Mann–Whitney U-test was conducted to assess whether gender differences existed with respect to prevalence of complex-repetitive patterns. Males' vocal fry productions were characterized by a significantly higher prevalence of complex-repetitive patterns than the females' ($U=15.5$, $p=0.007$). Of the 83% of the male productions that contained complex-repetitive patterns, 56% contained only doublet cycles and 44% contained a combination of doublet and triplet cycles. Of the 38% of the female productions that contained complex-repetitive patterns, 87% were characterized by the presence of only doublet cycles and 13% contained both doublet and triplet cycles.

D. Perceptual data

Responses from the 55 listeners were calculated and divided into two groups: scores for the modal productions and scores for the vocal fry productions. A total of 2200 responses were recorded from the listeners (55 listeners$\times$40 tokens). The individual listener scores on the 20 modal productions ranged from 16 to 20 correct responses ($M=19.04$, s.d.$=1.27$). Of the 1100 responses pertaining to the modal productions, 1050 were classified correctly as modal or 95.5% correct. Of the 1100 responses corresponding to the vocal fry production, 100% were correctly identified by the listeners as vocal fry.

Intra-judge reliability was determined by comparing responses for each listener on 10 randomly selected samples that were presented a second time. An intra-judge reliability score of 99.3% was obtained. Agreement was obtained on 546/550 responses.

III. DISCUSSION

The main purpose of this study was to compare acoustic, aerodynamic, EGG, and perceptual properties of vocal fold vibration during modal and vocal fry conditions. Each speaker acted as his/her own control. Comparisons of speech samples produced in the modal and the vocal fry registers were made for each speaker.

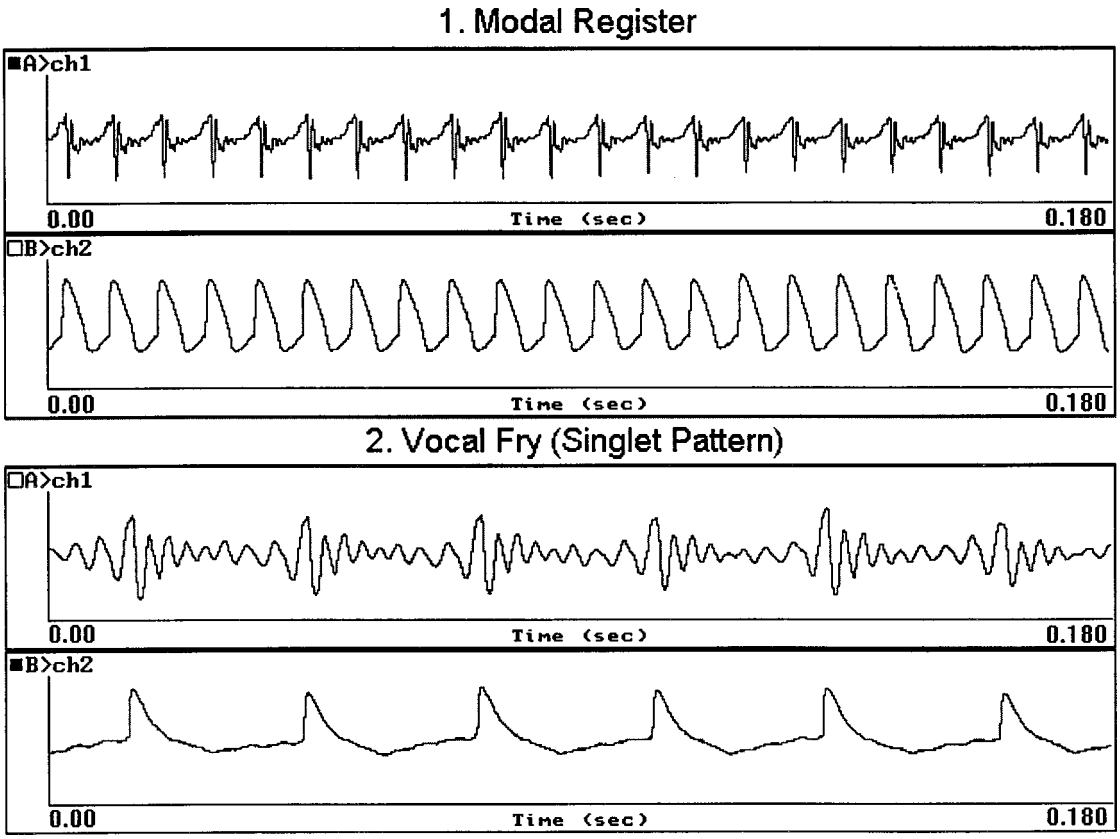


FIG. 1. Illustrations of modal and vocal fry production. Both frame 1 and frame 2 are of the same male speaker producing /ɑ/. Panel A (channel 1) in both the modal and vocal fry frames contain the acoustic signal. Panel B (channel 2) contains the EGG signal. In modal register the speaker's F_0 was 122 Hz while in vocal fry his F_0 was 33 Hz. The illustrated vocal fry production is characteristic of a single pulsing (singlet) pattern.

A. Acoustic measures

1. Fundamental frequency

Considerable differences between modal and vocal fry registers were observed in the present study. Both male and female speakers exhibited significantly reduced F_0 values in vocal fry compared to modal register. These findings agree with the results reported by Hollien and Wendahl (1968), McGlone (1967), and Murry (1971). No intra- or inter-speaker overlap in F_0 between modal and vocal fry was observed. These findings also agree with Hollien's (1974) description of voice register.

The female speakers exhibited significantly higher F_0 values than the male speakers in the modal register. However, the average F_0 values for the female speakers during vocal fry phonation were not significantly different from that of the male speakers. These findings are in agreement with those reported by McGlone (1967) and Hollien and Michel (1968), who suggested a distinct vocal fold vibratory behavior when comparing vocal fry and modal registers.

While general anatomical differences between male and female larynges is often assumed to account for differences in modal F_0, these anatomical differences may not be as relevant when considering vocal fry. For both male and female speakers' vocal fry production, it is possible that vocal fold vibration only occurs over a small portion of the folds. Zemlin (1988) reported that examination of vocal fry with high speed photography revealed ''the folds are approximated tightly, but at the same time they appear flaccid along their free borders, and subglottal air simply bubbles up between them at about the junction of the anterior two-thirds of the glottis'' (p. 166). If vocal fold vibration in vocal fry does indeed occur over a small portion of the total vocal fold length, then the intrinsic longitudinal length of the vocal folds may be less of a mitigating factor in determining the rate of vibration.

Although only part of vocal folds may be vibrating in vocal fry, the intrinsic mass of the vocal folds for male and female speakers may still be inherently different. One might, therefore, still predict gender differences in vocal fry F_0 based on vocal fold mass differences. An alternate ''vocal control'' theory may aid in explaining the absence of gender differences in vocal fry F_0 as found in the present study. Since vocal fry is not typically used by the speakers, gender differences may be obscured. Both the male and female subjects could have adopted the same external standard of a very low fundamental frequency. The similar vocal fry F_0 ranges for males and females may be influenced by the speakers attempt to attain the same task of vocalizating in the sub-60-Hz range.

2. Jitter, shimmer, and S/N ratio

Gender differences for signal-to-noise (S/N) ratio and shimmer values under both vocal fry and modal register conditions were not significant. However, the male subjects exhibited significantly higher jitter values in vocal fry production compared to the female subjects. Jitter factor values for

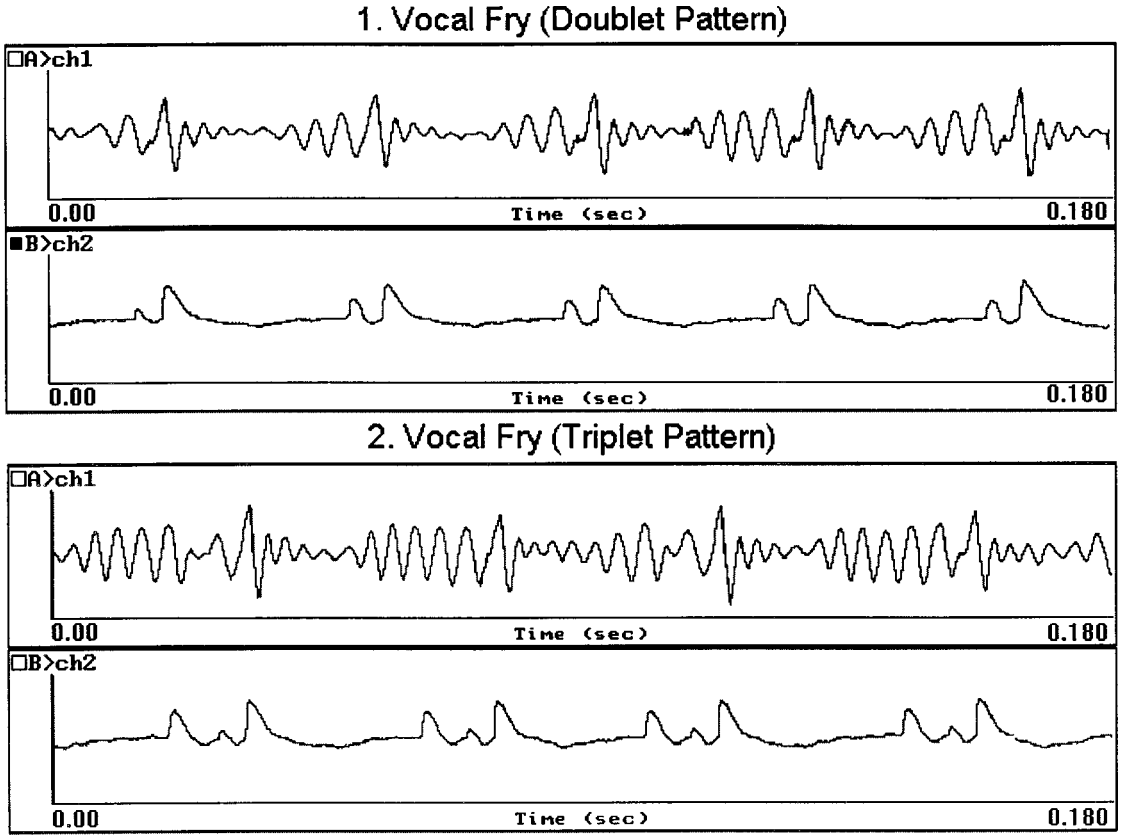


FIG. 2. Illustrations of complex-repetitive vocal fry waveforms. Frame 1 and frame 2 are of two different male speakers producing /ɑ/. Panels A (channel 1) contain the acoustic signal and panels B (channel 2) contain the EGG signal. In the upper frame, double pulsing (doublet) vibratory cycles are displayed at 28 Hz. In the lower frame, triple pulsing (triplet) cycles are displayed at 24 Hz.

modal register production in the present study (*M* males: 1.23%; *M* females: 1.79%) were slightly higher than those reported elsewhere, which are typically slightly less than 1% (Hollien *et al.*, 1973; Murry and Doherty, 1980; Sorensen and Horii, 1983). Shimmer values for modal register production in the present study (*M* males: 0.40 dB; *M* females: 0.38 dB) appeared very similar to previous studies. Horii (1980) reported an average shimmer value across adult males of 0.39 dB and Sorensen and Horii (1983) reported an average shimmer value for adult females of 0.25 dB.

Significant differences in S/N ratio, shimmer, and jitter values were present between the modal and fry registers. When compared to modal register, the vocal fry register was associated with significantly reduced S/N ratio and greater jitter and shimmer values. In vocal fry, higher jitter and shimmer values indicate decreased phonatory stability in the frequency and amplitude domains. Decreased S/N ratio in the vocal fry signal indicates a higher noise energy component relative to the periodic energy component when compared to modal register.

B. Aerodynamic measures

1. Rate of airflow

Rate of airflow was significantly reduced in the vocal fry register when compared to the modal register (see Table II). Rate of airflow in the modal register was reported previously to be 70–200 ml/s (Hirano, 1981). Modal register airflow rates in the present study were slightly higher (*M* males=214 ml/s; *M* females=155 ml/s) than those reported by Hirano (1981). Vocal fry airflow results in the present study (*M* males=70 ml/s; *M* females=58 ml/s) were also higher than values reported by McGlone and Shipp (1971) (*M*=40.4 ml/s).

The present results also indicate that the rate of airflow between male and female speakers was not significantly different in vocal fry. This appears to further support the notion of a similar vocal fold vibratory pattern for both male and female speakers during vocal fry when compared to modal voice production (Hollien *et al.*, 1966).

2. Air pressure

Significantly reduced subglottal pressure values were associated with the vocal fry condition for both males and females (see Table II). The present results are in agreement with the findings of Murry and Brown (1971b), who found lower subglottal pressure values for vocal fry than modal register in connected speech. The present subglottal pressure findings, however, do not support the results reported by Murry (1971). Murry (1971) found that subglottal pressure in sustained vocal fry production was significantly higher than that in modal register. The discrepancy in findings may be due to the different measurement techniques used. Murry (1971) used a tracheal puncture technique to measure subglottal pressure. In the present study, subglottal pressure was estimated for the vowel /i/ by measuring and interpolating the peak intraoral pressure during the stop phase of several

/pi/ syllable repetitions. Although one may argue for the increased validity of the more direct tracheal puncture technique, Murry's five subjects were tested in a nonstandard supine position, compared to a more natural upright speaking position. The absence of gravity on true and ventricular vocal folds in the rostral-caudal (inferior-superior) dimension may influence the vocal fold vibrating pattern.

As described above, vocal fry phonation has been characterized by a condition of closely adducted vocal folds with simultaneously (relatively) flaccid vocal fold margins (Zemlin, 1988). The vocal fold margins are flaccid because of decreased longitudinal tension. It is suggested that less air pressure is required to set the flaccid vocal fold margins into vibration when compared to the taut vocal margins in modal register. In addition, it was observed that the subjects in the present study tended to produce the ''best'' vocal fry phonation when their fry production was perceived to be relaxed with no perception of tension. The present findings of decreased subglottal pressure indirectly supports the notion that reduced vocal fold tension may be required to produce vocal fry.

Glottal resistance can also be used to characterize vocal fry and modal registers. Glottal resistance is calculated by dividing the transglottal pressure differential by glottal airflow ($R=\Delta P/Q$) (Smitheran and Hixon, 1981). In the present study, the average resistance values were 11.5-cm H_2O/l/s for males and 17.5-cm H_2O/l/s for females in modal register. In vocal fry, the average resistance values were 21.4-cm H_2O/l/s for males and 30.0-cm H_2O/l/s for females. The average glottal resistance values were higher in vocal fry than in modal register for both males and females. In order to increase glottal resistance, subglottal pressure increases while flow rate remains constant, or flow rate decreases while subglottal pressure remains constant. In the present study, resistance was increased due to a decrease in both subglottal pressure and flow rate, however, flow rate decreased to a greater extent than subglottal pressure.

C. EGG analysis

The results of the EGG waveform analyses revealed that the vibratory patterns associated with vocal fry may be characterized by single opening and closing, or multiple opening and closing patterns during a single vibratory cycle (doublet and triplet cycles). These findings support previous research indicating that the vocal folds are capable of multiple vibration patterns (Hollien and Wendahl, 1968; Moore and von Leden, 1958; Timcke *et al.*, 1959; Whitehead *et al.*, 1984).

Some researchers have suggested that doublet or triplet vibratory patterns occur as a result of ''mixed'' phonation that contains elements of both vocal fry and modal registers (Hollien *et al.*, 1977; Wendahl *et al.*, 1963). We would tend to argue against this position as the vocal fry productions in the present study were consistently identified by the judges in the perceptual task as vocal fry. In addition, our observed low mean frequencies for vocal fry (below 50 Hz for both males and females) mitigates against classification as a mixed type of phonation.

Vocal fry phonation appears to be associated with closely approximated vocal folds, which is supported by the present findings of lower airflow and increased glottal resistance when compared to modal register. During vocal fry, the arytenoid cartilages may be pressed anteriorly and medially causing the posterior portions of the vocal folds to be held together. The vocal folds become short and thick with only the anterior ligamental portions vibrating. The vocal fold margins, however, remain relatively flaccid because of decreased longitudinal tension (Zemlin, 1988, p. 166). Allen and Hollien (1973) and Hollien (1974) reported that mean vocal fold thickness was greater in vocal fry than in modal register. The thicker and more compact true vocal folds may account for the obtained complex-repetitive pattern in some of the vocal fry EGG waveforms. The doublet or triplet patterns may be representative of double or triple oscillations (per cycle) along the free margin of the vocal folds during vocal fry phonation.

Another explanation for the observed complex-repetitive patterns is based on the x-ray image findings of Allen and Hollien (1973). They found that the inferior surfaces of the ventricular folds may contact the superior surface of the true vocal folds during vocal fry phonation. The subsequent ''thick'' vibrating structure may contribute to the similarity in F_0 between males and females. In addition,the complex vibratory activity of the coupled ventricular and true vocal folds may also contribute to the complex-repetitive patterning.

D. Perceptual identification of vocal fry

In the perceptual study, 55 listeners correctly identified 95.5% of the experimentally classified modal productions of /ɑ/ and 100% of the vocal fry productions. These findings indicate that perceptual differences are discerned between the experimentally classified modal and vocal fry productions. The high agreement between the acoustic, aerodynamic, and EGG distinctions and the perceptual distinction of vocal fry and modal productions appear to add validity to the classification of these distinct registers.

IV. CONCLUSIONS

The findings in the present study attempted to increase present knowledge regarding phonation in vocal fry register. Acoustic findings indicated there was a significant difference in F_0 between males and females in modal register, while gender differences in vocal fry register were not observed. For both males and females, jitter and shimmer values in vocal fry were significantly higher than those in modal register indicating greater frequency and amplitude perturbation in vocal fry. The S/N ratio values in vocal fry register were significantly lower than those in modal register, possibly indicating less efficient valving of the vocal fold mechanism in the vocal fry register. In addition, differences in airflow, air pressure, glottal resistance, and EGG characteristics were identified between modal register and vocal fry register for both males and females.

Vocal fold vibration is the result of a delicate balance of subglottal air pressure driving the folds apart, and the mus-

cular, elastic, and Bernoulli forces that bring them together. In modal register, the vocal folds vibrate in a relatively consistent manner, with the movement of a wave along the vocal margin that propagates superiorly. These movements have been characterized by the two-mass model of vocal fold vibration (Ishizaka and Matsudaira, 1968; Titze, 1981). Vibration of the vocal folds in vocal fry appears to be more complex. While a two-mass model provides a succinct account of phonation in modal register, the complex differences observed in vocal fry phonation will likely require description using a multi-mass model. In order to model the complex vibration patterns in both the anterior/posterior plane and superior/inferior plane of the vocal folds, simultaneous recordings using EGG, acoustic, *and* direct visualization incorporating high-speed filming or stroboscopic techniques appears necessary.

The differences between modal and vocal fry registers with respect to acoustic, aerodynamic, and EGG parameters are believed to be due to two distinct aeromechanical modes of vocal fold vibration. During phonation in the modal register, the vocal folds appear to be tightly adducted along the anterior margins, with the presence of a commonly observed posterior glottal chink (Broad, 1973). In vocal fry, the vocal processes of the arytenoids appear to be pulled inward and forward. The posterior glottal chink may be greatly diminished in vocal fry. The thyroarytenoid muscles tend to shorten, and subsequently thicken the antero-posterior dimension of the vocal folds, allowing the anterior part of the vocal ligaments to vibrate (Allen and Hollien, 1973; Hollien, 1974). As the thyroarytenoid muscle shortens, the vocal fold margins also become more flaccid. The thickening of the vocal folds in combination with decreased vocal ligament stiffness during vocal fry may be two prime factors contributing to the reduced rate of vocal fold vibration, which consequently alter the characteristics of the vibratory cycle.

In the present study, vocal fold vibration in the modal register was maintained by subglottal pressure values of approximately 7.5-cm H_2O, which counteracts the longitudinal tension and medial compression of the thyroarytenoid muscles (Borden *et al.*, 1984). In vocal fry, however, the subglottal pressure values were approximately 5.5-cm H_2O, possibly indicating reduced tension and medial compression. The finding of increased glottal resistance in vocal fry may be explained by the increased vocal fold mass, rather than increased longitudinal tension and medial compression. If, indeed, the vocal folds are less tense in vocal fry than in modal register, then it would appear that less subglottal driving pressure is required to force them apart. We further suggest that reduced subglottal pressure is not just a consequence of vocal fry, but that individuals actively reduce subglottal pressure as an aerodynamic strategy to facilitate achieving fundamental frequencies below 60 Hz.

In conclusion, it was found that vocal fry was distinguishable from modal register production in acoustic, aerodynamic, physiologic, and perceptual measures. Future research regarding the precise degree of vocal fold adduction, degree of vocal fold margin stiffness, and the precise movements of the vocal fold margins and contribution of the ventricular folds will help to further describe the interesting laryngeal phenomenon of vocal fry.

ACKNOWLEDGMENTS

Portions of this paper were presented at the 133rd meeting of the Acoustical Society of America, June 1997 in State College, Pennsylvania. We wish to thank Dr. Anders Lofqvist, Dr. Michael Robb, and two anonymous reviewers for their excellent comments on an earlier version of this manuscript.

Allen, E., and Hollien, H. (**1973**). ''A laminagraphic study of pulse (vocal fry) phonation,'' Folia Phoniat. **25**, 241–250.

Borden, G. J., Harris, K. S., and Raphael, L. J. (**1984**). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech* (Williams & Wilkins, Baltimore, MD), 3rd ed.

Broad, D. J. (**1973**). ''Phonation,'' In *Normal Aspects of Speech, Hearing, and Language*, edited by F. D. Minifie, T. J. Hixon, and F. Williams (Prentice-Hall, Englewood Cliffs, NJ).

Childers, D. G., and Lee, C. K. (**1991**). ''Vocal quality factors: Analysis, synthesis, and perception,'' J. Acoust. Soc. Am. **90**, 2394–2410.

Colton, R. H., and Casper, J. K. (**1996**). *Understanding Voice Problems* (Williams & Wilkins, Baltimore, MD), 2nd ed.

Daniloff, R., Shuckers, G., and Feth, L. (**1980**). *The Physiology of Speech and Hearing* (Prentice-Hall, Englewood Cliffs, NJ).

Hirano, M. (**1981**). *Clinical Examination of Voice* (Springer-Verlag, New York).

Hollien, H. (**1974**). ''On vocal registers,'' J. Phon. **2**, 25–43.

Hollien, H., and Jackson, B. (**1973**). ''Normative data on the speaking fundamental characteristics of young adult males,'' Journal of Phonetics **1**, 117–120.

Hollien, H., and Michel, J. (**1968**). ''Vocal fry as a phonational register,'' J. Speech Hear. Res. **11**, 600–604.

Hollien, H., and Wendahl, R. W. (**1968**). ''Perceptual study of vocal fry,'' J. Acoust. Soc. Am. **43**, 506–509.

Hollien, H., Girard, G. T., and Coleman, R. F. (**1977**). ''Vocal fold vibratory patterns of pulse register phonation,'' Folia Phoniat. **29**, 200–205.

Hollien, H., Michel, J., and Doherty, E. (**1973**). ''A method for analyzing vocal jitter in sustained phonation,'' J. Phon. **1**, 85–91.

Hollien, H., Moore, P., Wendahl, R., and Michel, J. (**1966**). ''On the nature of vocal fry,'' J. Speech Hear. Res. **9**, 245–247.

Horii, Y. (**1980**). ''Vocal shimmer in sustained phonation,'' J. Speech Hear. Res. **23**, 202–209.

Ishizaka, K., and Matsudaira, M. (**1968**). ''Analysis of the vibration of the vocal cords,'' J. Acoust. Soc. Jpn. **24**, 311–312.

Krook, M. (**1988**). ''Speaking fundamental frequency characteristics of normal Swedish subjects obtained by glottal frequency analysis,'' Folia Phoniat. **40**, 82–90.

Lofqvist, A., Carlborg, B., and Kitzing, P. (**1982**). ''Initial validation of an indirect measure of subglottal pressure during vowels,'' J. Acoust. Soc. Am. **72**, 633–635.

McGlone, R. (**1967**). ''Air flow during vocal fry phonation,'' J. Speech Hear. Res. **10**, 299–304.

McGlone, R., and Shipp, T. (**1971**). ''Some physiologic correlates of vocal-fry phonation,'' J. Speech Hear. Res. **14**, 769–775.

Moore, G. P., and von Leden, H. (**1958**). ''Dynamic variations of the vibratory pattern in the normal larynx,'' Folia Phoniat. **10**, 205–238.

Murry, T. (**1971**). ''Subglottal pressure and airflow measures during vocal fry phonation,'' J. Speech Hear. Res. **14**, 544–551.

Murry, T., and Brown, W. S. (**1971a**). ''Regulation of vocal intensity during vocal fry phonation,'' J. Acoust. Soc. Am. **49**, 1905–1906.

Murry, T., and Brown, W. S. (**1971b**). ''Subglottal air pressure during two types of vocal activity: Vocal fry and modal phonation,'' Folia Phoniat. **23**, 440–449.

Murry, T., and Doherty, E. (**1980**). ''Selected acoustic characteristics of pathologic and normal speakers,'' J. Speech Hear. Res. **4**, 321–328.

Rothenberg, M. (**1973**). ''A new inverse-filtering technique for deriving the glottal air flow waveform during voicing,'' J. Acoust. Soc. Am. **53**, 1632–1645.

Rothenberg, M. (**1982**). ''Interpolating subglottal pressure from oral pressure,'' Journal of Speech & Hearing Disorders **47**, 218–224.

Smitheran, J. R., and Hixon, T. J. (**1981**). ''A clinical method for estimating air way resistance during vowel production,'' J. Speech Hear. Disord. **46**, 138–146.

Sorensen, D., and Horri, Y. (**1983**). ''Frequency and amplitude perturbation in the voices of female speakers,'' J. Commun. Disord. **16**, 57–61.

Stoicheff, M. L. (**1981**). ''Speaking fundamental frequency characteristics of nonsmoking female adults,'' J. Speech Hear. Res. **24**, 437–441.

Timcke, R., von Leden, H., and Moore, P. (**1959**). ''Laryngeal vibrations: measurement of the glottic wave, part 2. Physiological variations,'' American Medical Assoc. Archives of Otolaryngology **69**, 438–444.

Titze, I. R. (**1981**). ''Biomechanics and disturbed-mass models of vocal fold vibration,'' in *Vocal fold physiology, Proceedings of the vocal fold physiology conference,* edited by K. Stevens and M. Hirano (University of Tokyo, Tokyo, Japan), pp. 245–270.

Wendahl, R., Moore, P., and Hollien, H. (**1963**). ''Comments on vocal fry,'' Folia Phoniat. **15**, 251–255.

Whitehead, R. W., Metz, D. E., and Whitehead, B. H. (**1984**). ''Vibratory patterns of the vocal folds during pulse register phonation,'' J. Acoust. Soc. Am. **75**, 1293–1297.

Zemlin, W. R. (**1988**). *Speech and Hearing Science: Anatomy and Physiology* (Prentice-Hall, Englewood Cliffs, NJ).

Automatic parametrization of differentiated glottal flow: Comparing methods by means of synthetic flow pulses

Helmer Strik[a)]
University of Nijmegen, Department of Language and Speech, P.O. Box 9103, 6500 HD Nijmegen, The Netherlands

(Received 14 July 1997; accepted for publication 14 January 1998)

The automatic parametrization of the first derivative of glottal flow is studied. Representatives of the two types of methods used most often for parametrization were tested and compared. The chosen representatives are all based on the Liljencrants–Fant model. As numerous tests were needed for a detailed comparison of the methods, a novel evaluation procedure is used which consists of the following stages: (1) use the Liljencrants–Fant model to generate synthetic flow pulses; (2) estimate voice source parameters for these synthetic flow pulses; and (3) calculate the errors by comparing the estimated values with the input values of the parameters. This evaluation procedure revealed that in order to reduce the average error in the estimated voice source parameters, the estimation methods should be able to estimate noninteger values of these parameters. The proposed evaluation method was also used to study the influence of low-pass filtering on the estimated voice source parameters. It turned out that low-pass filtering causes an error in all estimated voice source parameters. On average, the smallest errors were found for a parametrization method in which a voice source model is fitted to the flow derivative, and in which the voice source model is low-pass filtered with the same filter as the flow derivative. [S0001-4966(98)03204-4]

PACS numbers: 43.70.Aj, 43.72.Ar [AL]

INTRODUCTION

The technique of inverse filtering has been available for a long time now. This technique, which was first described in Miller (1959), can be used to decompose the speech signal into two components: the voice source and the filter (the vocal tract). In this way an estimate of the glottal volume velocity waveform (U_g) or its first derivative (dU_g) is obtained. For many applications, estimating a voice source signal (either U_g or dU_g) is not enough and the glottal flow signals have to be parametrized. Parametrization of the voice source signals and evaluation of the parametrization methods have received far less attention in the past. That is why we focus on these aspects in this study.

Parametrization of U_g or dU_g can be done in several ways. Often landmarks (like minima, maxima, zero crossings) are detected in the signals (e.g., Sundberg and Gauffin, 1979; Gauffin and Sundberg, 1980, 1989; Alku, 1992; Alku and Vilkman, 1995; Koreman, 1996). Because these landmarks are estimated directly from the voice source signals, these methods will be called direct estimation methods.

Voice source parameters are also calculated by fitting a voice source model to the data (e.g., Ananthapadmanabha, 1984; Schoentgen, 1990; Karlsson, 1992; Strik and Boves, 1992; Fant, 1993; Milenkovic, 1993; Alku *et al.*, 1997). Many different voice source models have been proposed in the literature (see, e.g., Rosenberg, 1971; Fant, 1979; Ananthapadmanabha, 1984; Fant *et al.*, 1985; Fujisaki and Ljungqvist, 1986; Lobo and Ainsworth, 1992; Cummings and Clements, 1995). Because in estimation methods of this kind a model fitting procedure is used, they will be referred to as ''fit estimation'' methods.

Estimation of voice source parameters can be useful for many applications. Although speech synthesis is the application most mentioned, the estimated voice source parameters are also used for fundamental research on speech production (e.g., Ní Chasaide and Gobl, 1993; Strik, 1994; Koreman, 1996). Other applications for which methods to measure voice source behavior could be useful are clinical use, speech analysis, speech coding, automatic speech recognition, and automatic speaker verification and identification. Since most of these applications require that the methods be fully automatic, there is an increasing need for automatic parametrization methods (see, e.g., Fritzell, 1992; Fant, 1993; Ní Chasaide and Gobl, 1993).

The development of an automatic parametrization method constitutes the long term goal of our research. Both direct and fit estimation methods can be made completely automatic. For this reason, and because they are the methods used most often, a representative of the direct estimation method will be compared with a representative of the fit estimation method. The representatives chosen are described in Secs. I E and I F.

The goals of the research reported on in this article are to find out what the pros and cons of each method are, to get a better understanding of the problems involved in estimating voice source parameters, and finally to determine which method performs best. In order to make it easier to compare the two methods, the same voice source model is used in both methods. To this end we use the Liljencrants–Fant (LF) model (Fant *et al.*, 1985). The LF model and the reasons for choosing it are described in Sec. I B. The evaluation method

a)Electronic mail: strik@let.kun.nl

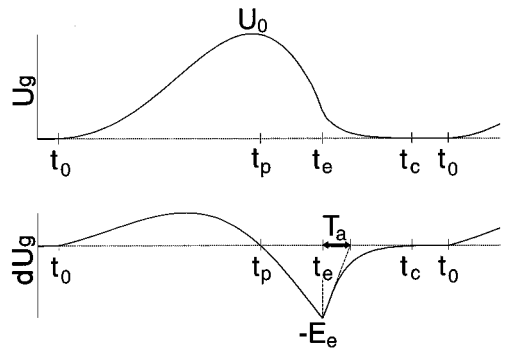


FIG. 1. Glottal flow (U_g) and glottal flow derivative (dU_g) with the parameters of the LF model: time of glottal opening (t_0); time (t_p) and value (U_0) of the maximum of U_g; time (t_e) and absolute value (Ee) of the minimum of dU_g; T_a describes the return phase, it is the length of the time interval between t_e and the projection of the tangent of dU_g in t_e; and the time of glottal closure (t_c).

and material are described in Secs. I C and I D, respectively. Because we want to focus on the parametrization method, we shall not evaluate inverse filtering in the current research. The performance of the parametrization methods is assessed in Secs. II and III. First, in Sec. II, it is studied how well the estimation methods succeed in estimating noninteger values of the parameters, which turned out to be a crucial property. Second, we focus on low-pass filtering in Sec. III. In Sec. IV the findings are discussed and some general conclusions are drawn.

I. GENERAL PROCEDURES

In this article two estimation methods used to parametrize dU_g are tested and compared. Before going on to describe these two methods (in Secs. I E and I F), we shall first give some definitions in Sec. I A, discuss the LF model in Sec. I B, and describe the method and material used for evaluation in Secs. I C and I D, respectively.

A. Definitions

In the current article it will be assumed that dU_g is a digital signal. In order to avoid confusion later on, we shall first define some terms related to sampling and quantization.

For all tests the sampling frequency $F_s=10$ kHz, the number of bits used for quantization $B_c=12$ and the amplitude range is $[-2048,2047]$. Consequently, the sampling time $T_s=1/F_s=1$ ms and the step size $\delta=4096/2^{B_c}=1$. Throughout this article a *time parameter* is said to have an integer value if its value is precisely an integer multiple of T_s. Likewise, an *amplitude parameter* is said to have an integer value if its value is exactly an integer multiple of δ.

B. Liljencrants–Fant model

In the current research the voice source model used is the LF model (see Fig. 1) because the LF model has the following advantages:

(1) In previous research the LF model has often been used to estimate voice source parameters, with manual or (semi-)automatic methods. This research has shown that it is a suitable model for description of the flow derivative (see, e.g., Fujisaki and Ljungqvist, 1986; Karlsson, 1992; Strik and Boves, 1992; Strik *et al.*, 1992, 1993; Childers and Ahn, 1995).

(2) Fujisaki and Ljungqvist (1986) compared several voice source models. Their results showed that the LF model and their own FL-4 model performed best (i.e., had the smallest prediction error).

(3) Previous research has also proven that the LF model is suitable for speech synthesis (see e.g., Carlson *et al.*, 1989).

(4) Due to all research already performed, the model and its behavior are well known.

The parameters shown in Fig. 1, in turn, can be used to derive many other parameters. For instance, the speed quotient is often calculated: $SQ=(t_p-t_0)/(t_c-t_p)$ (e.g., Alku and Vilkman, 1995). However, in our opinion these derived parameters are less suitable for evaluation of the parametrization methods, because whenever there is a change in a derived parameter, it is difficult to determine how this change came about (Strik, 1996). An increase in SQ could be the result of a larger t_p, a smaller t_0, a smaller t_c, or a combination of any of these three changes. On the other hand, whenever a derived parameter remains constant, this does not necessarily imply that the underlying parameters (i.e., the parameters which were used to calculate the derived parameters) remain constant. It is always possible that changes in these underlying parameters cancel each other out. Therefore, we prefer to use the LF parameters specified in Fig. 1 for the evaluation of estimation methods. Since the parameters E_e, t_0, t_p, t_e, and T_a give a complete description of an LF pulse, this set of parameters will be used in this article.

C. Evaluation method

Estimates of voice source parameters can be influenced by a large number of factors. So far, 11 of these factors have been studied: sampling frequency, number of bits used for quantization, position (shift) and amplitude (E_e) of the glottal pulses, t_c, T_0 (length of the fundamental period), signal-to-noise ratio (i.e., the effect of additive noise), phase distortion (which can be caused, e.g., by high-pass filtering), errors in the estimates of formant and bandwidth values during inverse filtering (which will bring about formant ripple in the estimated voice source signals), and low-pass filtering (Strik and Boves, 1994). We have performed over 1000 model fits for each of these 11 factors, making a total of much more than 11 000 model fits. The fact that so many tests had to be performed is the main reason for using the evaluation method described below (other reasons can be found in Strik, 1997).

In our experiments we first synthesize flow pulses (see Sec. I D). As we use the LF model for the fitting procedure, it is obvious that we also used the LF model to synthesize the flow pulses. Subsequently, the parametrization methods are used to estimate the voice source parameters. Finally, the estimated voice source parameters are compared with the correct values (used to synthesize the flow pulses), and the errors are calculated:

TABLE I. Values of t_p, t_e, and T_a (all in ms) for the 11 base pulses.

	Base pulse										
	1	2	3	4	5	6	7	8	9	10	11
t_p	14.0	14.0	16.0	16.0	16.0	16.0	14.0	14.0	15.2	15.2	15.2
t_e	15.2	15.2	17.2	17.2	18.8	18.8	16.0	16.0	17.2	17.2	17.2
T_a	0.4	1.6	0.4	1.6	0.4	0.8	0.4	1.6	0.4	1.0	1.6

$$\mathrm{ERR}(X)=|X_{\mathrm{est}}-X_{\mathrm{inp}}|/X_{\mathrm{inp}}, \quad \text{for } X=E_e$$

$$\mathrm{ERR}(Y)=|Y_{\mathrm{est}}-Y_{\mathrm{inp}}|, \quad \text{for } Y=t_0,\ t_p,\ t_e,\ \text{and } T_a.$$

The experiments were carried out for a number (say N) of test pulses. After calculating the errors in the estimates of the five LF parameters for each test pulse, the errors had to be averaged. This can be done in a number of ways. Generally, averaging was done by taking the median of the absolute values of the errors. Absolute values were taken because otherwise positive and negative errors could cancel each other. The median was taken because (compared to the arithmetic mean) it is less affected by outliers which are occasionally present in the estimates. This method of averaging is the default method in the current article. Whenever another way of averaging was used, this is explicitly mentioned in the text.

In all figures below, the errors are arranged in a similar fashion (see, e.g., Fig. 2). In the upper left corner are the errors for E_e (in%), in the middle row are the errors for t_0 and t_p and in the bottom row are the errors for t_e and T_a. The errors in the time parameters t_0, t_p, t_e, and T_a are expressed in μs.

D. Material

The estimation methods used in this study are pitch synchronous. Among the parameters that have to be estimated are t_0 and t_c. Because these two parameters are not known beforehand, the pitch period cannot be segmented exactly. In practice, we first locate the main excitations (i.e., t_e) and then use a window with a width larger than the length of the longest (expected) pitch period. Generally, the pitch period will be situated between two other pitch periods (except for UV/V and V/UV transitions). Therefore, for each experiment sequences of three equal LF pulses were used. Each time voice source parameters were estimated for the (perturbed) pulse in the middle. Another reason for not using a single glottal pulse for evaluation is that the effects of perturbations cannot always be studied by a single, isolated LF pulse.

Since the effect of a studied factor can depend on the shape of a flow pulse, LF pulses with different shapes were used. These pulses will be called the base pulses. The base pulses were obtained by using the LF model for different values of the LF parameters. The parameters of E_e, T_0, t_0, and t_c were kept constant at 1024, 10 ms, 10 ms, and 20 ms, respectively. The values given for t_0 and t_c are the values for the second of the three pulses. For the first pulse one should subtract 10 ms from the values of t_0 and t_c, and for the last pulse add 10 ms. T_0 and t_c were kept constant because the results of our experiments showed that varying these parameters had very little effect on the estimations. The effects of varying E_e and shift (which is strongly related to t_0) were studied separately (see Sec. II).

For defining the base pulses the values of t_p, t_e, and T_a were varied. Based on the data given in Carlson *et al.* (1989), and the data from previous experiments (Strik and Boves, 1992; Strik *et al.*, 1992, 1993; Strik, 1994) the 11 base pulses shown in Table I were defined.

Subsequently, these 11 base pulses were used to generate the test pulses. For instance, to study the influence of the factor low-pass filtering, the 11 base pulses were filtered with M low-pass filters in order to generate $M\times 11$ test pulses. Calculation of the base pulses and the test pulses was first done in floating point arithmetic. After the test pulses had been created, the sample values were rounded towards the nearest integer (as is done in straightforward A/D conversion).

E. Direct estimation method

In direct estimation methods, voice source parameters are calculated directly from dU_g or U_g by means of simple arithmetic operators like min, max, argmin, and argmax. These arithmetic operators are used to detect landmarks in the signals. Some examples of estimations used quite often are: $U_0=\max(U_g)$, $t_p=\mathrm{argmax}(U_g)$, $E_e=-\min(dU_g)$, and $t_e=\mathrm{argmin}(dU_g)$ (see, e.g., Sundberg and Gauffin, 1979; Ananthapadmanabha, 1984; Gauffin and Sundberg, 1980, 1989; Alku, 1992; Alku and Vilkman, 1995; Koreman, 1996). Except for the value and the place of a maximum or minimum, the place of a zero crossing is also used to estimate parameters. For instance, in this way t_0 and t_c can be estimated (see Fig. 1).

One of the aims of the research reported in this article is to compare the performance of a typical direct estimation method with that of a fit estimation method. To this end we chose the direct estimation method described in Alku and Vilkman (1995), primarily because these authors provide a fairly detailed description of their method (see especially page 765 of their article), and because with this method it was possible to estimate the LF parameters E_e, t_0, t_p, and t_e (for which they use the terms $A_{\min}$, t_0, t_m, and t_{dm}, respectively).

In their method Alku and Vilkman (1995) do not estimate T_a. They use the parameter t_{ret} to describe the return phase. Since T_a cannot be derived from t_{ret} and an LF model is not complete without T_a, another method had to be used to estimate T_a. For the current research all estimates were made in the time domain. Because it is very difficult to estimate T_a in the time domain with a direct estimation method, estimates of T_a were obtained by fitting the LF model to the

glottal pulse. More precisely, for given values of E_e, t_0, t_p, and t_e (made with the direct estimation method) the optimal value of T_a was estimated by fitting the LF model to the data. Therefore, strictly speaking, only E_e, t_0, t_p, and t_e can be said to be the result of the direct estimation method, while T_a is subsequently estimated with a fitting procedure. However, it is important to notice that the estimate of T_a does depend to a large extent on the estimates of E_e, t_0, t_p, and t_e made before with the direct estimation method. Furthermore, estimating one parameter (here T_a) with a fitting procedure, is a relatively simple operation. Consequently, the results showed that the error in the estimates of T_a is mainly the result of the errors in the estimates of E_e, t_0, t_p, and t_e made with the direct estimation method. For instance, if estimates of E_e and/or t_e are too large, the resulting estimates of T_a will generally be too small.

F. Fit estimation methods

In our fit estimation method five LF parameters (E_e, t_0, t_p, t_e, and T_a) are estimated for each pitch period. The method consists of three stages:

(1) initial estimate;
(2) simplex search algorithm;
(3) Levenberg–Marquardt algorithm (Marquardt, 1963).

The goal of the fit estimation method is to determine a model fit which resembles the glottal pulse as much as possible. This resemblance is quantified by means of an error function, which is calculated in the following way. The optimization procedure provides a set of LF parameters. These LF parameters and the analytical expression of the LF model are used to calculate a continuous LF pulse. The LF pulse is then sampled and zeros are added before t_0 and after t_c (until the length of the fitted signal is equal to that of the glottal pulse). These samples of the fitted signal together with the samples of the glottal pulse constitute the input to the error function that provides a measure of the difference between these samples. The fitting procedure tries to minimize this error.

We have experimented with several error functions which were defined either in the time domain, the frequency domain, or in both domains simultaneously. Defining a suitable error function in the frequency domain, for this automatic fitting procedure, turned out to be problematic. Probably the main reason is that the spectrum contains some details (e.g., the harmonics structure, the high-frequency noise) which need not be fitted exactly. With simple error measures, like, e.g., the root-mean-square (rms) error, we did not succeed in obtaining a reasonable model fit. More sophisticated error functions are needed for this task. A suitable error function should abstract away from the details which are not important, and emphasize the important aspects (e.g., the slope of the spectrum).

In the time domain it is much easier to obtain a fairly good model fit of dU_g. Here a simple rms error does yield plausible results. Still, also in the time domain some aspects of dU_g could be more important than others. It is likely that more sophisticated error functions could be defined which emphasize the relevant (e.g., perceptual) aspects. However, what is relevant depends on the application. In the current research we did not have a specific application in mind. The goal of this research was to develop a method for which the error in the estimated voice source parameters is small. Therefore, an important property of the error function is that it should decrease when the errors in the voice source parameters become smaller (this may sound trivial, but it is not). The rms error (defined in the time domain) did have this property and thus was suitable for this task, as our experiments revealed.

For the fitting procedure different nonlinear optimization techniques were tested: several gradient algorithms and some versions of a nongradient algorithm, i.e., the simplex search algorithm of Nelder and Mead (1964). Of the algorithms tested the simplex search algorithms usually came closer to the global minimum than the gradient algorithms. Owing to discontinuities in the error function, gradient algorithms are more likely to get stuck in local minima than simplex search algorithms are. Therefore the best version of the simplex search algorithm is used in the second stage of the fit estimation method. However, in the neighborhood of a minimum, the simplex algorithm may do worse (see Nelder and Mead, 1964). As a final optimization, the Levenberg–Marquardt algorithm (a gradient algorithm) is therefore used in the third stage (Marquardt, 1963).

In order to start the simplex search algorithm of stage 2 an initial estimate is required, which is made in the first stage. In principle, the best available direct estimation method should be used to provide the initial estimate. In this case the rms error for the fit estimation method can never be larger, and will almost always be smaller than the rms error for the direct estimation method used (because in stage 2 and 3 of our fit estimation method the rms error can never increase, and usually decreases gradually). Consequently, the errors in the voice source parameters estimated with the fit estimation method would almost always be smaller than those estimated with the direct estimation method used for initial estimation. Therefore, if we had used the direct estimation method described in the section above for initial estimation, the performance of this direct estimation method would probably have been worse than that of the fit estimation method. Because we considered this to be an unfair starting point, we decided to apply for initial estimation the routine used in our previous research (Strik *et al.*, 1993).

In Sec. III we will introduce a second version of this fit estimation method. This second version differs only slightly from the version described here. Together with the direct estimation method described in Sec. I E, the number of methods studied amounts to three.

Above we already mentioned that so far 11 different factors have been studied. In this article we shall confine ourselves to the most important results, namely those concerning the factors position (shift) and amplitude (E_e) (Sec. II) and those of low-pass filtering (Sec. III).

II. EXPERIMENT 1:SHIFT AND AMPLITUDE

A. Introduction

Direct estimation methods try to locate (important) events in the voice source signals. Thus the resulting esti-

mates are generally limited to the place or amplitude of samples in the discrete signals, i.e., they are integers. Our intention was to develop a fit estimation method that would make it possible to estimate noninteger values too. Here we shall test how well the fit estimation method succeeds in estimating noninteger values of the voice source parameters, and what the resulting errors are for the two estimation methods for different values of shift and amplitude.

B. Material

The definition of the 11 base pulses is such that all time parameters have an integer value (see Sec. I D). In order to create test pulses in which the time parameters did not have integer values, the 11 base pulses were shifted in steps of 0.01 ms, from 0.0 up to 0.1 ms (11 values). This variable will be called *shift*. For only two of the chosen 11 values of shift (i.e., shift=0.0 and 0.1), the time parameters will have an integer value, while for the other 9 values of shift all time parameters will have noninteger values.

In order to create test pulses in which the amplitude (E_e) does not have integer values the amplitude E_e was varied from 1023 to 1025 in steps of 0.2 (11 values). This makes a total of 1331 test pulses (11 base pulses$\times$11 shift values$\times$11 E_e values). Next, the direct estimation method and the fit estimation method were used to estimate the voice source parameters for these 1331 test pulses. The errors in these estimations were then calculated.

C. Results of the direct estimation method

First, the results of the direct estimation method are presented in Figs. 2 and 3. Each error in Fig. 2 is the median of 121 errors (11 base pulses$\times$11 E_e values), while each error in Fig. 3 is the median of another set of 121 errors (11 base pulses$\times$11 shift values).

Let us first look at the errors in Fig. 2. To estimate t_0 a threshold function is used in the direct estimation method. The consequence is that the estimate of t_0 is always much too large (on average about 820 μs; see Fig. 3). For a shift of 0.03 ms the average error in t_0 is minimal, while for a shift of 0.04 ms it suddenly becomes maximal. The reason is that this extra shift of 0.01 ms causes the threshold to be exceeded one sample later in many test pulses, and thus the average error in t_0 suddenly increases. The average errors of the other parameters all behave as expected: the average errors are zero for a shift of 0.0 and 0.1 ms and larger in between.

The errors in the estimates for different values of E_e are shown in Fig. 3. The errors in the time parameters t_0, t_p, and t_e obviously do not depend on the value of E_e. Therefore, the errors for these time parameters are constant. If a large number of moments is randomly distributed, the average error (both the arithmetic mean and the median) due to rounding toward the nearest sample would be $T_s/4=25$ μs. The average errors of t_p, t_e, and T_a do not deviate much from this theoretical average. The reason why the error in t_0 is much larger was already explained above.

The average errors in the estimates of E_e behave as was expected: the average errors are minimal for integer values of E_e, and are larger in between. The median error in E_e is never zero, because it is obtained by averaging over different values of shift, and for most values of shift the error in E_e is larger than zero. The estimate of T_a depends on the estimates of E_e and t_e, and thus is not constant as a function of E_e.

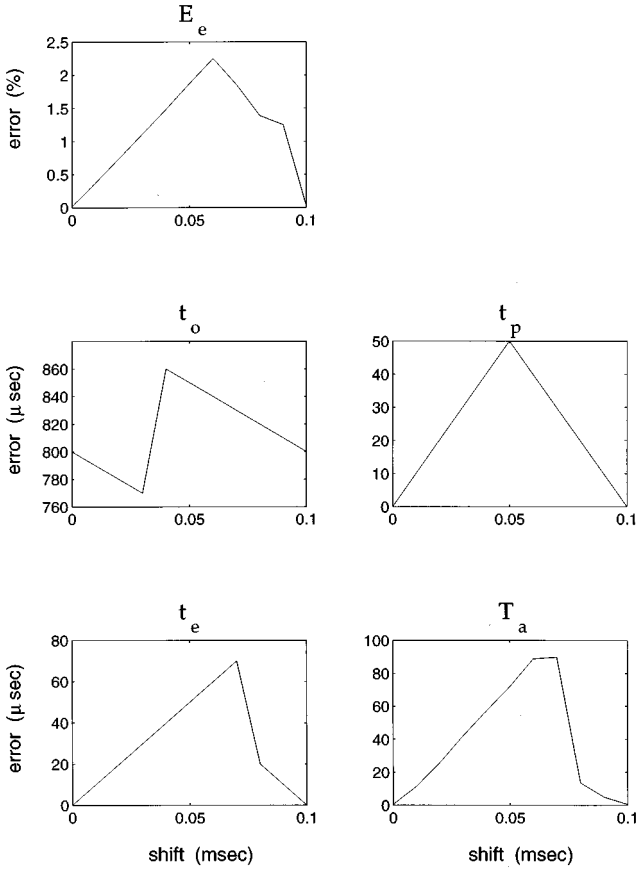


FIG. 2. Results of the direct estimation method: median error for the estimated parameters for different values of shift.

D. Results of the fit estimation method

The resulting average errors for the fit estimation method are shown in Figs. 4 and 5. In this case the errors were averaged by taking the mean value. This was done for two reasons: (1) since there are no outliers, median and mean values do not differ much; (2) by taking the mean it is also possible to calculate standard deviations. In turn, this makes it possible to test whether there is a significant difference between two mean values.

In this case for each value of shift the mean and standard deviation of 121 errors (11 base pulses$\times$11 E_e values) were calculated. The results are shown in Fig. 4. Likewise, for each value of E_e the mean and standard deviation of 121 errors (11 base pulses$\times$11 shift values) were calculated. The results are shown in Fig. 5.

In Figs. 4 and 5 one can observe that the mean errors do not differ significantly from each other. Furthermore, no trend can be observed in the errors. Put otherwise, the magnitude of the error in all estimated parameters does not depend on the value of the factors shift and E_e. Furthermore, all errors are very small, in general much smaller than the

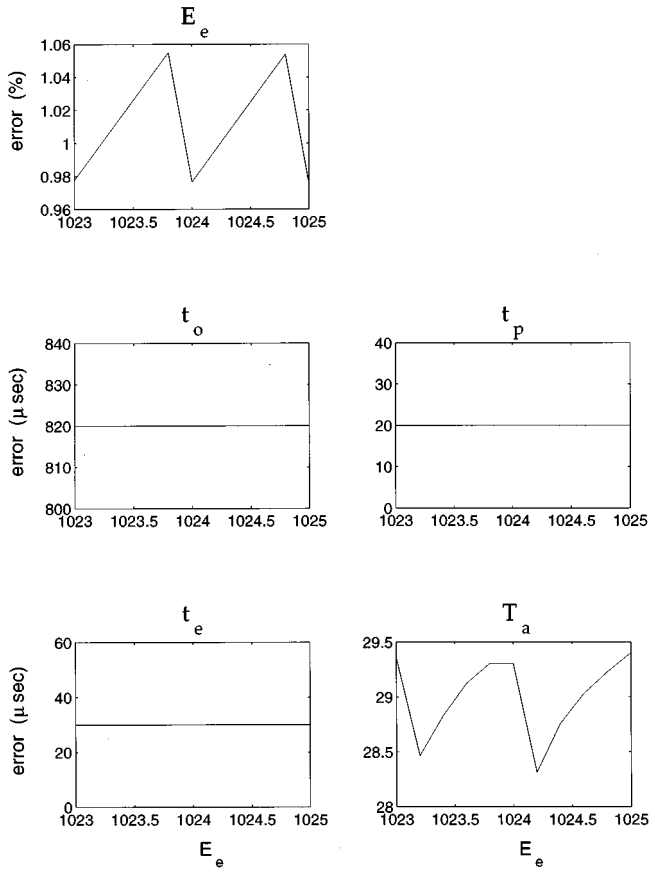


FIG. 3. Results of the direct estimation method: median error for the estimated parameters for different values of E_e.

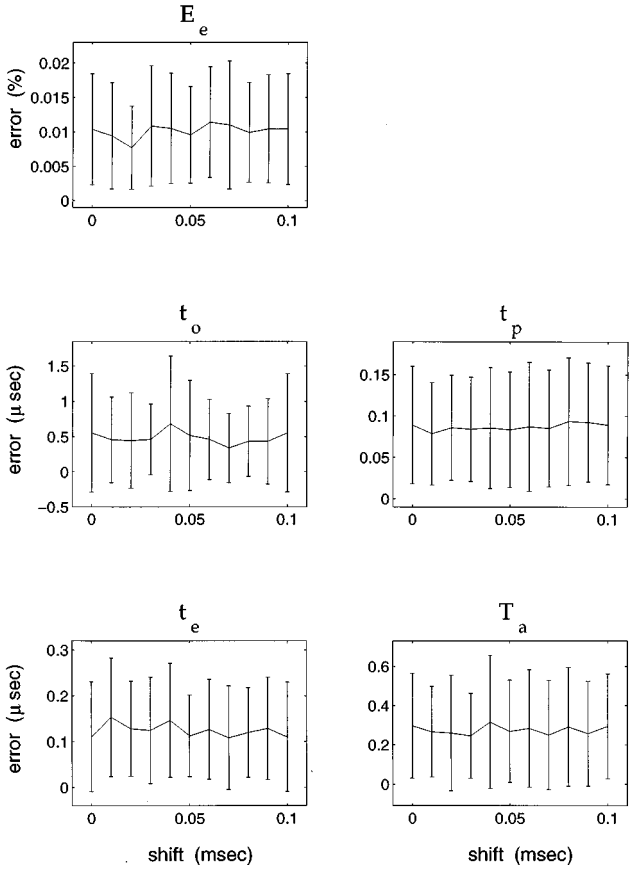


FIG. 4. Results of the fit estimation method: mean and standard deviation of the errors in the estimated parameters for different values of shift.

errors for the direct estimation method. Except of course for the cases in which all the LF parameters have an integer value. In the latter case the errors for the direct estimation method are zero, which is smaller still than the tiny errors found for the fit estimation method. However, it is clear that in practice the voice source parameters will seldom have exactly an integer value.

E. Conclusions

The conclusions that can be drawn from these tests are the following. The errors obtained with the fit estimation method are very small, in general much smaller than those for the direct estimation method. With the fit estimation method noninteger values can be estimated as accurately as integer values. Therefore, the quality of the model fit does not depend on the exact value of E_e and the position of the pulse (which is determined here by the variable shift). This explains why t_0 and E_e could be kept constant in the definition of the base pulses (see Sec. I D).

For the direct estimation method the average errors in t_0 are always larger than for the fit estimation method, because in the former a threshold function is used to estimate t_0. In fact, the error in t_0 can be substantially reduced, simply by subtracting a constant from its estimate. For the other parameters the estimation errors for the direct estimation method are zero if the parameters have exactly an integer value. Since in practice parameters rarely have an integer value, the estimates of the parameters will almost always contain an error due to this fact alone. These errors will be called the intrinsic errors, because they are intrinsic to the estimation methods. They will always be present, even if the glottal pulses are perfectly clean glottal pulses, as was the case in these tests. The results presented in this section make it possible to estimate what the average intrinsic errors are. For the direct estimation method the average error in the time parameters (except t_0) is about $T_s/4=25$ μs, which is the theoretical average for randomly distributed values, while for E_e it is about 1% (see Fig. 3). For the fit estimation method the average error in the time parameters is less than 0.5 μs, while the average error for E_e is about 0.01% (see Figs. 4 and 5).

III. EXPERIMENT 2: LOW-PASS FILTERING

A. Introduction

Before the glottal flow signals are parametrized, they are low-pass filtered at least once in all methods, viz., before A/D conversion. Often, they are low-pass filtered again after A/D conversion, usually to cancel the effects of formants that were not inverse filtered or to attenuate the noise component. The latter operation seems very sensible for direct estimation methods, because in these methods high-frequency disturbances can influence the estimated parameters to a large extent. Although parametrization of inverse filtered signals has been done in many studies for almost 40 years now (i.e.,

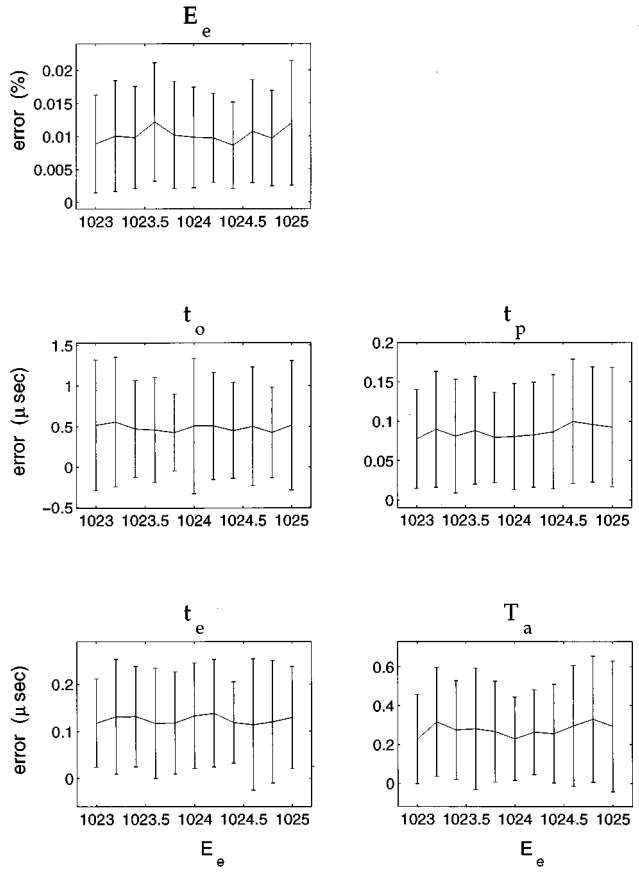


FIG. 5. Results of the fit estimation method: mean and standard deviation of the errors in the estimated parameters for different values of E_e.

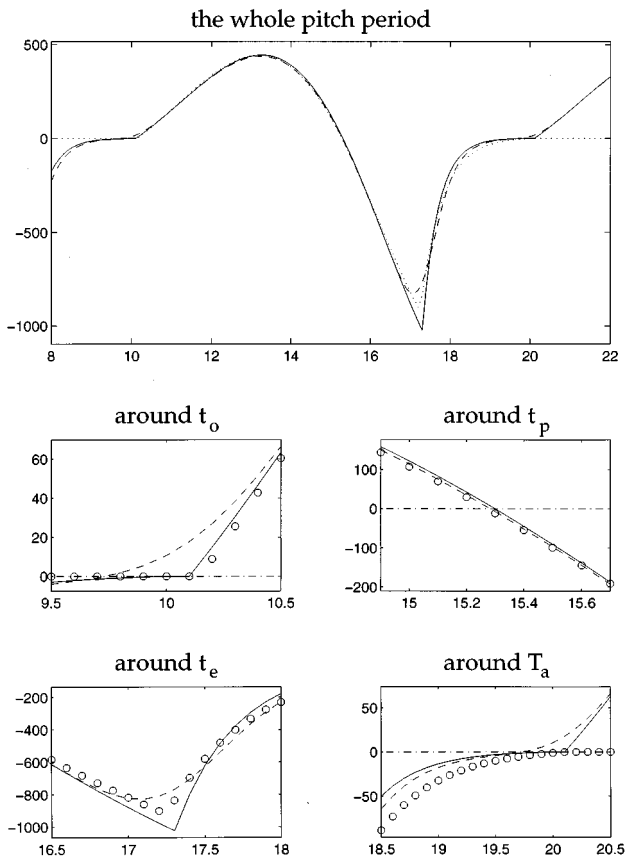


FIG. 6. An example of a differentiated flow pulse before (solid) and after (dashed) low-pass filtering, and a fit on the low-pass filtered pulse (dots in the top panel, open circles in the lower four panels). Shown are the whole pitch period, and some details around important events. For clarity, the zero line (dashed-dotted) has been omitted in the top panel.

since Miller, 1959), it has only recently been noted that low-pass filtering can influence the estimated voice source parameters (Strik *et al.*, 1992, 1993; Perkell *et al.*, 1994; Alku and Vilkman, 1995; Strik, 1996; Koreman, 1996). Thus it becomes very important to study what the effect of low-pass filtering exactly is. This will be done in the present section.

An example of the distortion of a differentiated flow pulse caused by low-pass filtering is given in Fig. 6. For low-pass filtering a convolution with a 19-point Blackman window was used. Shown are a base pulse before (solid) and after (dashed) low-pass filtering, and a model fit on the low-pass filtered pulse (dotted). Besides a picture of the three signals for the whole pitch period, some details around important events are also provided.

One can see in Fig. 6 that low-pass filtering does influence the shape of the pulse. From this figure one can deduce that the change in shape can have a large impact on the estimates obtained by means of a direct estimation method. This is most clear for the estimate of E_e, which will generally be too small, but the estimates of the other parameters will also be affected.

Low-pass filtering will also affect the estimates of a fit estimation method. After low-pass filtering the shape of the pulse is changed. The fitting procedure will try to find an LF pulse that resembles the filtered pulse as closely as possible. This is done by minimizing the rms error, which is a measure of the difference between the test pulse and the fitted LF pulse. The result is a fitted LF pulse that deviates from the original base pulse (see Fig. 6).

The distortion of the differentiated glottal flow signals depends on a number of factors, like, e.g., the type and the bandwidth of the low-pass filter, the frequency contents of the differentiated glottal flow signals, and the parametrization method used. We will study the effect of low-pass filtering for two parametrization methods (i.e., the direct estimation and the fit estimation method), for glottal pulses with different frequency contents (i.e., the 11 base pulses), and for different values of the bandwidth of the low-pass filter.

Low-pass filtering is done by means of a convolution with a Blackman window.[1] The bandwidth of this low-pass filter is varied by changing the length of the Blackman window (the longer the window, the smaller the bandwidth). This type of low-pass filtering was chosen because preliminary tests had shown that the error in the estimates induced by this filter was smaller than that of other tested filters. In part this can be explained by the fact that this low-pass filter does not have a ripple in its impulse response, while a ripple is present for many other low-pass filters. Therefore, for most other low-pass filters (including the generally used standard FIR filters) the estimation errors will be (much) larger than the errors presented below (Strik, 1996).

In the example provided in Fig. 6 the test signal is low-pass filtered. An LF model is then fitted to the low-pass

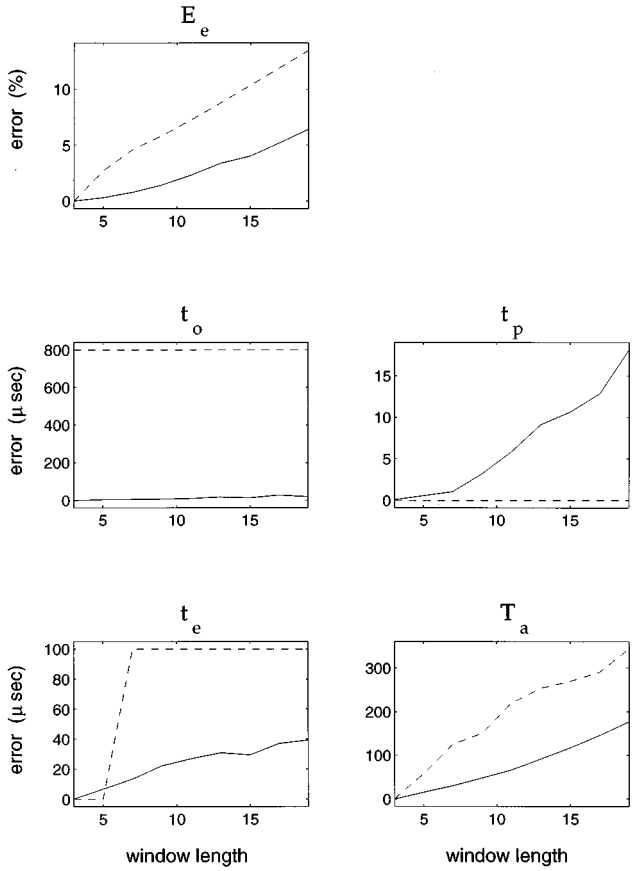


FIG. 7. Median errors in the estimated voice source parameters due to low-pass filtering by means of a convolution with a Blackman window. The length of the Blackman window varies from 3 to 19 in steps of 2. Shown are the errors for the direct estimation method (dashed) and for the first version of the fit estimation method (solid).

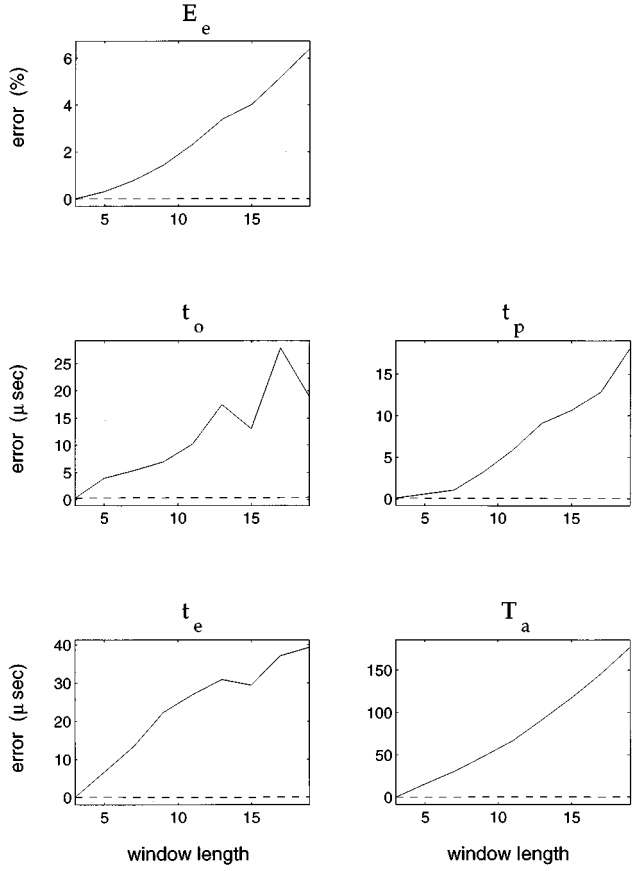


FIG. 8. Median errors in the estimated voice source parameters due to low-pass filtering by means of a convolution with a Blackman window. The length of the Blackman window varies from 3 to 19 in steps of 2. Shown are the errors for the first (solid) and the second (dashed) version of the fit estimation method. Note that the vertical scales are different from those in Fig. 7.

filtered test pulse. This seems the most obvious way to apply the fit estimation method, and will be called the first version of the fit estimation method. However, there is an alternative (which will be called the second version of the fit estimation method): apart from the test pulse one could also low-pass filter the fitted LF pulse. In this case, the test pulse and fitted LF pulse are altered in a similar fashion. In this way we hope to achieve that the error in the estimated parameters (which is due to low-pass filtering) will be smaller than when only the test pulses are low-pass filtered. It is obvious that the same procedure cannot be used in a direct estimation method, because in this case the parameters are calculated directly from the (low-pass filtered) signal.

B. Material

The 11 base pulses were low-pass filtered by means of a convolution with a Blackman window. The length of the window was varied from 3 to 19 samples in steps of 2 samples (9 lengths). For the resulting 99 test pulses (11 base pulses$\times$9 window lengths) the parameters were estimated with the direct estimation method and the fit estimation method. For each length of the Blackman window the results of the 11 base pulses were pooled and the median values of the absolute errors were calculated. These median values are shown in Figs. 7 and 8.

C. Results of the direct estimation method

In Fig. 6 one can see that low-pass filtering has most effect on the amplitude of the signal (E_e) and the shape of the return phase. Low-pass filtering causes the excitation peak to be smoother, and thus the estimate of E_e will be too small. Low-pass filtering also makes the return phase less steep, and therefore the estimate in T_a too large. These effects are enhanced if the length of the Blackman window increases (i.e., if the bandwidth of the low-pass filter is reduced). Therefore, the median errors of E_e and T_a increase with increasing window length.

Low-pass filtering does not have much influence on t_p (=the position of the zero crossing in dU_g; see Fig. 6). Therefore, in the majority of the cases the error in the estimates remains within half a sample, and the median of the errors is zero.

Usually, low-pass filtering causes the estimates of t_e to be too small (see Fig. 6). If the window length is 3 or 5, most of the errors in t_e remain within half a sample, and thus the median error is zero. However, for larger window lengths the errors in t_e become larger. As a result the median error increases too.

Finally, the error in t_0 remains constant, at the value of 820 μs (see also Fig. 3). This can be explained with the help

of Fig. 6. In this figure one can see that low-pass filtering has a large effect on the signal in the direct neighborhood of t_0, and that this effect diminishes away from t_0. If the threshold chosen is high enough (which is the case for the direct estimation method used in the current research), low-pass filtering will not have much influence on this estimate of t_0.

D. Results of the fit estimation method

In Fig. 7 not only the errors of the direct estimation method are presented, but also those of the first version of the fit estimation method (i.e., the version in which only the test pulses were low-pass filtered). If the median errors of the fit estimation method are compared with those of the direct estimation method, the following observations can be made:

(i) The median errors are larger for t_p for all window lengths, and for t_e for windows with a length of 3 or 5.

(ii) In all other cases the errors of the first version of the fit estimation method are smaller than those of the direct estimation method.

The fact that in certain cases the error of the direct estimation method is smaller than the error of the fit estimation method can be explained quite easily. If the effect of a studied phenomenon (here low-pass filtering) on an event (here t_p or t_e) is such that the event is shifted by less than half a sample, the error with the direct estimation method is zero, while that of the fit estimation method is larger than zero. However, one should keep in mind that this is only the case for pulses in which all events coincide exactly with a sample position, as is the case with the test pulses. Only in this case does rounding towards the nearest sample position mean rounding towards the correct value.

In Fig. 8 the results of the two versions of the fit estimation method are compared, i.e., the first version, in which only the test pulses are low-pass filtered (solid lines), and the second version, in which both test pulses and fitted LF pulses are low-pass filtered (dashed lines). Clearly, the errors for the second version are much smaller. The errors are not zero, as may seem to be the case from Fig. 8, but they are extremely small. The largest error observed in the time parameters is 1 μs, and the errors in E_e are always smaller than 0.03%.

E. Conclusions

From our research we can conclude that low-pass filtering changes the shape of the flow pulses, and thus affects the estimates of all voice source parameters. The error due to low-pass filtering does depend on a lot of factors, e.g., the shape of the flow derivative, the low-pass filter and the estimation method used. So even for a given low-pass filter and estimation method (i.e., within one experiment) the error is not constant, because the shape of the glottal pulses is generally not constant. Furthermore, for a low-pass filter with a ripple in its impulse response (like the often used standard FIR filters) the average errors will be larger than for the low-pass filter used in this study, i.e., a convolution with a Blackman window (Strik, 1996).

Generally, the errors for the direct estimation method are larger than those of the first version of the fit estimation method. In turn, these errors are larger than the errors of the second version of the fit estimation method. Therefore, the conclusion is that the second version of the fit estimation method is superior. Low-pass filtering both the test pulse and the fitted voice source model seems to be a very good way to reduce the error caused by low-pass filtering. Of course, it cannot be used in a direct estimation method (as was already noted above).

IV. DISCUSSION AND GENERAL CONCLUSIONS

Before we draw our conclusions regarding the comparison of the three estimation methods, we first discuss some aspects of the fit estimation methods used in this study. The first aspect is the voice source model used in the fit estimation method, in our case the LF model. In the literature several voice source models have been described (see, e.g., Rosenberg, 1971; Fant, 1979; Ananthapadmanabha, 1984; Fant *et al.*, 1985; Fujisaki and Ljungqvist, 1986; Lobo and Ainsworth, 1992; Cummings and Clements, 1995). All voice source models for which an analytical expression exists can be used with the proposed fit estimation method to parametrize either U_g or dU_g. In our program there is a subroutine which calculates the fitted signal. The model fit is now calculated with the LF model, but this part can easily be replaced by the analytical expression of any voice source model. Furthermore, any number of voice source parameters can be used for parametrization. However, increasing the number of parameters makes the optimization problem (i.e., the error space) more complex, thus increasing the probability that the fitting procedure gets stuck in a local minimum.

Using a voice source model for parametrization has some advantages, one of them being the possibility that the estimated voice source parameters can subsequently be used for speech synthesis. Of course, for fit estimation methods a voice source model is mandatory. However, probably the most important disadvantage of a voice source model used for this purpose is that it cannot describe all the observed glottal pulses. Although the LF model is capable of describing many different glottal pulse shapes, it cannot describe all details. Whether a voice source model is suitable for a certain type of research depends on the goals of this research. Above we explained that with our fit estimation method it is possible to use many voice source models. The reasons for choosing the LF model in this study are given in Sec. I B.

The second aspect of the fit estimation method we want to discuss concerns the properties of the LF routine, which is the routine used to calculate the LF pulses. The way in which the LF routine is implemented turned out to be extremely important. The first version of our LF routine was taken from Lin (1990). Since in this version all input parameters are rounded toward the nearest integer, the shapes of the resulting LF pulses do not change gradually but abruptly. The consequence is that also the calculated rms error jumps from one value to the next. Thus the error function has the shape of a staircase, which is problematic for many optimization algorithms: they often get stuck in a local minimum. This is especially the case for gradient algorithms, because the gradient is zero for each stair.

In the second version of the LF routine, oversampling was used within the LF routine. For instance, we tried over-

sampling by a factor 10. Thus not only integer values can be estimated, but also nine values between these integers. However, the error function still has the shape of a staircase. Since the stairs are ten times smaller (compared to the first version of the LF routine), the resulting estimates were better. Still, the optimization often did not come close to the global minimum.

Our conclusion is that oversampling can reduce the width of the stairs in the error function, and thus improve the estimates, but it can never take away the fundamental problem for optimization, i.e., that the error function is a staircase. That is why we tried to find an implementation of the LF routine for which the error function changes smoothly. This property will be called the ''smooth property.'' The third version of the LF routine, which is described in Sec. I F, did have this property. In this version the analytical expression of the LF model is used to calculate a continuous LF pulse, which is then sampled. An enormous improvement in the fit estimation method was observed when the third version of the LF routine was used (compared to the first and second version). The reason is that a smooth error function is an enormous advantage for both simplex search and gradient algorithms. All results presented in this article are obtained with the third version of the LF routine.

The third aspect of the fit estimation method which will be discussed is that no anti-aliasing low-pass filter is used. In the LF routine a continuous LF pulse is first calculated and is then sampled with the same sampling frequency (F_s) as the flow derivative which has to be parametrized (here, 10 kHz). We did not use an anti-alias low-pass filter here, because we wanted to be able to study each factor in isolation. If we had used an anti-alias low-pass filter, this factor (and its effect on the estimated voice source parameters) would always have been present, thus making it impossible to study it independently of other factors.

If no anti-aliasing low-pass filter is used, aliasing effects can be present in the digital signals. Careful inspection showed that this was not the case for the LF pulses used in this study. The dU_g signals on average have a slope of -6 dB/oct. The first fundamental is at 100 Hz, so at 5 kHz the attenuation is usually more than 30 dB. Using a F_s of 10 kHz made it possible to study the effect of the factor low-pass filtering independently of other factors (like, e.g., shift and E_e).

If aliasing is a problem (e.g., because F_s is smaller than 10 kHz), an anti-alias low-pass filter has to be used. The most straightforward way to do this is to sample the continuous LF signal first with a sampling frequency F_s, and next use a digital low-pass filter with a bandwidth smaller than $F_s/2$. However, in that case the smooth property is lost, and the error function (which quantifies the difference between the LF signal and the flow derivative) becomes a staircase. The result is that the average error in the estimated voice source parameters becomes larger, as mentioned above. A somewhat better solution is to oversample the LF signal before digital low-pass filtering. By oversampling noninteger values can also be estimated. Furthermore, the stairs of the staircase become smaller. Consequently, the average error in the estimated voice source parameters also becomes smaller. Probably the best solution would be to use the analytic anti-alias low-pass filter proposed by Milenkovic (1993), which can be applied in continuous time. In this way the smooth property is preserved, and the error function remains a function that changes smoothly (instead of being a staircase).

In the current study two factors were studied in detail. As parameters rarely have an integer value, we first estimated what the resulting intrinsic errors are for the two methods. For the direct estimation method they turned out to be much larger than for the fit estimation method.

Next, the effect of the factor low-pass filtering was studied independently, i.e., with all input parameters having an integer value. For low-pass filtering we found that the errors of the direct estimation method are sometimes smaller than those of the fit estimation method. However, if the important events had been positioned randomly, the errors of the fit estimation method would have been slightly larger while those of the direct estimation method would have been substantially larger. For a realistic comparison of the two methods the intrinsic errors should be added to the errors found for low-pass filtering alone. If this is done the average errors of the direct estimation method are always larger than those of the first version of the fit estimation method, and these in turn are larger than the average errors of the second version of the fit estimation method.

The conclusion which can be drawn on the basis of the tests presented in this article is that the second version of the fit estimation method is superior. However, the effect of more single factors and factors in combination should be studied to get a more thorough understanding of the intricacies of the various parametrization methods.

In order to test and compare the parametrization methods we have used a novel evaluation method in which synthetic test material is generated by a production model. Subsequently, the same production model is used to re-estimate the synthesis parameters. This evaluation method turned out to be useful for our research, e.g., it helped us find the importance of the properties of the implementation of the LF routine and the effects of the factor low-pass filtering. We are convinced that with other evaluation methods this would have been much more difficult or even impossible (see also Strik, 1996).

Since in the present research we want to focus on the estimation of voice source parameters from the flow derivative, without being distracted by the problems of inverse filtering, we use a voice source model (the LF model) as the production model. For other purposes a vocal tract model or a complete synthesizer could be used.

A similar method was used by McGowan (1994) to evaluate the estimation of vocal tract parameters. In our research, just as in McGowan's work (1994), all details of the generating procedure are explicitly known. We therefore agree with him that these kinds of studies should be regarded as best case studies which can be used to study the limitations of estimation procedures and to optimize these estimation procedures. There are two other reasons why the present study is a best case study. First of all, because the test signals are clean LF pulses, and besides the influence of low-pass filtering contain none of the other disturbances that are gen-

erally present in natural speech. And second, because for a standard FIR filter, which is used most often as a low-pass filter, the resulting average errors are larger than for the low-pass filter used in this study. Consequently, when estimation methods are used to parametrize inverse filtered natural speech signals, the errors in the resulting parameters will generally be (much) larger.

The final topic we want to discuss is how the proposed estimation methods can be used to estimate voice source parameters for natural speech. The answer is straightforward: first use inverse filtering to obtain estimates of the glottal flow signals, and then apply the estimation methods. In Strik and Boves (1992) and Strik *et al.* (1992) we showed that this is possible for previous versions of the fit estimation method. We only have to exchange the previous version of the fit estimation method with the new improved version. The best solution would be to take the second version of the fit estimation method, and in the error routine use the same low-pass filter as used during the inverse filter procedure.

ACKNOWLEDGMENTS

The research of Dr. H. Strik has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences. I would like to thank Loe Boves, Bert Cranen, and Jacques Koreman for fruitful discussions. Furthermore, I am grateful to Paavo Alku, Anders Löfqvist, and an anonymous reviewer for their useful comments on a previous version of this paper.

[1]This idea was suggested to me by Bert Cranen.

Alku, P. (**1992**). ''An automatic method to estimate the time-based parameters of the glottal pulseform,'' Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, San Francisco, CA, Vol. 2, 29–32.

Alku, P., and Vilkman, E. (**1995**). ''Effects of bandwidth on glottal airflow waveforms estimated by inverse filtering,'' J. Acoust. Soc. Am. **98**, 763–767.

Alku, P., Strik, H., and Vilkman, E. (**1997**). ''Parabolic Spectral Parameter—A new method for quantification of the glottal flow,'' Speech Commun. **22**, 67–79.

Ananthapadmanabha, T. V. (**1984**). ''Acoustic analysis of voice source dynamics,'' Speech Transmiss. Lab. Q. Prog. Stat. Rep. **2-3**, 1–24.

Carlson, R., Fant, G., Gobl, C., Granström, B., Karlsson, I., and Lin, Q. (**1989**). ''Voice source rules for text-to-speech synthesis,'' Proceedings of the IEEE International Conference on Acoustic Speech Signal Process, Glasgow, Scotland, Vol. 1, 223–226.

Childers, D. G., and Ahn, C. (**1995**). ''Modeling the glottal volume-velocity waveform for three voice types,'' J. Acoust. Soc. Am. **97**, 505–519.

Cummings, K. E., and Clements, M. A. (**1995**). ''Analysis of the glottal excitation of emotionally styled and stressed speech,'' J. Acoust. Soc. Am. **98**, 88–98.

Fant, G. (**1979**). ''Glottal source and excitation analysis,'' Speech Transmiss. Lab. Q. Prog. Stat. Rep. **1**, 70–85.

Fant, G. (**1993**). ''Some problems in voice source analysis,'' Speech Commun. **13**, 7–22.

Fant, G., Liljencrants, J., and Lin, Q. (**1985**). ''A four-parameter model of glottal flow,'' Speech Transmiss. Lab. Q. Prog. Stat. Rep. **4**, 1–13.

Fritzell, B. (**1992**). ''Inverse filtering,'' J. Voice **6**, 111–114.

Fujisaki, H., and Ljungqvist, M. (**1986**). ''Proposal and evaluation of models for the glottal source waveform,'' Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Tokyo, Japan, Vol. 4, 1605–1608.

Gauffin, J., and Sundberg, J. (**1980**). ''Data on the glottal voice source behavior in vowel production,'' Speech Transmiss. Lab. Q. Prog. Stat. Rep. **2-3**, 61–70.

Gauffin, J., and Sundberg, J. (**1989**). ''Spectral correlates of glottal voice source waveform characteristics,'' J. Speech Hear. Res. **32**, 556–565.

Karlsson, I. (**1992**). ''Analysis and synthesis of different voices with emphasis on female speech,'' Ph.D. dissertation, KTH, Stockholm.

Koreman, J. (**1996**). ''Decoding linguistic information in the glottal airflow,'' Ph.D. dissertation, University of Nijmegen.

Lin, Q. (**1990**). ''Speech production theory and articulatory speech synthesis,'' Ph.D. dissertation, KTH, Stockholm.

Lobo, A. P., and Ainsworth, W. A. (**1992**). ''Evaluation of a glottal ARMA model of speech production,'' Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, San Francisco, CA, Vol. 2, 13–16.

Marquardt, D. (**1963**). ''An algorithm for least-squares estimation of nonlinear parameters,'' SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. **11**, 431–441.

McGowan, R. (**1994**). ''Recovering articulatory movement from formant frequency trajectories using task dynamics and a genetic algorithm: Preliminary model tests,'' Speech Commun. **14**, 19–48.

Milenkovic, P. H. (**1993**). ''Voice source model for continuous control of pitch period,'' J. Acoust. Soc. Am. **93**, 1087–1096.

Miller, R. L. (**1959**). ''Nature of the Vocal Cord Wave,'' J. Acoust. Soc. Am. **31**, 667–677.

Nelder, J. A., and Mead, R. (**1964**). ''A simplex method for function minimization,'' Comput. J. (Switzerland) **7**, 308–313.

Ní Chasaide, A., and Gobl, C. (**1993**). ''Contextual variation of the vowel voice source as a function of adjacent consonants,'' Language and Speech **36**, 303–330.

Perkell, J. S., Hillman, R. E., and Holmberg, E. B. (**1994**). ''Group differences in measures of voice production and revised values of maximum airflow declination rate,'' J. Acoust. Soc. Am. **96**, 695–698.

Rosenberg, A. E. (**1971**). ''Effect of glottal pulse shape on the quality of natural vowels,'' J. Acoust. Soc. Am. **49**, 583–590.

Schoentgen, J. (**1990**). ''Non-linear signal representation and its application to the modelling of the glottal waveform,'' Speech Commun. **9**, 189–201.

Strik, H. (**1994**). ''Physiological control and behavior of the voice source in the production of prosody,'' Ph.D. dissertation, University of Nijmegen.

Strik, H. (**1996**). ''Comments on ''Effects of bandwidth on glottal airflow waveforms estimated by inverse filtering'' [J. Acoust. Soc. Am. **98**, 763–767 (1995)],'' J. Acoust. Soc. Am. **100**, 1246–1249.

Strik, H. (**1997**). ''Automatic parametrization of voice source signals: A novel evaluation procedure is used to compare methods and test the effects of low-pass filtering,'' Internal Report, University of Nijmegen (available at http://lands.let.kun.nl/TSpublic/strik/).

Strik, H., and Boves, L. (**1992**). ''On the relation between voice source parameters and prosodic features in connected speech,'' Speech Commun. **11**, 167–174.

Strik, H., and Boves, L. (**1994**). ''Automatic estimation of voice source parameters,'' Proc. Int. Conf. Spoken Language Process., Yokohama, Japan, Vol. 1, 155–158.

Strik, H., Cranen, B., and Boves, L. (**1993**). ''Fitting an LF-model to inverse filter signals,'' Proc. of the 3rd European Conf. on Speech Technology, Berlin, Germany, Vol. 1, 103–106.

Strik, H., Jansen, J., and Boves, L. (**1992**). ''Comparing methods for automatic extraction of voice source parameters from continuous speech,'' Proc. Int. Conf. Spoken Language Process., Banff, Canada, Vol. 1, 121–124.

Sundberg, J., and Gauffin, J. (**1979**). ''Waveforms and spectrum of the glottal voice source,'' in *Frontiers of Speech Communication Research*, Festschrift for Gunnar Fant, edited by B. Lindblom and S. Öhman (Academic, London), pp. 301–320.

"How to milk a coat:" The effects of semantic and acoustic information on phoneme categorization

Susan Borsky
Program in Complex Systems and Brain Sciences, Center for Complex Systems and Department of Psychology, Florida Atlantic University, Boca Raton, Florida 33431
and Department of Communicative Disorders, San Diego State University, San Diego, California 92182

Betty Tuller
Program in Complex Systems and Brain Sciences, Center for Complex Systems and Department of Psychology, Florida Atlantic University, Boca Raton, Florida 33431

Lewis P. Shapiro
Department of Communicative Disorders, San Diego State University, San Diego, California 92182

(Received 6 June 1997; revised 17 November 1997; accepted 19 December 1997)

This study examined the effect of sentence context and local acoustic structure on phoneme categorization. Target stimuli from a 10-step GOAT–COAT continuum, differing only on a temporal cue for voice onset time (VOT), were embedded in carrier sentences that biased interpretation toward either ''goat'' or ''coat.'' While subjects listened to the sentences they also responded as quickly as possible to a visual probe by indicating whether the probe matched the target stimulus they heard. Results showed that the interaction of VOT and sentence context significantly affected both identification and RT for stimuli near the perceptual boundary; the identification function showed a boundary shift in favor of the biased context and peak response times for each context reflected the shifted identification boundaries. In addition, response times were faster for identification of stimuli near the category boundary when responses were congruent, rather than incongruent with the sentence context. The response time differences for congruent versus incongruent responses in the boundary region are interpreted as depending on the results of initial phonological analysis; potentially ambiguous categorizations may be subject to additional evaluation in which a context-congruent response is both preferred and available earlier. © *1998 Acoustical Society of America.* [S0001-4966(98)02405-9]

PACS numbers: 43.71.An, 43.71.Es, 43.71.Hw [WS]

INTRODUCTION

The perception of speech sounds is an important part of understanding spoken language. A speaker's message must be recovered by the listener, even though the continuously changing acoustic signal that comprises the speech stream does not stand in one-to-one correspondence with the words and meanings of an utterance. Instead, perception flexibly adjusts to changing contexts so that different acoustic signals can yield equivalent categorizations and identical acoustic signals may be categorized differently (Liberman *et al.*, 1967, 1957). Experimental studies of nonlinguistic contextual influences on phoneme identification abound (see Repp and Liberman, 1990 for a good review), but less is known about the specific role of other linguistic information in phoneme identification.

Linguistic context alone can guide identification when acoustic information is absent, as in the phoneme restoration effect (Warren, 1970). When part of an utterance is replaced by noise, such as a cough, listeners automatically and unconsciously fill in the missing segment. In normal conversation, however, both the acoustic object and its context are usually available to the listener. The relative contribution of these two factors has been investigated by presenting speech sounds in varying signal-to-noise ratios under different context conditions and to different populations. For young adults, the presence of lexical, syntactic, and semantic constraints increases the probability of recognizing degraded speech sounds; more highly constrained (more predictable) contexts produce greater effects (Boothroyd and Nittrouer, 1988). Although children and older adults do not recognize speech in noise as accurately as do young adults, and they differ in the effectiveness of semantic and lexical constraints, all three groups improve when contextual constraints are available (Nittrouer and Boothroyd, 1990).

The demonstration of faster detection of a target phoneme in words than in nonwords (Rubin *et al.*, 1976) provided early evidence that linguistic context influences phoneme recognition even when the signal is not degraded. Phoneme categorization is also facilitated when the phoneme appears as part of a word. For example, in continua in which the endpoint stimuli begin with voiced or voiceless stop consonants, but only one forms a word (e.g., DASH–TASH, DASK–TASK), identifications of boundary stimuli show a strong word bias in that the same acoustic stimulus elicits a greater probability of identification as a word than a nonword (Ganong, 1980). Although average response times are faster to endpoint stimuli than to those at the boundary, response times to boundary stimuli are faster when the identifications result in words rather than nonwords (Connine and Clifton, 1987).

One possibility is that the effect of lexical status on phonemic processing actually reflects a general attentional effect, not a specifically linguistic one. Connine and Clifton (1987) examined identification and response times to nonword stimulus continua (e.g., DICEL–TICEL), but biased the voiced or voiceless response by a monetary payoff scheme. In the voiced-biased condition, for example, voiced responses were rewarded and voiceless responses were penalized. There was also a small penalty for responses that were ''acoustically incorrect'' according to the individual subjects' identifications on an unbiased pre-test of the same stimuli. The results were a shift in the identification boundary in favor of the biased response (i.e., the point on the voicing continuum that would correspond to a 50% identification of either percept is shifted away from the biased endpoint). However, unlike the effects of lexical bias, response times for monetary biased responses were faster than nonbiased ones only for endpoint stimuli that could be reliably categorized on the basis of acoustic information alone.

We assume that processes that theoretically require more ''processing resources'' or are more difficult to complete will be associated with longer response times; this assumption is well grounded in the experimental literature even for the same kinds of effects being investigated here (Clifton *et al.*, 1984; Fodor *et al.*, 1996; Shapiro *et al.*, 1993). Thus at the unambiguous endpoints, a phonological decision may be produced quickly on the basis of the acoustic signal alone, but at the potentially ambiguous boundary, lexical information influences phonological processing. Unlike the effects of lexical status, the influence of monetary payoff seems to be post-perceptual.

The most obvious difference between lexical bias and monetary payoff is that only the former is in the linguistic domain, as is phonological processing. Hence, one might expect the effects of other linguistic information, such as sentence context to resemble those of lexical status more than monetary payoff. Connine (1987) constructed word–word continua (e.g., GOLD–COLD) and embedded the stimuli in sentences biased toward one or the other interpretation. For example,

(1a) She told the client to sell all the [GOLD/COLD].
(1b) She closed the window to keep out the [GOLD/COLD].

Subjects were asked to identify the beginning sound of the last word in sentences like (1a) and (1b), then circle whether the word was ''sensible'' or ''anomalous'' in the given sentence frame. The identification functions showed the same boundary shift in favor of the biased response that had been observed with lexical status and monetary payoff. But when response times for sentence-congruent versus sentence-incongruent responses were compared, the congruency advantage appeared only at the endpoints, not near the category boundary. It was not the case that categorizations that ''make sense'' are simply faster because anomalous responses at the endpoints were faster than sensible responses near the boundary. Connine concluded that the similar response time patterns for monetary payoff and sentence bias were produced by similar post-perceptual processing that was different from the interactive processing of lexical bias.

Taken together, the results for biasing by lexical status, monetary payoff and sentence context present something of a puzzle. If it were the case that linguistic information is more available during the process of phoneme identification than nonlinguistic information, why would semantic sentence context be treated more like the nonlinguistic monetary payoff than like lexical status? The answer to that may lie in the degree of autonomy of different linguistic domains, but such an account would require the unlikely combination of a semi-autonomous phonological level that is both interactive with a word recognition system and impenetrable by semantic sentence context. An alternative explanation for the surprising difference between the results of biasing by lexical status and sentence context might lie in the methodology of the Connine (1987) experiment. First, seven different voicing continua were used, each with a different VOT range (e.g., 14–60 ms for DENT–TENT and 68–115 ms for DRIP–TRIP); there was, in fact a significant effect of continuum (''item'') in both the boundary and endpoint analyses as well as several interactions with item. These differences, as well as significant individual differences among subjects, required extensive adjustments in the analyses that may have muddied the results. Second, the critical acoustic target for identification was always the last word in the sentence, but context integration has been shown to occur at pauses or interruptions, even for tasks and stimuli for which context effects were not otherwise demonstrated (Nicol and Swinney, 1989; Nicol *et al.*, 1997). A third potential problem in the Connine experiment was the use of each sentence frame for every VOT value (14 unique sentence frames for seven voicing series×two contexts); recall of a prior response to an apparently identical sentence might affect a current identification.

The present experiment sought to clarify whether semantic sentence context influences phoneme categorization during the processing of an ongoing sentence and whether the congruency of a response has a similar effect for sentence context and lexical status. To this end, we examined the effects of acoustic parameter value and sentence context using (a) a single stimulus continuum, (b) a cross-modal ''word-matching'' task that requires identification judgments during the course of a sentence, and (c) a unique sentence frame for every trial. The critical tests for a context effect are a shift in the identification function in favor of the biased context and the presence of an interaction between the acoustic parameter value and the sentence bias. If the congruency of a response with a real word and with the meaning of a sentence have similar effects on response times, the boundary stimuli, rather than endpoint stimuli, should be preferentially influenced by sentence context.

I. METHOD

A. Subjects

Thirty-three undergraduates participated in this experiment in return for research credit in their introductory psychology class. They identified themselves as monolingual

TABLE I. GOAT-biased sentence components.[a]

	Sentence position				
	A	B	C	Target	D
1	The busy	dairyman	hurried to milk the	[GOAT/COAT]	in the kitchen garden
2	The handsome	farmer	forgot to feed the	[GOAT/COAT]	in the drafty barn
3	The laughing	children	loved to tease the	[GOAT/COAT]	in the grassy field
4	The fussy	zoo-keeper	wanted to cage the	[GOAT/COAT]	in the crowded pen
5	The lively	toddler	ran to chase the	[GOAT/COAT]	in the quiet pasture

[a]Sentence example: [The laughing]$_{A3}$ [farmer]$_{B2}$ [hurried to milk the]$_{C1}$ [GOAT/COAT]$_{Target}$ [in the crowded pen]$_{D4}$.

TABLE II. COAT-biased sentence components.[a]

	Sentence position				
	A	B	C	Target	D
1	The elderly	grandma	stopped to button the	[GOAT/COAT]	in the corner closet
2	The careful	laundress	had to dry-clean the	[GOAT/COAT]	in the back bedroom
3	The expert	seamstress	tried to shorten the	[GOAT/COAT]	in the tiny cloakroom
4	The cheerful	tailor	hurried to press the	[GOAT/COAT]	in the dusty storeroom
5	The foolish	cleaner	liked to wear the	[GOAT/COAT]	in the cluttered attic

[a]Sentence example: [The careful]$_{A2}$ [tailor]$_{B4}$ [had to dry-clean the]$_{C2}$ [GOAT/COAT]$_{Target}$ [in the dusty storeroom]$_{D4}$.

native speakers of American English with normal hearing, normal or corrected to normal vision with no known speech or language impairments.

B. Stimuli

1. Original sentences

Sentence meanings were biased toward either ''goat'' or ''coat'' as in (2a) and (2b).

(2a) GOAT-biased: The laughing dairyman hurried to milk the [GOAT/COAT] in the drafty barn.

(2b) COAT-biased: The expert tailor tried to shorten the [GOAT/COAT] in the cluttered attic.

Each of the sentences in each biasing condition contained a different, unique arrangement of the same set of components. The sentence components are listed in Table I (GOAT-biased) and Table II (COAT-biased). Twenty-one subjects participated in a paper and pencil ''fill-in-the-blank'' test in which they judged whether (a) only ''goat,'' (b) only ''coat,'' or (c) either ''goat'' or ''coat'' could complete 25 randomly selected sentences of each type, as well as 50 unbiased sentences. Results confirmed the sentence biases: 99.6% of GOAT-biased sentences and 99.5% of COAT-biased sentences were completed with the appropriate item (21 subjects each). The original sentences were recorded in a quiet room using a high fidelity table top microphone and a Marantz stereo cassette tape recorder (model PMC 420). The sentences were then digitized onto computer from the tape using a 16 bit audio card and a sampling rate of 22 050 Hz. The original sentences included the context appropriate word, ''goat'' or ''coat,'' although the original items were later replaced by the manipulated target stimuli described below.

2. Target words

A stimulus continuum was constructed by manipulating the recorded natural speech of the sentence reader, using a waveform editor. The sentence fragment ''the coat in the'' that contained a ''good'' /kot/ was excised from one of the COAT-biased sentences. This unmanipulated sentence fragment in which the aperiodic initial segment of /kot/ was 64 ms (measured from the beginning of the stop release burst to the onset of periodicity) was used as the longest VOT stimulus. The other tokens in the continuum were constructed by subtracting incremental segments (in 6-ms steps) from the midpoint of the initial aperiodicity of the original /k/ in /kot/ in order to preserve the natural variability in the energy associated with the consonant burst as well as the natural transitions between the different states. For example, the 58-ms stimulus value consisted of the first 29 ms of the original 64-ms stimulus, then the interval from 35 ms to 64 ms (see Fig. 1). In the continuum, the shortest VOT value corresponded to that measured in the natural ''goat'' (10 ms); the longest VOT value was equivalent to /k/ in the natural ''coat'' (64 ms). The ten resulting VOT values were: 10, 16, 22, 28, 34, 40, 46, 52, 58, and 64 ms.

3. Manipulated sentences

In each sentence, the fragment ''the goat/coat in the'' was replaced with one from the stimulus continuum using a waveform editor. The splices were placed at zero crossings

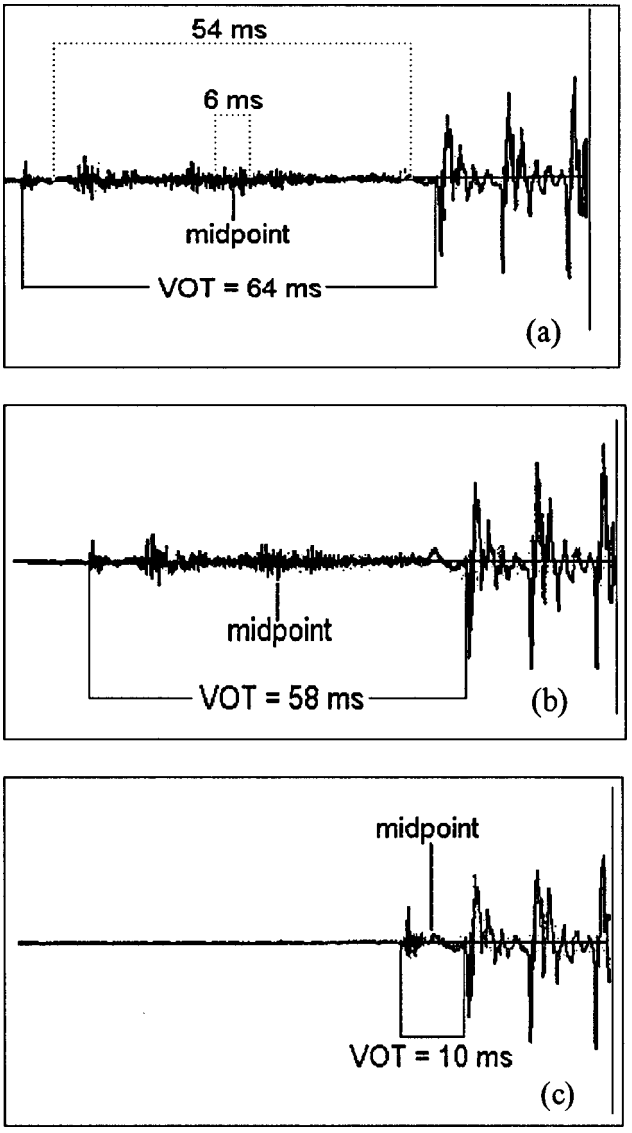


FIG. 1. The acoustic manipulation of the target GOAT/COAT stimuli by deleting segments from the midpoint of the 64-ms aperiodic portion of a ''natural'' /kot/. (a) VOT=64 ms: the voiceless endpoint stimulus that was the unaltered ''natural'' /kot/. The 6-ms and 54-ms segments that were deleted to produce VOT=58 ms and VOT=10 ms, respectively, are also indicated. (b) VOT=58 ms: The result of the 6-ms deletion shown above. The remainder of the stimuli were created by incrementing the value of the deleted segment in 6-ms steps until the result was (c) VOT=10 ms: The voiced endpoint where the deleted segment equaled 54 ms.

in the waveform to avoid acoustic discontinuities; also, any unnatural acoustic effects introduced by the manipulation would not be near the target word. Ten GOAT-biased sentences and ten COAT-biased sentences were constructed for each of the ten VOT values for a total of 200 sentences. A 1000-Hz tone inserted at the offset of the target stimulus was used to trigger the visual probe. This tone was placed on a second audio channel and was inaudible to the subjects. The manipulated sentences were then recorded back onto tape.

4. Visual probes

The visual probes consisted of the word ''goat'' or the word ''coat'' that appeared on the CRT at the offset of the target word in each sentence. The probes were balanced within each VOT×Context combination.

C. Procedure

Subjects, tested individually, were seated directly in front of a CRT in a quiet room. The sentences were presented binaurally over headphones. Subjects were instructed to look at the CRT while listening carefully to each sentence, paying attention to both sound and meaning. Comprehension was encouraged and assessed by testing recall for 20% of randomly selected trials. Subjects were instructed that when the word ''goat'' or ''coat'' appeared on the CRT, they should press, as quickly as possible, the button labeled YES (the word on the screen is the word they heard) or NO (the word on the screen is not the one they heard); the auditory sentence continued uninterrupted. Because the probes were balanced within each VOT×Context combination, a subject making, for example, the same ''goat'' identification response for every presentation of the same target stimulus would do so by responding ''yes'' to GOAT on one-half of the trials, and ''no'' to COAT on the other half. This strategy balanced any ''yes'' or ''no'' response biases and also prevented subjects from responding prior to the presentation of the embedded target word. The computer recorded both the responses and the response times, measured from the onset of the visual probe (at the offset of the target word) to the button press. The experiment consisted of a 20 sentence practice session and two 100 sentence blocks, with sentence order randomized, separated by a rest period. The criterion for including a subject in the data analysis was correct recall for 95% of the sentences.

II. RESULTS

The two dependent measures were the probability of a ''coat'' versus ''goat'' identification for each VOT×Context condition, and the response time for the decision task. If context influences phoneme categorization during sentence processing, a VOT×Context interaction for identification and a shift of the identification boundary in favor of the biased context should be observed. Similarly, response times should show a VOT×Context interaction and peak response time should shift with the identification function. If the congruency of a response with the sentence context acts like congruency with a real word (Connine and Clifton, 1987), the congruency advantage should be observed only for intermediate VOTs. In contrast, if the results confirm those of Connine (1987), the congruency advantage should be restricted to extreme VOTs.

Data from 30 subjects were included in all analyses (three subjects were excluded for failure to meet the sentence recall requirement). The effect of identification by a ''yes'' or by a ''no'' response was not significant ($p>0.8$), nor were there any significant interactions of ''yes'' or ''no'' responses with VOT or context. Therefore, subjects' ''yes'' and ''no'' responses will be referred to in terms of the stimulus identification.

For all 30 subjects, identification was near 100 percent ''goat'' at the shortest VOT and near 100 percent ''coat'' at the longest VOT. A shift of the identification boundary in favor of the biased context was observed for every subject as well as for the mean. The results of a 10×2 repeated-

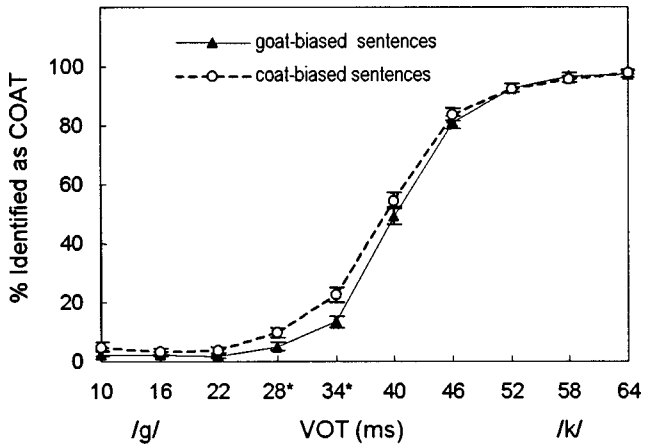


FIG. 2. The mean identification function by acoustic parameter and sentence context. Error bars indicate one standard error above and below the mean.

measures ANOVA, with VOT and biasing context as main effects were significant for VOT [$F(9,261)=388.79$, $p<0.01$], context [$F(1,29)=11.41$, $p<0.02$], and their interaction [$F(9,261)=2.37$, $p<0.01$].[1] The simple effect of context at each VOT was significant only for VOT=28 ms [$F(1,29)=4.12$, $p<0.05$] and VOT=34 ms [$F(1,29)=14.61$, $p<0.01$]. For both stimuli, a greater percentage of ''coat'' identifications were made in the COAT-biased context than in the GOAT-biased context (Fig. 2).

The response times for identification were then examined. ''Yes'' responses were faster overall [$F(1,29)=36.04$, $p<0.01$] but there were no ''yes'' vs ''no'' interactions with any other factors. Mean response times were greater for mid-range VOT and, like the identification functions, the peak response time for each biased context was shifted away from the biased endpoint. This pattern is shown in Fig. 3.

A 10×2 repeated measures ANOVA determined that VOT [$F(9,261)=15.32$, $p<0.01$], context [$F(1,29)=14.46$, $p<0.01$], and the VOT×Context interaction [$F(9,261)=2.03$, $p<0.03$] all influenced response times. The simple effect of context at each VOT was significant for VOT=10 ms [$F(1,29)=7.33$, $p<0.01$], VOT=16 ms [$F(1,29)=15.77$, $p<0.01$], and VOT=28 ms [$F(1,29)$

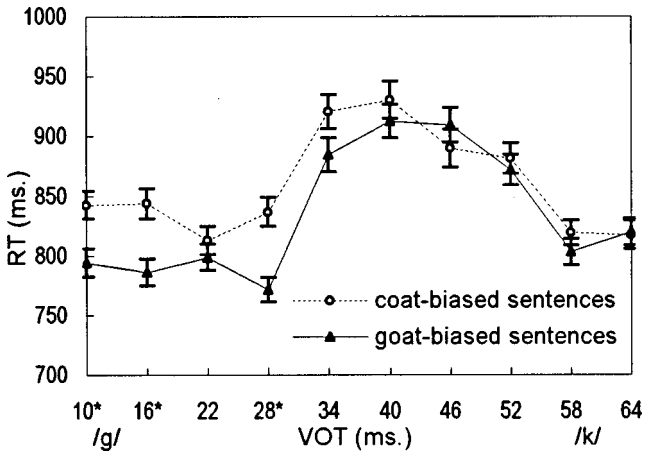


FIG. 3. Mean response time by acoustic parameter value and biasing context. Error bars indicate one standard error above and below the mean.

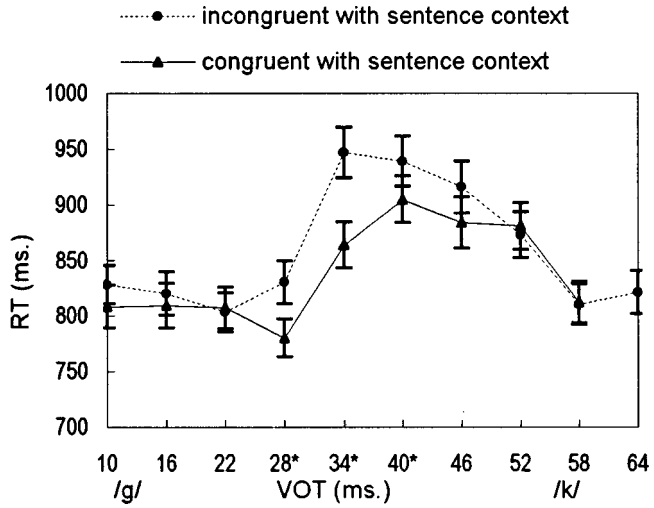


FIG. 4. Mean response time by acoustic parameter value and congruency of response with the biased sentence context. Congruent=GOAT identifications in the GOAT-biased context plus COAT identifications in the COAT-biased context; Incongruent=GOAT identifications in the COAT-biased context plus COAT identifications in the GOAT-biased context. Error bars indicate one standard error above and below the mean.

$=12.84$, $p<0.01$]. Additionally, at VOT=34 ms, response times differences due to sentence context were in the same direction as the group mean for 19 of the 30 subjects ($p=0.2$).

In order to analyze the data relative to the subject's response rather than the experimenter-determined condition, we partitioned the responses into two categories: a ''congruent'' category in which the identification matched the biasing context (i.e., ''goat'' identifications for GOAT-biased sentence contexts; ''coat'' identifications for COAT-biased sentence contexts) and an ''incongruent'' category in which identification and sentence context did not match (i.e., ''coat'' identifications for GOAT-biased sentences, etc.), see Fig. 4. A 10×2 ANOVA on response times showed main effects of VOT [$F(9,261)=15.32$, $p<0.01$] and congruency: $F(1,29)=7.61$, $p<0.01$, as well as their interaction [$F(9,261)=2.63$, $p<0.01$]. The simple effect of congruency at each VOT value was significant for three mid-range stimuli: VOT=28 ms [$F(1,29)=13.13$, $p<0.01$], VOT=34 ms [$F(1,29)=12.24$, $p<0.01$], and VOT=40 ms [$F(1,29)=6.36$, $p<0.02$].

An identification response could also be viewed in terms of congruency with acoustic information, but, because VOT values do not correspond to a single identification, we can only say that a ''goat'' identification is more congruent with the acoustic information at the shortest VOT endpoint than at the longest, whereas a ''coat'' identification is more congruent with the acoustics of the longest VOT value. For ''goat'' identifications, response times increased as the VOT got longer; but for ''coat'' identifications, the relationship was less clear. The mean response times for ''goat'' and ''coat'' identifications for each VOT and for the two different sentence contexts are shown in Fig. 5(a)–(b); note that the number of observations for each identification was far greater at the acoustically congruent endpoint than at its opposite. In fact, in Fig. 5(b) we see the root of the sentence context

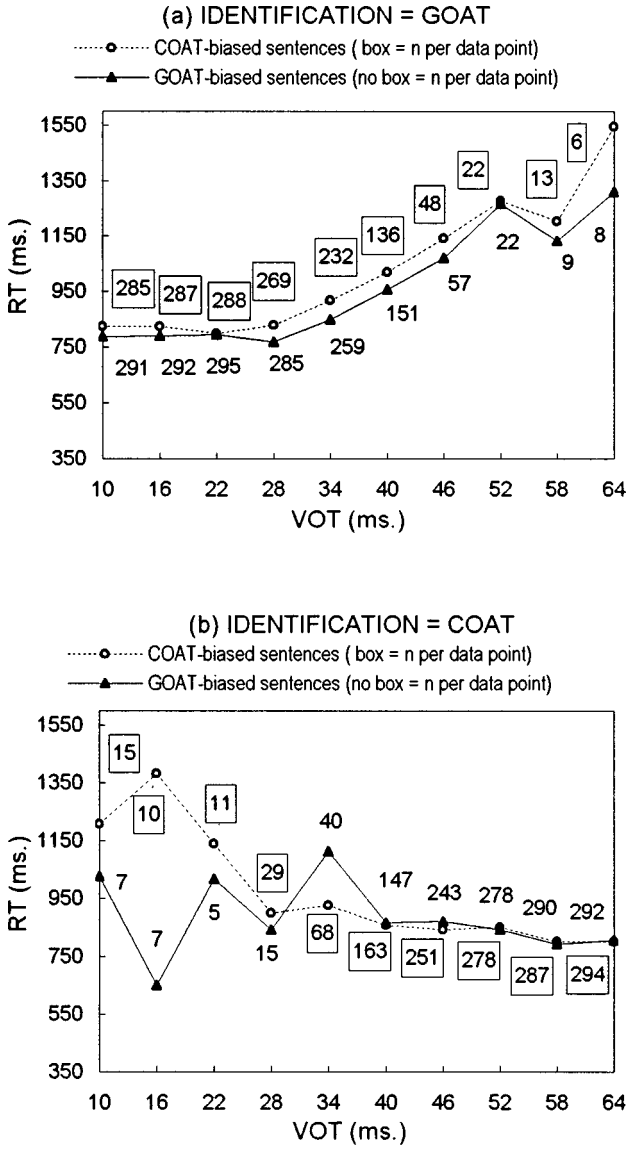


FIG. 5. Mean response time and number of observations (n) by VOT and context for (a) GOAT identifications only, and (b) COAT identifications only. Note that the n's are asymmetrical due to the identification function.

effect on RT for short VOT stimuli. That is, the increase in RT occurs only for the very few stimuli (<4% of the total) identified as ''coat'' despite very short VOTs. Although their numbers are small, the effect is still seen because their mean RT is over 300 ms slower than for ''goat'' responses to these same stimuli. Of those ''coat'' responses, twice as many occur in the COAT-biased context as in GOAT-biased context and their mean RT is fully 200 ms longer. When only ''goat'' responses to stimuli with 10-ms or 16-ms VOT are considered, the effect of sentence context on RT at the voiced endpoint disappears.

III. DISCUSSION

The results of this experiment confirm the influence of sentence context on phoneme categorization during sentence processing. Our data are much like the earlier results for biasing by lexical status (Connine and Clifton, 1987) and contrast with previous results for biasing by sentence context or monetary payoff (Connine, 1987; Connine and Clifton, 1987). Effects of local acoustic structure and sentence context on phoneme categorization confirmed the effectiveness of our stimuli and experimental task.

The observation that response times were also sensitive to local acoustic structure and sentence context confirmed their value as indicators of processing differences. If processing load increases with the degree of conflict between acoustic and contextual information then longer response times are predicted at the acoustically clear endpoints when contextually biased toward the opposite endpoint. This does not seem to be the case because endpoint response times are clearly fastest, regardless of context congruency, whether the biasing context is lexical status, monetary payoff, or sentence meaning. The longest response times are observed for intermediate VOT values which we interpret as an indication that more time or greater processing resources are needed for ambiguity resolution.

The principal criterion by which Connine (1987) and Connine and Clifton (1987) discriminated between early and late influence of context was the effect of the congruency of the response to the biasing context. When identifications are faster for congruent than for incongruent identifications only for boundary stimuli, the effect is thought to be early in perceptual processing. In contrast, when congruency influences only endpoint stimuli, the effect is thought to be post-perceptual. Both the present sentence context experiment and the Connine (1987) lexical status experiment showed an effect of congruency at the boundary, but not at the endpoints. In the monetary payoff experiment (Connine and Clifton, 1987), and the off-line sentence context experiment (Connine and Clifton, 1987), a congruency advantage was found at the endpoints, but not near the boundary. However, as Connine and Clifton point out in their discussion, the expected result of a ''pure'' post-perceptual response bias would be a congruency advantage for all stimuli. Whether the congruency effect represents a speeding up of congruent responses, or a slowing down of incongruent responses, the finding of any interaction between congruency and acoustic stimulus argues against a simple anomaly effect (longer times for ''incongruent'' responses in every case) because the balanced context conditions in each experiment and the relatively small identification differences, mean that ''mismatches'' are almost evenly distributed throughout the stimulus range. Our data suggest that sentence context acts in much the same way as lexical status and that both of these linguistic factors influence phonological processing.

The implications of these results for a larger theory of language comprehension depend mainly on what claims we are prepared to make about the place of our identification task in the time course of normal sentence processing. In a modular account, a sentence is constructed on-line by rapid, automatic sub-processes that use only domain specific information for initial processing (e.g., Fodor, 1983; Forster, 1979). A number of on-line studies have shown modularity for at least some language operations including access to multiple noun meanings (Swinney, 1979), access to multiple argument structures (Shapiro *et al.*, 1987, 1989), and gap filling (Swinney and Osterhout, 1990; Lewis *et al.*, 1998).

Since the words that are identified by phoneme categorization need to be available for ''higher level'' processes such as lexical access, structural analysis and semantic interpretation, according to this account, initial phonological analysis must be automatic and immediate, and should not be influenced by sentence context.

In an interactive account of sentence comprehension, word identification can change and develop over time as new information becomes available (for example, as in McClelland and Elman's 1986 TRACE model). If such a process is viewed as the satisfaction of multiple constraints (MacDonald *et al.*, 1994), probabilistic information from the acoustic characteristics of a stimulus could interact both with lexical characteristics of the word specified by a particular phoneme categorization and with contextual constraints on word selection. In this case, initial phonological analysis could be influenced by sentence context.

One important issue in the modularity-interactionism debate was not resolved by the present study: When in the time course of sentence processing did our observed context effects occur? In other words, did the results of our identification task indicate immediate, interactive context effects? An interactive account of initial phonological processing is possible in that biasing information from the sentence context was available prior to the presentation of the embedded target stimulus and could theoretically modify the criteria for phoneme categorization. On the other hand, identification tasks may not be sensitive to fast-acting, automatic processing of acoustic information because they require an explicit categorization decision prior to response; integration of contextual information before identification may be a consequence of task demands. In an ongoing experiment with the same stimuli, we use a cross modal lexical decision (CMLD) to an unrelated probe word that does not require an explicit decision about the target stimulus. Finding a VOT×Context interaction for CMLD times at the offset of the target stimulus would lend additional support to an interactive account of sentence processing in which context directly influences initial perception. Finding a VOT×Context interaction a few syllables after the target stimulus, but not at the target itself, would support a more modular account in which context exerts its effect only after initial context-independent phonological processing. Nevertheless, the results presented here are suggestive that sentence context can guide phonological processing during ongoing sentence analysis.

ACKNOWLEDGMENTS

This research was supported in part by NIMH Grant No. MH42900 and NIH Grant No. DC00494. This manuscript was prepared while the second author was a Visiting Fellow at the Neurosciences Institute, San Diego, California.

[1]The few (1%) contrary responses at the acoustically clear endpoints were interpreted as response errors and were not included in the ANOVA for identification by VOT and sentence context. The results of the same analysis with the 1% contrary endpoint responses included were: context, $F(1,29)=9.89$, $p<0.01$; VOT, $F(9,261)=359.51$, $p<0.01$; VOT ×Context, $F(9,261)=1.83$, $p<0.06$. The significant simple effects were identical.

Boothroyd, A., and Nittrouer, S. (**1988**). ''Mathematical treatment of context effects in phoneme and word recognition,'' J. Acoust. Soc. Am. **84**, 101–114.

Clifton, C., Frazier, L., and Connine, C. (**1984**). ''Lexical expectations in sentence comprehension,'' Journal of Verbal Learning and Verbal Behavior **23**, 696–708.

Connine, C. M. (**1987**). ''Constraints on interactive processes in auditory word recognition: The role of sentence context,'' Journal of Memory and Language **26**, 527–538.

Connine, C. M., and Clifton, C., Jr. (**1987**). ''Interactive use of lexical information in speech perception,'' Journal of Experimental Psychology: Human Perception and Performance **13**, 291–299.

Fodor, J. A. (**1983**). *The Modularity of Mind* (MIT, Cambridge, MA).

Fodor, J. A., Ni, W., Crain, S., and Shankweiler, D. (**1996**). ''Tasks and timing in the perception of linguistic anomaly,'' Journal of Psycholinguistic Research **25**, 25–57.

Forster, K. I. (**1979**). ''Levels of structure of the language processor,'' in *Sentence Processing*, edited by W. Cooper and E. Walker (Erlbaum, Hillsdale, NJ).

Ganong III, W. F. (**1980**). ''Phonetic categorization in auditory word perception,'' Journal of Experimental Psychology: Human Perception and Performance **6**, 110–125.

Lewis, J. R., Shapiro, L. P., Tuller, B., and Afton, R. (**1998**). ''Local contextual impenetrability of lexical access and gap-filling,'' Poster presented at the Tenth Annual CUNY Conference on Human Sentence Processing, Santa Monica, CA.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (**1967**). ''Perception of the speech code,'' Psychol. Rev. **74**, 431–461.

Liberman, A., Harris, K. S., Hoffman, H. S., and Griffith, B. C. (**1957**). ''The discrimination of speech sounds within and across phoneme boundaries,'' J. Exp. Psychol. **54**, 358–368.

McClelland, J. L., and Elman, J. L. (**1986**). ''The TRACE model of speech perception,'' Cogn. Psychol. **18**, 1–86.

MacDonald, M. C., Pearlmutter, N. J., and Seidenberg, M. S. (**1994**). ''Lexical nature of syntactic ambiguity resolution,'' Psychol. Rev. **101**, 676–703.

Nicol, J., and Swinney, D. (**1989**). ''The role of structure in coreference assignment during sentence comprehension,'' Journal of Psycholinguistic Research **18**, 5–19.

Nicol, J., Swinney, D., Love, T., and Hald, L. (**1997**). ''Examination of sentence processing with continuous vs. interrupted presentation paradigms.'' Center for Human Information Processing Technical Report, #97-2, University of California, San Diego.

Nittrouer, S., and Boothroyd, A. (**1990**). ''Context effects in phoneme and word recognition by young children and older adults,'' J. Acoust. Soc. Am. **87**, 2705–2715.

Repp, B. H., and Liberman, A. M. (**1990**). ''Phonetic category boundaries are flexible,'' in *Categorical Perception: The Groundwork of Perception*, edited by S. Harnad (Cambridge U.P., Cambridge, MA), pp. 89–112.

Rubin, P., Turvey, M. T., and Van Gelder, P. (**1976**). ''Initial phonemes are detected faster in spoken words than in spoken nonwords,'' Percept. Psychophys. **19**, 394–398.

Shapiro, L. P., Nagel, H. N., and Levine, B. A. (**1993**). ''Preferences for a verb's complements and their use,'' Journal of Memory and Language **32**, 96–114.

Shapiro, L. P., Zurif, E. B., and Grimshaw, J. (**1987**). ''Sentence processing and the mental representation of verbs,'' Cognition **27**, 219–246.

Shapiro, L. P., Zurif, E. B., and Grimshaw, J. (**1989**). ''Verb processing during sentence comprehension: Contextual impenetrability,'' Journal of Psycholinguistic Research **18**, 223–243.

Swinney, D. A. (**1979**). ''Lexical access during sentence comprehension: (Re)consideration of context effects,'' Journal of Verbal Learning and Verbal Behavior **18**, 645–659.

Swinney, D. A., and Osterhout, L. (**1990**). ''Inference generation during auditory language comprehension,'' in *The Psychology of Learning and Motivation, Vol. 25*, edited by A. Graesser and G. Bowers (Academic, New York), pp. 17–33.

Warren, R. M. (**1970**). ''Perceptual restoration of missing speech sounds,'' Science **267,** 392–393.

Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration

Ken W. Grant, Brian E. Walden, and Philip F. Seitz
Walter Reed Army Medical Center, Army Audiology and Speech Center, Washington, DC 20307-5001

(Received 19 March 1997; revised 10 December 1997; accepted 24 January 1998)

Factors leading to variability in auditory-visual (AV) speech recognition include the subject's ability to extract auditory (A) and visual (V) signal-related cues, the integration of A and V cues, and the use of phonological, syntactic, and semantic context. In this study, measures of A, V, and AV recognition of medial consonants in isolated nonsense syllables and of words in sentences were obtained in a group of 29 hearing-impaired subjects. The test materials were presented in a background of speech-shaped noise at 0-dB signal-to-noise ratio. Most subjects achieved substantial AV benefit for both sets of materials relative to A-alone recognition performance. However, there was considerable variability in AV speech recognition both in terms of the overall recognition score achieved and in the amount of audiovisual gain. To account for this variability, consonant confusions were analyzed in terms of phonetic features to determine the degree of redundancy between A and V sources of information. In addition, a measure of integration ability was derived for each subject using recently developed models of AV integration. The results indicated that (1) AV feature reception was determined primarily by visual place cues and auditory voicing+manner cues, (2) the ability to integrate A and V consonant cues varied significantly across subjects, with better integrators achieving more AV benefit, and (3) significant intra-modality correlations were found between consonant measures and sentence measures, with AV consonant scores accounting for approximately 54% of the variability observed for AV sentence recognition. Integration modeling results suggested that speechreading and AV integration training could be useful for some individuals, potentially providing as much as 26% improvement in AV consonant recognition. [S0001-4966(98)02305-4]

PACS numbers: 43.71.An, 43.71.Es, 43.71.Kg, 43.71.Ma [WS]

INTRODUCTION

For all but the most profoundly hearing-impaired individuals, auditory-visual (AV) speech recognition has consistently been shown to be more accurate than auditory-only (A) or visual-only (V) speech recognition. Although this is especially true when the auditory signal is distorted (e.g., due to hearing loss, environmental noise, or reverberation), the influence of visual cues on speech recognition is not limited to conditions of auditory distortion. Even with fully intact speech signals, visual cues can have an impact on recognition. McGurk and MacDonald (1976) demonstrated that when the auditory production of one consonant is synchronized with the visual production of another consonant, most observers will perceive a third consonant that is not represented by either auditory or visual modality. For example, when auditory /bɑ/ is presented with visual /gɑ/, the perceived result is often /dɑ/. This illusion, known as the ''McGurk Effect,'' occurs even when the auditory signal is perfectly intelligible.

Another example of the influence of visual speech cues on intact auditory signals occurs when listeners are asked to repeat unfamiliar phrases, as when learning a second language or when presented with grammatically complex passages. For instance, when asked to shadow the speech of a native French speaker, students with four years of French training performed significantly better when they were able to see the speaker's face as compared to when they could only hear his voice (Reisberg *et al.*, 1987). Similarly, when native speakers of English were asked to shadow passages from Kant's *Critique of Pure Reason* spoken by another native English speaker, performance was significantly better when visual cues were available (Reisberg *et al.*, 1987).

These brief examples suggest that the perception of speech is inherently multimodal. Cues extracted from both auditory and visual sources are integrated early in the perceptual analysis of speech, and this information is used without much regard to the sensory modality of origin (Massaro, 1987; Summerfield, 1987). Yet, in spite of the superior status of auditory-visual speech recognition relative to auditory-only speech recognition, recent explications of speech perception often omit, or only mention briefly, the role of visual speech information (Diehl and Kluender, 1989; Klatt, 1989; Stevens, 1989; Halle and Stevens, 1991; Greenberg, 1995; Lindblom, 1996; Ohala, 1996). In contrast, recent studies of automatic speech recognition have begun to emphasize the importance of incorporating visual speech information along with acoustic information to improve recognition performance, especially in noisy environments (Stork and Hennecke, 1996).

From an applied perspective, a theory of AV speech perception would be an extremely valuable tool for addressing the communication problems encountered by hearing-impaired individuals or by normally hearing individuals in

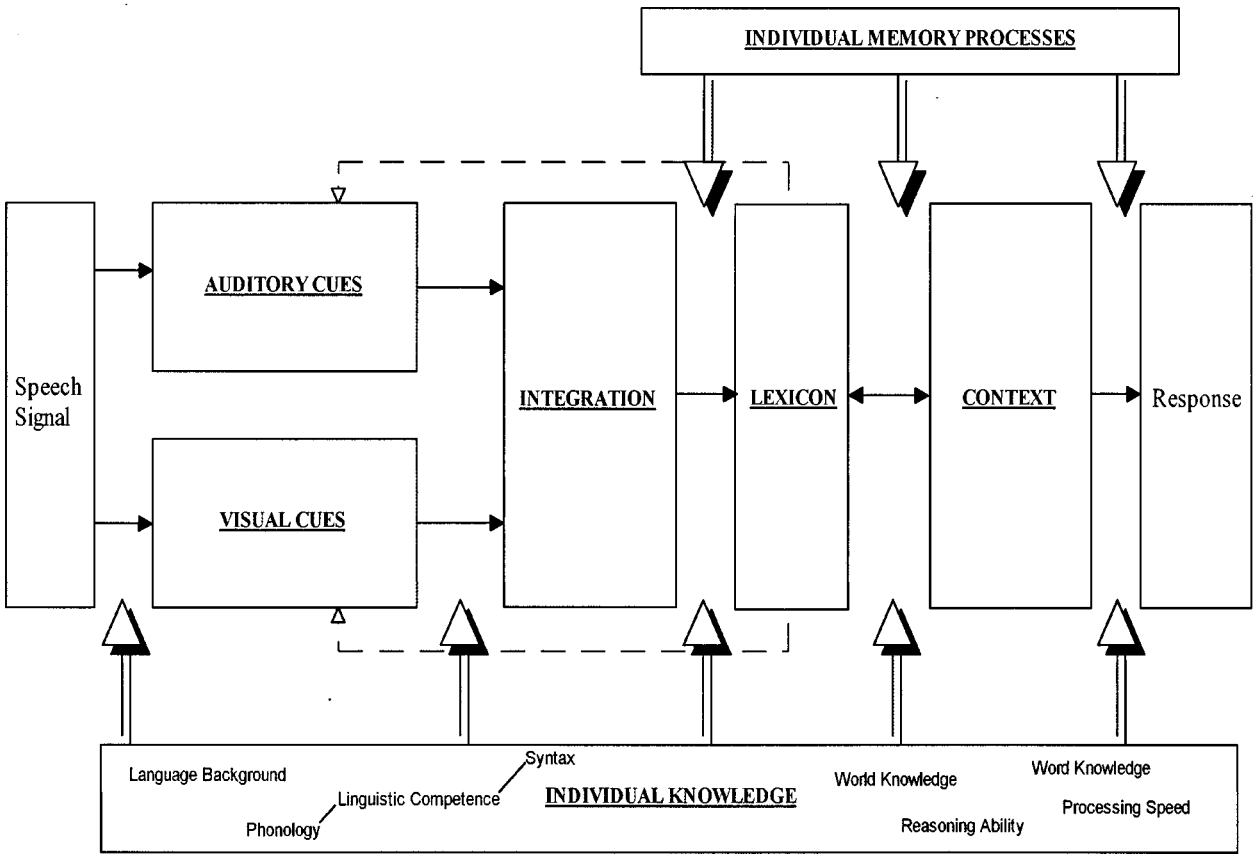


FIG. 1. Schematized framework of auditory-visual speech recognition.

noisy or reverberant environments. Such a theory would provide a conceptual framework that could serve as a guide for developing signal-processing strategies and/or rehabilitation programs when AV speech perception is less than perfect. Consider the simple conceptual framework shown in Fig. 1. A speech signal composed of both optical and acoustic information is presented. The listener-observer extracts signal-related segmental and suprasegmental cues from each modality, integrates these cues, and applies top-down semantic and syntactic constraints in an effort to interpret the message before making a response. The basic components, bottom-up signal-related cue extraction, integration, and top-down linguistic processes, are common to most speech perception theories (e.g., Liberman *et al.*, 1967; Stevens and House, 1972; Studdert-Kennedy, 1974). The major distinction drawn here from auditory-only theories of speech perception is that in an audiovisual communication environment, cues from the visual modality must be considered, and the integration of A and V cues, both within and across modalities, must be described (Massaro, 1987).

From this perspective, consider a hearing-impaired individual whose AV recognition of words and sentences is less than perfect. In order to evaluate the exact nature of the communication problem, it is necessary to determine whether the communication failure is due to poor reception of auditory and/or visual cues, difficulty in integrating A and V cues, difficulty in applying linguistic and contextual constraints, reduced working memory capacity, or a combination of these factors. If the problem is determined to be primarily difficulty in receiving A or V cues, signal-processing strategies to enhance the relevant cues may be used. If, on the other hand, the problem is determined to be difficulty in integrating A and V cues or difficulty in applying top-down language processing rules, then training and practice techniques may be applied. Simply knowing the individual's AV sentence or word recognition performance is not sufficient for determining a plan for rehabilitation.

In order to use the framework displayed in Fig. 1 as a guide for rehabilitation, three questions must be addressed: (1) What are the most important cues for AV speech recognition that can be extracted from acoustic and visual speech signals? (2) Is it possible to measure an individual's ability to integrate auditory and visual cues separate and apart from their ability to recognize syllables, words, and sentences? (3) What are the most important nonsignal-related ''top-down'' processes that contribute to individual variability in AV speech recognition? A brief discussion relating to each of these questions is presented below.

1. What are the most important cues for AV speech recognition that can be extracted from acoustic and visual speech signals?

Clearly, the answer to this question cannot come from studies of auditory speech perception alone. For example, auditory place-of-articulation cues, which are extremely important for A-only speech recognition, are not as important in AV speech recognition because this information is readily

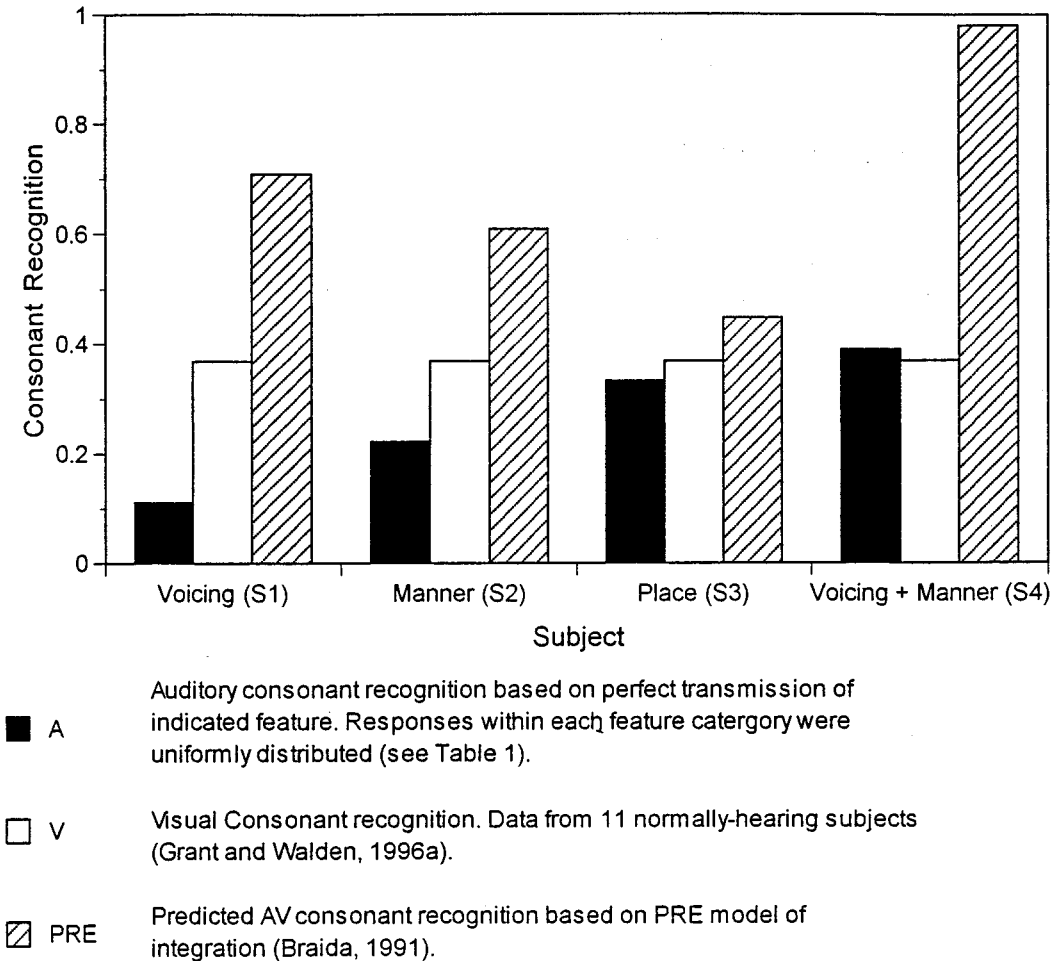


FIG. 2. Scores for A, V, and (predicted) AV conditions for four hypothetical subjects with average speechreading ability. Auditory-visual predictions were made with the PRE model using an average normal-hearing V consonant confusion matrix and an appropriate A confusion matrix with perfect recognition of the specified speech feature. Each set of bars represents a subject with perfect auditory feature recognition for voicing (S1), manner-of-articulation (S2), place-of-articulation (S3), and voicing-plus-manner (S4), respectively.

available through the visual channel. On the other hand, auditory voicing and manner-of-articulation cues are extremely important in AV speech recognition because this information is difficult to obtain visually. A careful accounting of the cues extracted from the separate unimodal conditions can reveal the degree of redundancy or complementarity between A and V cues. If A and V cues are redundant, there will be less benefit resulting from their combination than if the cues are complementary (Walden *et al.*, 1977; Grant and Walden, 1996a).

The relevancy of cue redundancy between A and V conditions as a predictor of AV performance is important for the present discussion because it forces us to recognize that when comparing two different listening conditions (or two listeners with different degrees of hearing loss, or two different hearing aids), the condition resulting in the higher rate of auditory intelligibility need not result in the highest AV score or the greater improvement when combined with speechreading. This point is illustrated graphically in Fig. 2.

Figure 2 shows predicted AV consonant recognition results for four different hypothetical subjects. For each hypothetical subject, three bars are shown. The black bar represents the auditory recognition score on a consonant identification task that would result from perfect recognition of either voicing (S1), manner-of-articulation (S2), place-of-articulation (S3), or the combination feature, voicing-plus-manner (S4). Table I shows the feature classification for the set of consonants used in this example. The auditory recognition scores (black bars) were calculated assuming perfect recognition of each specified feature and a uniform distribution of responses within each feature category. For example, for the hypothetical subject who receives perfect voicing information (S1), the confusion matrix shows a uniform distribution of responses among all voiced consonants and a uniform distribution of responses among unvoiced consonants. Since the subject is perfect in recognizing the voicing feature, there are no cases of voiced–voiceless confusions. Once constructed in this manner, the confusion matrix can then be scored for overall accuracy by simply observing the number of correct responses (lying along the main diagonal of the confusion matrix) divided by the total number of trials.

The unfilled bars are the same for each subject and represent consonant recognition performance of a typical speechreader (Grant and Walden, 1996a).[1] The shaded bars represent AV predictions made by Braida's Pre-Labelling Integration Model (1991).[2] Note that a subject who has perfect voicing recognition (S1) derives substantially more AV benefit (e.g., difference between AV and A recognition) accord-

TABLE I. Feature classification for voicing, manner, and place categories.

VOICING					
Voiced				Unvoiced	
b,d,g,m,n,v,ð,z,ʒ,dʒ				p,t,k,f,θ,s,ʃ,tʃ	
MANNER OF ARTICULATION					
Stop		Nasal	Fricative		Affricate
b,p,g,k,d,t		m,n	v,f,ð,θ,z,s,ʒ,ʃ		dʒ,tʃ
PLACE OF ARTICULATION					
Bilabial	Lingua-Velar	Lingua-Alveolar	Lingua-Dental	Lingua-Palatal	Labio-Dental
b,p,m	g,k	d,t,n,s,z	ð,θ	ʒ,ʃ,dʒ,tʃ	v,f

ing to Braida's model than subjects who obtain either perfect manner (S2) or-perfect place recognition (S3), even though these latter subjects have higher overall A recognition scores. Finally, a subject who obtains perfect voicing and manner recognition (S4) is predicted to have a nearly perfect AV score. This implies that hearing-impaired listeners with average speechreading skills who can extract voicing and manner information from the A condition may be expected to have high AV consonant scores, regardless of their overall A recognition score. Figure 2 also demonstrates that improvements in absolute unimodal recognition scores that may result from various forms of signal processing (e.g., hearing aids) may not produce the best bimodal recognition score.

2. Is it possible to measure an individual's ability to integrate auditory and visual cues separate and apart from their ability to recognize syllables, words, and sentences?

The recent development of quantitative models of multisensory integration (Massaro, 1987; Blamey *et al.*, 1989; Braida, 1991) has made it possible to derive measures of AV integration that are independent of the individual's ability to extract cues from the A and V modalities. These models make predictions of AV performance based on speech feature recognition or from the detailed patterns of confusions obtained for each separate modality. If these predictions are based on optimum processing models, then the extent to which the actual obtained AV performance approximates the predicted performance may be used as an index of AV integration skill. That is, subjects whose AV performance is well-predicted by the models are better at integrating A and V cues than subjects whose actual AV performance is overpredicted by the models.

Some of the integration models considered in this study are not optimum processor models, and observed AV scores can be higher than predicted scores. Nevertheless, deviations between predicted and obtained AV performance may still be used as an index of integration skill. For example, suppose that predictions from a particular model were based only on a subject's overall score from each separate modality and did not take into account the pattern of confusions generated in the A and V conditions. Even though such a model would likely underpredict real performance, the extent to which subjects' AV recognition scores exceed model predictions might indicate better integration abilities.

3. What are the most important nonsignal-related "top-down" processes that contribute to individual variability in AV speech recognition?

Individuals interpret speech signals in conjunction with stored linguistic knowledge (Lindblom, 1996). This knowledge base includes the individual's vocabulary and other properties of the lexicon, their knowledge of semantics and syntax, their use of word-level and sentence-level context to compensate for misperceptions and impoverished acoustic and visual information, and memory processes and strategies for lexical access based on partial information in the signal. Measures of variability for each of these sources of linguistic knowledge are likely to reveal interesting subject differences which may impact on AV recognition of speech. Unfortunately, this potentially important source of variability in AV performance continues to receive little attention in the literature.

The experiments reported here represent the first in a series of studies to evaluate the importance of these different factors (cue extraction, integration, top-down processing) in determining AV speech recognition. In the present study, we focus mainly on the first two factors and describe individual variability in AV consonant and sentence recognition by hearing-impaired persons in terms of two factors: the ability to recognize consonants auditorily and visually, and the ability to integrate A and V cues. Admittedly, consonant recognition represents only one class of several potentially important "bottom-up signal-related" factors (as compared to the recognition of consonant clusters, vowels, word and sentence stress, syllabification, etc.). One reason for choosing to investigate consonant recognition is that in English, consonants account for a large proportion of the total segmental information making up lexical items (Kucera and Francis, 1967). Furthermore, many of the acoustic cues that help distinguish one consonant from another are often of low intensity and involve rapid, short-duration spectral transitions making them difficult to hear under conditions of hearing-impairment or environmental distortion (e.g., noise or reverberation). These two attributes of consonants (frequency of occurrence in the language and acoustic characteristics) lead us to hypothesize that variability in consonant recognition will contribute strongly to variability in word and sentence recognition.

This study is divided into three main parts: obtaining measures of auditory (A), visual (V), and auditory-visual

(AV) consonant recognition, deriving a measure of auditory-visual integration ability from A and V consonant recognition data, and obtaining measures of A, V, and AV sentence recognition. All measures were obtained on individual hearing-impaired patients. The overall objective was to relate the variability in consonant recognition and AV integration ability to variability in AV sentence recognition.

I. METHODS

A. Subjects

Twenty-nine subjects between the ages of 41 and 88 years (mean age=65.0) were recruited from the patient population of the Army Audiology and Speech Center, Walter Reed Army Medical Center. All had acquired sensorineural hearing losses due primarily to noise exposure. The average three-frequency (500, 1000, and 2000 Hz) pure-tone threshold for the better ear was 33 dB HL (range: 0–63.3 dB HL re; ANSI, 1989). The average two-frequency (2000 and 4000 Hz) pure-tone threshold was 53.5 dB HL (range: 20–77.5 dB HL). No additional audiometric criteria were imposed in subject selection. Because variability in the patterns of auditory consonant confusions was important to the planned analyses, subjects with a variety of configurations and severity of hearing loss were included. In order to include a wide range of speech recognition abilities, potential subjects were initially screened on their ability to recognize IEEE Harvard sentences (IEEE, 1969) in noise [signal-to-noise ratio (S/N) of 0 dB]. Screening of subjects continued until 4–5 subjects with scores at each performance level of approximately 10%, 20%, 30%, 40%, 50%, and 60% correct were identified (i.e., 5 subjects at 10%, 5 subjects at 20%, etc.). All subjects were native speakers of American English and had normal or corrected-to-normal vision. None had prior experience as subjects of experiments involving visual or bisensory speech recognition. Although many of the subjects were experienced hearing-aid users, all testing was conducted binaurally under headphones with speech levels approximating 85 dB SPL. When eligible subjects were paid for their participation in the screening process, and if selected, during the main experiment.

B. Procedure

1. Consonant recognition

Consonant recognition was measured separately for A, V, and AV presentations. Speech materials consisted of eighteen medial consonants (/p,b,t,d,k,g,m,n,s,z,f,v,θ,ð,ʃ,ʒ,tʃ,dʒ/) surrounded by the vowel /ɑ/. Ten productions of each /ɑ/–consonant–/ɑ/ (aCa) stimulus were spoken by a female talker of American English and recorded on optical disk (Panasonic TQ-3031F). The audio portion of each production was digitized (16-bit A/D, 20-kHz sampling rate), normalized in level, and stored on computer (Vanguard 486). For auditory-visual presentations, the digitized computer audio and optical disk video portions of each production were realigned using custom auditory-visual control software. Alignments were checked using a dual trace oscilloscope to compare the original and digitized productions of each utterance and were found to be within ±2 ms.

Prior to testing, subjects were familiarized with the consonant set and the use of a touch screen in making responses. Care was taken to ensure that each subject understood the task and could make use of the eighteen consonant category labels. Auditory-visual practice blocks of 36 trials, each with trial-by-trial feedback, were presented binaurally through headphones (Beyer DT-770) at a comfortable listening level appropriate for each subject. Speech materials were presented in quiet with subjects seated in a sound-treated room facing a 19-in. color video monitor (SONY PVM 2030) situated approximately 5 ft from the subject. After one or two initial practice blocks, subjects were required to achieve performatice levels on three additional consecutive practice blocks of 92% correct consonant recognition or better. Only one subject (out of 30 tested) failed to meet this requirement and was eliminated from further testing.

Following familiarization training, data were obtained on consonant recognition in noise. On each test trial, a single aCa production was presented. The overall level of the speech signal was approximately 85 dB SPL. The aCa utterances were mixed with a speech-shaped noise matched to the average long-term spectrum of the stimulus set and presented at a S/N ratio of 0 dB. This level of noise was chosen to insure a sufficient number of recognition errors for the planned analyses. A different sample of noise was used for each of the 180 tokens (18 consonants ×10 productions). Each speech token began 50 ms after the noise started and ended 50 ms before the noise finished. Subjects were required to identify the consonant presented, selecting their response from a touch screen terminal displaying all 18 possible consonant labels. Subjects were tested in blocks of 72 trials. Ten blocks each were presented in the A, V, and AV conditions, yielding a total of 40 trials per consonant per condition. The order of A, V, and AV conditions was randomized for each subject. No feedback was provided.

2. Auditory-visual integration

Three recent models of auditory-visual speech recognition (Massaro, 1987; Blamey *et al.*, 1989; Braida, 1991) were used for estimating the efficiency with which individual subjects integrate A and V cues. Each of these models predicts AV recognition scores from A and V scores alone. Differences between predicted and obtained AV scores were used as a measure of each subjects' integration efficiency. Below, a brief description of each model is presented along with the methods used to derive individual integration measures.

a. Braida (1991)—PRE. In the Prelabeling Model of Integration, confusion matrices from A and V consonant recognition are subjected to a special form of multidimensional scaling (MDS) and interpreted within a Theory of Signal Detection (e.g., Green and Swets, 1966; Macmillan *et al.*, 1988). The model provides a spatial interpretation of the ability to distinguish between consonants analogous to that derived from traditional multidimensional scaling (Borg and Lingoes, 1987). However, unlike traditional MDS, the scaled distances between consonants in the separate A and V spaces are converted to a common metric, d', explicitly reflecting the correctness of responses. The decision process assumes a

comparison between stimulus attributes (modeled as a multidimensional vector of cues, $\vec{X}$) and prototypes or response centers ($\vec{R}$) in memory. Subjects are assumed to respond R_k if and only if the distance from the observed vector of cues $\vec{X}$ to $\vec{R}_k$ is smaller than the distance to any other prototype. A subject's sensitivity $d'(i,j)$ in distinguishing stimulus S_i from stimulus S_j is given by

$$d'(i,j) = \|\vec{S}_i - \vec{S}_j\| = \sqrt{\sum_{k=1}^{D} (S_{ik} - S_{jk})^2}, \tag{1}$$

where $\|\vec{S}_i - \vec{S}_j\|$ is the distance between the D-dimensional vector of cues generated by stimuli S_i and S_j.

In the present study, estimates of stimulus and response centers that best fit a given confusion matrix were obtained iteratively using a KYST procedure (Kruskal and Wish, 1978). For the first iteration, $\vec{S}$ and $\vec{R}$ are assumed to be aligned. Subsequent iterations attempted to improve the match between predicted and obtained matrices (using a x^2-like measure) by displacing slightly both stimulus and response centers. Each iteration assumed 5120 presentations per consonant token yielding a total of 92 160 trials per matrix. This number was selected to reduce the stimulus variability in each simulation to approximately 1/10th of the variability in the data. The MDS fits were further optimized by choosing either two- or three-dimensional solutions depending on which gave the best fit to the unimodal matrix.

Prelabeling model predictions for AV performance are made solely on the basis of unimodal performance. Assuming that A and V cues are combined optimally, the decision space for the AV condition is the Cartesian product of the space for the A condition and the space for the V condition. Thus the relation between a subject's sensitivity in the AV condition and the corresponding unimodal sensitivities, assuming no perceptual interference (e.g., masking or distraction) across modalities, is given by

$$d'_{\mathrm{AV}}(i,j) = \sqrt{d'_{\mathrm{A}}(i,j)^2 + d'_{\mathrm{V}}(i,j)^2}. \tag{2}$$

Predictions of AV consonant recognition were made using the A and V consonant recognition data described above and compared to AV consonant recognition obtained by individual subjects. Since Braida's Prelabeling Model is an optimum processor model, predicted AV scores will always equal or exceed observed AV scores. A subject's integration efficiency, as predicted by the model, was given by the difference between predicted and observed AV recognition scores (with zero difference indicating perfect integration).

b. Massaro (1987)—FLMP. In the Fuzzy Logical Model of Perception (FLMP), auditory and visual channels are considered to be independent sources of information about the identity of the AV stimulus. Continuously valued features are evaluated, integrated, and matched against prototype descriptions in memory, and an identification decision is made on the basis of the relative goodness of match of the stimulus information with the relevant prototype descriptions. In the multimodal case, each speech segment possesses a set of feature values for the auditory channel and a set of feature values for the visual channel. For example, in the syllable /bɑ/, the lips coming together represent a visual feature, whereas second and third formant transitions appropriate for /bɑ/ represent acoustic features. To relate these sets of informational values to each other, a common metric based on fuzzy truth values (Zadeh, 1965) is used. Truth values ranging between 0 and 1 are used to represent the goodness-of-fit between each particular feature in the information set and the prototypical feature value in memory. Although feature values are considered the basis for the match between a given speech signal and its prototypical representation, these values are not observed directly. Instead, they are inferred from the response patterns obtained from separate auditory-only and visual-only recognition experiments. To predict bimodal performance, auditory-visual feature values are assumed to be proportional to the product of unimodal feature values. According to Massaro (1987) and Massaro and Friedman (1990) a multiplicative combination of feature values predicts an optimal level of performance given multiple sources of information. The multiplicative rule for AV integration used by Massaro is given by

$$P_{\mathrm{AV}}(R_j|S_i) = \frac{P_{\mathrm{A}}(R_j|S_i) \times P_{\mathrm{V}}(R_j|S_i)}{\sum_{k=1}^{N} P_{\mathrm{A}}(R_k|S_i) P_{\mathrm{V}}(R_k|S_i)}, \tag{3}$$

where $P_{\mathrm{AV}}(R_j|S_i)$ is the conditional probability that response R_j is given when the AV stimulus S_i is presented, and $P_{\mathrm{A}}(R_j|S_i)$ and $P_{\mathrm{V}}(R_j|S_i)$ are the conditional probabilities for the A and V unimodal cases, respectively. The denominator is the sum of the conditional probabilities in the ith row of the confusion matrix and normalizes the feature values for the N responses to conditional probabilities.

In our implementation of the FLMP, a fixed form of the model was used where the parameter values representing the goodness-of-fit of any particular token (presented in either the A or V modality) to its prototypical memory representation are assumed to be equal to the unimodal performance level for that token.[3] In other words, the truth values were derived directly from the original confusion matrices for the unimodal conditions. For example, if a subject correctly identified auditory /bɑ/ 85% of the time, then $P_{\mathrm{A}}(R_{\mathrm{ba}}|S_{\mathrm{ba}}) = 0.85$. Similarly, if the subject responded ''va'' 10% of the time given this same auditory stimulus, then $P_{\mathrm{A}}(R_{\mathrm{va}}|S_{\mathrm{ba}}) = 0.10$, and so on. Thus the set of A and V conditional probabilities derived from the A and V confusion matrices were used directly with the multiplicative rule to predict AV responses.

Braida (1991) has shown that the fixed FLMP model makes reasonably good predictions of multimodal scores without systematically under- or over-predicting observed accuracy. However, the model has a tendency to underestimate AV performance when A performance is particularly poor and overestimate AV performance when A performance is relatively good. As with the PRE model, integration efficiency was defined as the difference between predicted and observed AV consonant recognition scores.

c. Blamey et al. (1989)—PROB. According to the model proposed by Blamey *et al.* (1989), AV speech recognition errors occur only when there are simultaneous errors in *both* A and V recognition. Model predictions are stated in

terms of percentage of information transmitted (%IT). Thus given the probability of an auditory error [1−%IT(A)] and visual error [1−%IT(V)], the AV error rate is given by

$$AV_{ERR}=(1-\%IT_A)(1-\%IT_V), \quad (4)$$

and the AV score is given by

$$1-(AV_{ERR}). \quad (5)$$

The advantage of this simple probabilistic model (PROB) is that it can be used to predict a wide range of speech recognition scores, including information transmission scores for speech features (e.g., voicing, manner-of-articulation, etc.), words, and sentence recognition scores. The disadvantage is that human observers often do better than the model predicts (Blamey *et al.*, 1989), suggesting that the model is not optimal. The Blamey *et al.* model represents a probabilistic combination of the A recognition response and the V recognition response made independently of each other. However, Braida (1991) showed that predicted AV recognition scores are generally higher when intormation from the two modalities are integrated before response labels are assigned. Thus actual observed AV scores that exceed those predicted by Blamey *et al.* (1989) are likely the result of the observers' use of prelabeling integration processes, and, the greater the deviation from predicted score, the greater the integration efficiency.

3. Sentence recognition

As with consonant recognition, sentence recognition was measured separately for A, V, and AV presentations. Speech materials consisted of the IEEE/Harvard (1969) sentences. These are 720 phonetically balanced low-context sentences each containing five key words (e.g., ''The *birch canoe slid* on the *smooth planks*''). The sentences are organized into 72 lists with 10 sentences in each list. These stimulus materials were filmed at the Massachusetts Institute of Technology using the same female talker used for the VCV materials and dubbed onto an optical disc (Panasonic TQ-3031F). The audio portion of the sentences were treated in the same manner as the consonants.

A and AV sentence recognition was measured at a S/N ratio of 0 dB. For AV presentations, the subjects viewed a 19-in. color video monitor (SONY PVM 2030) from a distance of approximately 5 ft.

The subjects were asked to write down what they perceived in A, V, and AV presentation modes. Subjects were encouraged to respond with as much of the sentence as they could, and to guess whenever they were uncertain. All sentences were scored for the percentage of key words correctly identified. A strict scoring criterion was used in which all tense (e.g., ''–ed'') and number (e.g., ''–s'') affixes were required to be correct for the key word to be correct. On the other hand, homophones were counted as correct (e.g., ''read'' for ''red'').

For each test condition (A, V, and AV), five lists of sentences (ten sentences/list) were presented. Thus each score was based on 250 words (5 words per sentence, 10 sentences per list, 5 lists per condition). The order of the test conditions and lists were randomized for each subject. No feedback regarding the correctness of responses was provided.

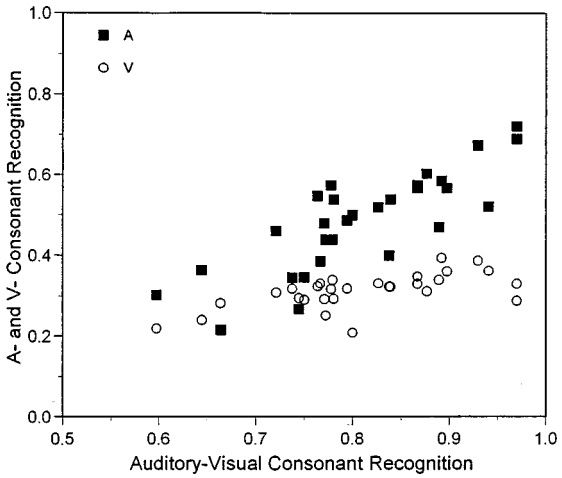


FIG. 3. A and V consonant recognition as a function of AV consonant recognition.

4. Summary

The specific question addressed by the various speech measures obtained in this study was how much of the individual variability in AV sentence recognition could be accounted for by variability in AV consonant recognition? Further, it was assumed that AV consonant recognition is determined primarily by auditory and visual cue extraction abilities and the efficiency at which the unimodal cues could be integrated. The ability to extract auditory and visual cues were determined by analyses of error patterns made in separate A and V consonant recognition tests. To measure integration ability, three models of integration (PRE, FLMP, and PROB) were used to predict AV consonant recognition scores from independent measures of A and V consonant recognition. Integration efficiency was defined as the difference between predicted and observed AV consonant recognition scores. This is the first report, as far as we are aware, to describe individual differences in AV integration efficiency, and how these differences may relate to speech recognition performance.

II. RESULTS

A. Consonant recognition

Variability in AV consonant recognition was examined in terms of A and V consonant recognition. In addition, consonant confusions made in each of the unimodal conditions were examined to determine the relation between AV recognition and unimodal feature recognition.

Focusing initially on overall recognition accuracy, Fig. 3 shows A, V, and AV consonant recognition scores for each of the 29 subjects. Auditory-visual recognition scores (displayed along the abscissa) ranged from 60% to 98% (mean =81%, s.d.=9.5%), whereas A recognition scores ranged from 20% to 74% (mean=49%, s.d.=12.3%). Speechreading (i.e., V-only) scores were less variable across subjects and

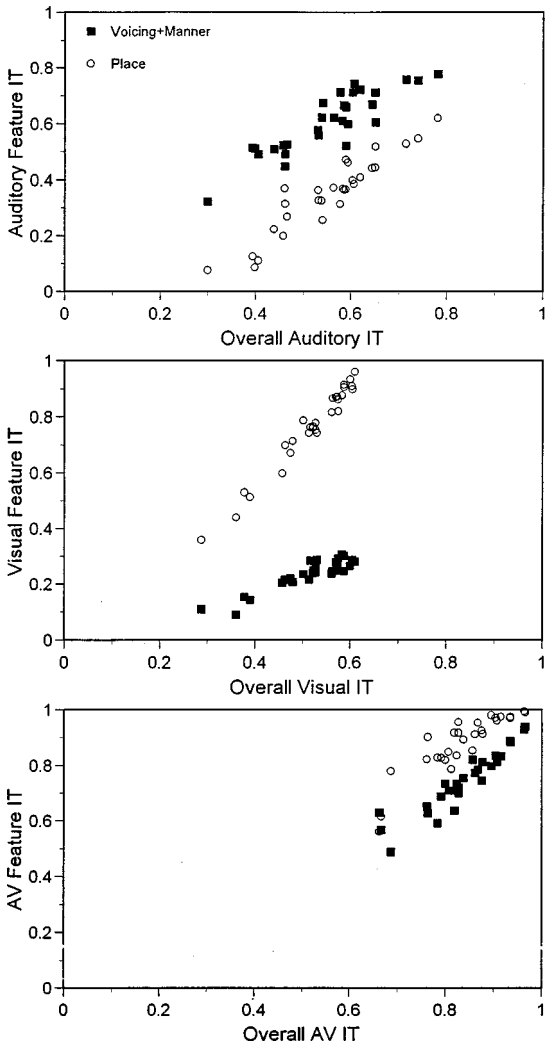


FIG. 4. Voicing-plus-manner and place information transmission for A, V, and AV modalities as a function of overall information transmitted (IT).

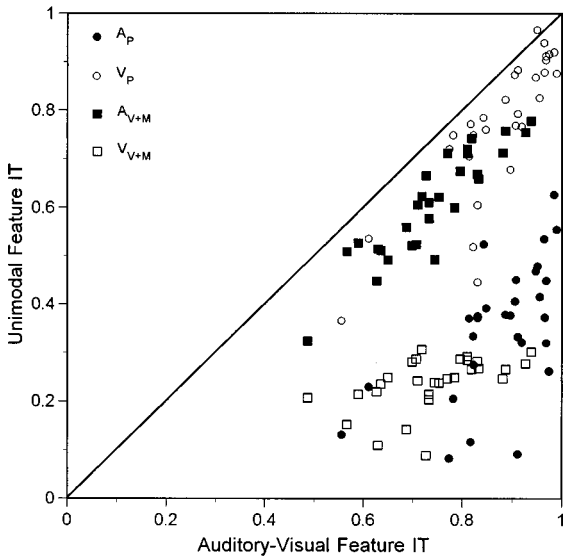


FIG. 5. A and V feature transmission scores as a function of AV feature transmission scores.

ranged from 21% to 40% (mean=31%, s.d.=4.3%). All subjects were able to benefit substantially from the combination of audition and speechreading with the mean performance difference between AV and A test conditions equal to 32% (s.d.=7%). For these materials, this amount of gain is roughly equivalent to a 5–6 dB improvement in signal-to-noise ratio (Grant and Walden, 1996a). Also apparent from the figure is the strong relation between AV and A scores ($r=0.82$) and the weaker relation between AV and V scores ($r=0.63$). Both of these correlations were highly significant ($p<0.001$).

The correspondence between unimodal and bimodal performance shown in Fig. 3 is based solely on overall recognition accuracy in each of the three receiving conditions. It is likely that even higher correlations between unimodal and bimodal performance would result if the patterns of A and V confusions were considered. Previous data obtained by Grant and Walden (1996a), as well as the predicted AV scores shown in Fig. 2, suggest that AV consonant recognition performance is determined in large part by the accuracy with which auditory voicing plus manner cues and visual place cues are extracted from the unimodal conditions. To investigate this, a feature-based information analysis (Miller and Nicely, 1995) was performed on the unimodal and bimodal confusion matrices of the 29 subjects. The accuracy of information transmission (IT) for a variety of speech features was determined, including a 7-category voicing+manner (V+M) feature and a 6-category place (P) feature as described in Table I. The three panels displayed in Fig. 4 show the results for the V+M and P features as a function of the overall IT for A (top), V (middle), and AV (bottom) conditions, respectively. As expected for auditory presentations (top panel) given the subjects' hearing losses and the relatively poor S/N, A_{V+M} cues (mean=0.61) were received more accurately than A_P cues (mean=0.35). In contrast (middle panel), V_P cues (mean=76%) were received more accurately than V_{V+M} (mean=0.24). In the AV condition (bottom panel), AV_P cues (mean=0.88) were received with greater accuracy than AV_{V+M} cues (mean=0.74). This is noteworthy in that place cues, as opposed to voicing or manner cues, have been shown consistently to be more susceptible to the deleterious effects of noise and hearing loss (Miller and Nicely, 1955; Rabinowitz *et al.*, 1996). The implications of this observation for hearing-aid development and rehabilitation will be discussed later.

An attempt was made to relate unimodal feature recognition to AV feature recognition to explore further the connection between A and V segmental cues and AV recognition. Figure 5 shows scatter plots of four unimodal-to-bimodal relations: AV_P vs A_P (filled circles), AV_P vs V_P (open circles), AV_{V+M} vs A_{V+M} (filled squares), and AV_{V+M} vs V_{V+M} (open squares). All four correlations were significant ($r \geqslant 0.53$, $p \leqslant 0.003$). However the strongest correlations were between AV_P recognition and V_P ($r=0.83$) and between AV_{V+M} recognition and A_{V+M} ($r=0.91$). This is not to imply that A_P cues did not contribute to AV_P recognition or that V_{V+M} cues did not contribute to AV_{V+M} recognition. In fact, by combining V_P and A_P scores into a single P score (e.g, by using the PROB model), the correlation with AV_P ($r=0.90$) is significantly higher than the correlation between V_P and AV_P. A similar analysis regarding

AV_{V+M} recognition showed that combining the two unimodal V+M scores hardly improved the correlation previously obtained with the A_{V+M} score alone ($r=0.93$). Therefore, in describing the AV feature IT for the 29 hearing-impaired subjects, AV_{V+M} recognition was determined primarily by A_{V+M} whereas AV_P was determined by the combined influences of V_P and A_P.

B. AV consonant integration

Auditory-visual consonant recognition is assumed to be determined primarily by the amount and type of A and V cues that can be extracted from the speech signal (unimodal cue extraction), and the efficiency with which these cues can be combined across the two modalities (bimodal integration). In order to separate these two factors, a measure of AV integration must be devised. In this section, three models capable of predicting AV consonant recognition from A and V consonant recognition are used to derive such measures. Each model is used to predict AV consonant identification accuracy for individual subjects. The differences between obtained and predicted scores are then used as measures of individual integration skill. It should be noted that subjects with relatively good integration skills need not achieve better overall AV recognition scores than subjects with relatively poor integration skills. As we are using the term, integration only refers to the ability to combine the cues derived from the separate modalities. Therefore, subjects with very poor A and/or V cue resolution might still be excellent integrators if they use these cues to their fullest potential in bimodal perception. The net result, depending on the amount of A and V cues recognized, might still be a relatively poor overall AV score. Conversely, subjects with excellent cue resolution but relatively poor integration skills may end up with a relatively high AV score by virtue of a high A or V score alone. These examples suggest that integration skill might be better related to AV benefit than to AV score. For the purpose of this study, AV benefit was defined as (AV-A)/(100-A), or the recognition improvement relative to the total possible improvement given an individual's A-alone score (Sumby and Pollack, 1954).

Predicted versus observed AV consonant recognition scores are shown in Fig. 6. AV predictions made by the PRE, FLMP, and PROB models of integration are shown separately in the top, middle, and bottom panels of the figure. The Pearson correlations between observed and predicted AV recognition scores were fairly similar for the three models: 0.89 for the PRE, 0.83 for the FLMP, and 0.89 for the PROB. As indicated by the position of the data points relative to the line $AV_{OBSERVED}=AV_{PREDICTED}$ shown in each panel, predictions by the PRE model were either equal to or greater than obtained performance, whereas predictions by the FLMP and PROB models generally underpredicted AV performance.[4] This was especially true for the FLMP when the input unimodal scores were low. In the fixed FLMP model, stimuli identified correctly in one modality but incorrectly in the other are predicted to be incorrect in the combined AV condition [see Eq. (3)]. As Braida (1991) noted, the fixed FLMP model does not properly account for structured errors and relies too heavily on unimodal accuracy. In contrast, the PRE model focuses more on the consistency of unimodal responses (as determined by MDS) and not necessarily on accuracy. Thus the PRE model makes a prediction of optimal performance (not necessarily optimal fit) and would therefore be expected to over predict observed scores.

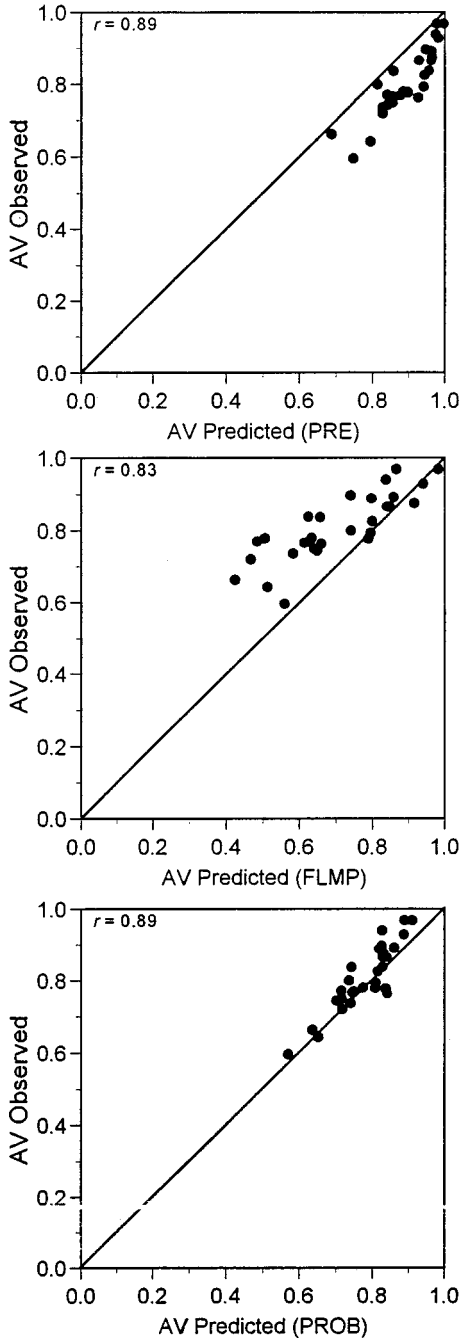


FIG. 6. PRE, FLMP, and PROB model predictions of AV consonant recognition.

In Fig. 7, the derived measure of integration skill (i.e., the difference between predicted and observed scores for individual subjects) is shown with respect to the amount of relative AV benefit obtained. When compared in terms of the ability to predict the relative benefit provided by speechreading, the derived measure based on PRE model predictions accounted for approximately 56% of the variance in relative

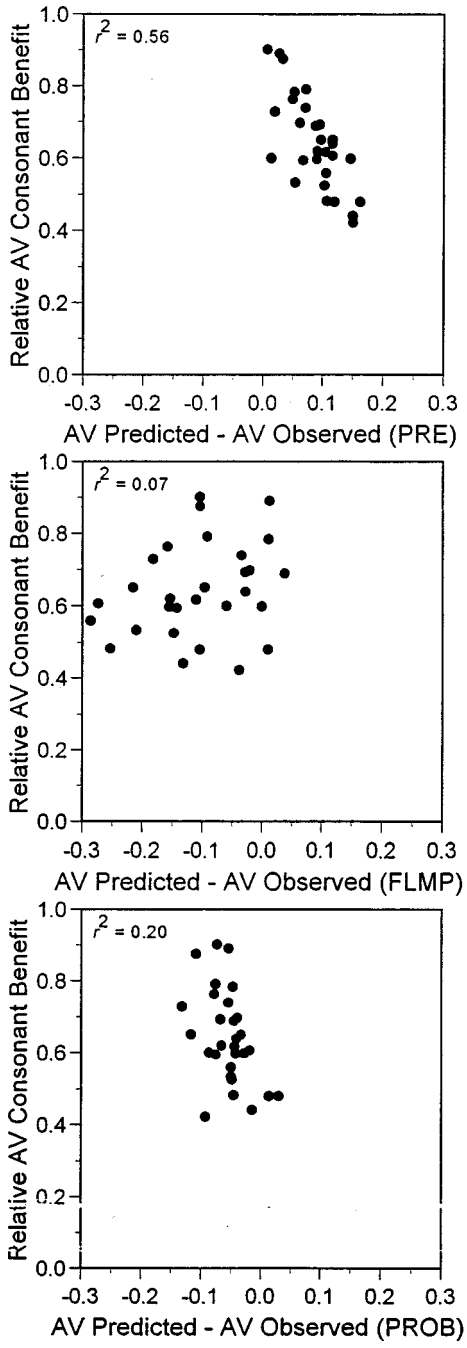


FIG. 7. Relation between derived integration measure ($AV_{PREDICTED} - AV_{OBSERVED}$) and relative AV benefit for consonants.

AV benefit. In comparison, the PROB model accounted for roughly 20% of the variance, whereas the FLMP accounted for about 7% of the variance. The correlations between derived measures of integration efficiency and relative AV benefit were significant for the PRE and PROB models, but not for the FLMP. Empirically, the PRE model appears to provide a better estimate of integration efficiency than either the FLMP or PROB model.

Because the PRE model predicts optimum AV recognition performance, anything less than perfect AV consonant recognition for subjects who are well predicted by the PRE model is probably due to poor cue extraction and not to poor integration. On the other hand, subjects whose AV consonant

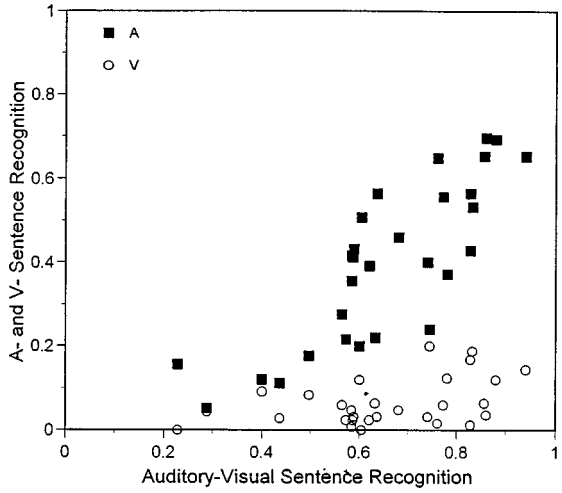


FIG. 8. A and V sentence recognition as a function of AV sentence recognition.

recognition scores fall below PRE model predictions may not be integrating all the unimodal cues available to them. For these subjects, improved AV speech recognition scores may be achieved by focusing on integration training in addition to cue extraction. The amount of AV improvement in consonant recognition that could be expected with integration training is indicated by the performance difference between predicted and observed AV scores shown along the abscissa in the top panel of Fig. 7. The average difference was 8.5% and the maximum difference was 16.1%.

C. Sentence recognition

Recognition results for IEEE/Harvard sentences are shown in Fig. 8. As in Fig. 3, A and V recognition results are shown in relation to AV recognition scores. Auditory-visual sentence recognition of key words (shown along the abscissa) ranged from 23% to 94% (mean=65.5%; s.d. =17.5%). Auditory sentence recognition ranged from 5% to 70% (mean=39.7%; s.d.=19.1%) and visual sentence recognition scores ranged from 0% to 20% (mean=6.5%; s.d. =5.6%). Although there was no appreciable correlation between the subject's A and V scores, and in spite of the very low speechreading scores overall, all subjects benefited from the addition of visual cues. Every subject's AV score was better than his/her A score, even when the V score was 0% correct. A strong correlation was evident between A and AV sentence recognition ($r=0.82$; $p<0.001$). The correlation between V and AV sentence recognition was also significant, but weaker ($r=0.44$; $p<0.02$).

The benefit obtained from combining A and V cues in sentence recognition varied substantially across individuals. To address this issue, relative AV benefit scores, (AV-A/(100-A), were computed for each subject. The average relative benefit was 44% (s.d.=17.8%) with a maximum of 83% and a minimum of 8.5%. The large individual differences in the amount of relative benefit received were only weakly related to the subject's A performance ($r=0.34$). A much better accounting of individual differences in relative AV benefit for sentences was provided by subjects' speechreading scores, in that better speechreaders obtained more AV

benefit for sentences ($r=0.72$).[5] The relatively strong correlation between V sentence recognition and relative AV sentence benefit, in conjunction with the weaker relation observed between visual consonant recognition and relative AV consonant benefit ($r=0.55$), suggests that better speechreaders of sentence materials are able to extract more visual information than the poorer speechreaders with regard to certain cues such as word segmentation and stress, both of which would prove useful in the AV sentence condition (Risberg, 1974; Risberg and Lubker, 1978; Grant and Walden, 1996b). Obviously, these cues would play a much smaller role in consonant recognition than in sentence recognition which might help explain the apparent greater significance of speechreading for AV sentences than for AV consonants.

A second possibility may be that the better speechreaders engage useful higher-level cognitive skills in addition to the bottom-up information extracted from the visual speech signal. With meaningful sentence materials, speechreading requires that subjects not only extract signal cues from visual speech movements, but form linguistic wholes from perceived fragments as well. This ability to perform *perceptual closure* (Wertheimer, 1938; O'Neill and Oyer, 1961; Watson *et al.*, 1996) is useful regardless of modality, especially when the input signals are ambiguous. Therefore, it is possible that better speechreaders of sentence materials have better perceptual closure skills than poorer speechreaders, which would facilitate AV sentence benefit.

D. Relations among consonant and sentence recognition

According to Fig. 1, auditory-visual consonant recognition is primarily a function of A and V cue extraction and cue integration. AV sentence recognition, on the other hand, is a function of A and V signal-cue extraction, cue integration, lexical processes, memory processes, and the use of linguistic and world knowledge in conjunction with semantic and syntactic context. If this conceptual framework is correct, then the proportion of variance in AV sentence recognition data not accounted for by AV consonant recognition is most likely due to individual differences in *top-down* speech recognition processes.

Figure 9 shows the relation between consonant recognition and sentence recognition for A, V, and AV conditions. In all three receiving modalities, the correlations were highly significant ($p<0.007$). For the A and AV conditions, respectively, 52% and 54% of the variance observed in sentence recognition was accounted for by consonant recognition. For the V condition, 25% of the variance in visual sentence recognition could be accounted for by visual consonant recognition. This reduction in the amount of variance explained for the V modality (compared to either A or AV modalities) is probably due to the relatively narrow range of scores observed for visual consonant recognition (mean $=31.2\%$, s.d.$=4.3\%$). In general, however, the data show that variability in A and AV sentence recognition is determined to a great extent by *bottom-up* processes related to cue extraction and cue integration.

With regard to individual differences in the benefit received from combining A and V cues in sentence recognition, the results suggested that speechreading ability, measured either with nonsense syllables or with meaningful sentences, was the most consistent unimodal predictor of AV sentence benefit. For example, measures of relative AV sentence benefit correlated significantly with visual consonant recognition ($r=0.66$, $p<0.001$) and with visual sentence recognition ($r=0.72$, $p<0.001$). In contrast, relative AV benefit was not significantly correlated with auditory consonant or auditory sentence recognition. Finally, although the correlations between measures of relative AV benefit for consonants and relative AV benefit for sentences was significant ($p=0.02$), relative AV consonant benefit accounted for only 18% of the variance in relative AV sentence benefit.

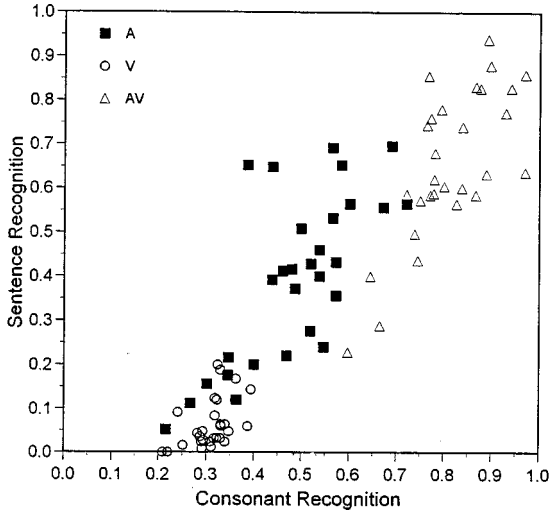


FIG. 9. Relation between consonant recognition and sentence recognition for A, V, and AV modalities.

III. DISCUSSION

The conceptual framework schematized in Fig. 1 highlighted a variety of *bottom-up* and *top-down* sources of information that are likely to be important in AV speech recognition. The general purpose of this work was to further delineate the various factors and perceptual processes that determine individual variability in AV sentence recognition. In this study, we focused on consonant recognition and AV integration at the segmental level and the relation between consonant recognition and the recognition of words in sentences.

Auditory-visual consonant recognition can be reasonably well described as a simple combination of visual place-of articulation cues and auditory manner-plus-voicing cues. This rather simple conception of the problem of predicting AV consonant recognition in individual listeners is supported by the results of Massaro's FLMP and Braida's Prelabeling Model of integration, and by earlier work by Grant and Walden (1996a). As shown in Fig. 2, model predictions of AV recognition given average speechreading skill, depend primarily on which speech features are resolved in the auditory condition, and not on the overall accuracy of auditory speech recognition. For example, individuals who receive only voicing or manner cues through audition are predicted to have higher AV consonant recognition scores

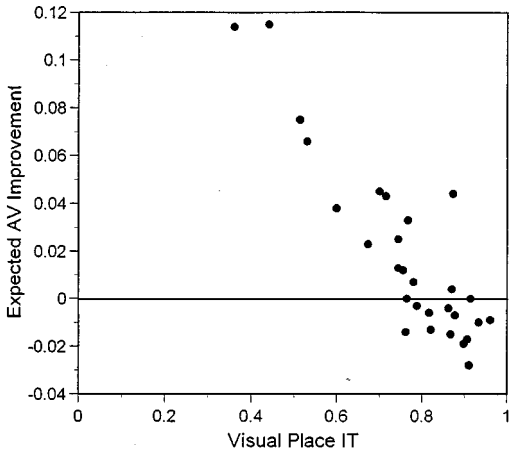


FIG. 10. Predicted AV consonant improvement that would likely result from speechreading training.

than individuals who receive only auditory place-of-articulation cues. These model predictions support the idea that integration of complementary A and V speech features results in higher AV recognition scores than integration of redundant features, as is the case when the A modality contributes primarily place-of-articulation cues. A similar conclusion can be drawn from the study by Grant and Walden (1996a). In that study, normal-hearing listeners were presented with filtered bands of speech in combination with speechreading. Low-frequency bands of speech, which conveyed primarily voicing and manner information, resulted in higher AV recognition of consonants than high-frequency bands which conveyed mostly place information. This was the case, even though the high-frequency bands were often more intelligible than the low bands when presented without speechreading.

In the present study, the ability to extract V+M and P cues from the A and V conditions was quite variable across subjects. For example, when compared to previous speechreading results obtained with normal-hearing subjects (Grant and Walden, 1996a), the HI subjects displayed a significantly wider range of performance for the recognition of V_P cues (36%–96% for the HI subjects as compared to 73%–94% for the NH subjects). For HI subjects with relatively low V_P recognition (below 70%) speechreading training would seem appropriate. Using Braida's model we can estimate the amount of AV improvement expected to result from such training. Model fits were computed for each HI subject using the average visual confusion matrix from Grant and Walden (1996a) in combination with individual auditory confusion matrices and compared to original model fits obtained with the subject's own visual confusion matrix. This average or *generic* visual matrix had a V_P score of 0.87 and represents conservatively what we feel can be achieved through speechreading training (Walden *et al.*, 1977). Figure 10 shows the predicted AV improvements obtained for each subject. According to this analysis, subjects with V_P recognition less than 0.7 would show improvements between 3% and 11.5%.

With regard to auditory consonant features, it is clear that strategies for improving the information transmission of the V+M feature need to be developed, either through auditory training or signal processing. For our subjects, the A_{V+M} feature was received more accurately than the A_P feature, whereas the AV_{V+M} feature was received less accurately than the AV_P feature. Note that the finding for AV feature recognition is precisely the opposite result of classic feature studies of auditory-only listening conditions (e.g., Miller and Nicely, 1955). Manner and voicing information can be represented by time-intensity envelope cues which tend to be more resistant to the effects of hearing loss or background noise than are the rapid spectral transitions and spectral shapes associated with A_P cues (Van Tasell *et al.*, 1987; Rosen, 1989). It is reasonable, therefore, for researchers and clinicians concerned with improving auditory speech recognition to advocate hearing aid fitting and other rehabilitative strategies to improve A_P reception (Rabinowitz *et al.*, 1992). However, to improve AV speech recognition, improving A_{V+M} reception (as opposed to A_P reception) needs to be the focus of rehabilitative efforts, assuming that the individual has at least average speechreading and integration abilities.

The integration processes by which A and V speech cues are combined are not well understood. The results of our efforts to measure this ability across HI listeners suggests that there is substantial intersubject variability in integration efficiency and that better integrators are likely to derive more AV benefit than poorer integrators, at least for consonant recognition.

The integration abilities of individual listeners currently are not subject to direct observation. Instead, they must be derived from speech recognition performance obtained from unimodal and bimodal receiving conditions, and models of the integration process. The approach used in this study was an attempt to partial out the contributions of information extraction and information processing as they relate to AV speech recognition. Braida's PRE model is a theoretically optimal integrator which produces the best possible recognition score for each subject, given their A and V consonant recognition data. The predictions of this model can be used, among other clinical applications, to estimate the performance that may be possible if the subject was able to integrate the available information from the two modalities perfectly. The difference between predicted and observed AV scores suggests that, on average, a gain of 8.5% (maximum gain=16.1%) could be expected for the subjects of this study with appropriate integration training. This model also allows us to test the effectiveness of certain rehabilitation strategies or signal processing algorithms with respect to predicted AV gain. For example, if we eliminated a particular auditory consonant confusion (e.g., /ɑbɑ/ vs /ɑvɑ/) through training or signal processing, it would be possible to calculate the predicted improvement in overall AV recognition score. Similarly, one could estimate the overall benefit of combining several rehabilitation strategies, such as speechreading training and integration training. For example, for the 14 subjects who either received less than 70% place information in the V condition or who were over-predicted by the PRE model by more than 10%, we would expect a potential combined training benefit (speechreading training and integration training) of 9% to 26% depending on the individual subject.[6]

In comparing speech recognition results with nonsense aCa syllables and IEEE sentences, significant correlations were found within receiving condition (i.e., $A_{CONS} \propto A_{SENT}$, $V_{CONS} \propto V_{SENT}$, and $AV_{CONS} \propto AV_{SENT}$). Especially noteworthy was the finding that over 50% of the variability in AV (or A) sentence recognition could be accounted for by AV (or A) consonant scores. In addition, the relative AV benefit for consonants was significantly correlated with the relative benefit for sentences ($r = 0.43$, $p = 0.02$). Thus AV recognition of IEEE sentences and the amount of benefit provided by visual cues to audition is determined, in large measure, by the recognition of medial consonants, which can be fairly well predicted by separate measures of A and V consonant recognition and a measure of AV consonant integration.

When discussing the relationship between AV consonant recognition and AV sentence recognition, it is important to consider the many differences between these two sets of materials. It is well known that fluent speech productions like the IEEE sentences and carefully articulated nonsense syllable productions like the vCv consonant set differ with regard to phonetic (speaking rate, segment duration, etc.), phonological (vowel neutralization, flapping, etc.), lexical, morphosyntactic, and semantic factors. These differences are likely to weaken the association between segment articulation scores and sentence intelligibility scores. In addition, they suggest several possible processes important for fluent speech recognition, such as the ability to use lexical, semantic, and grammatical constraints (Fletcher, 1953; Boothroyd and Nittrouer, 1988), short-term memory, and the processing speed required to make lexical decisions (Seitz and Rakerd, 1996), that may vary significantly across individuals. Given this array of different factors separating the recognition of nonsense medial consonants in a single vowel context from that of meaningful sentence materials, the strength of the observed association across HI subjects found in this study is quite remarkable.

In summary, individual performance on AV speech recognition tasks involving words or sentences ultimately depends on how lexical access is affected by information provided by auditory and visual sources, the processes by which information is integrated, and the impact of top-down contextual constraints and memory processes. Our efforts thus far to evaluate these factors in individual subjects have focused mainly on nonsense syllable recognition (consonant recognition and the recognition of selected speech features), the recognition of certain prosodic contrasts (Grant and Walden, 1996b), and segmental integration skills. These studies have shown that the benefit of visual cues in AV speech recognition can be readily interpreted in terms of the degree of redundancy between A and V cues, with greater redundancy leading to smaller benefits. Furthermore, there appear to be substantial individual differences regarding the efficiency with which A and V segmental cues are integrated. These differences in integration abilities can also lead to fairly large differences in the amount of AV benefit observed. Taken together, an accounting of the redundancy between A and V features and a measure of integration efficiency can account for the bulk of the variability in AV segment recognition. Additional work is required, however, to more fully account for the variability observed in the recognition of words and sentences. Ongoing efforts to expand this work to include the recognition of additional speech segments in different phonetic environments, multi-syllabic consonant and vowel sequences (to mimic speech rates found in fluent productions), measures of AV integration in connected speech, and measures of lexical redundancy and semantic context usage across individual subjects, will no doubt improve our overall understanding of AV speech recognition.

ACKNOWLEDGMENTS

This research was supported by Grant Nos. DC 00792 and DC 01643 from the National Institute on Deafness and Other Communication Disorders to Walter Reed Army Medical Center, and by the Clinical Investigation Service, Walter Reed Army Medical Center, under Work Unit # 2528. The authors would like to thank Dr. Louis Braida for fitting our consonant data with the Prelabeling Model of Integration. We would also like to thank Drs. Dominic Massaro and Michael Cohen for help in implementing the fixed form of the FLMP, and for numerous helpful discussions regarding theories of information extraction and information processing. Helpful comments on a earlier draft of the manuscript were provided by Dr. Winifred Strange and by two anonymous reviewers. All subjects participating in this research provided written informed consent prior to beginning the study. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

[1]The visual confusion matrix used for these model fits was obtained using the same consonant stimuli used in the present study. The results of 11 normal-hearing subjects were pooled to form a single confusion matrix. The overall percent information transmitted (IT) was 55%. The percent place IT was 87%.

[2]AV predictions based on Massaro's (1987) Fuzzy Logical Model of Perception (FLMP) produced nearly identical results as those shown in the figure. Descriptions of the PRE and FLMP model are presented in the methods section.

[3]In Massaro's evaluation of the FLMP, response probabilities obtained from the unimodal A and V confusion matrices are treated as approximations to the auditory and visual truth values. The program STEPIT (Chandler, 1969) is used to derive a set of optimal feature values which are then used in conjunction with Eq. (3) to predict the AV response. This iterative process is accomplished with prior knowledge of the AV response probabilities by minimizing the average root mean square deviation between the predicted and obtained AV responses, the original and adjusted A responses, and the original and adjusted V responses. Thus by manipulating the original A and V response patterns by small amounts, application of the multiplicative rule results in a better fit than if the multiplicative rule were applied to the original unimodal matrices. This version of the FLMP is known as the variable FLMP. The variable FLMP although providing excellent fits to AV data, can fail to demonstrate differences among subjects with similar A and V matrices but different AV matrices. Such subjects are likely to be fit equally well by the variable FLMP model which, according to Massaro, suggests that the integration processes are similar. However, for the purposes of measuring individual differences in AV integration, the model needs to distinguish among subjects according to the efficiency at which A and V cues are integrated, even when subjects use similar integration processes.

[4]The number of subjects whose observed AV consonant scores exceeded their predicted scores was 25 and 27 for the FLMP and PROB models, respectively.

[5]Although the average sentence speechreading score was extremely low, there was nevertheless a range of scores between 0% and 20% correct to

support the correlations between speechreading performance and relative AV sentence benefit. It should be noted that with low-context IEEE materials, speechreading scores of 15%–20% reflect excellent speechreading skills.

[6]Estimates of training benefit assume that speechreading training would increase the number and recognition accuracy of place-of-articulation categories to that of the average normal-hearing subjects studied by Grant and Walden (1996a). Further, it was assumed that integration training would result in integration efficiency ratios (i.e., $AV_{OBSERVED}/AV_{PREDICTED}$) between 0.9 and 1.0.

ANSI (**1989**). ANSI S3.6–1989, ''Specifications for audiometers'' (ANSI, New York).

Blamey, P. J., Cowan, R. S. C., Alcantara, J. I., Whitford, L. A., and Clark, G. M. (**1989**). ''Speech perception using combinations of auditory, visual, and tactile information,'' J. Rehab. Res. Dev. **26**, 15–24.

Boothroyd, A., and Nittrouer, S. (**1988**). ''Mathematical treatment of context effects in phoneme and word recognition,'' J. Acoust. Soc. Am. **84**, 101–114.

Borg, I., and Lingoes, J. (**1987**). *Multidimensional Similarity Structure Analysis* (Springer-Verlag, New York).

Braida, L. D. (**1991**). ''Crossmodal integration in the identification of consonant segments,'' Q. J. Exp. Psych. **43**, 647–677.

Chandler, J. P. (**1969**). ''Subroutine STEPIT—Finds local minima of a smooth function of several parameters,'' Behav. Sci. **14**, 81–82.

Diehl, R., and Kluender, K. (**1989**). ''On the objects of speech perception,'' Ecological Psychol. **1**, 121–144.

Fletcher, H. (**1953**). *Speech and Hearing in Communication* (Krieger, Huntington, NY).

Grant, K. W., and Walden, B. E. (**1996a**). ''Evaluating the articulation index for auditory-visual consonant recognition,'' J. Acoust. Soc. Am. **100**, 2415–2424.

Grant, K. W., and Walden, B. E. (**1996b**). ''The spectral distribution of prosodic information,'' J. Speech Hear. Res. **39**, 228–238.

Green, D. M., and Swets, J. A. (**1966**). *Signal Detection Theory and Psychophysics* (Wiley, New York).

Greenberg, S. (**1995**). ''The ears have it: The auditory basis of speech perception,'' in *Proceedings of the XIIIth International Congress of Phonetic Sciences*, edited by K. Elenius and P. Branderud, ICPhS 95, Stockholm, Sweden, Vol. 3, pp. 34–41.

Halle, M., and Stevens, K. N. (**1991**). ''Knowledge of language and the sounds of speech,'' in *Music, Language, Speech, and Brain*, edited by J. Sundberg, L. Nord, and R. Carlson (Macmillan, Basingstoke, Hampshire), pp. 1–19.

IEEE. (**1969**). *IEEE Recommended Practice for Speech Quality Measurements* (Institute of Electrical and Electronic Engineers, New York).

Klatt, D. H. (**1989**). ''Review of selected models of speech perception,'' in *Lexical Representation and Process*, edited by W. Marslen-Wilson (MIT, Cambridge, MA), pp. 169–226.

Kruskal, J. B., and Wish, M. (**1978**). *Multidimensional Scaling* (Sage, Beverly Hills, CA).

Kučera, F., and Francis, W. (**1967**). *Computational Analysis of Present Day American English* (Brown U.P., Providence, RI).

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (**1967**). ''Perception of the speech code,'' Psychol. Rev. **74**, 431–461.

Lindblom, B. (**1996**). ''Role of articulation in speech perception: Clues from production,'' J. Acoust. Soc. Am. **99**, 1683–1692.

Macmillan, N. A., Goldberg, R. F., and Braida, L. D. (**1988**). ''Vowel and consonant resolution: Basic sensitivity and context memory,'' J. Acoust. Soc. Am. **84**, 1262–1280.

Massaro, D. W. (**1987**). *Speech Perception by Ear and Eye: A Paradigm for Psychological Inquiry* (Earlbaum, Hillsdale, NJ).

Massaro, D. W., and Friedman, D. (**1990**). ''Models of integration given multiple sources of information,'' Psychol. Rev. **97**, 225–252.

McGurk, H., and MacDonald, J. (**1976**). ''Hearing lips and seeing voices,'' Nature (London) **264**, 746–748.

Miller, G. A., and Nicely, P. E. (**1955**). ''An analysis of perceptual confusions among some English consonants,'' J. Acoust. Soc. Am. **27**, 338–352.

Ohala, J. J. (**1996**). ''Speech perception is hearing sounds, not tongues,'' J. Acoust. Soc. Am. **99**, 1718–1725.

O'Neill, J. J., and Oyer, H. J. (**1961**). *Visual Communication for the Hard of Hearing* (Prentice-Hall, Englewood Cliffs, NJ).

Rabinowitz, W. M., Eddington, D. K., Delhorne, L. A., and Cuneo, P. A. (**1992**). ''Relations among different measures of speech reception in subjects using a cochlear implant,'' J. Acoust. Soc. Am. **92**, 1869–1881.

Reisberg, D., McLean, J., and Goldfield, A. (**1987**). ''Easy to hear but hard to understand: A lip-reading advantage with intact auditory stimuli,'' in *Hearing by Eye: The Psychology of Lipreading*, edited by B. Dodd and R. Campbell (Erlbaum, Hillsdale, NJ), pp. 97–113.

Risberg, A. (**1974**). ''The importance of prosodic speech elements for the lipreader,'' in *Visual and Audiovisual Perception of Speech VI. Danavox Symposium*, edited by H. B. Nielson and B. Klamp, Scand. Audiol. Suppl. **4**, 153–164.

Risberg, A., and Lubker, J. L. (**1978**). ''Prosody and speechreading,'' Speech Transmission Lab—Quarterly Progress Status Report, **4**, 1–16.

Rosen, S. (**1989**). ''Temporal information in speech and its relevance for cochlear implants,'' in *Cochlear Implant Acquisitions and Controversies*, edited by B. Fraysse and N. Cochard (Cochlear AG, Basel), pp. 3–26.

Seitz, P. F., and Rakerd, B. (**1996**). ''Hearing impairment and same-different reaction time,'' J. Acoust. Soc. Am. **99**, 2602(A).

Stork, D. G., and Hennecke, M. E. (Eds.). (**1996**). *Speechreading by Humans and Machines* (Springer-Verlag, New York), Models, Systems, and Applications Proceedings of the NATO Advanced Study Institute on Speechreading by Man and Machine, held in Castera-Verzudan, France, 28 August–8 September 1995, NATO ASI Series F: Computer and Systems Sciences, Vol. 150, 1996.

Stevens, K. N. (**1989**). ''On the quantal nature of speech,'' J. Phon. **17**, 3–45.

Stevens, K. N., and House, A. S. (**1972**). ''Speech perception,'' in *Foundations of Modern Anditory Theory* (Vol. II), edited by J. Tobias (Academic, New York), pp. 3–62.

Studdert-Kennedy, M. (**1974**). ''The perception of speech,'' in *Current Trends in Linguistics*, edited by T. A. Sebeok (Mouton, The Hague), pp. 2349–2385.

Sumby, W. H., and Pollack, I. (**1954**). ''Visual contribution to speech intelligibility in noise,'' J. Acoust. Soc. Am. **26**, 212–215.

Summerfield, Q. (**1987**). ''Some preliminaries to a comprehensive account of audio-visual speech perception,'' in *Hearing by Eye: The Psychology of Lip-Reading*, edited by B. Dodd and R. Campbell (Lawrence Erlbaum, Hillsdale, NJ), pp. 3–52.

Van Tasell, D. J., Soli, S. D., Kirby, V. M., and Widin, G. P. (**1987**). ''Speech waveform envelope cues for consonant recognition,'' J. Acoust. Soc. Am. **82**, 1152–1161.

Walden, B. E., Prosek, R. A., Montgomery, A. A., Scherr, C. K., and Jones, C. J. (**1977**). ''Effect of training on the visual recognition of consonants,'' J. Speech Hear. Res. **20**, 130–145.

Watson, C. S., Qiu, W. W., Chamberlain, M. M., and Xiaofeng, L. (**1996**). ''Auditory and visual speech perception: Confirmation of a modality-independent source of individual differences in speech recognition,'' J. Acoust. Soc. Am. **100**, 1153–1162.

Wertheimer, M. (**1938**). ''Laws of organization in perceptual forms,'' in *A Sourcebook of Gestalt Psychology*, edited by W. D. Ellis (Routledge & Kegan Paul, London), pp. 71–88.

Zadeh, L. A. (**1965**). ''Fuzzy sets,'' Information and Control **8**, 338–353.

An analysis of the effect of basilar membrane nonlinearities on noise suppression

Jayadev Billa
BBN Systems and Technologies, Cambridge, Massachusetts 02140 and Department of Electrical Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

Amro El-Jaroudi
Department of Electrical Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

(Received 12 September 1996; accepted for publication 4 February 1998)

Computational models of the peripheral auditory system have largely modeled basilar membrane (BM) mechanics as a linear filter-bank-like entity. Recent mathematical work on the nature of auditory system noise suppression allows us to analyze and argue for the incorporation of BM nonlinearities into these models. This analysis shows that vowel perception improves with increasing presence of BM nonlinearities whereas consonant perception degrades with increasing influence of BM nonlinearities. Experimental results on tones and real speech corrupted by noise corroborate the analysis as well as suggest a novel approach to speech processing. © *1998 Acoustical Society of America.* [S0001-4966(98)04405-1]

PACS numbers: 43.72.Dv, 43.72.Ne, 43.71.Cq, 43.64.Wn [JLH]

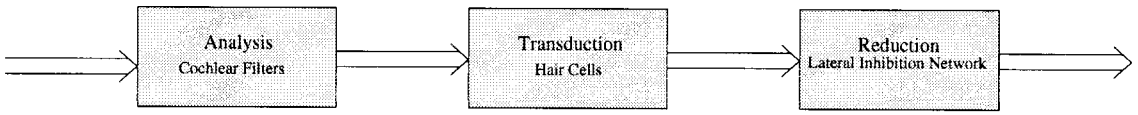
INTRODUCTION

Recent studies have conclusively shown the presence of nonlinearities in the basilar membrane (BM) [e.g., see Robles (1986)]. Nonlinearities in the basilar membrane result in a loss of sensitivity, i.e., broadening of basilar membrane filtering with increasing stimulus levels and has been observed by numerous groups [e.g., see review by Ruggero (1992)]. Although the presence of nonlinearities has long been known, early models essentially assumed the BM to be linear with the nonlinearities concentrated elsewhere (Seneff, 1986; Shamma *et al.*, 1986; Meddis and Hewitt, 1991). As a consequence, most models of the auditory periphery followed a linear filtering stage with the various nonlinearities assumed to be present within the auditory periphery. This type of modeling was able to capture most of the observed experimental results from studies of neural representation of sounds within various species. Work by Deng *et al.* (Deng and Geisler, 1987; Deng *et al.*, 1988) showed that the presence of a nonlinearity within the BM could capture certain salient characteristics of experimental neural data which cannot be captured by a linear BM model followed by nonlinearities modeled at the transduction stage. There is still much controversy over exactly where and how these nonlinearities occur within the BM. For this discussion, we limit ourselves to the effects of such nonlinearities, specifically the observed broadening of BM filtering with increasing stimuli levels.

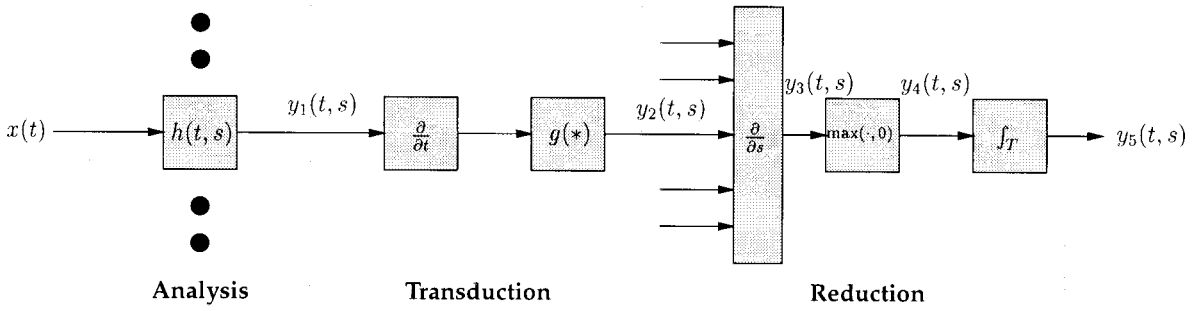
The noise suppression abilities of the peripheral auditory system (PAS) has been demonstrated by many groups, both in physiological experiments (Delgutte and Kiang, 1984a; Geisler and Gamble, 1989; Sachs *et al.*, 1983) and in the responses of models of the PAS (Deng and Geisler, 1987; Deng *et al.*, 1988; Deng and Kheirallah, 1993; Ghitza, 1988; Payton, 1988). This noise ''robustness'' is commonly attributed to the filtering action of the BM. Linear time-invariant filtering approaches have allowed several groups (Ghitza, 1988; Payton, 1988; Ghitza, 1992; Seneff, 1986) to obtain encouraging results for speech processing. Some researchers (Ghitza, 1993, 1994; Giguère and Woodland, 1994a, b) have incorporated BM nonlinearities into their models to accurately model data from physiological experiments. However, the lack of any compelling evidence that such an inclusion would help in applications, such as speech processing, provides little motivation for their integration into such models.

Recent work, initially by Yang *et al.* (1992) and later by Wang and Shamma (Wang *et al.*, 1993; Wang and Shamma, 1994), on the representation of sounds within the early auditory system provides insight into the nature of noise suppression by the PAS. Their work accounted only for the case of a linear time-invariant BM. In this work we account for the effects of basilar membrane nonlinearities, in particular, the broadening of BM filtering with increasing stimulus levels. We proceed by defining a suitable measure that responds to this broadening of BM filtering and accounting for it in our analysis. This analysis shows that in empirical models of the auditory periphery, BM nonlinearities act in a manner that strengthens perception of vocalic sounds such as vowels in noise but degrades perception of consonants such as fricatives supporting evidence from physiological studies such as Delgutte and Kiang (Delgutte, 1984; Delgutte and Kiang, 1984a, b, c, d) and Geisler and Gamble (1989). This extension is the essential point of this correspondence.

The organization of the paper is as follows: Section I briefly reviews the early auditory system model used in Wang and Shamma's (1994) original analysis. In Sec. II, we outline Wang and Shamma's analysis forming the background for our extension. Our analysis extending Wang and Shamma's work to account for basilar membrane nonlinearities proceeds in Sec. III. Demonstrations of the effect of nonlinear BM filtering on tone, natural vowel stimuli, and natural consonant stimuli are provided in Sec. IV. Finally in Sec. V we conclude with a brief summary and possible implications of this work in areas such as speech processing and recognition.



(a) Macroscopic view.



(b) Expanded view.

FIG. 1. The early auditory system model. A functional view of the peripheral auditory model is shown in (a), and an expanded view is presented in (b). The analysis stage corresponds to basilar membrane filtering modeled as a filter bank, the transduction stage to transduction of basilar membrane motion into neural firings in the auditory nerve and the final reduction stage corresponding to information extraction in higher areas of the auditory nervous system. See text for a detailed description.

I. REVIEW OF THE EARLY AUDITORY SYSTEM MODEL

The early auditory system constitutes the system that performs the initial transformation of auditory information (sound) from pressure differentials in air to the initial neural representations in the auditory nerve. These initial neural representations are then further processed at higher regions of the auditory nervous system. The interested reader is referred to a standard text such as Pickles (Pickles, 1988) for an in-depth introduction to the anatomy and physiology of the auditory system. Typically, models of the early auditory system are composed of three cascaded stages: analysis, transduction, and reduction as shown in Fig. 1(a). The model used in our analysis strongly follows the one described in Wang and Shamma (1994). The only modifications are that certain stages, which are not central to the analysis, have been discarded.

A. Analysis

Sound, in the form of pressure differentials in the surrounding air, impinges on the tympanic membrane setting up sympathetic vibrations which are then transmitted to the inner ear (cochlea) via the middle ear ossicles. The cochlea can be likened to a fluid-filled spiral tube separated into two primary ducts by the cochlear partition which consists of the basilar membrane (BM) and associated structures. The stiffness and width of the BM varies significantly from a narrow and stiff base to a much broader and compliant apex. The change in stiffness causes the BM to respond selectively to different frequencies, the base being most responsive to high frequencies and the apex most sensitive to low frequencies. This sensitivity results in a filtering action that decomposes the acoustic signal into spectral bands akin to a filter-bank decomposition but differing in the nonlinear intensity-dependent nature of BM filtering. The cochlea thus decomposes the input signal into constituent spectral components at different spatial locations in an orderly manner. See Pickles (Pickles, 1988) for a more detailed and in-depth exposition.

Functionally, one may model the BM as a linear time-invariant (LTI) filter bank and this is indeed the most common approach in computational models used in applications other than the actual modeling of the PAS. Such an approach has the advantage that it is relatively simple to implement, but suffers in that BM nonlinearities are completely ignored. These nonlinearities may be important in some applications like speech processing, indeed some models incorporating BM nonlinearities have demonstrated significant noise robustness, e.g., Deng *et al.* (Deng and Geisler, 1987; Deng *et al.*, 1988).

Our current model differs from many models of the auditory periphery in that the assumption of a linear and broad BM filtering has been discarded in favor of an explicit attempt to crudely model the BM filter bank as a dynamic system that changes its characteristics with sound-pressure level (SPL). This has allowed us to investigate in a tractable manner the effect that such BM nonlinearities has on the resulting peripheral auditory representation. We have indicated in the Introduction that BM nonlinearities affect filtering at high SPLs and at low SPLs the filtering is essentially linear. This may not be a commonly accepted view. However, due to the phenomenological nature of our model this course does not affect the analysis presented here, an argument for BM nonlinearities at low SPLs and progressive linearization of the BM still results in the same analysis.

Ideally one would like to model the filter bank as a dynamic system that changes its characteristics with sound-

pressure level. In practice, however, this is not a trivial issue [e.g., see Carney (1993) for one approach]. For the work presented, we make the reasonable assumption that the BM is LTI at a *particular* sound level. Thus in our implementation we use three different filter banks each of which represent the state of the BM filtering at one of three different sound levels. This approach is similar to that of Jenison *et al.* (1991) in that we try to capture the effect of BM nonlinearities by the use of several linear time-invariant filter banks. In later work, Jenison (1991) used a dynamic model which incorporated nonlinearities based on the responses of single auditory-nerve fibers directly into the model. We have, however, decided to follow the original idea in Jenison *et al.* (1991) in view of its simplicity. The choice of implementation of the nonlinear BM filter bank has no effect on the analysis itself, the phenomenological nature of analysis renders it valid for any implementation scheme.

If we denote the sound signal as $x(t)$, from Fig. 1(b), the output of the BM model is given as[1]

$$y_1(t,s) = x(t) *_t h(t,s), \tag{1}$$

where $*_t$ represents convolution in time and $h(t,s)$ the response of the cochlear filter at location s on the basilar membrane at any chosen sound level.

The filters chosen in our implementation were designed to be highly asymmetric and to resemble basilar membrane filters at high, moderate, and low sound-pressure levels. The filter for moderate sound-pressure levels is virtually identical to the one described in Wang and Shamma (1994) which was designed to produce the effects of upward masking and two-tone suppression. The filters for low and high sound levels were designed to follow the reasoning in Sec. III, i.e., to have no change in the high-frequency edge of cochlear filters and a marked change in the low-frequency edge of the cochlear filters. These filters were not designed with any considerations other than the ones described above. This is an overtly simplified model of the filtering stage, intended solely to provide simulation results in a tractable manner. As previously stated, the analysis presented here is independent of the implementation and addresses only the effects of the broadening of BM filtering without regard to the nature or physical process by which this broadening is achieved.

The cochlear filters responses are also related by a simple dilation (Wang and Shamma, 1994; Jenison *et al.*, 1991). This allows us to simply design a single seed filter and to obtain the remaining filters of the filter bank by dilating it. This results in filters that are equally spaced on a logarithmic scale. Mathematically:

$$H(\omega,s) = H_{\text{seed}}(a^s \omega) \tag{2}$$

for some constant a and seed filter $H_{\text{seed}}(\omega)$.

B. Transduction

Basilar membrane motion is transformed into neural firings through its coupling to the auditory nerve. Initially, the motion of the basilar membrane causes deflections in inner hair cell cilia. This ciliary displacement is transformed into a modulation of receptor potential which, in turn, modulates the release of a chemical neurotransmitter. The neurotransmitter release then results in depolarization of the auditory-nerve fibers innervating the inner hair cells.

These stages of fluid-cilia coupling, cilia deflection-electrical potential transformation, and neurotransmitter release can be effectively modeled as a three-stage model of transduction: a velocity coupling stage (a time derivative), a memoryless, compressive nonlinearity (a sigmoid-like function), and a low-pass filter in cascade.

The low-pass filter essentially removes frequencies above 4 kHz. For frequencies below 4 kHz, the low-pass filter is in its passband and phase locking is in effect. Under phase locking, neural firing patterns are largely governed by the more intense frequency components of the input stimulus. Phase locking is observed only for frequencies below 4 to 5 kHz, beyond which the inner hair cell receptor potentials can no longer effectively modulate the neurotransmitter release. Due to the high correlation between signal harmonics and neural firing under phase locking, the fine temporal variations in the structure of the cochlear filtered input signal are preserved. This retention of temporal variations by phase locking is important in the subsequent enhancement of phase cues by the lateral inhibition network (LIN) described in Sec. I C. As this analysis mostly concerns itself with the lower frequencies where phase locking is in effect and the low-pass filter is mostly in its passband we have ignored the filter in the current treatment.

The output of this stage can be written as

$$y_2(t,s) = g(\partial_t y_1(t,s)). \tag{3}$$

Consistent with the earlier analysis of Wang and Shamma, we consider the simplifying assumption of a high-gain nonlinearity, in which case, the nonlinearity reduces to a step function.

The justification for this essentially one of simplification. The original formulation allowed for a compressive nonlinearity to model the effects of the nonlinear transduction of ciliary displacement to ionic conductance (Shamma *et al.*, 1986). Further, this nonlinearity was described by a monotonically nondecreasing sigmoidal function (Shamma *et al.*, 1986; Wang and Shamma, 1994). The hair cell does have a finite dynamic range of about 30–40 dB. Below this range, the gain is negligible and, above it, there is negligible change in gain due to saturation of the hair cell response. Our approach here of replacing the sigmoidal nonlinearity with a high gain nonlinearity, the step function, allows us to develop a more tractable analysis. It should be noted that at most sound levels the nonlinearity is driven into its high gain region and thus this assumption does not significantly affect the analysis.

A comprehensive treatment of the model and background and justification for the assumptions can be found in Shamma *et al.* (1986); Yang *et al.* (1992); Wang and Shamma (1994).

C. Reduction

The neural firings produced by the depolarization of the auditory-nerve fibers are then passed on to the cochlear nucleus. The cochlear nucleus and in turn the higher levels of

the auditory nervous system integrate and process the information in the auditory nerve before finally extracting the biologically important acoustic features.

The current model uses a lateral inhibition network (LIN) to model the information enhancement and extraction process. Lateral inhibition networks are found in many biological systems and can be modeled (in their simplest form) as a single layer feedforward or recurrent neural network with mutually inhibitory weighting. A comprehensive overview of the present discussion can be found in Shamma (1985a, b, 1988) and Wang and Shamma (1994).

Essentially the LIN seeks to extract spatial cues and enhance the resulting representation by means of a spatial derivative operation which simply subtracts the output of adjacent channels. At low frequencies the average rates and synchrony patterns provide the spatial cue and for higher frequencies this is primarily due to the rate profiles which display peaks at the characteristic frequencies (CFs) that correspond to stimulus harmonics. The LIN presented in this paper was initially proposed by Shamma (1985b, 1988) and was able to extract key formant and harmonic features from neural firing patterns of the cat auditory nerve. The current model follows the one described in Wang and Shamma (1994). This spatial differential approach to enhancing the representation has been advocated in various forms by other authors. Holton *et al.*'s ''local time-domain phase-correlation'' approach suggests using a ''spatial filter'' on the output of local synchrony detectors to enhance formant information of vocalic segments (Holton *et al.*, 1992). This approach of synchrony detection and spatial filtering closely resembles the LIN based spatial cue extraction scheme presented here. Also relevant here is Huggins' phase model (Huggins, 1952) which suggested that the phase cues could be an effective information carrier. Following from the fact that for linear filters, rapid changes in amplitude response also correspond to rapid changes in phase response, Huggins' phase model shows that frequency components can be sharpened and extracted from the phase of damped input stimuli (Huggins, 1952). This is very similar to the current treatment whereby the LIN uses local phase shifts, which perform the function of frequency markers marking key frequency components, to extract spatial cues.

The current inhibitory network is implemented as a two-stage process. An initial spatial derivative stage that models the lateral coupling/inhibition within the LIN neuron layer and a half-wave rectifier that models the neuronal nonlinearity. This is followed by a temporal integration that mimics the inability of the auditory nervous system to follow rapid temporal variations. This deviates slightly from the original model as described in Wang and Shamma (1994) in that the spatial derivative is considered ''perfect'' as opposed to the ''leaky'' nature in the original formulation. Though this is a slightly unrealistic assumption it greatly simplifies the analysis and implementation and does not detract from the basic thrust of the analysis presented here. In fact the implementation of Wang and Shamma also makes this assumption.

Mathematically the LIN operation can be written as

$$y_3(t,s) = \partial_s y_2(t,s) \tag{4}$$

$$= g'(\partial_t y_1(t,s))\partial_s \partial_t y_1(t,s), \tag{5}$$

$$y_4(t,s) = \max(y_3(t,s),0), \tag{6}$$

$$y_5(t,s) = y_4(t,s) *_t \Pi(t), \tag{7}$$

where the final temporal integration stage has been denoted as a time convolution with a low-pass filter $\Pi(t)$. The spatial difference $\partial_s y_2(\cdot)$ is essentially a filtering operation. The notion of the LIN as a spatial difference filter will be used throughout this paper to further the analysis presented here.

The output of this stage $y_5(t,s)$, allowing for all our approximations and possible modeling errors, represents the sound as it would be portrayed to the higher regions of the auditory nervous system. This output is referred to as the auditory spectrum.

As mentioned earlier, under phase locking, fine temporal variations in the response are retained as phase cues. The LIN subsequently extracts and enhances these phase cues as a spatial operation (Shamma, 1985b). In regions where phase locking is no longer in effect, these cues are produced by the rate profile of neural responses, which show peaks at the appropriate CF (Shamma, 1985b). The auditory spectrum can be described as an enhanced measure of the synchrony content of the stimulus. Thus it retains the salient aspects of the incident sound, much like that provided by the Fourier transform but in a completely different manner.

The auditory spectrum described here is similar to two other approaches: the Ensemble Interval Histogram (EIH) approach (Ghitza, 1992) and the ''dominant frequency'' approach (Carlson *et al.*, 1975). The EIH is formed by passing sound through a set of cochlear filters each of which is followed by level crossing detectors, the output of the ensemble of level-crossing detectors then represents the firing activity in the auditory-nerve fiber array (Ghitza, 1992). The ''dominant frequency'' approach similarly possesses a cochlear filter bank but follows each channel filter with zero-crossing counting averaged over a 100-ms interval converted to frequency (Carlson *et al.*, 1975). A histogram formed by counting the number of channels which have the same frequency within a quantized range of 75 Hz then produces the ''spectrum'' (Carlson *et al.*, 1975). All these approaches deliver a representation of the frequency content by means of a measure of coherent neural activity in the auditory nerve. The notable difference between the current approach and that of others is the latter's emphasis on the temporal details conveyed by the nerve fiber without regard to the spatial details conveyed in the nerve fiber identity, i.e., ''temporal-nonplace'' approach versus the current ''temporal-place'' advocated here. The temporal-place approach is not unique to this work and has been described by others (Young and Sachs, 1979; Shamma, 1985b; Holton *et al.*, 1992).

The next section reviews the stochastic analysis of the model due to Wang and Shamma (1994) allowing for our modifications to the model as described above.

II. NOISE SUPPRESSION IN THE EARLY AUDITORY SYSTEM ASSUMING A LINEAR BM

If the sound signal is modeled as a zero mean random process it follows, from the analysis by Wang and Shamma

(1994), that all the outputs at various stages $y_1(t,s)\cdots y_5(t,s)$ will also be random processes that are not necessarily zero mean. For such a situation, assuming the properties of the filter $\Pi(\cdot)$ are known, it can be argued that $y_5(t,s)$ is a first-order statistic of the auditory spectrum defined as the ergodic mean of $y_4(t,s)$. The auditory spectrum is then characterized by $E[y_4(t,s)]$, which, for a monotonically nondecreasing nonlinearity $g(t)$, is given by Wang and Shamma (1994):

$$E[y_4(t,s)]=E[g'(U)E(\max(V,0)|U)], \tag{8}$$

where $U=\partial_t y_1(t,s)$ and $V=\partial_s\partial_t y_1(t,s)$.

For our prior choice of hair cell nonlinearity as a step function, (8) can be rewritten as in Wang and Shamma (1994):

$$E[y_4(t,s)]=E(\max(V,0)|U=0)f_u(0), \tag{9}$$

where $f_u(0)$ is defined to be the probability density of $U=0$.

Further, if we let the sound signal $x(t)$ be weakly stationary zero mean Gaussian and define r, σ_u, σ_v to be the correlation coefficient and the standard deviations of U and V, respectively, one can further reduce (9) to

$$E[y_4(t,s)]=\frac{1}{2\pi}\frac{\sigma_v}{\sigma_u}\sqrt{1-r^2}. \tag{10}$$

If we denote the spectral density of input sound $x(t)$ as $S_x(\omega)$ we also have

$$\sigma_u^2=\frac{1}{2\pi}\int|\omega H(\omega,s)|^2 S_x(\omega)d\omega, \tag{11}$$

$$\sigma_v^2=\frac{1}{2\pi}\int|\omega\partial_s H(\omega,s)|^2 S_x(\omega)d\omega, \tag{12}$$

$$r=\frac{\sigma_{uv}}{\sigma_u\sigma_v}, \tag{13}$$

$$\sigma_{uv}=\frac{1}{2\pi}\int\omega^2 H^*(\omega,s)\partial_s H(\omega,s)S_x(\omega)d\omega, \tag{14}$$

where from (2),

$$\partial_s H(\omega,s)=(\log a)\omega\partial_\omega H(\omega,s). \tag{15}$$

Now to estimate the noise-suppression properties of the auditory spectrum, consider the case of white Gaussian noise as input [i.e., $S_x(\omega)=1$]. We can write (10) as[2]

$$E[y_4(t,s)]=\frac{1}{2\pi}\frac{\sigma_{v|n}}{\sigma_{u|n}}\sqrt{1-r_n^2}, \tag{16}$$

where $\sigma_{v|n}$ and $\sigma_{u|n}$ are the standard deviations of U and V, respectively, in the presence of Gaussian white noise.

Now consider the presence of a tone of frequency ω_0 in the input with the noise, i.e., $S_x(\omega)=\sigma^2\pi\delta(\omega\pm\omega_0)+1$, σ^2 being equal to the signal-to-noise ratio (SNR). For this case, (10) reduces to

$$\sigma_v^2=\sigma_{v|n}^2,$$

$$\sigma_u^2=\sigma_{u|n}^2+\sigma^2|\omega_0 H(\omega_0,s)|^2=\frac{\sigma_{u|n}^2}{\beta^2},$$

$$r=r_n\beta,$$

with β defined as

$$\beta=\frac{\sigma_{u|n}}{\sigma_u}=\left(1+\left|\frac{\sigma\omega_0 H(\omega_0,s)}{\sigma_{u|n}}\right|^2\right)^{-1/2}. \tag{17}$$

All of which yields, for a tone in noise,

$$E[y_4(t,s)]=\frac{1}{2\pi}\frac{\sigma_{v|n}}{\sigma_{u|n}}\cdot\beta\cdot\sqrt{1-r_n^2\beta^2}. \tag{18}$$

From (17) and (18) one can see that for filter channels far from the tone frequency, the filter removes most of the spectral energy around the tone frequency, i.e., $H(\omega_0,s)\approx 0$ and β varies little, but for channels whose filtering retains most of the spectral energy at the tone frequency, β is significantly smaller than in the prior case. At the channel corresponding to the tone frequency itself, the tone falls in the passband of the differential filter, thus, $\sigma_{v|n}/\sigma_{u|n}$ is large relative to the neighboring channels. As a result the tone passes through the model retaining most of its spectral characteristics. The net effect on the auditory spectrum is then to ''suppress'' channels close to the tone frequency ω_0 while at the same time retaining the key spectral features, in this case the features of the tone.

Alternatively, from (10) a more intuitive characterization is possible. Noting that σ_u^2 and σ_v^2 can be associated with the spectral energy in the cochlear and differential bands, $E[y_4(t,s)]$ represents the self-normalized spectral profile and hence auditory spectrum. The term *self-normalized* arises from the fact that the spectral profile, the output of the BM, is divided by a smoothed spectral profile of itself, the output of the differential component (Wang and Shamma, 1994). This normalization is not easily quantified, as it depends very much on the spectral profile of the input signal. The enhancement (normalization) received by spectral peaks resolved by the narrow-band differential filter will be high, resulting in a ''heightening'' of spectral peaks, as the cochlear filter integrates energy in the neighboring valleys as well as at the peak. This results in an overall enhancement of the spectral representation of the stimuli.

In the presence of noise, it should be clear from the previous argument that enhancement occurs for SNRs that *at least* locally exceed 0 dB. This is not such a stringent requirement as it would seem, even at global SNRs$<$0 dB, the vocalic portions typically exceed the 0 dB SNR level. For example, consider a tone in noise, even when the global SNRs$<$0 dB, if we take any period over which the tone is present the spectrum will still show a strong peak at the tonal frequency. This peak will be enhanced more if we had band-passed the signal around the tonal frequency, thus locally increasing a relatively low global SNR. Unvoiced portions would, however, suffer significant degradation and would typically not be as enhanced as the vocalic portions. This is as much a function of the structural differences between voiced and unvoiced sounds as that of the signal energies. Most voiced stimuli have strong frequency components with relatively large signal energy due to their underlying periodicity whereas most unvoiced stimuli have weak frequency

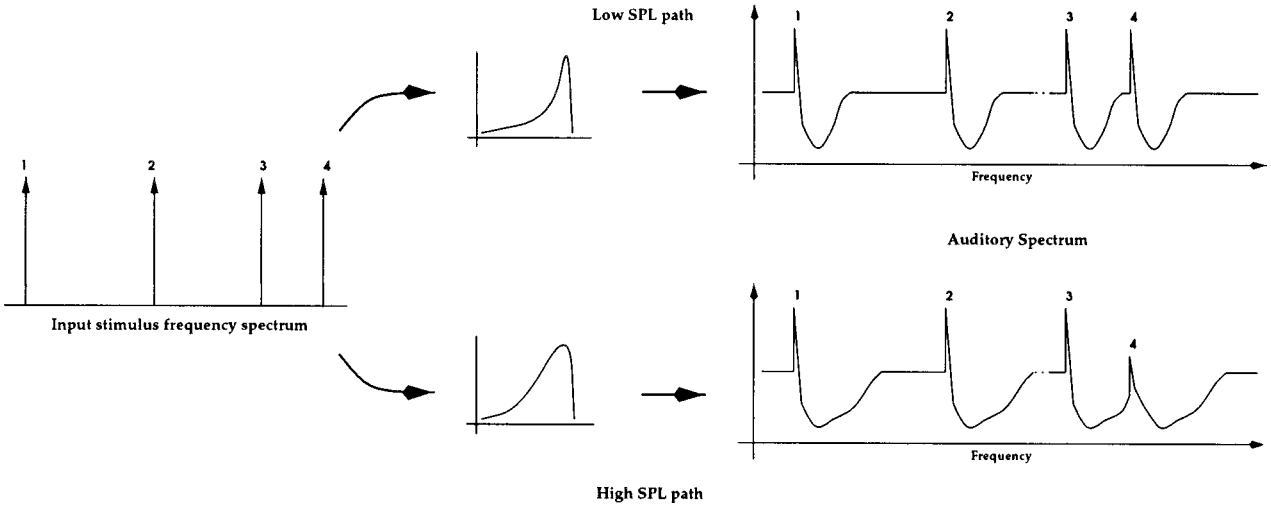


FIG. 2. An exaggerated demonstration of the auditory spectrum of a contrived input stimulus at low and high SPLs. The top path shows the auditory spectrum for low SPL input stimulus, the lower path, the auditory spectrum for a high SPL input stimulus.

components with relatively low signal energy spread out over the frequency range due to exactly the opposite, i.e., the lack of such periodicity.

Of interest here is the relationship with Young and Sachs' average localized synchronized rate (ALSR) representation (Young and Sachs, 1979). Both the current auditory spectrum and ALSR approaches use a spatial tonotopic representation to describe the stimulus spectrum. Briefly, to obtain a measure of frequency content at a particular frequency the ALSR algorithm ''filters'' the spectrum obtained via a Fourier transform of the period histogram of each fiber, by forming an average of the response of fibers whose CFs lie in a narrow band around that frequency. Their approach is to focus on the encoding potential in a single channel, the ro-

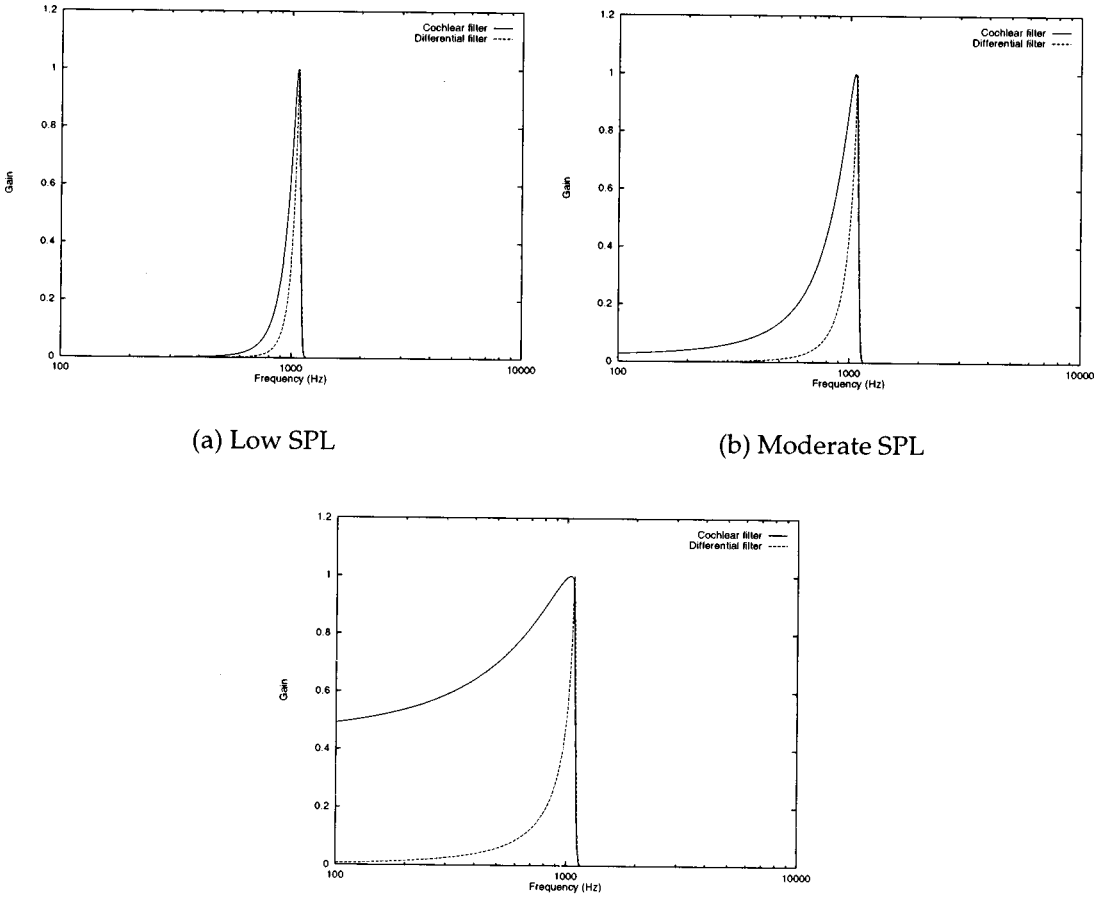


FIG. 3. Seed filters. The solid line is the magnitude response of the cochlear filter, the dotted line is the corresponding differential filter.

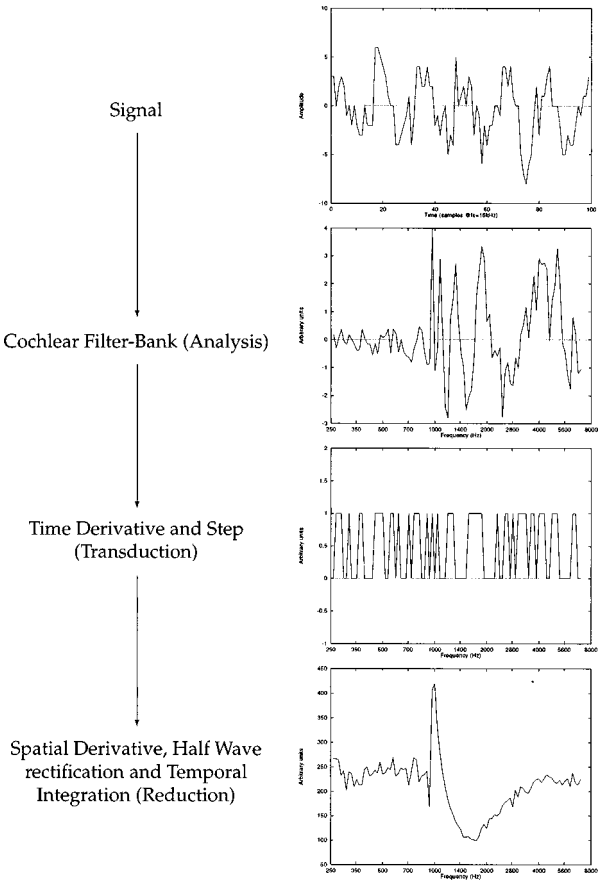


FIG. 4. A pictorial description of the procedure used to generate the auditory spectrum shown in Fig. 5(c) for a tone in noise at 0 dB SNR. The top panel shows the input stimulus, a 1-kHz tone in noise at 0 dB SNR. The next three panels show the output at each of the major stages of the model plotted as the CF's of cochlear channel (x axis) versus their output (y axis).

bustness arising from ''suppressing'' the encoding potential of other channels. The current analysis describes essentially the same idea with its *channel specific* self-normalization. There are also some key differences between the ALSR approach and the current treatment, the most important of which is the emphasis on cross-channel correlation in the current approach which the ALSR essentially ignores. The effect of the cross-channel correlation is to emphasize the parts of the spectral profile that show the most change.

The expressions above, due to Wang and Shamma (1994), form the basis of our analysis in the next section. Our goal is to describe the effect of basilar membrane nonlinearities, specifically those that result in the broadening of BM filtering, on the above analysis.

III. NOISE SUPPRESSION IN THE EARLY AUDITORY SYSTEM ASSUMING A NONLINEAR BM

In order to describe the effect of a change in filtering characteristics of the basilar membrane we need to define a suitable characteristic that responds to this change. One such characteristic is the average power of the filter channel $H(\omega,s)$ defined as

$$E[|h(t,s)|^2]=\int |H(\omega,s)|^2\, d\omega. \tag{19}$$

Further, let us define a *filter form factor*, Γ, as the ratio of the root average power of the filter and a fixed reference filter, i.e.,

$$\Gamma_s=\left(\frac{\int|H(\omega,s)|^2\, d\omega}{\int|H^{\mathrm{ref}}(\omega,s)|^2\, d\omega}\right)^{1/2}, \tag{20}$$

where the explicit dependence of Γ on the channel for which it is computed is shown.

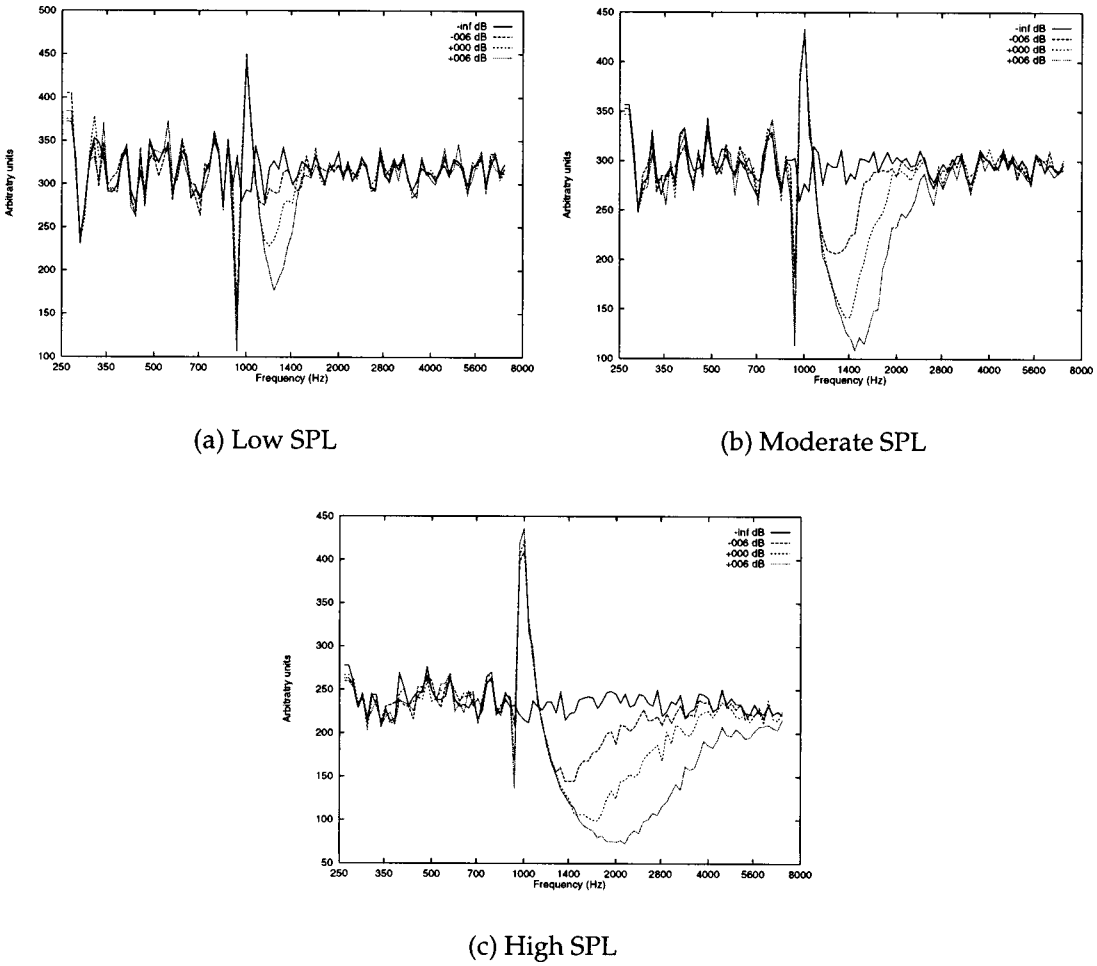
With this definition we can condense the complex nonlinear amplitude dependent behavior of BM filtering into a simple amplitude dependent constant, Γ. Though this simplification ignores the richness of BM nonlinearities it does, however, allow us to develop a reasonable and tractable analysis of its behavior.

We now contend that, for cochlear filters, the overlap is such that the differential filter is essentially independent of the originating cochlear filters, i.e., the nonlinearities only affect the cochlear filter. First, at least up to about 2 kHz, there is the observed lack of any appreciable change in the high-frequency edge of cochlear filters [e.g., the discussion in Secker-Walker and Searle (1990)]. Second, there is a high degree of overlap between the responses of adjacent channels. Assuming 3000 channels and equal logarithmic distribution of CFs in a human cochlea, the overlap between adjacent channels will still not exceed 100 Hz, compared to a maximum BW of approximately $\frac{1}{10}$ of an octave at normal conversational sound levels (Evans, 1982). Clearly a differential filter would possess a bandwidth (BW) significantly smaller than the originating cochlear filters at almost all sound levels. For these reasons we will associate the factor Γ exclusively with the increase in BW of the cochlear filter, all other effects such as those due to the increased energy in the channel due to increased sound level are completely ignored. This is not a strict assumption as the self-normalization eventually cancels out this additional channel energy. Also, following from fact that the LIN is a phase detector and from our notion of the LIN as a differential filter (see Sec. I C), if the phase is stable over a wide range of intensity it implies that the differential filter changes little over a wide range of intensity. Indeed, the phase of basilar membrane motion has been shown to be stable over wide ranges of intensity (Huggins, 1952; Kim *et al.*, 1980). This further supports our contention that the differential filter is essentially independent of the originating cochlear filters.

With these approximations, consider now the effect on r, σ_u, and σ_v for the case of a tone in noise. Only σ_u is directly influenced by our discussion and is specified by

$$\sigma_u^2=\frac{1}{2\pi}\int |\Gamma\cdot\omega H^{\mathrm{ref}}(\omega,s)|^2 S_x(\omega)\, d\omega, \tag{21}$$

σ_v remains unchanged following our arguments above. σ_{uv} remains unaffected as it is determined primarily by the differential filter. This follows from our prior discussion in that the cochlear filter is essentially constant over the region where the differential filter has appreciable response. The correlation r changes due to its dependence on σ_u. If we use superscript r^{ref} to denote the original parameters to which we would like to compare the changed parameters with, we have



(a) Low SPL

(b) Moderate SPL

(c) High SPL

FIG. 5. Noise suppression by a tone in noise.

$$\sigma_u = \Gamma \sigma_u^{\mathrm{ref}}, \tag{22}$$

$$\sigma_v = \sigma_v^{\mathrm{ref}}, \tag{23}$$

$$\sigma_{uv} = \sigma_{uv}^{\mathrm{ref}}, \tag{24}$$

and

$$r = \frac{\sigma_{uv}}{\sigma_u \sigma_v} = \frac{r^{\mathrm{ref}}}{\Gamma}. \tag{25}$$

Finally consider the effect on β,

$$\beta = \frac{\sigma_{u|n}}{\sigma_u} = \left(1 + \left| \frac{\Gamma \cdot \sigma \omega_0 H^{\mathrm{ref}}(\omega_0, s)}{\Gamma \cdot \sigma_{u|n}^{\mathrm{ref}}} \right|^2 \right)^{-1/2} = \left(1 + \left| \frac{\sigma \omega_0 H^{\mathrm{ref}}(\omega_0, s)}{\sigma_{u|n}^{\mathrm{ref}}} \right|^2 \right)^{-1/2}. \tag{26}$$

It is interesting to note that β is *independent of* the BM nonlinearities. This is understandable because β is simply a measure of the *relative* energy passing through the cochlear filter compared to the energy passing through the filter in the presence of noise. As such, the change in filter characteristics does affect both these energies, following our earlier discussion.

Now if we bring all these together we finally arrive at the following expression for $E[y_4(t,s)]$

$$E[y_4(t,s)] = \frac{1}{2\pi} \frac{\sigma_{v|n}^{\mathrm{ref}}}{\Gamma \cdot \sigma_{u|n}^{\mathrm{ref}}} \cdot \beta^{\mathrm{ref}} \cdot \sqrt{1 - \frac{r_n^{\mathrm{ref}^2} \beta^{\mathrm{ref}^2}}{\Gamma^2}}. \tag{27}$$

Now in this expression if we associate the Γ with β^{ref}, then even though β by itself does not change, the effect of the nonlinearities is such that it scales β in the resulting auditory spectrum.

Now for small β, $E[y_4(t,s)]$ is essentially proportional to β and from (27) the suppression is enhanced by a factor Γ, representing the increase in the filter bandwidth. It is therefore readily apparent that broad filters enhance peaks better than relatively narrow filters. Intuitively, in terms of phase locking, at low SPLs, where the filter bandwidths are mostly narrow, phase locking is significantly less widespread. This is appropriate when frequency content is more widely distributed or is present at frequencies beyond phase locking ($\geqslant$4 kHz) such as in voiceless fricatives. At higher SPLs, under broader BW filtering the opposite will hold and sounds with more localized frequency content such as vowels will benefit from the more widespread phase locking.

We can now quantify the relationship between BM nonlinearities with increasing stimulus levels. Consider the auditory spectrum of an input signal that has been presented at two sound-pressure levels: A spectral peak resolved by the differential filter will be enhanced, compared to one ob-

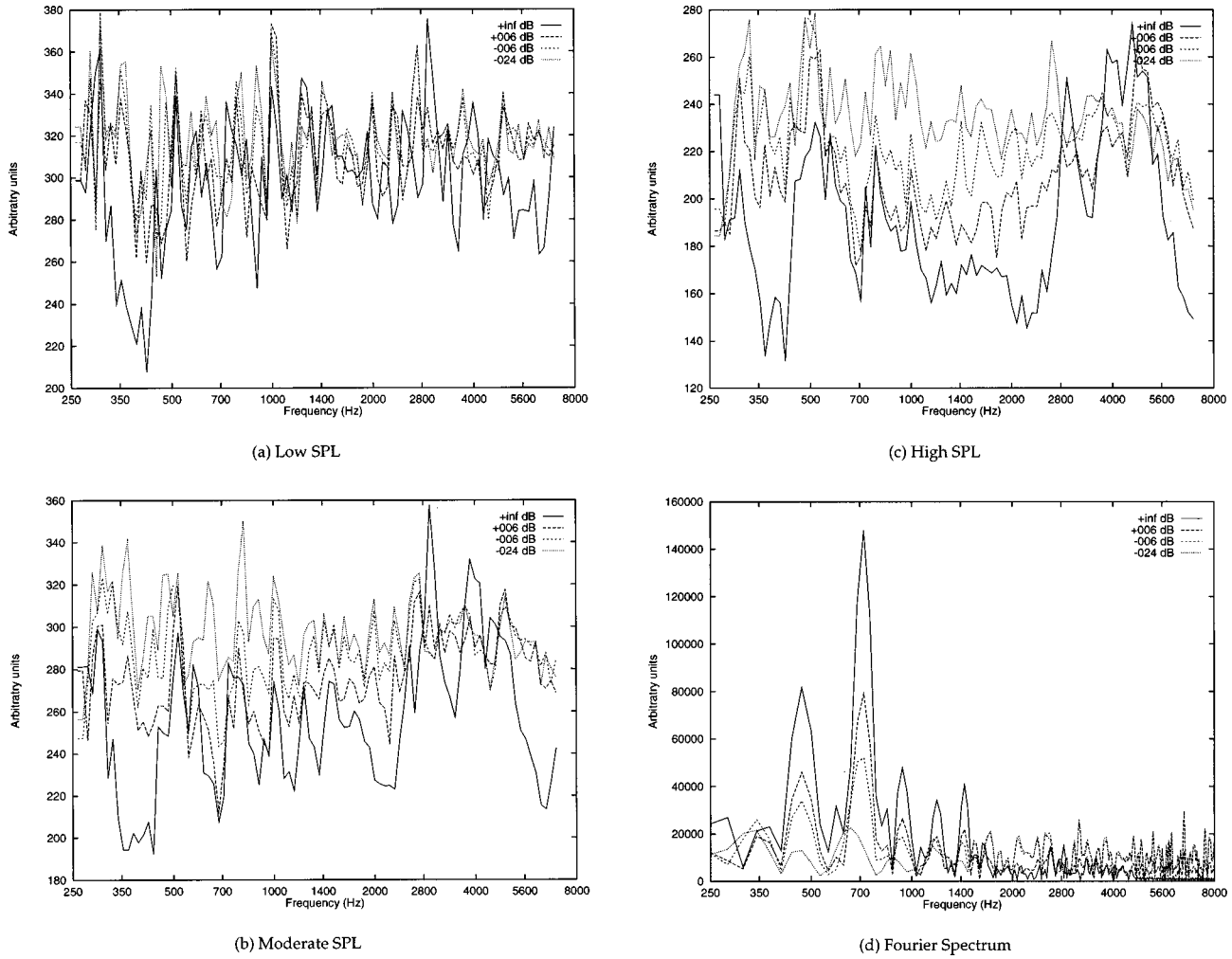


FIG. 6. Noise robustness of the auditory spectrum for vowel-like stimulus /aa/.

served at the lower sound-pressure level, if the increase of energy of the output of the wider cochlear filter does not result in inclusion of neighboring peaks. However, if the wider cochlear filter does result in inclusion of neighboring peaks the resultant spectrum will slightly attenuate the resolved spectral peak. Figure 2 presents a pictorial version of this argument, for an input signal consisting of four peaks, numbered 1–4. Peaks 1, 2, and 3 are relatively far apart compared to peaks 3 and 4. The top path shows the auditory spectrum if a low-SPL version of this input is presented to the model. There is a suppression region immediately following each of the peaks. The suppression regions are, however, too narrow to influence the neighboring peaks, even peaks 3 and 4. The peaks are thus, apart from the enhancement, faithfully reproduced in the resultant auditory spectrum. The lower path, describing the auditory spectrum at high-SPL shows a considerably more interesting picture. The suppression regions are noticeably larger. The first three peaks, 1–3, are unaffected as the suppression region, though large, is insufficient to influence them. However, peak 4 falls within the suppression region of peak 3, thus attenuating the peak, as described above. Such behavior is understandable considering that our perception of most voiced sounds (spectrally peaky) improves with increasing intensity levels even in the presence of noise. In addition, one can argue that with increasing stimulus levels, the considerable spread of spectral information of the spectrum results in a consistently ''whiter'' spectrum thus possibly accounting for the loss in perception of unvoiced sounds such as most consonants at high intensities. This is supported by physiological experiments such as Geisler and Gamble (1989), wherein for certain consonant-vowel syllables at various levels of noise, certain parts of the responses to consonant segments were more affected by noise than responses to vowel segments.

An interesting corollary follows from our discussion above, that of the relationship between the typical sound-pressure levels of various speech classes, Greenberg (1995) has recently suggested that the amplitude of spoken speech classes is governed by the robustness of the resulting auditory representation. Our argument above essentially indicates that consonants would suffer at higher intensities whereas vowel perception improves, thus tending to confirm his suggestion.

IV. EXPERIMENTAL RESULTS

To support the above analysis several experiments were conducted by passing various stimuli, including a pure tone in noise, a natural vowel, and several natural consonants, through the model. The input stimuli were preprocessed by

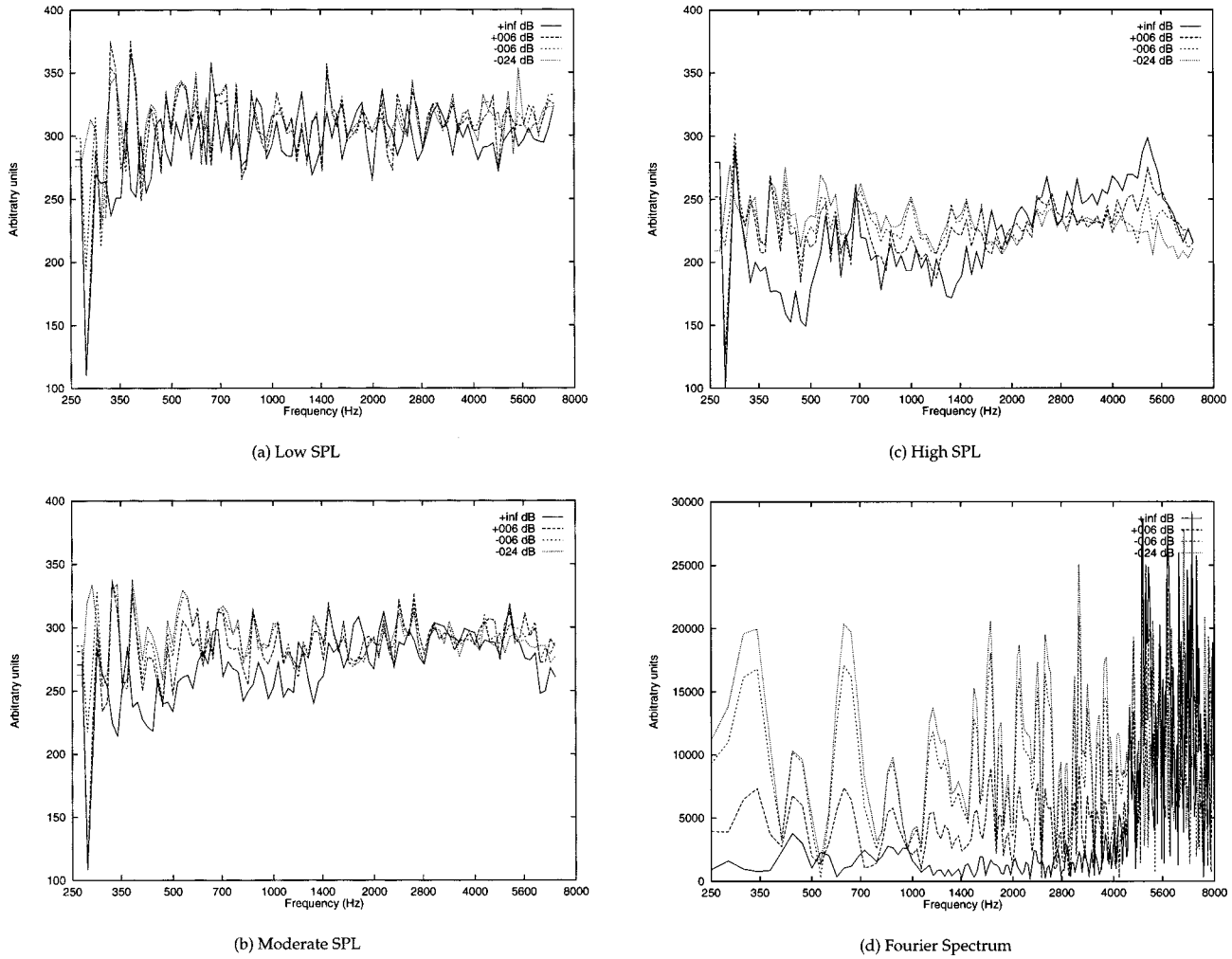


FIG. 7. Noise robustness of the auditory spectrum for fricative stimulus /s/.

adding Gaussian white noise to yield segments at varying SNRs covering the range $+\infty$ to -24 dB. The output of the model for the three cases of low, moderate, and high sound-pressure levels (SPLs) was obtained by passing the stimulus through the appropriate filter bank. The resulting auditory spectrum is then plotted with the cochlear filter CFs on the x axis and their output on the y axis for various stimuli. The auditory spectrum at the output of the model then provides a measure of nerve synchrony across the tonotopic array in response to such signals.

A. Implementation

Experiments were carried out on an implementation of the model described in Sec. II. The three filter banks used in our implementation are generated by the dilation of three seed filters. The three seed filters are assumed to represent the state of basilar membrane filtering at low, moderate, and high SPLs. Figure 3 shows the three seed filters used in this study, as well as their differential filters [using (15)]. As mentioned earlier in Sec. II A, the filter for moderate sound levels is virtually identical to the one described in Wang and Shamma (1994) and was designed to produce the effects of upward masking and two-tone suppression. The filters for low and high SPLs were designed to follow the reasoning in Sec. III, i.e., to have no change in the high-frequency edge of cochlear filters and a marked change in the low-frequency edge of the cochlear filters. While this is understandably a suboptimal representation of BM filtering, the simulation results do substantiate the analysis presented in the previous section.

Ninety-six filters, corresponding to dilation of the seed filters 2 oct down and 2.8 oct up, were generated to cover the frequency range extending from 250 to 7000 Hz.

A check was made on the viability of our analysis for these filters. Taking the seed filter for moderate sound levels as our reference filter, we found that for white noise, i.e., $S_x(\omega)=1$, σ_v of the filters at low and high SPLs was within $\pm 5\%$ of $\sigma_v^{\text{ref}} \cdot \sigma_u$ of the filters, however, was slightly less accurate, showing 7% deviation for the filter at low SPLs and 19% deviation for the filter at high SPL.

The remainder of the model is a straightforward implementation of a time difference, step, spatial difference, half-wave rectification, and integration in cascade following closely the details presented in Sec. I. Figure 4 provides a pictorial depiction of the process. The signal is first passed through a cochlear filter bank, the output of which is shown in the second panel of Fig. 4. The filter bank is prespecified, i.e., there is no automatic selection of which filter bank is to be used. A time difference of the filter-bank output is then

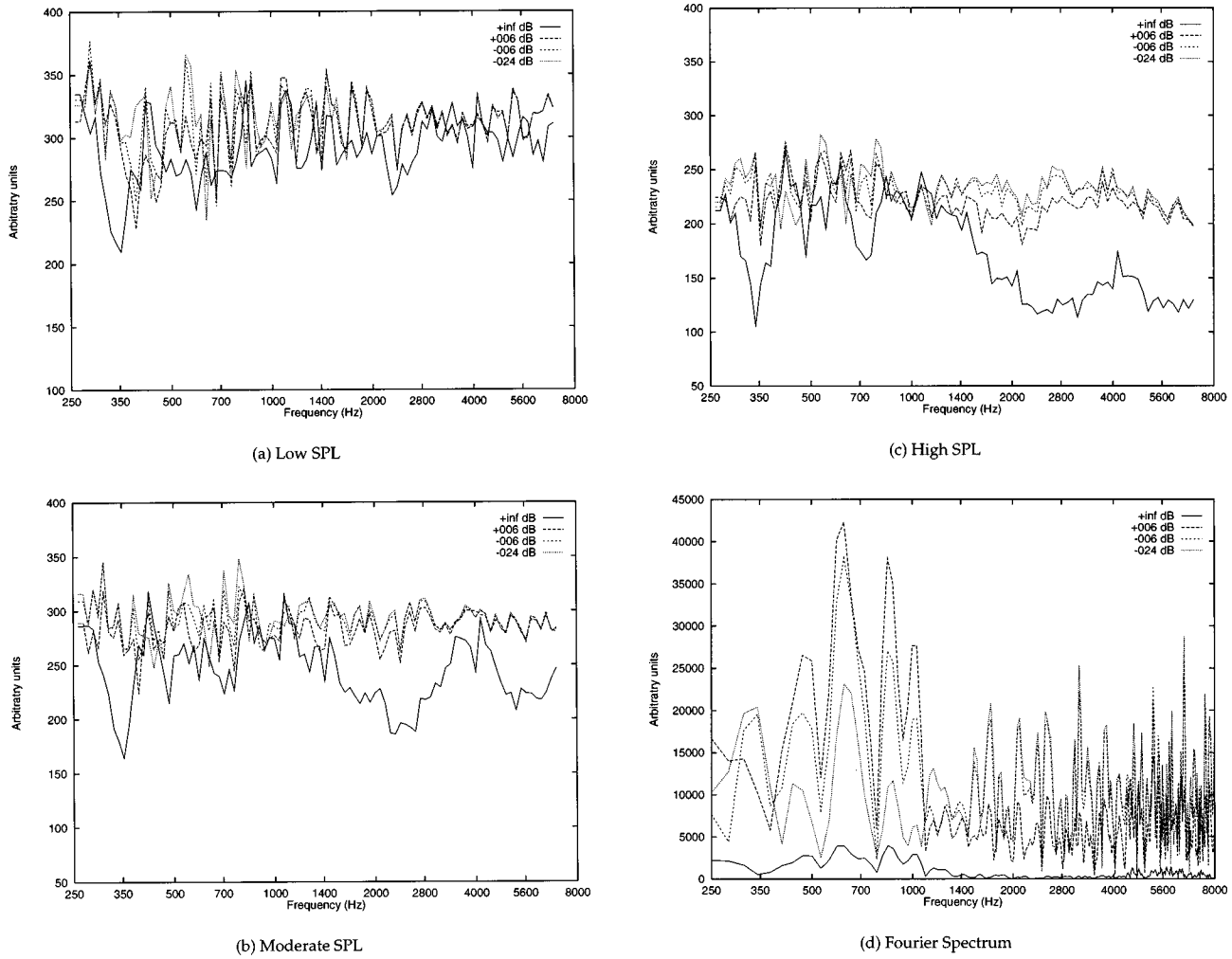


FIG. 8. Noise robustness of the auditory spectrum for stop stimulus /k/.

passed through a step. The output of this stage is shown in the third panel of Fig. 4. The spatial derivative is taken by subtracting the output of adjacent channels. Half-wave rectification and temporal integration of the output of each channel then yields the final auditory spectrum depicted in the bottom panel of Fig. 4. This corresponds to the 0-dB tone in noise auditory spectrum shown in Fig. 5(c).

B. Tone in noise

A tone of frequency 1 kHz was presented in a background of Gaussian white noise. The SNR of the combined signal was then adjusted in steps from $-\infty$ to 6 dB. Figure 5 shows the results of these experiments arranged according to their increasing filter form factor, Γ. The suppression regions are clearly visible as depressions following each of the tonal peaks. As expected from the analysis the suppression increased going from the filter with low Γ, the narrow bandwidth filter bank [Fig. 3(a)], to the filter with high Γ, the broad bandwidth filter bank [Fig. 3(c)]. The suppression is more evident on the high-frequency side of the tone due to the asymmetrical shape of the filters chosen in this study.

C. Vowel in noise

A segment of natural speech containing the phoneme segment /aa/ was extracted from the TIMIT corpus. This segment was then processed through the model at varying SNRs covering the range $+\infty$ to -24 dB. Figure 6 shows the auditory spectrum at a time corresponding to the middle of /aa/. Also shown is the Fourier spectrum of the segment for the same SNRs. Figure 6(c) shows the response at high SPLs, i.e., with broad BW filters, and shows its robustness with increasing noise levels. This corresponds well with physiological studies on temporal representations of such signals (Geisler and Gamble, 1989; Rhode and Greenberg, 1994).

For any particular set of filters the auditory spectrum is a fairly robust representation of this sound. Robustness being described as the auditory spectrum's ability to retain key synchrony features from $+\infty$ to -24 dB. Looking across the filters with increasing Γ, the robustness becomes more apparent. For the broadest filters the auditory spectrum retains its characteristics over all the SNRs considered, whereas for the narrowest filters the auditory spectrum shows much more variation in its characteristics.

This is consistent with reported results for synthesized

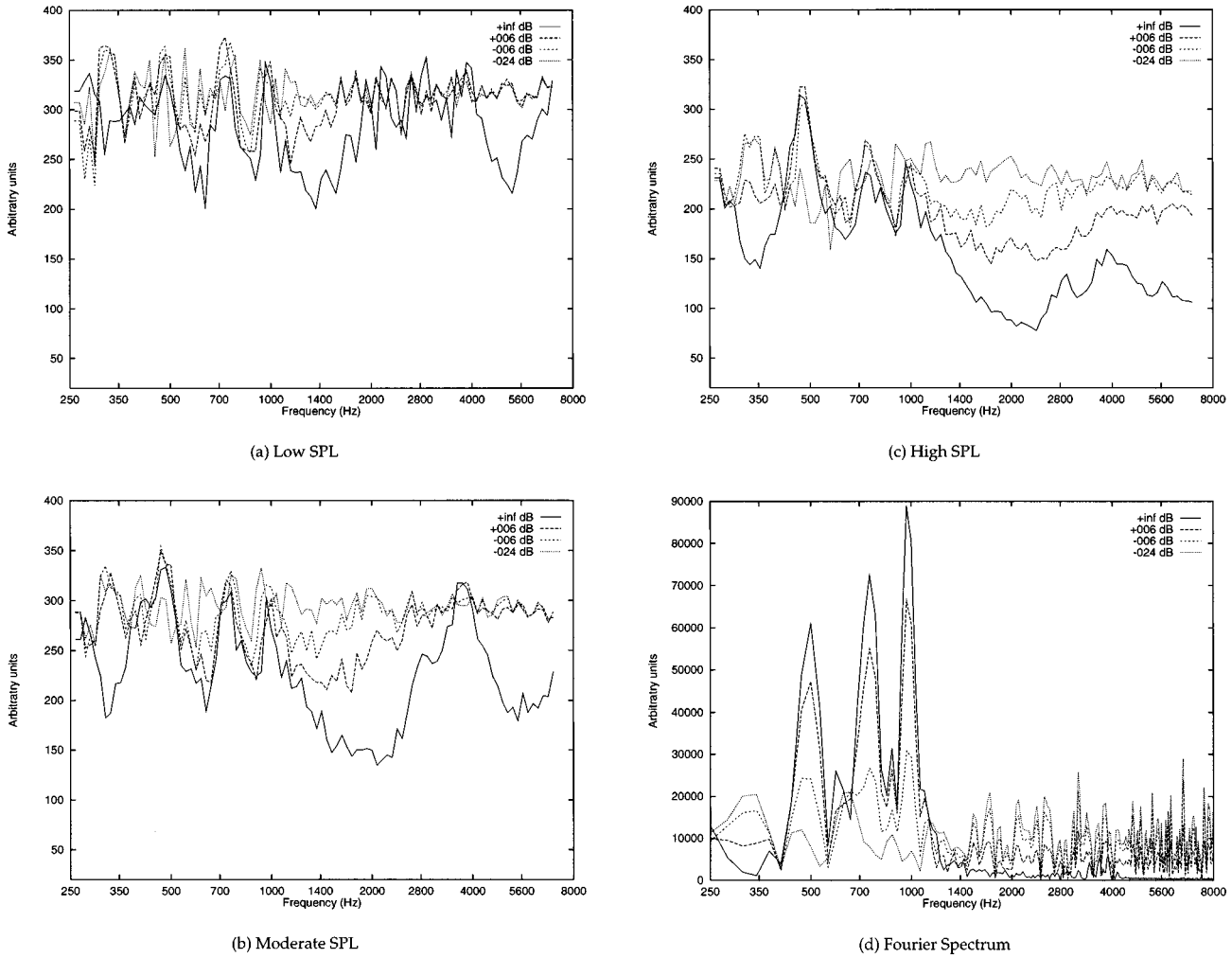


FIG. 9. Noise robustness of the auditory spectrum for liquid stimulus /l/.

vowels and vowel-like sounds at different SPLs in the presence of noise (Sachs *et al.*, 1983; Delgutte and Kiang, 1984a). Both these studies indicate that for increasing vowel SPL, the effect of noise on the temporal representation is small except at very low SNR levels. However, due to the differences in both the presented sounds and background noise a direct comparison cannot be made. The results do show strong encoding of the key frequency components of the acoustic signal. This is more so in the case of the response at moderate and high SPL levels [Fig. 6(b) and (c)].

D. Consonants in noise

Similar to the procedure followed for the vowel we extracted segments of natural speech corresponding to consonant classes: fricatives, stops, liquids, and nasals. The phonemes /s/, /k/, /l/, and /n/, respectively, were selected as exemplars of these segments from an utterance in the TIMIT corpus. The same utterance was used in the prior section. This was done to minimize any effects due to speaker variability. The segment was passed through the model after appropriate addition of noise to vary the SNRs through the range $+\infty$ to -24 dB. The auditory spectrum at the middle of these consonants is plotted in Figs. 7–10.

Fricatives, depending on whether they are voiced or unvoiced, have aperiodic, low intensity waveforms with most of the energy at high frequencies and some periodicity with formant characterization for voiced fricatives. Stops, on the other hand, are transient phonemes with dynamic spectra and cannot be easily described in terms of steady-state spectra. For the two exemplars chosen, /s/ and /k/, /s/ shows a lack of energy at low frequencies and possesses most of its energy at higher frequencies whereas /k/ due to its transient nature shows a varying spectral profile with an initial low-energy silence part followed by a noise burst. In view of this it is expected that broader filters will degrade the auditory spectrum by resulting in flat and smooth profiles with no distinguishing features for the fricative /s/ and for stop /k/. The output of the model for the three filter banks for fricative /s/ and stop /k/ are shown in Figs. 7 and 8, respectively. For fricative /s/, the auditory spectrum at the output of the low SPL model [Fig. 7(a)], shows much of the same characteristics with increasing noise level. However, the high SPL model [Fig. 7(c)] shows a progressively flatter profile with increasing noise level. This is also the case for stop /k/, where once again the high SPL model [Fig. 8(c)] shows a progressively flatter profile with increasing noise level com-

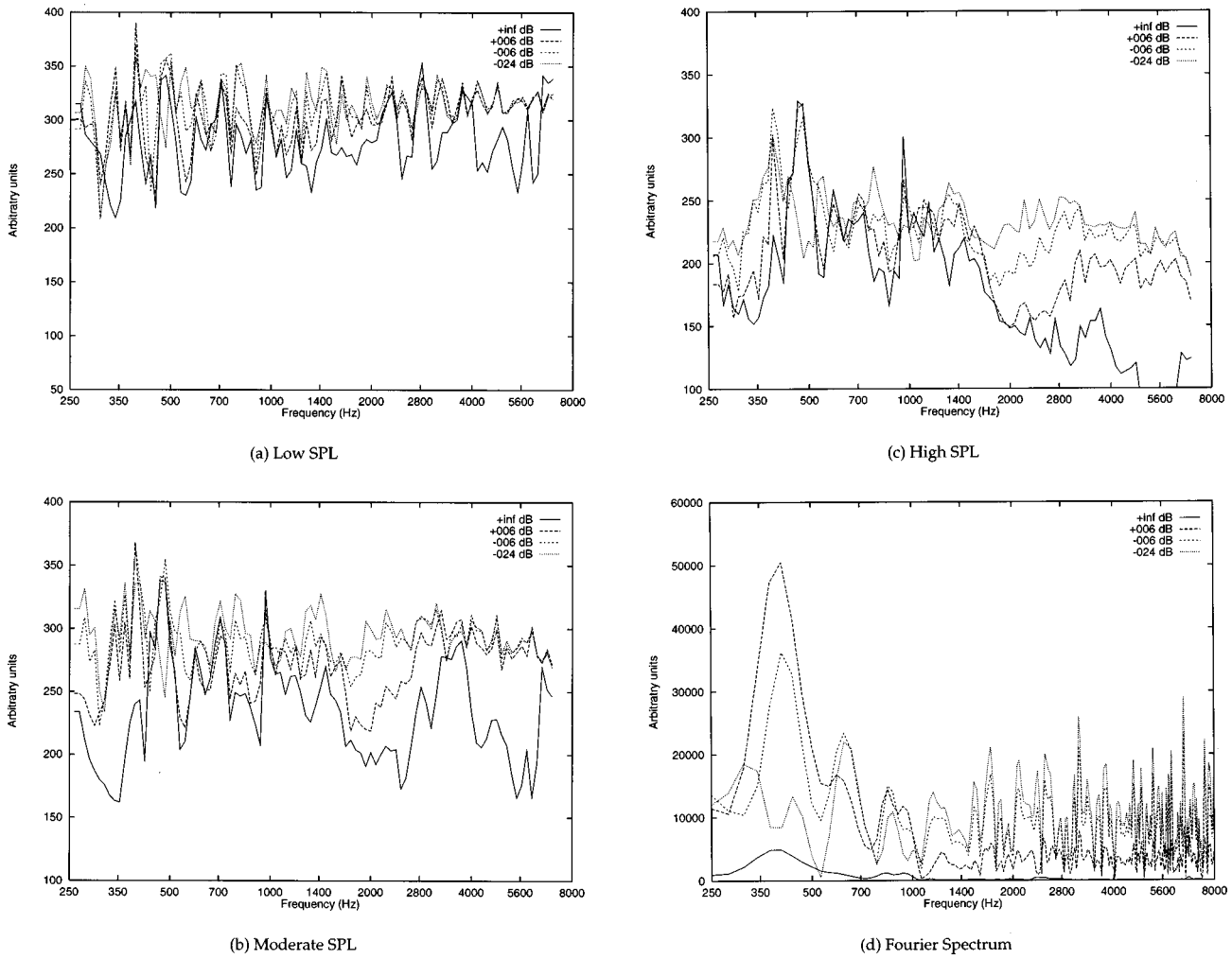


FIG. 10. Noise robustness of the auditory spectrum for nasal stimulus /n/.

pared to a more stable profile from the low SPL model [Fig. 8(a)].

Liquids are very similar to vowels in that they possess intense, periodic waveforms with most of the energy concentrated in the lower formants. However, they are weaker than vowels. Based on our discussion in Sec. III, signals with strong synchrony content will be enhanced with increasing Γ, i.e., by broader BM filters compared to narrower BM filters. Figure 9 shows the auditory spectrum output of the model at various SPLs and SNRs with the corresponding segmental Fourier spectrum. The formant information is clearly visible in the output of the broadest filter bank [Fig. 9(c)] at all but the highest noise level (-24 dB) while the narrowest filters [Fig. 9(a)] show a progressively less informative picture with increasing noise level. Again, the auditory spectrum is fairly robust at moderate and high SPLs due to the strong synchrony in the signal.

Nasals resemble vowels but are significantly weaker due to attenuation within the nasal cavity. The simulation results are shown in Fig. 10. As expected the auditory spectra show much of the same characteristics as vowels with strong formant encoding. The prominent first formant is visible at high SNRs at all SPLs, with the performance under noise degrading for filters with smaller Γ, i.e., narrower filters.

V. CONCLUSIONS

We have presented an extension to an earlier analysis of noise suppression within the early auditory system by Wang and Shamma. This extension accounts for the variation of basilar membrane filtering due to the nonlinearities of the BM with increasing SPLs. The validity of the approach was demonstrated by several experiments on real and synthetic sound stimuli.

Based on the analysis in Secs. II and III and the results presented in the previous section, it seems that one can conclude that sounds with strong periodicity such as vowels and vocalic consonants are enhanced with broadly tuned filters, the opposite holding true for aperiodic sounds (fricatives/stops). The analysis interestingly affords an explanation for why the sound levels of vowels and consonants tend to be separable by energy. It could be theorized that this difference in SPLs, and hence energy, could be a result of evolutionary constraints on biological auditory systems. This idea has been recently put forth by Greenberg (Greenberg, 1995). It would be interesting to see whether psychoacoustical studies confirm the analysis presented here. This effect may be an important cue in speech perception. The differences in encoding of sounds at different SPLs may indicate that there

are many cues available with the most informative being selected or emphasized over other cues under any given set of circumstances.

This work presents an analytical approach to the analysis of basilar membrane nonlinearities. The resulting analysis shows that vowel perception in noise improves as the BM filtering shows progressively broader tuning with increasing sound levels whereas consonant perception degrades with increasing sound levels. This conflict suggests that improved robustness can be achieved by the use of dual filter banks in applications such as speech recognition, an approach that is being actively pursued within our group.

ACKNOWLEDGMENTS

We deeply appreciate the help extended to us by Li Deng, Steven Greenberg, and Kuansan Wang by patiently answering our questions. We would also like to thank Steven Greenberg and Li Deng for their invaluable comments on earlier drafts of this paper.

[1]We have attempted to follow closely the notation of Wang and Shamma to clearly demonstrate the nature of our extension to their work.

[2]We concern ourselves only with $E[y_4(t,s)]$ as with the common assumption of short-time stationarity of speech, Wang and Shamma's and our choice of $\Pi(\cdot)$, $E[y_5(t,s)]$ and $E[y_4(t,s)]$ are related by a simple scaling constant.

Carlson, R., Fant, G., and Granstörm, B. (**1975**). ''Two-formant models, pitch and vowel perception,'' in *Auditory Analysis and Perception of Speech*, edited by G. Fant and M. Tatham (Academic, London), pp. 55–82.

Carney, L. H. (**1993**). ''A model for the responses of low-frequency auditory-nerve fibers in cat,'' J. Acoust. Soc. Am. **93**, 401–417.

Delgutte, B. (**1984**). ''Speech coding in the auditory nerve: II. Processing schemes for vowels-like sounds,'' J. Acoust. Soc. Am. **75**, 879–886.

Delgutte, B., and Kiang, N. Y. S. (**1984a**). ''Speech coding in the auditory nerve: I. Vowel-like sounds,'' J. Acoust. Soc. Am. **75**, 866–878.

Delgutte, B., and Kiang, N. Y. S. (**1984b**). ''Speech coding in the auditory nerve: III. Voiceless fricative consonants,'' J. Acoust. Soc. Am. **75**, 887–896.

Delgutte, B., and Kiang, N. Y. S. (**1984c**). ''Speech coding in the auditory nerve: IV. Sounds with consonant-like dynamic characteristics,'' J. Acoust. Soc. Am. **75**, 897–907.

Delgutte, B., and Kiang, N. Y. S. (**1984d**). ''Speech coding in the auditory nerve: V. Vowels in background noise,'' J. Acoust. Soc. Am. **75**, 908–918.

Deng, L., and Geisler, C. D. (**1987**). ''A composite auditory model for processing speech sounds,'' J. Acoust. Soc. Am. **82**, 2001–2012.

Deng, L., Geisler, C. D., and Greenberg, S. (**1988**). ''A composite model of the auditory periphery for the processing of speech,'' J. Phonetics **16**, 93–108.

Deng, L., and Kheirallah, I. (**1993**). ''Dynamic formant tracking of noisy speech using temporal analysis on outputs from a nonlinear cochlear model,'' IEEE Trans. Biomed. Eng. **40**, 456–467.

Evans, E. F. (**1982**). ''Functional anatomy of the auditory system,'' *The Senses*, edited by H. B. Barlow and J. D. Mollon (Cambridge U. P., Cambridge, England), Chap. 14, pp. 251–306.

Geisler, C. D., and Gamble, T. (**1989**). ''Responses of ''high-spontaneous'' auditory-nerve fibers to consonant-vowel syllables in noise,'' J. Acoust. Soc. Am. **85**, 1639–1652.

Ghitza, O. (**1988**). ''Temporal non-place information in the auditory-nerve firing patterns as a front-end for speech recognition in a noisy environment,'' J. Phonetics **16**, 109–123.

Ghitza, O. (**1992**). ''Auditory representations as a basis for speech processing,'' in *Advances in Speech Signal Processing*, edited by S. Furui and M. M. Sondhi (Marcel Dekker, New York), pp. 453–485.

Ghitza, O. (**1993**). ''Adequacy of auditory models to predict human internal representation of speech sounds,'' J. Acoust. Soc. Am. **93**, 2160–2171.

Ghitza, O. (**1994**). ''Auditory models and human performance in tasks related to speech coding and speech recognition,'' IEEE Trans. Speech Audio Process. **2**, 115–132.

Giguère, C., and Woodland, P. C. (**1994a**). ''A computational model of the auditory periphery for speech and hearing research. I. Ascending path,'' J. Acoust. Soc. Am. **95**, 331–342.

Giguère, C., and Woodland, P. C. (**1994b**). ''A computational model of the auditory periphery for speech and hearing research. II. Descending paths,'' J. Acoust. Soc. Am. **95**, 343–349.

Greenberg, S. (**1995**). ''Auditory processing of speech,'' in *Principles of Experimental Phonetics*, edited by N. J. Lass (Mosby, New York), Chap. 10, pp. 362–407.

Holton, T., Love, S. D., and Gill, S. P. (**1992**). ''Formant and pitch-pulse detection using models of auditory signal processing,'' in *Proceeding of the International Conference on Spoken Language Processing*, Banff (University of Alberta, Edmonton, Alberta, Canada), pp. 81–84.

Huggins, W. H. (**1952**). ''A phase principle for complex-frequency analysis and its implications in auditory theory,'' J. Acoust. Soc. Am. **24**, 582–589.

Jenison, R. L. (**1991**). ''A dynamic model of the auditory periphery based on the response of single auditory-nerve fibers,'' Ph.D. thesis, University of Wisconsin—Madison.

Jenison, R. L., Greenberg, S., Kluender, K. R., and Rhode, W. S. (**1991**). ''A composite model of the auditory periphery for the processing of speech based on the filter response functions of single auditory-nerve fibers,'' J. Acoust. Soc. Am. **90**, 773–786.

Kim, D. O., Molnar, C. E., and Mathews, J. W. (**1980**). ''Cochlear mechanics: Nonlinear behavior in two-tone responses as reflected in cochlear-nerve fiber response and in ear-canal sound pressure,'' J. Acoust. Soc. Am. **67**, 1704–1721.

Meddis, R., and Hewitt, M. J. (**1991**). ''Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification,'' J. Acoust. Soc. Am. **89**, 2866–2882.

Payton, K. L. (**1988**). ''Vowel processing by a model of the auditory periphery: A comparison to eighth-nerve responses,'' J. Acoust. Soc. Am. **83**, 145–162.

Pickles, J. O. (**1988**). *An Introduction to the Physiology of Hearing* (Academic, London), 2nd ed.

Rhode, W. S., and Greenberg, S. (**1994**). ''Lateral suppression and inhibition in the cochlear nucleus of the cat,'' J. Neurophysiol. **71**, 493–519.

Robles, L., Ruggero, M. A., and Rich, N. C. (**1986**). ''Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input-output functions, tuning curves, and response phases,'' J. Acoust. Soc. Am. **80**, 1364–1374.

Ruggero, M. A. (**1992**). ''Responses to sound of the basilar membrane of the mammalian cochlea,'' Curr. Opin. Neurobiol. **2**, 449–456.

Sachs, M. B., Voigt, H. F., and Young, E. D. (**1983**). ''Auditory nerve representation of vowels in background noise,'' J. Neurophysiol. **50**, 27–45.

Secker-Walker, H. E., and Searle, C. L. (**1990**). ''Time-domain analysis of auditory-nerve-fiber firing rates,'' J. Acoust. Soc. Am. **88**, 1427–1436.

Seneff, S. (**1986**). ''A computational model for the peripheral auditory system: Application to speech recognition research,'' in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing, Tokyo* (IEEE, New York), pp. 1983–1986.

Shamma, S. (**1988**). ''Spatial and temporal processing in central auditory networks,'' in *Methods in Neuronal Modeling: From Synapses to Networks*, edited by C. Koch and I. Segev (MIT, Cambridge, MA), Chap. 8, pp. 247–289.

Shamma, S. A. (**1985a**). ''Speech processing in the auditory system I: The representation of speech sounds in the responses of the auditory nerve,'' J. Acoust. Soc. Am. **78**, 1612–1621.

Shamma, S. A. (**1985b**). ''Speech processing in the auditory system II: Lateral inhibition and the central processing of speech evoked activity in the auditory nerve,'' J. Acoust. Soc. Am. **78**, 1622–1632.

Shamma, S. A., Chadwick, R. S., Wilbur, W. J., Morrish, K. A., and Rinzel, J. (**1986**). ''A biophysical model of the cochlear processing: Intensity dependence of pure tone responses,'' J. Acoust. Soc. Am. **80**, 133–145.

Wang, K., and Shamma, S. (**1994**). ''Self-normalization and noise-robustness in early auditory representations,'' IEEE Trans. Speech Audio Process. **2**, 421–435.

Wang, K., Shamma, S. A., and Byrne, W. J. (**1993**). ''Noise robustness in the auditory representation of speech signals,'' in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing, Paris* (IEEE, New York), Vol. 2, pp. 335–338.
Yang, X., Wang, K., and Shamma, S. A. (**1992**). ''Auditory representations of acoustic signals,'' IEEE Trans. Inf. Theory **38**, 824–839.
Young, E. D., and Sachs, M. B. (**1979**). ''Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of the auditory nerve fibers,'' J. Acoust. Soc. Am. **66**, 1381–1403.

Jet-wave amplification in organ pipes[a)]

Shigeru Yoshikawa[b)]
3-12-2-6304 Nagase, Yokosuka, 239, Japan

(Received 25 August 1997; revised 29 January 1998; accepted 5 February 1998)

An envelope-based method to estimate both the jet-wave amplification factor and the mouth-field strength in organ pipes is developed by using flow visualization of a smoked jet with a high-speed digital video camera. A theoretical envelope of wave growth, which is approximated using a negative displacement model of the jet drive, is compared with an experimental envelope derived from superposing many instantaneous shapes of jet deflection in the steady-state oscillation. The estimation results are presented in dimensional terms with respect to two particular models, where the flue-to-edge distances are, respectively, 15.8 and 10.2 mm, with a common flue thickness of 2.2 mm. In our experiment the jet velocity ranges from 7 to 33 m/s, the Reynolds number from 1000 to 5000, and the sounding frequency from 130 to 580 Hz. The amplification factor of organ pipe jets, estimated to lie in the 0.18–0.26-mm^{-1} range, tends to decrease and saturate with increasing blowing velocity in each oscillation mode; the mouth-field strength defined as the acoustic displacement amplitude, roughly estimated to be 0.5–1.5 mm, tends to increase and saturate with increasing blowing velocity. A hot-wire anemometer is then used to measure the mouth-field strength, whose value shows a good agreement with the estimated one. This result confirms the validity of our envelope-based method. A dimensionless representation of the experimental data is used to compare wave characteristics between an organ pipe jet and an acoustically perturbed free jet. The applicability of the spatial and temporal theories of jet instability is discussed to analyze them. If we can assume a Poiseuille flow at the flue exit and a subsequent Bickley jet, the spatial theory seems to be relevant to our organ pipe jets. However, for lack of a reliable experimental measurement of the jet half-thickness we cannot draw a definite conclusion about the wave characteristics of organ pipe jets.
[S0001-4966(98)04305-7]

PACS numbers: 43.75.Np [WJS]

INTRODUCTION

The physics of the sounding mechanism in organ flue pipes is still a challenging subject despite its long research history dating back to the pioneering works of Helmholtz[1] and Rayleigh.[2] The organ pipe shows the sounding characteristics as depicted in Fig. 1 of Coltman's paper.[3] The sounding frequency depends on the velocity (or the blowing pressure) of the jet issuing from the flue slit, and the stable frequency range for playing is restricted to a region of proper jet velocities. This relation between the jet velocity and the resulting tone frequency in the steady-state oscillation is well understood and established by many theoretical and experimental investigations.[4] However, how other jet characteristics, such as oscillation amplitude, are determined is not yet fully understood.

A few essential problems remain unsolved: (1) How is the sounding amplitude determined in the steady state? The jet drive model of Fletcher,[5] which is widely accepted, seems to be too simple to give an accurate value of the acoustic pressure amplitude in the pipe. Fabre, Verge *et al.*[6–10] took a more fluid-dynamical approach and took into account the effect of vortex shedding on the saturation. However, the calculation of the sounding amplitude in the steady states from the presaturated to the saturated regime is still incomplete. (2) How does the splitting between the pipe tone and the edge tone with increasing velocity cause the transition of the pipe tone to a steady state? (3) Which physical quantity stimulates and sustains the sounding process throughout? These problems have been attacked by several researchers, particularly by Coltman,[3,11] but we have no definite conclusions as yet.

The objective of this paper is to reconsider problem (1) above. The other problems will be approached in the future. The jet, which is the source of acoustic energy, exhibits a wavy motion over the mouth due to its inherent instability and the acoustic flow of the pipe resonance. However, we do not have sufficient knowledge of the amplification characteristics of the jet wave despite its crucial importance in determining the sounding amplitude. We do have experimental results obtained from an acoustically perturbed free jet,[12] where the pipe and the edge are excluded, but no experimental results on an organ pipe jet are available. So, more specifically speaking, the purposes of this paper are (a) the presentation of an experimental method and measurement data on jet-wave growth in organ pipes and (b) discussion of whether an acoustically perturbed free jet exhibits the same wave characteristics as an organ pipe jet or not.

[a)]A part of this paper was presented at the third joint meeting of the Acoustical Society of America and the Acoustical Society of Japan, 2–6 December 1996, Honolulu, Hawaii [J. Acoust. Soc. Am. **100**, 2843–2844(A) (1996)].

[b)]Present address: Department of Acoustical Design, Kyushu Institute of Design, 4-9-1 Shiobara, Minami-ku, Fukuoka, 815-0032, Japan; Electronic mail: shig@kyushu-id.ac.jp

The former (a) is given in dimensional terms using two organ pipe models (Secs. II–V), while the latter (b) is treated in a dimensionless way (Sec. VI). This is because the amplification factor (given in mm^{-1} throughout this paper) is not a universal quantity nor a simple constant for all organ pipe and free jets but depends on the frequency, the blowing pressure, and the jet dimensions, and also varies along the jet. If the amplification factor, nondimensionalized by the half-thickness of the jet, is expressed as a function of the Strouhal number, we may compare our results on organ pipe jets with previous results on free jets[12,13] as well as with the spatial[14,15] and the temporal analyses[16,17] of jet instability.

I. THEORETICAL APPROXIMATION

The jet-wave amplification factor in organ pipes will be evaluated using the negative displacement model proposed by Fletcher.[5] It is assumed that the jet deflection may be expressed as a superposition of a progressive wave due to jet instability and a spatially uniform oscillation displacement induced by the acoustical flow through the mouth:

$$\xi(x)=(v/\omega)\{-\cosh(\mu x)\sin[\omega(t-x/U_{\mathrm{ph}})]+\sin(\omega t)\}, \tag{1}$$

where $\xi(x)$ denotes the transverse displacement (z direction) of the jet at distance x from the flue exit. Also, t and ω are the time and angular frequency, respectively, and μ and U_{ph} denote the amplification factor and phase speed of the instability wave, respectively, while v is the transverse acoustic velocity of the mouth field. Spatial growth is here modeled as $\cosh(\mu x)$ instead of $\exp(\mu x)$ because the former gives a smoother curve at the flue, but neither expression is exactly correct. Moreover, both μ and U_{ph} are functions of x in a rigorous sense.

Equation (1) may be a good model for a free jet that passes through a standing wave field formed between two loudspeakers with opposite phases. Although the validity of Eq. (1) has not been confirmed experimentally for real organ pipe jets, Eq. (1) is a reasonably simple and practical representation of the jet deflection that describes our current knowledge. Therefore we postulate in this paper that Eq. (1) is valid.

Equation (1) yields the following expression for the envelope of positive and negative peak displacements:

$$\xi_{\mathrm{env}}(x)=\pm(v/\omega)[\cosh^2(\mu x) - 2\cosh(\mu x)\cos(\omega x/U_{\mathrm{ph}})+1]^{1/2}. \tag{2}$$

Approximating the second term in square brackets of Eq. (2) by $2\cosh(\mu x)$, we obtain the following expression:

$$\xi_{\mathrm{env}}(x)\approx\pm(v/\omega)[\cosh(\mu x)-1]. \tag{3}$$

While the spatial dependence of the envelope of Eq. (2) depends on the spatial wave number $k=\omega/U_{\mathrm{ph}}$ of the instability wave, that of Eq. (3) is simply governed by $\cosh(\mu x)-1$ apart from the v/ω term which is the absolute displacement corresponding to the strength of the mouth field. In other words, approximation (3) enables one to treat the amplification and phase delay of the jet wave separately.

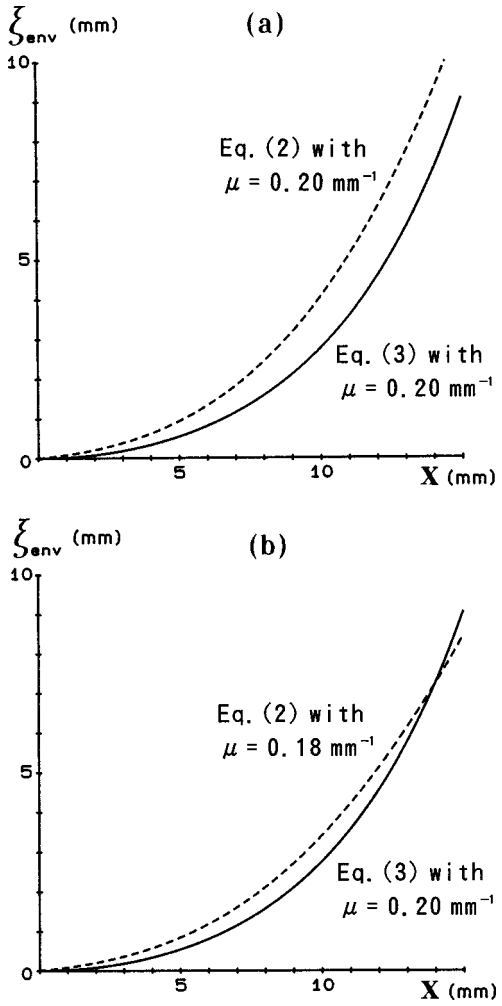


FIG. 1. Comparison of jet-wave envelopes. (a) Original Eq. (2) with $\mu=0.20\ \mathrm{mm}^{-1}$ (broken line) versus approximate Eq. (3) with $\mu=0.20\ \mathrm{mm}^{-1}$ (solid line); (b) Eq. (2) with $\mu=0.18\ \mathrm{mm}^{-1}$ (broken line) versus Eq. (3) with $\mu=0.20\ \mathrm{mm}^{-1}$ (solid line).

We will estimate the amplification factor μ based on this simpler Eq. (3) by comparing with the envelope shape obtained from our visualization experiment as discussed in Secs. III and IV. However, before proceeding to the experimental investigation, we need to evaluate the error arising from approximating Eq. (2) by (3). The absolute value of Eq. (2) is always larger than that of Eq. (3) for the same value of μ, while the $\cos(kx)$ dependence diminishes as μx becomes larger than one.

An example illustrating the approximation error is given in Fig. 1(a), where the mouth-field strength v/ω is assumed to be 1 mm for the comparison. Moreover, it is assumed that $\mu=0.20\ \mathrm{mm}^{-1}$ and $k=0.073+0.01x\ \mathrm{mm}^{-1}$ from our experimental results of an organ pipe model whose length and flue-to-edge distance are 120 cm and 15.8 mm, respectively. The value of $0.073\ \mathrm{mm}^{-1}$ in the above expression for k is calculated using $\omega=2\pi\cdot 133$ rad/s and $U_{\mathrm{ph}}=U_0=11.5$ m/s (measured at $x=0$). Also, the slowdown of U_{ph} is assumed so that k may be proportional to x (measured in mm) in order to make calculation easier. A proportionality constant of 0.01 is roughly estimated from our experiment. The k equals $0.231\ \mathrm{mm}^{-1}$ at the edge ($x=15.8$ mm), and implies U_{ph}

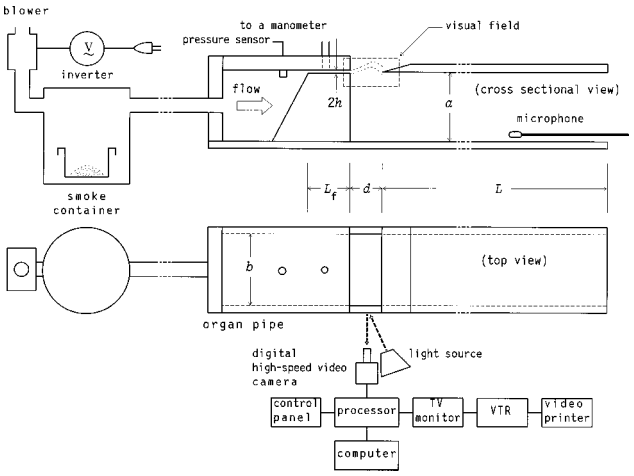


FIG. 2. Experimental setup for smoked-jet visualization and organ pipe geometry.

$=0.32\,U_0$ at the edge. The phase speed averaged over the mouth is reported to be about 0.4 to 0.5 times the jet speed at the flue in Refs. 4, 12, and 18, therefore the above expression for k is reasonable. The example of Fig. 1(a) indicates that approximation (3) gives an envelope magnitude which roughly corresponds to 0.75 times the magnitude given by Eq. (2) near $x=10$ mm (i.e., $\mu x=2.0$).

A closer match between the two is obtained if we apply a lower amplification factor ($\mu=0.18\ \mathrm{mm}^{-1}$) to Eq. (2) as indicated in Fig. 1(b). This means that the estimate of μ based on the theoretical envelope of Eq. (3) tends to be a little larger than that based on Eq. (2) for the same experimental envelope. However, the estimation error, about 10%, due to Eq. (3) is roughly equivalent to the resolution of the experimental data shown in Secs. III and IV. Thus we feel free to use the theoretical envelope of Eq. (3) to estimate the μ value of jet waves in organ pipes.

Of course, it is possible to directly estimate μ from Eq. (1). However, such an approach needs exact information about U_{ph} and a time (or phase) reference. Our envelope-based method using Eq. (3) is superior because of its independence of such information.

II. EXPERIMENTAL PROCEDURES

A. Flow visualization

Our experimental setup for jet-flow visualization is very simple and is shown in Fig. 2. The blowing pressure of a low-noise blower (Fuji Electric, type VFC206AN) is adjusted with an inverter (Fuji Electric, FVR-E9S), and steady pressures are obtained in the range 20–700 Pa. The smoke in a container, generated by many incense sticks and cigarettes, is driven by the blower. The smoked air is introduced into the foot of a model organ pipe through a rubber tube. The blowing pressure in the steady state is measured at this side of the flue slit using a manometer. The acoustic pressure is also measured inside the pipe at a position that is 7/12 of a pipe length from the farther pipe end.

The smoked jet that issues from the flue is observed with a high-speed digital video camera (Photron, HVC-11B or FASTCAM-hvc2) by shining the visual field with a 500-W spotlight. One camera (HVC-11B) can take digital pictures with a speed of 1297 pps (pictures per second) and stores a total of 2048 pictures in a 1.579-s duration. The other one (FASTCAM-hvc2) takes up to 1440 pps and can store 8192 pictures in a 5.69-s duration. The resulting picture consists of 128×64 pixels. Also, the latter one supports digital signal processing of the picture frames on a personal computer such as NEC PC-9801. The appropriate shutter speed is 1/5000 to 1/10 000 s for the former, while a speed of 1/110 000 s is automatically selected for the latter. The digitized data are recorded in a built-in memory, and replayed on a TV monitor with a speed of 2, 5, 10, 15, or 30 pps just after the recording. We can thus view slow-motion pictures of the jet waves at our convenience. Moreover, the interesting frames can be recorded on a VTR (8-mm video cassette, VHS video cassette, etc.) as analog data, and each frame can easily be printed out with a video printer.

B. Pipe geometry

Two experimental organ pipe models A and B are made of transparent acrylic resin (Plexiglas). A cross section of the models is shown in Fig. 2. Table I summarizes the basic model geometries. Both models have the same flue thickness (jet thickness) $2h=2.2$ mm and the same mouth width (jet width) $b=50.0$ mm. Model A has a longer cutup length (jet length) $d=15.8$ mm, while model B has a shorter one, $d=10.2$ mm. These relatively long cutup lengths are chosen to see wavy motion of the jet more easily (a very short cutup length of 4.02 mm was chosen by Fabre, Verge *et al.*[7–10] to investigate the vortex shedding in steady oscillation). Models A and B, whose physical lengths are 120.1 and 50.0 cm, have central frequencies of 134 and 286 Hz in their fundamental modes, respectively. The pipe cross section is almost square, however, its vertical side, a, is slightly different from the lateral side (mouth width) b. The offset h_0 between the flue centerplane and the pipe edge is nearly zero, but it is very difficult to accurately measure h_0 because of finite thickness of the edge tip. The flue channel length L_f is 100 mm ($L_f/2h=45.5$), and the edge angle is 16° for both models. Ears are not employed since they preclude the smoked-jet visualization.

III. WAVE CHARACTERISTICS OF THE JET

Wave characteristics of organ pipe jets are well known,[4] so only a brief explanation of them is given based on the examples of our visualization. Instantaneous shapes of jet deflections are visualized in Fig. 3 for the first mode of model A, while Fig. 4 shows the same for the second and third modes of model A. The first, second, and third modes, about 133, 268, and 406 Hz, are excited with blowing pressures of about 80, 240, and 640 Pa, respectively. They are visualized at intervals of 1/1297 s=0.771 ms. Also, note that ''0.00 ms'' in these figures is chosen arbitrarily during the period of steady-state oscillation.

Figure 5(a) indicates the centerplane of the planar jet shown in Fig. 3. Labels #1 to #7 in Fig. 5(a) correspond to those in Fig. 3. The centerplane is easily estimated as the midpoint of the thickness along the visualized jet. A thinner

TABLE I. Important dimensions of experimental organ pipe models.

Model		A	B
Cutup length	d (mm)	15.8	10.2
Flue thickness	$2h$ (mm)	2.2	2.2
Flue channel length	L_f (mm)	100.2	100.1
Mouth width	b (mm)	50.0	50.0
Offset	h_0 (mm)	~ 0	~ 0
Pipe length	L (cm)	120.1	50.0
Pipe side	a (mm)	49.0	49.2

smoke such as used in Fig. 3 made this estimation easier. Also, a smooth connection of these midpoints was applied to #4 to #7 in Fig. 5(a) because a small collapse was seen at some of the larger amplitudes.

Similarly, jet centerplanes at the second mode were estimated from the visualized results shown in Fig. 4 and superposed on Fig. 5(b). However, all of the centerplanes in Fig. 5(b) are not temporally consecutive. Only #3 and #4 in Fig. 5(b) correspond to those in Fig. 4. Others, #1′ to #3′, are consecutive, although they are taken from another sequence of the visualized result. Two remaining lines with no label are also taken from a different sequence.

A. Phase speed

The phase speed U_{ph} of the jet wave depicted in Fig. 5 is estimated as the speed with which the wave crest or the zero-crossing point progresses. For example, the wave crests of #1 to #5 in Fig. 5(a) are near the positions $x=4.3$, 7.8, 10.0, 12.0, and 13.5 mm, respectively. Therefore the jet wave progresses with the spatial steps of 3.5, 2.2, 2.0, and 1.5 mm per 0.771-ms time interval. The resulting phase speeds are thus 4.5, 2.9, 2.6, and 1.9 m/s within each time step. The initial jet velocity U_0 at the flue exit may be estimated from Bernoulli's law to be

$$P=(1/2)\rho U_0^2, \tag{4}$$

where P and ρ denote the blowing pressure and the air density, respectively. The blowing pressure was 80 Pa in the case of Fig. 5(a), so $U_0=11.5$ m/s. Consequently, we may consider that U_{ph} gradually decreases as $0.39U_0$, $0.25U_0$, $0.23U_0$, and $0.17U_0$ when the jet approaches the edge.

It should be noted that Eq. (4) is not valid for arbitrary configurations. If the flue channel is too long, there can be a significant friction loss which reduces U_0. In our case the channel length L_f ($=100$ mm) is quite long, although the pressure measurement was carried out at a distance of 10 mm from the flue exit. The pressure loss ΔP along a narrow channel can be roughly estimated as

$$\Delta P=[12\eta/(2h)^2]U_0\Delta l \tag{5}$$

assuming incompressible steady flow between stationary parallel plates.[19,20] In Eq. (5) η denotes the shear viscosity ($\eta=1.82\times10^{-5}$ Nsm^{-2} for the air at 20 °C) and Δl ($=10$ mm) is the distance between the position of pressure measurement and the flue exit. If we set $U_0=11.5$ m/s and $2h=2.2$ mm in Eq. (5), we have $\Delta P=5.2$ Pa $=0.065$ P for P=80 Pa. Also, the pressure measurement error is larger than 2 to 3 Pa depending on the steadiness of the air flow.

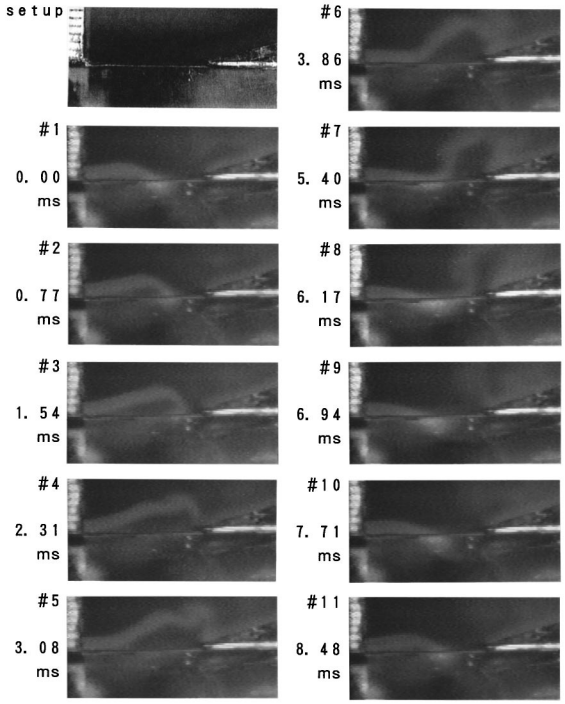


FIG. 3. Instantaneous shapes of the deflecting organ jet in steady-state oscillation (first mode of model A). The blowing pressure: 80 Pa; the sounding frequency: 133 Hz; the flue-to-edge distance: 15.8 mm; the flue thickness: 2.2 mm. As frame ''setup'' indicates, the flue exists at the left and the pipe edge at the right. A measure graduated in millimeters is attached on the outer wall above the flue. The time reference is arbitrary.

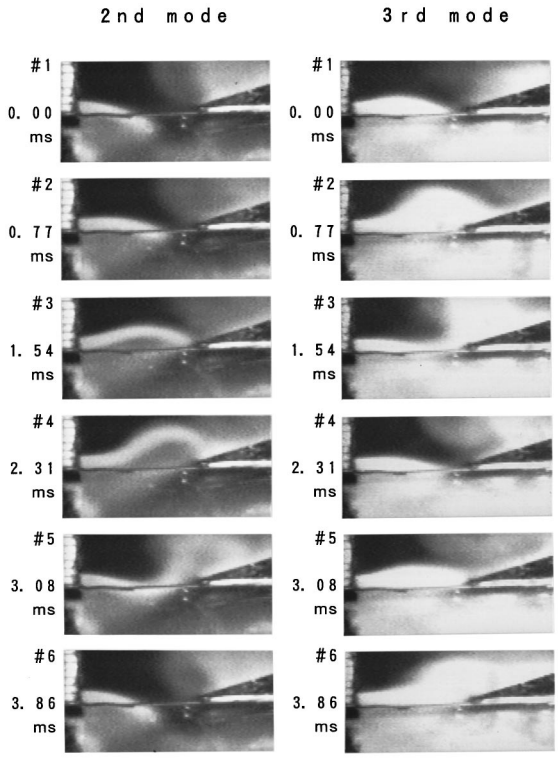


FIG. 4. Instantaneous shapes of the deflecting organ jet in steady-state oscillation (higher modes of model A). Left: second mode at the blowing pressure of 240 Pa and the frequency of 268 Hz; right: third mode at the blowing pressure of 640 Pa and the frequency of 406 Hz.

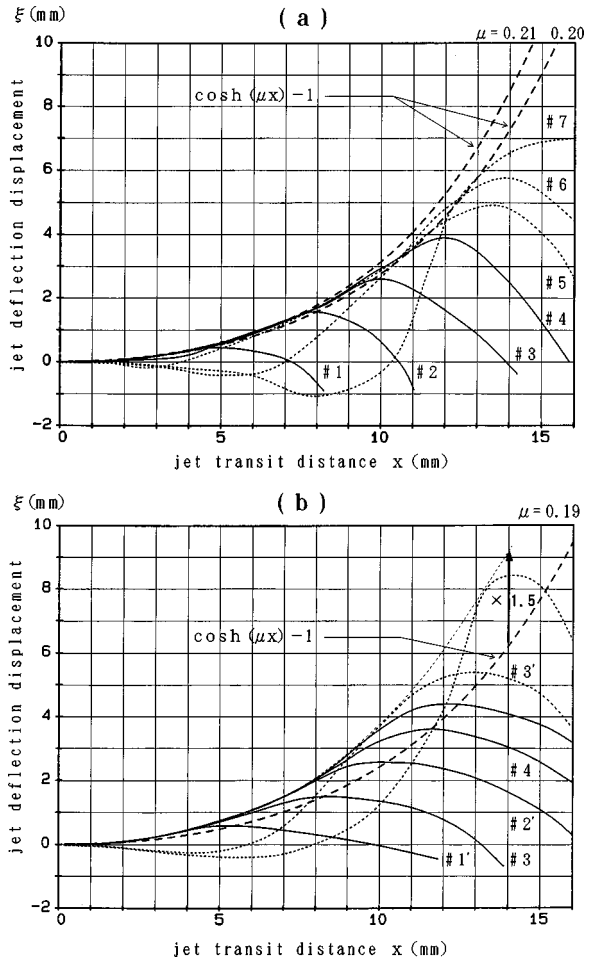


FIG. 5. Superposition of the deflecting jet centerplanes. (a) Superposition of frames #1 to #7 in Fig. 3; (b) superposition of various frames in the second mode (only labels #3 and #4 correspond to those in the left of Fig. 4). The thick broken line gives the theoretical envelope cosh $(\mu x)-1$; the thin broken line in (b) gives the experimental envelope. The unit of μ is mm^{-1}.

Therefore we can neglect the effect of viscous friction loss along the flue channel for the blowing pressure range in our experiment.

If we consider the zero-crossing point (that is, the crossing of the jet centerplane and the x axis of $\xi=0$), we have $x=7.2$, 10.6, 13.8, and 15.8 mm for #1 to #4 in Fig. 5(a), respectively. This yields decreases in U_{ph} of $0.38U_0$, $0.36U_0$, and $0.23U_0$. It is clear that U_{ph} considerably decreases as the jet approaches the edge. On the other hand, U_{ph} should be close to U_0 near the flue because there is no wavy crest. Therefore U_{ph} halfway between the edge and flue may be considered as an average phase speed, which is roughly estimated to be $0.40U_0-0.35U_0$ using the result of Fig. 5(a). Note that the expression for the spatial wave number k described in Sec. I B is based on the above experimental result and its simplification.

Similarly, we may estimate U_{ph} from Fig. 5(b). The wave crests of #3 and #4 give $U_{\mathrm{ph}}=0.24U_0$, where $U_0=20$ m/s. The crests of #1′ to #3′ give $U_{\mathrm{ph}}=0.31U_0$ and $0.21U_0$. These values based on the crest movement seem to be relatively small. However, the values based on the zero crossing are much larger. For example, #1′ and #2′ give a speed of about 8.4 m/s, which corresponds to $U_{\mathrm{ph}}=0.42U_0$. Since the temporal change of the acoustic pressure introduced into the pipe probably depends more on the zero crossing than on the crest, the phase speed of the jet wave may be estimated to be about $U_{\mathrm{ph}}=0.4U_0$ for the case of Fig. 5(b).

The estimate of U_{ph} from other data confirms that $U_{\mathrm{ph}}=0.4U_0$ holds approximately if we use the speed of zero crossing. This coincides with the past experimental results.[4,12,18] We will not discuss the phase speed of the jet wave any further in this paper.

B. Amplification

It is well understood from Fig. 5(a) that the jet wave forms an envelope between two bold broken lines, which are defined as $\cosh(\mu x)-1$ with amplification factors of $\mu=0.21$ and 0.20 mm^{-1}, respectively. The envelope seems to shift from $\cosh(0.21x)-1$ at the flue side to $\cosh(0.20x)-1$ at the edge side. This example shows a weak decrease of μ along the jet. Incidentally, the value of v/ω of the envelope in Fig. 5(a) may be estimated to be 1.0 mm.

Generally, we obtain results such as shown in Fig. 5(b) for the envelope. A thin broken line, which gives the envelope, does not fit any curve of $\cosh(\mu x)-1$ but does closely fit $\cosh(0.19x)-1$ if we multiply the envelope by 1/1.5. This means that the envelope can be expressed as $1.5[\cosh(0.19x)-1]$ mm. The value of 1.5 mm defines the mouth-field strength v/ω for the particular case of Fig. 5(b) when the x dependence of μ is ignored.

These two examples of the superposed jet waves suggest the approximate validity of Eq. (3) for the wave envelope. Hence we may estimate μ and v/ω at the same time by comparing the experimental envelope with the theoretical one approximated by Eq. (3).

IV. DERIVATION OF JET ENVELOPE FUNCTION

The above method of superposing the jet centerplanes is effective in estimating the jet envelope function defining the amplification factor μ and the mouth-field strength v/ω according to Eq. (3), but the process is rather tedious and time-consuming. Another simpler method is introduced in this section. Since the frames showing the jet waves in Fig. 5 are digitally memorized, the OR summation of these digital data yields a more direct superposition of jet waves. Examples are shown in Figs. 6 and 7 for models A and B, respectively.

Digitally superposed jet waves are shown in frames (a)–(f) for various blowing pressures in Figs. 6 and 7. Frames (g) and (h) on the right of Figs. 6 and 7 illustrate how the jet envelope function is derived. In frames (a)–(f) the x axis is drawn straightforwardly from the flue center labeled as ''0,'' and label ''e'' indicates the edge position.

The envelope of jet centerplanes (indicated by dotted lines in Figs. 6 and 7) is almost parallel to the outer fringe of the superposed jet waves as long as significant spreading of the jet can be neglected by using thin smokes for visualization. However, since it was not easy to make thin smokes for higher blowing pressures, jet spreading was appreciable for those pressures. This jet spreading is indicated by the convex shape of the jet fringes as shown in Fig. 6(e) and (f), which

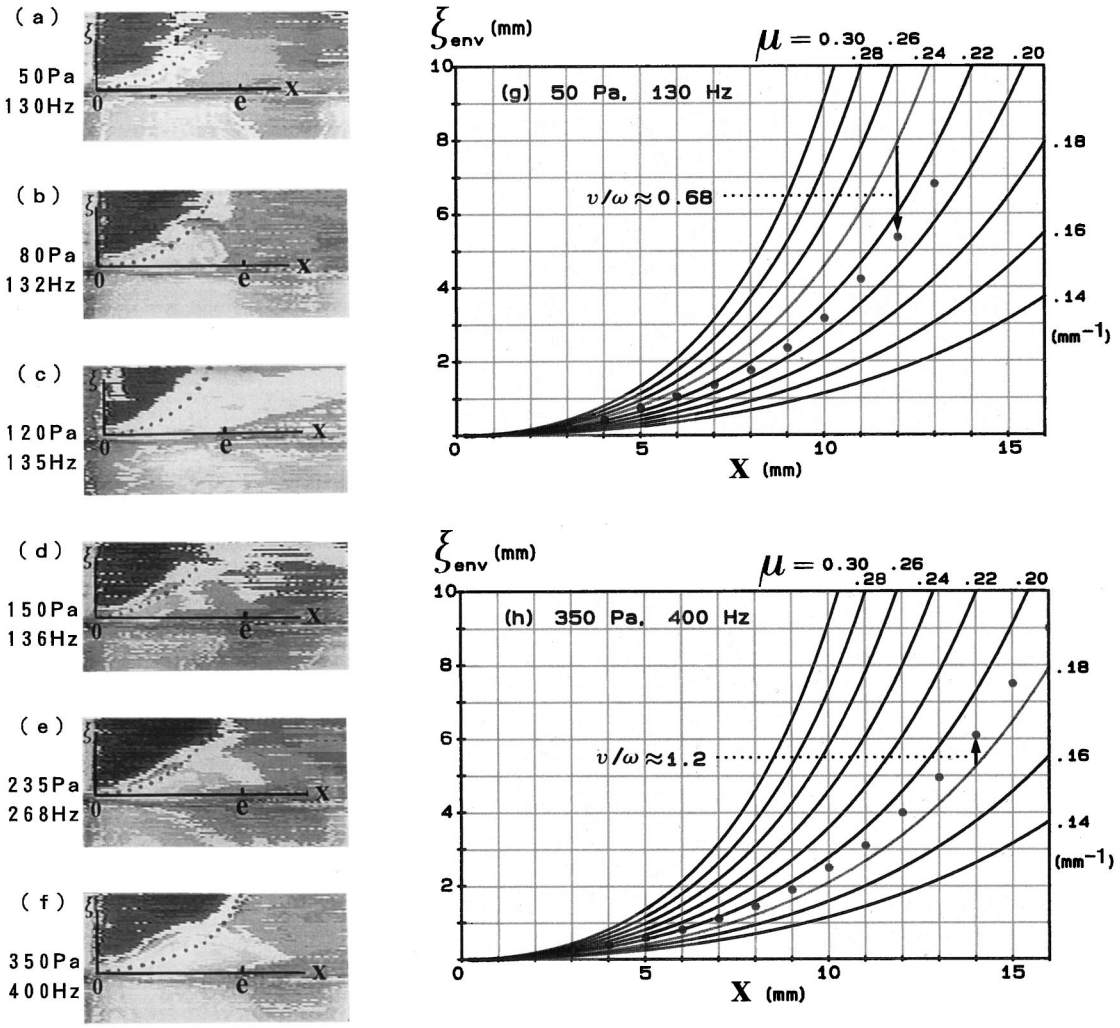


FIG. 6. Digitally superposed jet waves for various blowing pressures [(a)–(f)] and illustrations of how to derive the amplification factor μ and the mouth-field strength v/ω [(g) and (h)]. The envelope of jet centerplane is estimated by the dotted line. Organ pipe model A is used. The distance from ''0'' to ''e'' corresponds to 15.8 mm.

give examples of the second and third modes of excitation, respectively. However, it was possible to estimate the envelope of centerplanes because a darker portion was usually formed just below the fringe and it was used as a guide to guess the envelope. It should be also noticed that the roll-up of the jet often happens when the jet changes the direction of its transverse deflection after having reached maximum deflection. This roll-up was significant for low pressures as shown in Fig. 6(a), so we took care not to superpose rolled-up jets. Such neglect of the roll-up yielded a concavity of the fringe as shown in Fig. 6(d) and (e). However, this concavity did not cause any trouble in estimating the envelope of centerplanes. Of course, we may always refer back to the original jet waves when we run into trouble.

In Fig. 6(g) and (h) the estimated envelope of Fig. 6(a) and (f) are, respectively, shown by dots on a template, that is, curves of jet envelope function $\cosh(\mu x)-1$ for various μ values assuming that $v/\omega=1$. Figure 6(g) suggests a close fit of the dotted line to an envelope curve with $\mu=0.24\ \mathrm{mm}^{-1}$ when the magnitude ratio of the dotted line to that curve is about 0.68. This ratio determines $v/\omega=0.68$ mm. Also, Fig. 6(h) suggests that $\mu=0.18\ \mathrm{mm}^{-1}$ and $v/\omega=1.2$ mm. Following these procedures, we can derive the jet envelope functions of models A and B for various blowing pressures, using Figs. 6 and 7, respectively.

The results are summarized in Figs. 8 and 9 for models A and B, respectively. The abscissa denotes the jet blowing velocity U_0, which is given from the blowing pressure P according to Eq. (4). The values of μ and v/ω estimated from our flow visualization are indicated by an open circle and an open square, respectively. On the other hand, a closed square indicates a value of v/ω that was directly measured by a hot-wire anemometer. Our hot-wire anemometry method and result are described in the next section.

Flow visualization suggests the following general trends from the results shown in Figs. 8 and 9 concerning particular experimental models:

(1) The amplification factor μ tends to decrease and saturate to a given value as the oscillation of each mode shifts toward higher blowing velocities. (Although the limited data on the 3rd mode give the same value of $0.18\ \mathrm{mm}^{-1}$

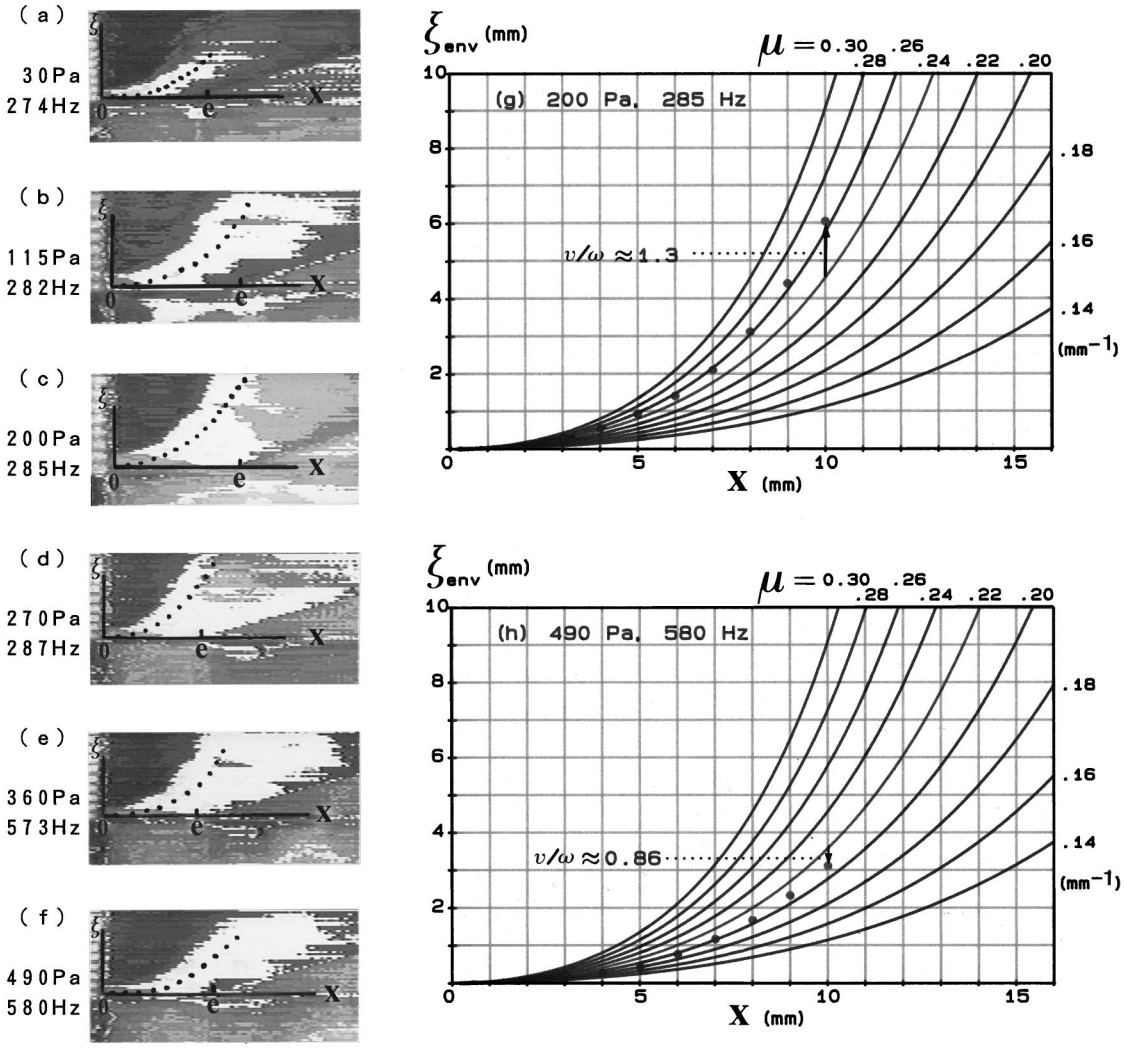


FIG. 7. Digitally superposed jet waves for various blowing pressures [(a)–(f)] and illustrations of how to derive the amplification factor μ and the mouth-field strength v/ω [(g) and (h)]. The envelope of jet centerplane is estimated by the dotted line. Organ pipe model B is used. The distance from ''0'' to ''e'' corresponds to 10.2 mm.

as shown in Fig. 8, they might be considered as a saturated state. Also, the data on the second mode of model B are not sufficient.)

(2) The averaged amplification factor is roughly estimated as 0.21 and 0.24 mm^{-1} for the jet of model A and B, respectively.

(3) The mouth-field strength v/ω, which means the displacement amplitude of the acoustic field at the mouth, tends to increase and saturate to a given value as the oscillation of each mode shifts toward higher blowing velocities. (An inclination to saturation is not so clear concerning the first and second modes of model A.)

(4) The averaged mouth-field strength seems to be almost mode independent and takes an approximate value of 1.2 mm for model A, while that for model B seems to be mode dependent and has values of 1.2 and 0.8 mm for the first and second modes, respectively.

Finally we would like to make a brief comment on the difference in the results from the superposition of the estimated jet centerplanes (cf. Sec. III) and from the direct OR summation of the digitally memorized jets (explained above). Both Fig. 5(a) given by the former method and Fig. 6(b) given by the latter method are based on data with a blowing pressure of 80 Pa, however, the sounding frequency is different by about 1 Hz between them due to a slight variation in the room temperature at which the experiments were performed. The former gave $\mu=0.20$ to 0.21 mm^{-1} and $v/\omega=1.0$ mm as indicated in Fig. 5(a); the latter gave $\mu=0.22$ mm^{-1} and $v/\omega=1.1$ mm as indicated in Fig. 8. Also, both Figs. 5(b) and 6(e) are for data at the same sounding frequency of 268 Hz. However, the blowing pressure is different by 5 Pa between them. This pair gave almost the same result of $\mu=0.19$ mm^{-1} and $v/\omega=1.5$ mm. The difference between these two examples might depend on the sounding condition. That is, the first example is in the presaturated condition, while the second one is almost in the saturated condition and should maintain a stable sounding. If these circumstances are considered, we may understand that both methods yield highly consistent estimations of the jet

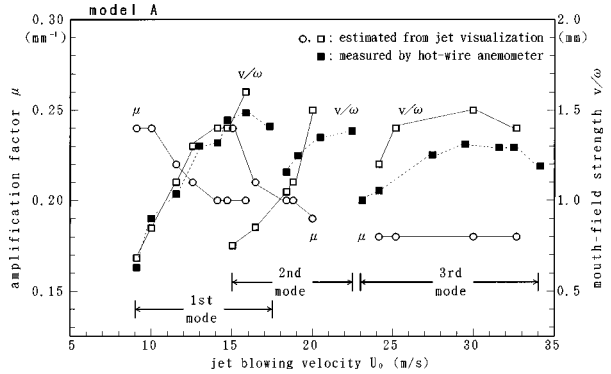


FIG. 8. The amplification factor μ and the mouth-field strength (absolute acoustic particle displacement at the mouth) v/ω of the organ pipe jet as functions of the jet velocity U_0 at the flue exit of model A.

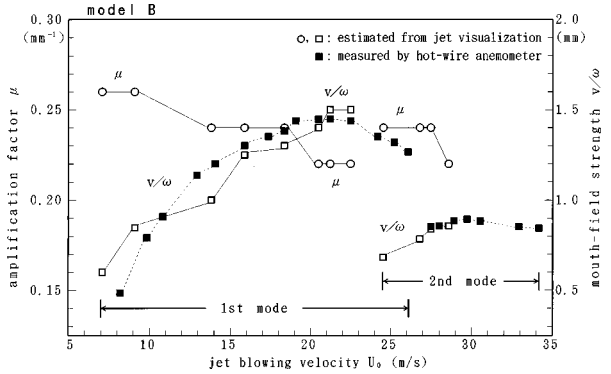


FIG. 9. The amplification factor μ and the mouth-field strength (absolute acoustic particle displacement at the mouth) v/ω of the organ pipe jet as functions of the jet velocity U_0 at the flue exit of model B.

envelope function. The expected estimation error is less than 10% according to the first example.

V. HOT-WIRE ANEMOMETRY OF MOUTH-FIELD STRENGTH

In order to confirm the validity of our OR summation method explained in the previous section, the mouth-field strength v/ω was determined from measurements of the acoustic particle velocity with a hot-wire anemometer and of the sounding frequency with a frequency counter. Since the position of a hot-wire probe at the mouth is very important, first we have to explain how we selected the probe position.

A. Preliminary experiment to select proper probe position

The position of a hot-wire filament, whose sensing part is 1-mm long and 5 μm in diameter, can be represented as its center position (x, y, z). The origin of the coordinate system is placed at the center of the flue slit. The x axis points towards the edge, the y axis along the flue width, and the z axis along the vertical of the mouth opening surface. Therefore the mouth opening is expressed as $0 \leqslant x \leqslant d$ and $-b/2 \leqslant y \leqslant b/2$ at $z=0$, the flue slit as $-b/2 \leqslant y \leqslant b/2$ and $-h \leqslant z \leqslant h$ at $x=0$ (cf. Table I).

Since the aim of the experiment is to measure the acoustic particle velocity, we have to avoid exposing the probe to the air-jet flow. However, the jet is deflecting with large amplitudes as shown in Figs. 3–7. Particularly, from Figs. 6 and 7 we know that proper positions for our measurement are limited to the flue side. On the other hand, if the probe is located far away from the mouth opening, the level of the acoustic particle velocity will decrease with increasing distance.

A constant-temperature type hot-wire anemometer (Kanomax model 1010) was used with a linearizer (model 1013), a dc voltmeter (model 1008), and a rms voltmeter (model 1015). Selectable bandpass filters are involved in the rms voltmeter. The output voltage of a hot-wire probe was calibrated using a small flow tank (probe calibrator, Kanomax model 1065) and a low-noise blower that was usedfor flow visualization. A linearized calibration curve was obtained with the aid of the linearizer. As the result, the output voltage of the anemometer from 0 to 5.0 V directly corresponded to the flow velocity from 0 to 5.0 m/s with sufficient accuracy. Such very low flow velocities were used to measure the acoustic particle velocity.

The data were taken while shifting the probe position step by step under a constant blowing pressure of 120 Pa and with the frequency of 133.7 Hz used for model A. The rms output of our hot-wire anemometer $V_{\rm rms}$ (bandpass filtration was applied between 100 and 200 Hz) is shown in Fig. 10, which gives the data when the probe was moved on the central x-z plane at $y=0$. It is clear from Fig. 10 that $V_{\rm rms}$ is strongly affected by the jet flow, whose speed is known from the dc voltmeter, when the probe is located at $z<6$ mm or $x>5$ mm. On the other hand, when the probe is located at $z>14$ mm, $V_{\rm rms}$ decreases as z increases. Also, when the probe is very close to the flue slit as $x=1$ mm, $V_{\rm rms}$ shows stronger z dependence than that at $x=3$ mm. This might be due to the entrainment of the flow along the wall surface above the flue. Consequently, an adequate region to measure the rms acoustic particle velocity $v_{\rm rms}$ at the mouth should be near $x=3$ mm and $z=8-10$ mm.

More data were taken as the probe was moved along the y axis with fixed $x=5$ mm and $z=4$ mm. The result indicated a reliable region of $-15 \leqslant y \leqslant 15$ mm. As the probe approached the pipe wall, $V_{\rm rms}$ clearly decreased and took a value of about 30% below the value in the reliable region when $y=\pm 25$ mm. Therefore the plane at $y=0$ should be the most reliable plane for the probe position. As the result, we located the probe at $(x,y,z)=(3,0,10)$ in mm units. At this position the rms output voltage $V_{\rm rms}$ should correspond to the rms value of the acoustic particle velocity $v_{\rm rms}$ concerning the first-mode oscillation of model A.

Similarly, other locations for the probe were sought depending on the oscillation mode and organ pipe model. The results gave almost the same position as that for the first-mode oscillation of model A. Therefore the probe was fixed to $(x,y,z)=(3,0,8)$ in mm units for the second and third modes of model A and the first and second modes of model B. A little bit smaller z position seemed to give a better result if a more rapid decrease in $v_{\rm rms}$ along z axis was considered because of shorter wavelengths in these oscillation modes.

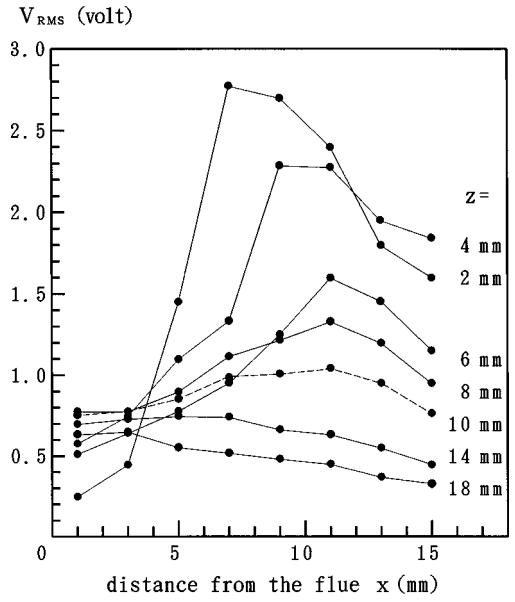


FIG. 10. The rms output voltage $V_{\rm rms}$ of the hot-wire anemometer at the mouth of model A. A hot-wire probe is shifted on the x-z plane at $y=0$. The blowing pressure of the jet is kept at 120 Pa, and the resulting tone frequency is 133.7 Hz.

B. Acoustic particle velocity at the mouth

According to the procedures above, the rms acoustic particle velocity at the mouth $v_{\rm rms}$ was measured as the rms output voltage $V_{\rm rms}$ of the hot-wire anemometer. The magnitude v, which was used to define the strength of mouth field v/ω in Eqs. (1) to (3), is given by multiplying $v_{\rm rms}$ by $\sqrt{2}$. The results of this v/ω are indicated by the closed square in Figs. 8 and 9 for models A and B, respectively. Note that a bandpass filtration was applied to detect only the fundamental component of the acoustic particle velocity in each oscillation mode.

Comparing the values of v/ω measured using the hot-wire anemometer with those estimated from flow visualization, we may see a good agreement between them except for the third mode in model A. This agreement implies that the method of deriving the envelope function of the jet wave from flow visualization (described in Secs. III and IV) is valid and is sufficiently accurate. An apparent overestimation by the flow-visualization method in the third mode of model A is possibly due to the spreading of thickly smoked jets.

The measurement result by the hot-wire anemometer well indicates the saturation tendency of the mouth-field strength. Although the measured v/ω of the fundamental component tends to decrease after having reached a peak level as shown in Figs. 8 and 9, the total level including higher harmonics (measured without filtration) is still increasing slightly. This possibly implies that the acoustic vortex shedding at the edge is a significant source of dissipation for the fundamental and is also a source of higher harmonics at saturation.[7–10] We had particular difficulty in obtaining a smooth jet-wave envelope near saturation, and the estimated data of μ and v/ω were lacking in Figs. 8 and 9. This was due to other jet waves which were generated by the second harmonic and superposed upon the jet waves generated by the fundamental.

VI. DISCUSSION

An envelope-based method to estimate both the jet-wave amplification factor and the mouth-field strength in organ pipes from flow visualization was demonstrated in previous sections. The experimental results were presented in dimensional terms for particular models. However, since the amplification factor is not a universal quantity as noted in the Introduction, we will use a dimensionless representation of data in this section to compare wave characteristics between organ pipe jets and free jets.

A. Organ pipe jet, free jet, and jet instability

Fletcher and Thwaites[1,2] did pioneering work on the measurement of the amplification factor on acoustically perturbed free jets. Their data are summarized in Fig. 11. Its abscissa is the Strouhal number,

$$\mathrm{St}=\omega h^*/U_0, \tag{6}$$

where h^* denotes the half-thickness of the jet. Note that their definition of the half-thickness b corresponds to our h^*, and that they used $U_{\rm ph}$ instead of our U_0. Therefore their kb corresponds to our 2 St if we assume that $U_{\rm ph}=U_0/2$ for simplicity. The data in Fig. 11 are based on such an assumption. It should be also noted that h^* is determined from the jet velocity profile in the cross-stream direction z. In addition, theoretical curves of the temporal analysis[18] (using the complex frequency) and spatial analysis[14] (using the real frequency) of the jet instability are shown in Fig. 11, where the velocity profile $\mathrm{sech}^2\,(z/h^*)$ of a Bickley jet is supposed. The data are closer to the curve of the temporal analysis for $\mathrm{St}<0.3$, although some data show anomalously large values for $\mathrm{St}>0.3$, which is contrary to the theoretical decrease.

On the other hand, very recently Nolle[13] investigated the characteristics of a free jet concerning the phase velocity and amplification factor using synchronous hot-wire measurements. His data on μh^* follows the spatial-analysis curve very closely over the range $0.03<\mathrm{St}<1$. Hence the acoustically perturbed free jet seems to share wave characteristics with the jet instability if it is described by the spatial analysis. Thus Nolle's result seems to endorse Verge's suggestion[8] that the spatial analysis[14] should be more relevant than the temporal analysis.[18]

However, it seems difficult to explain the discrepancy between the results of Nolle[13] and Fletcher and Thwaites[12] on the free jet within the framework of our current knowledge, although Nolle[13] realized a Bickley jet in his experiments, while Fletcher and Thwaites[12] had the profile only roughly approximating that of a Bickley jet in theirs. Also, it should be noted that the Bickley jet is the similarity solution for a fully developed laminar jet.[17] Therefore the Bickley jet cannot be well established when the jet transitions to turbulence.

The data shown in Figs. 8 and 9 are nondimensionalized and plotted in Fig. 12 except for a few to avoid overlapping. Since the measurement of the jet velocity profile was not done, h^* is supposed to be $h(=1.1\ \mathrm{mm})$ arbitrarily in Fig. 12. Note that the sounding frequencies of the models were as

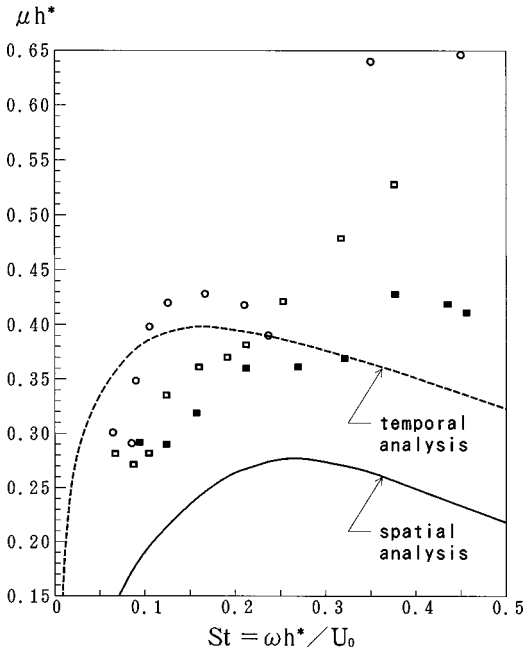


FIG. 11. Comparison of the dimensionless amplification factor μh^* of free jets (Ref. 12) with linear stability theory for a Bickley jet. The solid line: spatial theory; the broken line: temporal theory. Experimental data (Ref. 12) are shown by: ○ ($h^* \approx 0.7$ mm, $P = 100$ Pa), □ ($h^* \approx 0.45$ mm, $P = 80$ Pa), and ■ ($h^* \approx 0.45$ mm, $P = 40$ Pa).

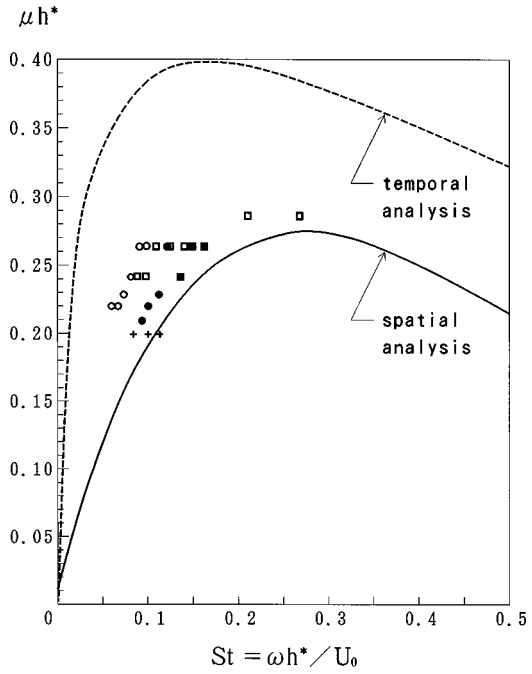


FIG. 12. Comparison of the dimensionless amplification factor μh^* of organ pipe jets with linear stability theory for a Bickley jet. The solid line: spatial theory; the broken line: temporal theory. The data of model A are shown by: ○ (first mode), ● (second mode), and + (third mode); the data of model B are shown by: □ (first mode) and ■ (second mode). The data are nondimensionalized supposing $h^* = h$.

follows: 130–136 Hz for the 1st mode of model A; 264–268 Hz for the second mode of model A; 400–406 Hz for the third mode of model A; 274–288 Hz for the first mode of model B; and 573–580 Hz for the second mode of model B. Also, note that the Reynolds number $\mathrm{Re} = 2hU_0/\nu$ (ν is the kinematic viscosity of the air, 1.51×10^{-5} m^2/s at 20 °C) extends from about 10^3 to about 5×10^3 in our experiments. We may thus consider that our jets are between laminar and turbulent jets.

The data of model A (open circle, closed circle, and ''plus'' for the first, second, and third mode, respectively) are distributed over a narrow range of St (from about 0.05 to 0.12) and indicate a sharp increase with increasing St except for the data of the third mode. On the other hand, the data of model B (open square and closed square for the first and second mode, respectively) extend over a wider range of St (from about 0.08 to 0.26) and indicate a more gradual increase with increasing St. The differences of data between two models may be attributed to the difference of d (the flue-to-edge distance) or nondimensionalized $d/2h$. Model A gives $d/2h = 7.2$ and model B gives $d/2h = 4.6$. Generally, a larger d requires a larger U_0 and a smaller ω (a larger pipe length L), that is, a smaller St. Unfortunately, the data for $\mathrm{St} > 0.3$ were not obtained because an experimental model of much smaller $d/2h$ blown with lower pressures was not provided.

It is important that our argument of dimensionless data analysis (shown in Fig. 12) is strongly influenced by the evaluation of h^*, which depends on the jet velocity profile at the flue exit and its subsequent change in the cross-stream direction.[8,9] Since we are lacking experimental results of jet velocity profile, we try to examine the sensitivity of the result shown in Fig. 12 for a change in h^*. Our organ pipe models have a long flue channel ($L_t = 100$ mm, $L_f/2h = 45.5$), so we probably have a Poiseuille profile with the parabolic distribution at the exit, which gives

$$h^* = (2/5)2h = 4h/5 \tag{7}$$

with the assumption that a subsequent velocity profile is given by a Bickley jet.[8,9] In Eq. (7) $2h$ defines the flue thickness, while Verge *et al.*[8] used h for the flue thickness and used b for h^*.

Sato[17] had experimental results showing $h \sim 0.8h^*$ for $x/2h < 7$ on the jets emerging from long channels ($L_t = 1100$ mm, $2h = 6$ and 10 mm, $U_0 = 10$ m/s, and $\mathrm{Re} = 4.0 \times 10^3$ and 6.6×10^3). Also, Ishizaka and Matsudaira[20] indicated that a Poiseuille profile was established as the distance from the channel inlet x' exceeded $x'_p = 0.028(2h)$ Re. So, we have $x'_p = 62$ mm for our organ pipe jet of $2h = 2.2$ mm and $\mathrm{Re} = 10^3$. Although the Reynolds number Re should be lower than 1.6×10^3 when the flue channel of our organ pipe models satisfies $x'_p < L_f (= 100\ \mathrm{mm})$, our organ pipe jet ($10^3 < \mathrm{Re} < 5 \times 10^3$) may be considered to have a Poiseuille-like profile at the flue exit and the relation of Eq. (7) holds as long as x is smaller than $7(2h) = 15.4$ mm.

On the other hand, if we have a uniform flow at the exit and a subsequent Bickley jet,[8,9]

$$h^* = (3/4)2h = 3h/2. \tag{8}$$

Fletcher and Thwaites[12] obtained $h^* \approx 1.4$ h for $P = 100$ Pa and $h^* \approx 1.8$ h for $P = 80$ and 40 Pa based on the measurement of the jet velocity profile. Their higher values of the ratio h^*/h might be attributed to a uniform flow at the exit. Consequently, it seems to be adequate to discuss the sensi-

tivity of our organ pipe jets to changes in h^* as given by Eqs. (7) and (8).

Figure 13 shows the result, where the data indicated in Fig. 12 ($h^*=h$) is roughly approximated by the thick solid line connecting the three points $(\mathrm{St},\mu h^*)=(0.05,\ 0.20)$, (0.18, 0.29), and (0.30, 0.29). The thick dotted and broken lines indicate the results when Eqs. (7) and (8) are applied to the data represented by the thick solid line, respectively.

Interestingly enough, the thick broken line ($h^*=3h/2$) in Fig. 13 roughly reflects the tendency of the experimental data in Fig. 11 except for a few anomalous data. On the other hand, the thick dotted line ($h^*=4h/5$) is close to the spatial-analysis curve for the jet instability. Theoretically, this thick dotted line seems to match our organ pipe jets because our long flue channel probably forms a Poiseuille flow at the exit. However, we do not know the exact changes of the subsequent velocity profile and of the velocity on the jet axis along our jets. Jet deflection with large amplitudes might also cause drastic changes in them. Hence we cannot draw a definite conclusion from Fig. 13 on whether an organ pipe jet shares wave characteristics with an acoustically perturbed free jet or not until we can obtain reliable experimental results along the jet.

Particularly, we should bear in mind the large deflection amplitudes of the organ pipe jets shown in Figs. 3–7 compared with the small free-jet amplitudes, which are at most equal to the jet thickness. The spatial and temporal theories of jet instability, which are linearized on the assumption of small amplitudes, may thus be applicable to free jets apart from the problem of which theory is more relevant. However, both theories might be inapplicable to organ pipe jets because of their nonlinear nature arising from their large amplitudes.

B. Stimulating source of jet deflection

Coltman[11] drew a seemingly contradictory conclusion about the source which initiates and sustains the oscillation of free and organ pipe jets. According to his result, the dominant source stimulating the jet wave is the transverse disturbance velocity in the flute and the transverse pressure gradient in free jet. Velocity and pressure-gradient models for organ pipes have been originally proposed by Coltman[18] and Yoshikawa and Saneyoshi,[21] respectively. Later Elder[22] developed a pressure-gradient model on the basis of the vibrating-diaphragm concept of the jet. Coltman[11] also rejected the negative displacement model of Fletcher.[5] See Ref. 13 for a more detailed and updated discussion on this problem.

However, the negative displacement model has been successfully applied in this paper, although its application is restricted to the jet envelope function in exact sense. This success might be illusive because Eq. (1) is only a semi-empirical relation. But, a good agreement of the mouth-field strength between envelope-based estimation and hot-wire measurement encourages us to reconsider the acoustic displacement as well as the acoustic velocity and the pressure gradient as the source of stimulation. We have to develop a radical new theory of the organ pipe jet deflection in the future based on fluid dynamics.

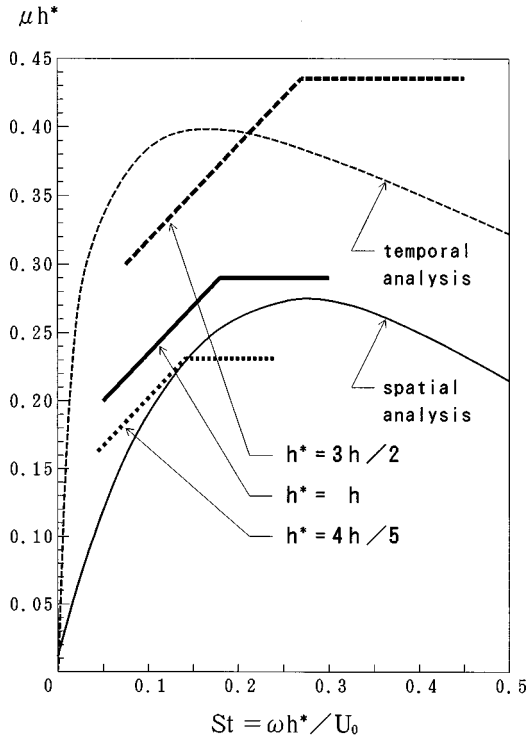


FIG. 13. Sensitivity of the experimental data shown in Fig. 12 (approximated by the thick solid line) for a change in the jet half-thickness h^*. The thick dotted line is for $h^*=4h/5$ (a Poiseulle channel flow and a Bickley jet) and the thick broken line for $h^*=3h/2$ (a uniform channel flow and a Bickley jet).

VII. CONCLUSIONS

The amplification factor of a growing jet wave in organ pipes was successfully estimated from flow-visualization data acquired with a high-speed digital video camera. The estimation was done by comparing the theoretical and experimental envelopes of jet waves in steady states. The theoretical envelope was approximated using the negative displacement model[5] to separate the amplification factor from the phase velocity and time reference of the jet wave. The experimental envelope was extracted by superposing jet waves visualized as instantaneous shapes (called streaklines) of lateral jet deflection. Theoretical approximation of the envelope tended to give an estimate of the amplification factor that was about 10% higher than the experimental one. We speculated that experimental error was below 10%.

Our envelope-based method gave an estimation of the mouth-field strength (defined as the absolute amplitude of the acoustic displacement) as well as the amplification factor. A good agreement of the mouth-field strength was obtained between envelope-based estimation and hot-wire measurement. This agreement confirmed the validity and reliability of our envelope-based method. As the result, the amplification factor was roughly estimated to be in the $0.18-0.26\ \mathrm{mm}^{-1}$ range, while the mouth-field strength was 0.5–1.5 mm in our two experimental models.

A dimensionless representation of our experimental data on the amplification factor was employed to discuss the

wavy nature of an organ pipe jet and an acoustically perturbed free jet. The importance of accurately estimating the jet half-thickness from the velocity profile was rerecognized when examining the sensitivity of our organ pipe jets for a change in the half-thickness. If it is valid to assume a Poiseulle flow at the flue exit and a subsequent Bickley jet, the spatial theory on jet instability seems to be relevant to our organ pipe jets.

However, definitive conclusions about the difference between organ pipe and free jets must be postponed until we can obtain experimental data on the half-thickness to determine the amplification factor more exactly. The spatial and temporal theories on jet instability, which are valid for small amplitudes of disturbance less than the jet thickness, might be inapplicable to large amplitudes of an organ pipe jet wave, which call for a nonlinear approach to proper analysis. Also, a brief comment on the source of stimulation of jet deflection was given, and our reconsideration of the acoustic displacement[5] as well as the acoustic velocity[18] and the acoustic pressure gradient[21,22] was encouraged.

ACKNOWLEDGMENTS

The author would like to show his sincere appreciation for helpful comments and kind suggestions for the manuscript and a great deal of encouragement from Professor Neville H. Fletcher of Australian National University, Professor A. Wilson Nolle of the University of Texas at Austin, Professor Samuel A. Elder of the U.S. Naval Academy, Professor Abraham Hirschberg of Technishe Universiteit Eindhoven, and Dr. Marc-Pierre Verge of IRCAM. The author is grateful to Professor James B. Cole of Tsukuba University for his careful editing of the English expression. The author is also grateful to Professor Keinosuke Nagai of Tsukuba University for his long-term support of hot-wire measurement. Finally, the author would like to express his acknowledgments to the anonymous reviewers, each of whom contributed relevant comments and valuable suggestions.

[1] H. L. F. Helmholtz, *On the Sensations of Tone*, translated by A. J. Ellis (Dover, New York, 1954), pp. 88–94.
[2] Lord Rayleigh (J. W. S. Strutt), *The Theory of Sound* (Dover, New York, 1945), Vol. 2, Chap. 21.
[3] J. W. Coltman, ''Jet drive mechanisms in edge tones and organ pipes,'' J. Acoust. Soc. Am. **60**, 725–733 (1976).
[4] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments* (Springer-Verlag, New York, 1991), Chap. 16.
[5] N. H. Fletcher, ''Sound production by organ flue pipes,'' J. Acoust. Soc. Am. **60**, 926–936 (1976).
[6] B. Fabre, ''La production de son dans les instruments a embouchure de flute: Modele aero-acoustique pour la simulation temporelle,'' Ph.D. thesis, Universite du Maine, Le Mans, France, 1992.
[7] B. Fabre, A. Hirschberg, and A. P. J. Wijnands, ''Vortex shedding in steady oscillation of a flue organ pipe,'' Acust. Acta Acust. **82**, 863–877 (1996).
[8] M. P. Verge, R. Caussé, B. Fabre, A. Hirschberg, A. P. J. Wijnands, and A. van Steenbergen, ''Jet oscillation and jet drive in recorder-like instruments,'' Acust. Acta Acust. **2**, 403–419 (1994).
[9] M. P. Verge, ''Aeroacoustics of confined jets with application to the physical modeling of recorder-like instruments,'' Ph.D. thesis, Technische Universiteit Eindhoven, Eindhoven, The Netherlands, 1995.
[10] M. P. Verge, B. Farbe, A. Hirschberg, and A. P. J. Wijnands, ''Sound production in recorderlike instruments. I. Dimensionless amplitude of the internal acoustic field,'' J. Acoust. Soc. Am. **101**, 2914–2924 (1997).
[11] J. W. Coltman, ''Jet behavior in the flute,'' J. Acoust. Soc. Am. **92**, 74–83 (1992).
[12] N. H. Fletcher and S. Thwaites, ''Wave propagation on an acoustically perturbed jet,'' Acustica **42**, 323–334 (1979).
[13] A. W. Nolle, ''Acoustically driven sinuous instability of a planar air jet'' (unpublished), pp. 1–41.
[14] G. E. Mattingly and W. O. Criminale, Jr., ''Disturbance characteristics in a plane jet,'' Phys. Fluids **14**, 2258–2264 (1971).
[15] A. Michalke, ''On spatially growing disturbances in an inviscid shear layer,'' J. Fluid Mech. **23**, 521–544 (1965).
[16] P. G. Drazin and L. N. Howard, ''Hydrodynamic stability of parallel flow of invisid fluid,'' Adv. Appl. Mech. **9**, 1–89 (1966).
[17] H. Sato, ''The stability and transition of a two-dimensional jet,'' J. Fluid Mech. **7**, 53–80 (1960).
[18] J. W. Coltman, ''Sounding mechanism of the flute and organ pipe,'' J. Acoust. Soc. Am. **44**, 983–992 (1968).
[19] T. E. Faber, *Fluid Dynamics for Physicists* (Cambridge U. P., New York, 1995), pp. 213–214.
[20] K. Ishizaka and M. Matsudaira, *Fluid Mechanical Considerations of Vocal Cord Vibration* (SCRL Monograph No. 8. Speech Communications Research Laboratory, Inc., Santa Barbara, California, 1972).
[21] S. Yoshikawa and J. Saneyoshi, ''Feedback excitation mechanism in organ pipes,'' J. Acoust. Soc. Jpn. (E) **1**, 175–191 (1980).
[22] S. A. Elder, ''The mechanism of sound production in organ pipes and cavity resonators,'' J. Acoust. Soc. Jpn. (E) **13**, 11–23 (1992).

Acoustic anisotropy in bovine cancellous bone

A. Hosokawa and T. Otani
Department of Electrical Engineering, Faculty of Engineering, Doshisha University, 610-0321 Kyotanabe-shi, Japan

(Received 8 August 1997; accepted for publication 22 January 1998)

This paper presents the experimental results on the acoustic anisotropy in bovine cancellous bone. The propagation of both fast and slow longitudinal waves in bovine cancellous bone was experimentally examined in relation to the structural anisotropy, or the trabecular arrangement. Propagation speeds of the fast and slow waves were measured as a function of the propagation angle to the trabecular alignment, and theoretically estimated by use of Biot's theory for an isotropic medium. © *1998 Acoustical Society of America.* [S0001-4966(98)01605-1]

PACS numbers: 43.80.Cs [FD]

INTRODUCTION

Cancellous bone is comprised of a connected network of trabeculae, and can be considered an anisotropic medium due to the arrangement of the trabecular pattern. It is predicted that cancellous bone has an acoustic anisotropy in relation to the trabecular arrangement. In our previous paper,[1] the propagation of both fast and slow longitudinal waves in bovine cancellous bone was clearly identified when the acoustic wave travels parallel to the direction of the trabecular alignment. The purpose of the present paper is to reveal the relation between the propagation of the two waves and the structural anisotropy in cancellous bone.

An ultrasonic immersion technique is used to determine the propagation speeds of the fast and slow waves as a function of the propagation angle to the trabecular alignment. Theoretical discussion is performed by applying Biot's theory on the acoustic propagation in a fluid saturated porous solid.

I. THEORY

Biot proposed a phenomenological theory of acoustic propagation in a porous elastic solid saturated by a viscous fluid,[2,3] by considering a coupling between the solid and fluid. In Biot's theory, the average motions of both the solid and fluid parts are separately described. At high frequencies sufficient enough that the viscous skin depth $(2\eta/\rho_f\omega)^{1/2}$ is much smaller than the pore size a_0 (η is the fluid viscosity, ρ_f is the density of the fluid, and ω is an angular frequency), the two longitudinal waves have nondispersive speeds $\nu_{\text{fast,slow}}$ of the form[4,5]

$$\nu^2_{\text{fast,slow}}=2(HM-C^2)/\{(M\rho+Hm-2C\rho_f)\mp[(M\rho+Hm-2C\rho_f)^2-4(HM-C^2)(m\rho-\rho_f^2)]^{1/2}\}. \tag{1}$$

Here, H, C, and M are generalized elastic coefficients which can be related to the bulk moduli of the solid K_s, the pore fluid K_f, the bulk K_b, the shear moduli μ of the skeletal frame, and the porosity β.[6,7] β is expressed as

$$\beta=\frac{\rho_s-\rho}{\rho_s-\rho_f}, \tag{2}$$

where ρ is the total density and ρ_s is the density of the solid. In Eq. (1), m is the density parameter, which is used to account for the fact that not all of the pore fluid moves in the direction of the pressure gradient because of tortuosity or sinuosity, and is defined by the equation

$$m=\alpha\rho_f/\beta. \tag{3}$$

Here, α is the structure factor and it is determined by the relation given by Berryman[8] as

$$\alpha=1-r(1-1/\beta), \tag{4}$$

where r is a variable calculated from a microscopic model of a frame moving in the fluid.

According to Gibson,[9] the bulk K_b and shear moduli μ of the trabecular frame of cancellous bone are functions of the bone volume fraction $V_f(=1-\beta)$ and expressed as

$$K_b=\frac{E_s}{3(1-2\nu_b)}V_f^n, \tag{5}$$

$$\mu=\frac{E_s}{2(1+\nu_b)}V_f^n, \tag{6}$$

where n is a variable depending on the geometrical structure of the trabecular frame[9] and ν_b is the Poisson's ratio of the frame.

II. EXPERIMENTS

A. Specimen preparation and experimental method

Eleven cancellous bone specimens, about 20×20 mm in size and 9 mm in thickness, were cut from the distal epiphysis of bovine femora, with soft tissue (bone marrow) *in situ*. The density ρ of each specimen was measured using Archimedes' principle, and then the porosity β was calculated by Eq. (2). The density ρ_s of the solid bone was given by Lang as 1960 kg/m^3,[10] and are equivalent in value to bovine cortical bone. (The value of $\rho_s=2000\pm50$ kg/m^3 was measured by ourselves.) The density $\rho_f=930$ kg/m^3 of bone marrow was taken from our own work.[1] The trabecular arrangement was estimated by x-ray photography. An x-ray photograph of a typical trabecular structure in cancellous bone used in our experiments and its local coordinate system are shown in Fig. 1. The axes x_1, x_2, and x_3 are equivalent

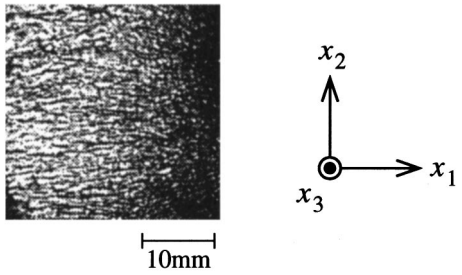


FIG. 1. X-ray photograph of bovine cancellous bone.

to the longitudinal, transverse, and sagittal axes in the general anatomic coordinates. The x_1 axis is in the main direction of the trabecular alignment, and x_2 and x_3 are perpendicular to the major direction of trabeculae. One of the eleven specimens had trabecular structure aligned parallel to the thickness direction and the others in the perpendicular direction. As we predict that air penetrates the narrow spaces between the trabeculae and soft tissue during the cutting and forming processes of the specimens, prior to experiments all specimens were saturated with water and decompressed to remove air bubbles.

The experimental arrangement of the ultrasonic immersion technique is shown in Fig. 2. A wide band (0.1–10 MHz) PVDF transmitter and a hydrophone were submerged in distilled water at $23.0 \pm 0.5\,^{\circ}\mathrm{C}$ at a distance of 90 mm apart. The specimen was then mounted between the transmitter and hydrophone. The transmitter was excited to emit an ultrasonic pulse wave with a center frequency of 1 MHz and the associated pulse waves, that propagate through the water/specimen/water system, were detected by the hydrophone. The output signal was amplified by a 40-dB preamplifier and displayed on a digital storage oscilloscope. The specimen was held at an angle θ_i to the ultrasonic beam as shown in Fig. 3. At oblique incidence on the surface of the specimen, an ultrasonic beam shift occurs in the y direction because of refraction, and has a displacement of

$$\Delta y = \frac{\sin(\theta_r - \theta_i)}{\cos\,\theta_r}\, d, \tag{7}$$

where θ_r is the angle of refraction and d is the thickness of the specimen. From this, time difference Δt of the signal

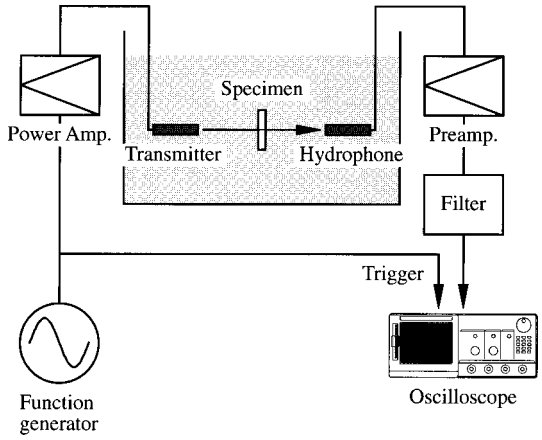


FIG. 2. Experimental arrangement of the ultrasonic immersion technique.

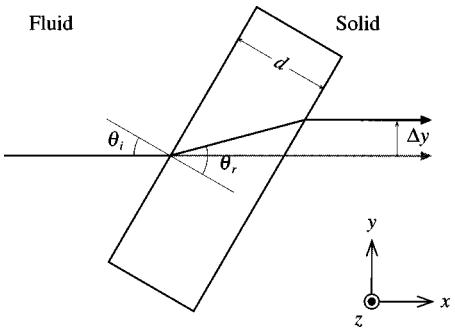


FIG. 3. Schematic drawing of refraction at fluid/solid/fluid interfaces.

with and without a specimen in the path of the ultrasonic beam is then

$$\Delta t = \left[\frac{\cos(\theta_r - \theta_i)}{\cos\,\theta_r}\, d\right] \frac{1}{\nu_0} - \frac{d}{\cos\,\theta_r}\,\frac{1}{\nu}, \tag{8}$$

where ν is the propagation speed of ultrasonic wave in the specimen and ν_0 the wave speed in water. The propagation angle, θ, to the trabecular direction equals the refracted angle θ_r if the trabeculae of the specimen are aligned parallel to the thickness direction. The propagation speed $\nu(\theta)$ in the specimen at an angle θ can be derived by measuring the ultrasonic beam shift Δy and the time difference Δt with Eqs. (7) and (8).

B. Experimental results

Figure 4 shows a pulsed waveform traveling only in water, and Fig. 5 pulsed waveforms traveling through the specimen of density $\rho = 1120\ \mathrm{kg/m^3}$ (porosity $\beta = 0.82$) with trabeculae aligned parallel to the thickness direction, at incident angles of $\theta_i = 0°$, 10°, 20°, and 30°. At incident angles of $\theta_i = 0°$, 10°, 20°, and 30°, the fast wave propagates at angles of $\theta = 0°$, 16°, 31°, and 44° to the trabecular alignment by refraction and the slow wave at $\theta = 0°$, 9°, 18°, and 27°. Both fast and slow longitudinal waves can be clearly observed at $\theta_i = 0°$ (or the x_1 direction shown in Fig. 1). As the angle θ_i of incidence increases (i.e., as the angle θ of propagation increases), the propagating time of the fast wave between the transmitter and hydrophone becomes slower. On the other hand, little change is observed in the propagating time for the slow wave. Accordingly, the fast and slow waveforms widely overlap in time at incident angles of $\theta_i = 20°$

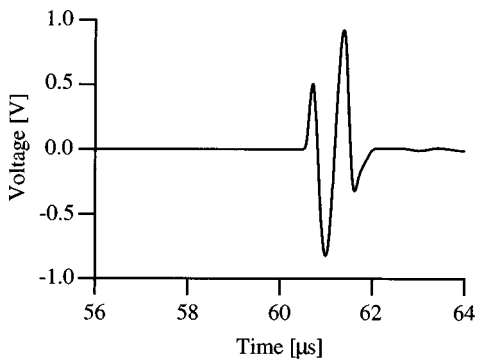


FIG. 4. Pulsed waveform with a center frequency of 1 MHz traveling in water.

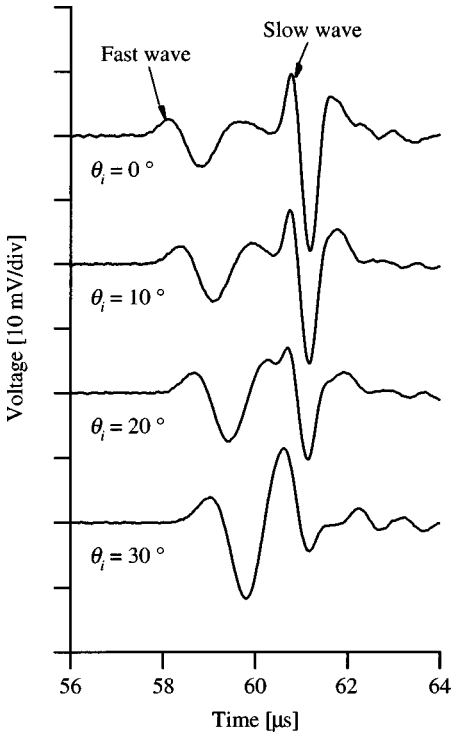


FIG. 5. Pulsed waveforms traveling through a cancellous bone with trabeculae aligned parallel to the thickness direction, at various angles θ_i of incidence.

and 30°. At oblique incidence on the solid surface, the shear wave can be generated in the solid by mode conversion. Although the experiments were carefully performed, the shear wave propagating through cancellous bone could not be observed at any angle of incidence. We therefore consider that the observed results are not affected by the generation of the shear wave. Figure 6 shows a pulsed waveform traveling through another specimen ($\rho = 1120$ kg/m^3) with trabeculae aligned in the perpendicular direction at normal incidence (or propagating at an angle $\theta = 90°$ to the trabeculae). In Fig. 6, the fast and slow waves completely overlap and are observed as if a single wave propagated. Both the fast and slow waves generally exist on propagating through cancellous bone.

The propagation speeds of the fast and slow waves were measured as a function of the propagation angle θ to the trabecular alignment. The specimen of porosity $\beta = 0.82$ with trabeculae aligned in the parallel direction was then used.

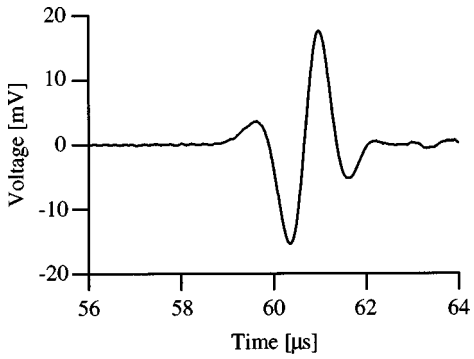


FIG. 6. Pulsed waveform traveling through a cancellous bone with trabeculae aligned perpendicular to the thickness direction, at normal incidence.

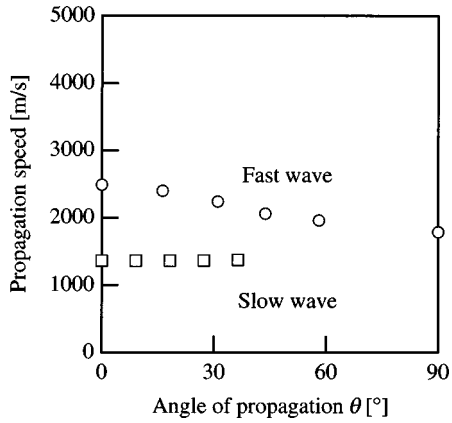


FIG. 7. Propagation speeds of fast and slow waves in cancellous bone as a function of propagation angle θ to the trabecular alignment.

For the measurement of the slow wave, a pulse wave with a center frequency of 3 MHz was used. The fast wave propagating in bovine cancellous bone is highly attenuated over 1.5 MHz.[1] At a frequency of 3 MHz, the signal of the fast wave is much weaker than that of the slow wave, allowing easy separation of the slow wave. Figure 7 shows the measured results. In Fig. 7, the propagation speed of the slow wave is almost constant, but the speed of the fast wave decreases and approaches close to that of the slow wave as the angle θ increases. Figure 8 shows the speed of the fast wave in the perpendicular direction to the trabeculae (or the x_2 direction) as a function of bone volume fraction V_f. Ten cancellous bone specimens with trabeculae aligned in the perpendicular direction were then used. No data for the slow wave was obtained in this case because the fast waveform completely overlaps that of the slow waveform. For comparison, our previous data[1] for propagation in the parallel direction (the x_1 direction) is reproduced in Fig. 9. In the range $V_f = 0.1-0.3$, the propagation speeds of the fast wave propagating in the perpendicular direction is much slower than that in the parallel. Slight scatters of the measured values are observed in both Figs. 8 and 9 because of the heterogeneity of the trabecular structure.

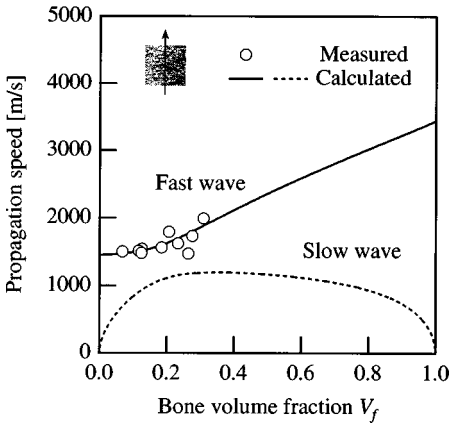


FIG. 8. Propagation speeds of fast and slow waves in cancellous bone with trabeculae aligned perpendicular to the direction of propagation, as a function of bone volume fraction V_f.

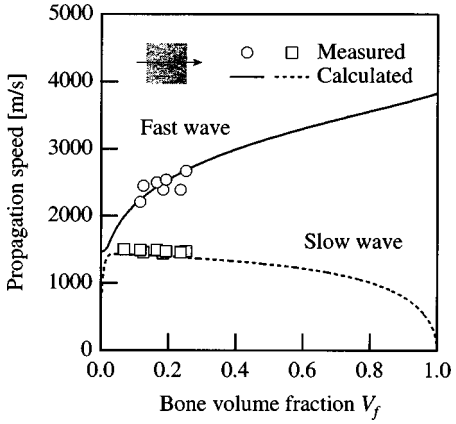


FIG. 9. Propagation speeds of fast and slow waves in cancellous bone with trabeculae aligned parallel to the direction of propagation, as a function of bone volume fraction V_f. This data first appeared in the Journal of the Acoustical Society of America (Hosokawa and Otani, 1997).

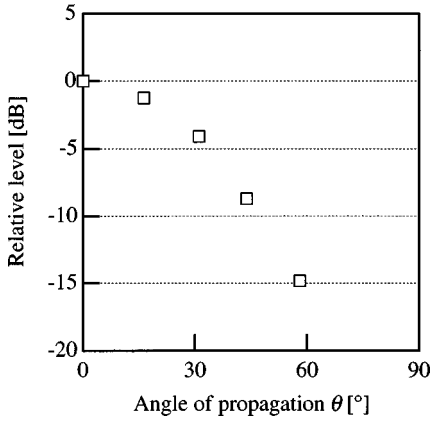


FIG. 10. Change in amplitude at 3 MHz of the slow wave as a function of propagation angle θ to trabecular alignment.

III. DISCUSSION

In our previous paper,[1] we attempted to apply Biot's theory for an isotropic medium to the ultrasonic wave propagation in bovine cancellous bone. By considering the trabecular structure in the direction of propagation, we could then apply Biot's theory for an isotropic medium to determine the propagation speeds of the fast and slow waves along the trabecular alignment. In the present paper, for comparison with our previous results, we attempt to investigate the wave speeds propagating in the perpendicular direction to the trabecular alignment by the same method.

Table I lists the values of the input parameters used in Biot's theory. The Young's modulus $E_s = 20$ GPa of the solid bone (cortical bone) were taken from work of Williams and Johnson.[11] The bulk modulus $K_f = 2.0$ GPa of bone marrow were taken from our own work.[1] The Poisson's ratio of the trabecular frame ν_{b21} (1 and 2 correspond to the x_1 and x_2 direction shown earlier in Fig. 1) was taken from the work of Williams and Johnson[11] and equals 0.23. The value of $r = 0.25$ was chosen by Williams.[12]

The values of n in Eqs. (5) and (6), which depends on the trabecular arrangement in cancellous bone, were estimated to fit the experimental data of the speed of the fast wave as a function of bone volume fraction V_f. This resulted in a value of $n = 2.14$ in case of propagation in the perpendicular direction to the trabecular alignment. For propagation parallel to the trabeculae, a value of $n = 1.46$ was obtained by the same method.[1] The propagation speeds of the fast and slow waves were calculated as a function of bone volume fraction V_f, using Eqs. (1)–(6). [For the viscosity of bone marrow, a value of $\eta = 1.5$ N s/m^2 was taken from the paper of McKelvie and Palmer.[13] Equation (1) is applicable to the propagation speeds in cancellous bone because the value of the viscous skin depth $(2\eta/\rho_f\omega)^{1/2}$ is 0.14 mm at 1 MHz and is much smaller than the pore size a_0 of cancellous bone used in our experiments, which vary from about 0.5 to 1.5 mm.] The calculated results for propagation in the perpendicular direction are shown in Fig. 8, with measured data. The results in the parallel direction are also shown in Fig. 9. We consider that in both Figs. 8 and 9 the measured values correspond well to the calculated values; with due consideration to the heterogeneity of the trabecular structure; although a slight scatter in the measured values is observed.

TABLE I. Input parameters of cancellous bone in Biot's theory.

Young's modulus of solid bone E_s	20 GPa
Density of solid bone ρ_s	1960 kg/m^3
Bulk modulus of bone marrow K_f	2.0 GPa
Density of bone marrow ρ_f	930 kg/m^3
Poisson's ratio of trabecular frame ν_{b21}	0.23
Variable r	0.25

As stated previously, the value of n depends on the direction of loading (or the direction of propagation) to the trabecular arrangement. Gibson derived by analyzing the mechanical behavior of cancellous bone that n has a value of 1 when cancellous bone is loaded along the direction of trabecular alignment, and a value between 2 and 3 in the transverse (or perpendicular) direction.[8] The values of n obtained in our experiments, $n = 2.14$ for propagation in the perpendicular direction to the trabeculae and $n = 1.46$ in the parallel direction, are slightly larger than Gibson's analytical results. This is because the trabecular orientations of the specimens used were partially heterogeneous. Here, the parameter n is the exponent of the power law for the elastic moduli of cancellous bone [see Eqs. (5) and (6)]. Accordingly, Biot's theory for an isotropic medium is applicable to the propagation speeds of both fast and slow waves in cancellous bone with anisotropic structure by considering the elastic moduli in the direction of propagation.

The decreasing amplitude of the slow wave with increasing incident angle θ_i (the propagation angle θ to the trabeculae) could clearly be observed in Fig. 5. Figure 10 shows the change in amplitude at 3 MHz of the slow wave as a function of the propagation angle θ. The propagation of the slow wave, which is related to the motion of the fluid in the pore spaces relative to the solid frame, significantly depends on the structure of the solid frame.[2,3] The fluid motion in the spaces parallel to the trabeculae is much smoother than the motion in the perpendicular spaces. From this observation of the slow wave, it is thus shown that the slow wave in cancellous bone is effectively generated by movement along the trabecular alignment.

IV. CONCLUSIONS

We have experimentally examined the acoustic anisotropy in bovine cancellous bone. It was shown that both fast and slow longitudinal waves generally propagate in cancellous bone. The propagation speed of the fast wave tends to be similar to that of the slow wave when the acoustic wave propagates at high angles to the trabecular alignment. The propagation speeds of the two waves propagating through a cancellous bone with trabeculae parallel and perpendicular direction could be estimated using Biot's theory for an isotropic medium by considering the elastic moduli of cancellous bone. Furthermore, the experimental results showed that the amplitude of the slow wave depends on the trabecular alignment, which means the trabecular structure contributes to the generation of the slow wave.

We conclude that cancellous bone has a strong acoustic anisotropy which is significantly related to the trabecular arrangement.

ACKNOWLEDGMENTS

Part of this work was supported by Grant-in-Aid for Scientific Research (B) from the Ministry of Education, Science, Sports and Culture. We wish to thank Daikin Industries, Inc. for supplying piezoelectric polymer films.

[1] A. Hosokawa and T. Otani, ''Ultrasonic wave propagation in bovine cancellous bone,'' J. Acoust. Soc. Am. **101**, 558–562 (1997).
[2] 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).
[3] 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).
[4] D. L. Johnson, ''Equivalence between fourth sound in liquid He II at low temperatures and the Biot slow wave in consolidated porous media,'' Appl. Phys. Lett. **37**, 1065–1067 (1951).
[5] D. L. Johnson and T. J. Plona, ''Acoustic slow waves and the consolidation transition,'' J. Acoust. Soc. Am. **72**, 556–565 (1982).
[6] M. A. Biot, ''Generalized theory of acoustic propagation in porous dissipative media,'' J. Acoust. Soc. Am. **34**, 1254–1264 (1962).
[7] 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.
[8] J. G. Berryman, ''Confirmation of Biot's theory,'' Appl. Phys. Lett. **37**, 382–384 (1980).
[9] L. J. Gibson, ''The mechanical behaviour of cancellous bone,'' J. Biomech. **18**, 317–328 (1985).
[10] S. B. Lang, ''Ultrasonic method for measuring elastic coefficients of bone and results on fresh and dried bovine bones,'' IEEE Trans. Biomed. Eng. **BME-17**, 101–105 (1970).
[11] J. L. Williams and W. J. H. Johnson, ''Elastic constants of composites formed from PMMA bone cement and anisotropic bovine tibial cancellous bone,'' J. Biomech. **22**, 673–682 (1989).
[12] J. L. Williams, ''Ultrasonic wave propagation in cancellous and cortical bone: Prediction of some experimental results by Biot's theory,'' J. Acoust. Soc. Am. **91**, 1106–1112 (1992).
[13] M. L. McKelvie and S. B. Palmer, ''The interaction of ultrasound with cancellous bone,'' Phys. Med. Biol. **36**, 1331–1340 (1991).

LETTERS TO THE EDITOR

Comments on "Biot model of sound propagation in water-saturated sand" [J. Acoust. Soc. Am. 97, 199–214 (1995)]

R. D. Stoll
Lamont–Doherty Earth Observatory of Columbia University, Palisades, New York 10964

(Received 31 January 1997; accepted for publication 13 June 1997)

Input parameters to the Biot theory, used by N. P. Chotiros [J. Acoust. Soc. Am. **97**, 199–214 (1995)] in an effort to claim that a slow wave propagating at 1200 m/s was experimentally observed in near-bottom sand, are shown to be inconsistent with generally accepted results from prior theoretical and experimental studies of granular media. © *1998 Acoustical Society of America.* [S0001-4966(98)06405-4]

PACS numbers: 43.30.Ma [DLB]

INTRODUCTION

In a paper published in this *Journal* (Chotiros, 1995), the author claims to have experimentally detected a Biot wave of the second kind propagating at a velocity near 1200 m/s in near-bottom sand. He justifies this claim by presenting numerical results obtained using the Biot theory with input parameters which purportedly model the physical conditions of the experiment. However, it appears that the author, in his efforts to obtain an output from the theory that matches the experimental results, has chosen certain parameters that are not justified on the basis of either theoretical considerations or historical experimental results. The purpose of this commentary is to discuss these inconsistencies and their consequences.

I. BIOT PARAMETERS

The Biot parameters used by the author are shown in Table I. In this table the column labeled ''Historical'' summarizes the range of parameters used by several different investigators who gave examples of Biot models for sand confined at low effective stress levels. Their work is described in Refs. 9–13 of the subject paper. In the column labeled ''Chotiros,'' the range of values chosen by the author in an effort to match the data at three different experimental sites is given.

The most striking differences between the historical parameters and the parameters chosen by the author are in the value of the frame bulk modulus, which is at least one order of magnitude larger than the largest historical value ($5.30E+09$ compared to $4.36E+08$ Pa) and the value of the bulk modulus of the grains, which is a factor of 5 smaller than the historical values. When the historical values of the parameters are used in the Biot equations to calculate the wave speeds of compressional waves of the first and second kind (Biot fast and slow waves) the slow wave, velocities never reach values larger than about 400 m/s (see Fig. 12 of the subject paper) which is much lower than the 1200 m/s which the author later attempts to associate with the Biot slow wave. Not surprisingly, when the revised parameters chosen by the author are used in the calculations, the predicted slow wave velocities increase to about 1200 m/s as can be seen from Fig. 14 of the subject paper. Unfortunately, some of the input parameters chosen by the author to force this result are inconsistent with essentially all of the prior theoretical and experimental studies regarding the response of granular media.

As a first example, if the author's value of frame bulk modulus, K_b, of $5.30E+09$ Pa and frame shear modulus, G, of $2.61E+07$ Pa are used to calculate the dilatational wave speed of the dry frame using

$$\nu_p = \sqrt{\frac{K_b + 4/3G}{\rho}}, \tag{1}$$

where ρ is the mass density of the dry frame given by $\rho = (1-\beta)\rho_r$, we obtain a value of 1833 m/s. This is larger than the speed of sound in water and much too large for any granular material confined at low effective stress levels such as those that would exist with only a meter or so of overburden pressure. In fact it is in the range one would expect for weak sandstone or sand subjected to an effective stress caused by several hundred meters of overburden. Moreover, if these values of the bulk and shear moduli are used to evaluate Poisson's ratio of the skeletal frame, a value of 0.498 is obtained whereas typical values of Poisson's ratio

TABLE I. Biot parameters used in subject paper.

		Historical	Chotiros
Bulk parameters			
porosity, β	···	0.36 to 0.47	0.36 to 0.40
grain density, ρ_r	kg/m^3	2650	2650
liquid density	kg/m^3	1000	1000
grain bulk modulus, K_r	Pa	3.60E+10 to 4.00E+10	7.00E+09
liquid bulk modulus, K_f	Pa	2.00E+09 to 2.30E+09	2.25E+09
Liquid motion			
viscosity	kg/m-s	1.00E−03	1.00−03
permeability	m^2	1.00E−10 to 6.49E−12	2.15E−10 to 4.99E−11
pore size	m	1.00E−05 to 5.00E−05	1.09E−04 to 4.99E−05
virtual mass constant	···	1.00 to 1.25	1.750 to 1.889
Frame response			
frame shear modulus, G	Pa	2.61E+07 to 1.19E+08	2.61E+07
shear log decrement	···	0.00 to 0.15	0.15
frame bulk modulus, K_b	Pa	4.36E+07 to 4.36E+08	5.30E09
bulk log decrement	···	0.00 to 0.15	0.075 to 0.15

from experiments where both shear and dilatational wave speeds (or shear and extensional or rod wave speed) are measured as a function of effective confining stress are usually less than about 0.2 for dry granular materials at relatively low confining stress [e.g., see Hardin (1978) and Gardner *et al.* (1964)]. A typical set of experimental data that illustrates the kind of response to be expected in both dry and water saturated sand is shown in Fig. 1. Numerous data sets similar to the one shown in Fig. 1 are scattered throughout the geophysical and geotechnical literature (e.g., see Lambe and Whitman, 1969, Fig. 12.8).

In Fig. 1 the curves for the dry-sediment response show that both the compressional-wave and shear moduli are affected in a similar manner by an increase in confining stress. As the macroscopic confining stress increases, intergranular forces at grain contacts increase with an attendant increase in the size of the minute areas of contact. This changes both the normal and tangential compliance at each contact, and hence, the macroscopic bulk and shear moduli of the assemblage of particles are changed. Theoretical studies of the response of granular media have been carried out by many investigators (e.g., see Deresiewicz, 1958 and Duffy and Mindlin, 1957) and they all predict a small-strain, dynamic response essentially similar to that shown by the experimental data of Fig. 1, with shear and dilatational wave velocity consistent with small values of Poisson's ratio. Minor differences between experiment and theory are due to the fact that it is very difficult to describe mathematically the precise geometry of a random array of particles of irregular shape that accurately describes real materials. However, the important thing to remember is that once the primitive geometric variables such as porosity, gradation and grain shape are specified, the macroscopic moduli of the skeletal frame (shear modulus, Young's modulus, bulk modulus, etc.) are all interrelated in a specific albeit complicated way and they cannot be chosen independently in an arbitrary way if they are to represent a specific granular aggregate. Since all of these moduli depend on *both* the normal and tangential compliance at the intergranular contacts, when either of these compliances is modified for whatever reason, all of the macroscopic moduli will be affected.

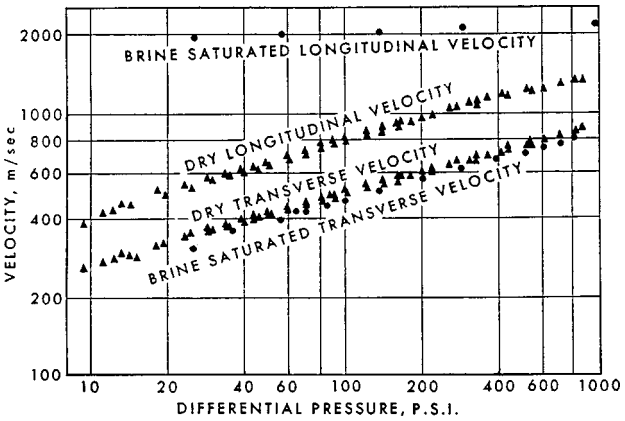


FIG. 1. Velocity of compressional and shear waves versus pressure in dry and brine saturated coarse sand (Hamilton, 1971).

The effects of fluid saturation are also illustrated by the results shown in Fig. 1. In this case, where the fluid saturant is water, there is very little change in the shear wave velocity indicating that the presence of fluid in the void space does not cause significant changes in the interparticle compliances and therefore we should also not expect large changes in the frame bulk modulus in the presence of water. On the other hand, there is a large change in the p-wave velocity because the voids are now filled with a relatively incompressible fluid and so the overall bulk modulus is also much higher. This kind of response is precisely what is predicted by the Biot theory in the general case or by the Gassmann equations (Gassmann, 1951), when fluid motion relative to the skeletal frame is neglected. In the latter case the bulk modulus of the two-component, water-saturated sediment acting as a closed system, is given by

$$K_{\text{Gassmann}}=\frac{(K_r-K_b)^2}{D-K_b}+K_b,$$

where

$$D=K_r(1+\beta(K_r/K_f-1)).$$

When this overall bulk modulus or an equivalent one from the more general Biot formulation is substituted into Eq. (1), it can be seen that the dilatational wave speed of a fully saturated sediment depends on the three bulk moduli K_b, K_r, and K_f as well as the porosity, β, and the mass density of the fluid and grains. In typical, near-bottom sediments with no cementation between grains, the dilatational wave velocity varies from values slightly less than water velocity to values in the range of 1600 to 1750 m/s depending largely on the porosity and on the compressibility of the fluid, with the shearing stiffness and compressibility of the skeletal frame playing a subordinate role. Moreover, in granular materials such as sand, it is impossible to accurately measure *in situ* porosity so that measurements of dilatational wave velocity alone generally do not lead to viable estimates of any of the input parameters. However in the present case, the author claims to have chosen K_b by back-figuring from a dilatational wave speed ''in the region of 1700 m/s'' and yet, in describing the experimental work, he estimates experimental variability to be as high as ± 300 m/s and makes no comment as to how porosities used in the calculations were determined and what variability might be expected. Finally it is easy to show that any number of different combinations of parameters will result in fast wave velocities in the region of 1700 m/s. In fact this is clearly illustrated by the author's own Fig. 12 where the computed velocities of the fast wave for all of the historical parameter sets fall within the right range without any resort to anomalously high values of frame modulus. Hence the choice of K_b resulting in a slow wave speed of 1200 m/s has the consequence of producing a model which has no physical meaning.

One additional point should be made regarding the experimental data presented by the author. Calculations of the amplitude of reflected and refracted waves for a p-wave incident to a sandy sea floor using the Biot theory (Stoll, 1977) show that the amount of energy partitioned into Biot waves of the second kind is very small (e.g., calculations for the sand model in Ref. 9 of the subject paper (Stoll and Kan, 1981) but never included in the final manuscript, indicate refraction amplitudes of the order of -40 dB relative to the displacement amplitude of the incident wave at normal incidence and even less at smaller grazing angles at a frequency of 1 kHz). This taken together with the fact that attenuation of slow waves is predicted by the theory to be very high, suggests that the amplitude of any slow waves reaching a sensor, even though it is only a few meters from the source, will be too small for reliable measurement given typical noise levels that would be produced by the prior passage of faster waves of the first kind. Hence the strong arrival that the author is trying to identify as a slow wave is further at odds with the predictions of the theory.

II. SUMMARY AND CONCLUSIONS

The Biot theory which describes the response of a porous medium saturated with a fluid is a phenomenological theory that correctly predicts the overall response of the medium based on the known response of the two separate components. It does not provide a direct link from microstructure to macroscopic behavior but rather requires the user to input a consistent and physically realistic phenomenological model for both the porous medium and the fluid as well as certain coupling ''factors'' that are inherently difficult to quantify. This puts a heavy burden on the user of the theory in that many of the thirteen or more input parameters are interrelated in a specific way and cannot be chosen independently without regard to the physics of deformation of the porous component, as appears to have been done in the present case.

It is the opinion of the writer that the author of the subject paper is not justified in his claim that a Biot wave of the second kind, with a wave speed of 1200 m/s, has been detected in a near-bottom sand deposit. This opinion is based on the fact that the author has used an unacceptable value for the bulk modulus of the skeletal frame and a questionable value for the bulk modulus of the individual grains. The very high value of frame modulus is completely incompatible with the shear modulus adopted and not justifiable on the basis of any accepted theory or experimental measurement. In addition, the grain bulk modulus obtained by the author experimentally (Molis and Chotiros, 1992) is significantly lower than all prior published measurements for typical polycrystalline materials obtained dynamically at low strain levels consistent with those in the experiments being discussed. Thus, while the values chosen for these two parameters force the Biot equations to predict a speed of 1200 m/s for the slow wave, they certainly are not representative of the material being modeled and therefore the calculated speed cannot be used to substantiate the author's interpretation of field experiments.

Chotiros, N. P. (**1995**). ''Biot model of sound propagation in water-saturated sand,'' J. Acoust. Soc. Am. **97**, 199–214.

Deresiewicz, H. (**1958**). ''Mechanics of granular matter,'' in *Advances in Applied Mechanics* (Academic, New York), Vol. 5.

Duffy, J., and Mindlin, R. D. (**1957**). ''Stress-strain relations and vibrations of a granular medium,'' J. Appl. Mech. **79**, 585–593.

Gardner, G. H. F., Wyllie, M. R. J., and Drochak, D. M. (**1964**). ''Effects of pressure and fluid saturation on the attenuation of elastic waves in sands,'' J. Pet. Technol. **16**, 189–198 (February).

Gassmann, F. (**1951**). ''Uber die Elastizitat Poroser Medien,'' Vierteljahrschr. Naturforch. Ges., Zurich **96**, 1–23.

Hamilton, E. L. (**1971**). ''Elastic properties of marine sediments,'' J. Geophys. Res. **76**, 579–605.

Hardin, B. O. (**1978**). ''The nature of stress-strain behavior for soils,'' in *Earthquake Engineering and Soil Dynamics*, Proc. ASCE Specialty Conference, Vol. 1 (Am. Soc. of Civil Engineers, New York), pp. 3–90.

Lambe, T. W., and Whitman, R. V. (**1969**). *Soil Mechanics* (Wiley, New York).

Molis, J. C., and Chotiros, N. P. (**1992**). ''A measurement of grain bulk modulus of sands,'' J. Acoust. Soc. Am. **91**, 2463A.

Stoll, R. D. (**1977**). ''Acoustic waves in ocean sediments,'' Geophysics **42**, 715–725.

Stoll, R. D., and Kan, T. K. (**1981**). ''Reflection of acoustic waves at a water-sediment interface,'' J. Acoust. Soc. Am. **70**, 149–156.

Response to: "Comments on 'Biot model of sound propagation in water-saturated sand'" [J. Acoust. Soc. Am. 103, 2723–2725 (1998)]

N. P. Chotiros
Applied Research Laboratories, The University of Texas at Austin, Austin, Texas 78713-8029

(Received 7 July 1997; accepted for publication 24 November 1997)

The approach advocated by Prof. Stoll, which lead to the parameter values marked ''Historical,'' has worked well for porous materials with cemented frames, such as fused glass beads and sandstone, but it does not work for sand. The parameter values, marked ''Chotiros,'' are uniquely able to match all four. The main issue lies with the values of the frame and grain moduli. © *1998 Acoustical Society of America.* [S0001-4966(98)06505-9]

PACS numbers: 43.30.Ma [DLB]

INTRODUCTION

The approach advocated in by Prof. Stoll,[1] which lead to the parameter values marked ''Historical,'' has worked well for porous materials with cemented frames, such as fused glass beads and sandstone, but it does not work for sand. Out of four measurable parameters, namely fast and slow compressional wave speeds, shear wave speed, and normal incidence reflection loss, the ''Historical'' parameter set is only able to match two, i.e., the fast compressional and shear wave speeds. It is unable to match measured values of either the slow compressional wave speed or the reflection loss. The parameter values, marked ''Chotiros,'' are uniquely able to match all four. The three wave speeds were addressed in the subject paper[2] and the reflection loss in a later publication.[3] The main issue lies with the values of the frame and grain moduli.

I. FRAME MODULI

The first part of the comments in Ref. 1 concerns the frame moduli, and it is based on the assumption that the bulk and shear moduli of the solid frame in the saturated condition must be the same as that of the dry condition. For unconsolidated sand the assumption is invalid, as shown below using the experimental data from Hamilton,[4] which is reproduced in Fig. 1 of Ref. 1. The Hamilton data contain the compressional (fast) and shear wave speeds of a sand specimen as a function of differential pressure, in both the dry and saturated conditions, which is a very rare and valuable combination of measurements. The differential pressure is defined as the external pressure applied to the grains, measured relative to the pressure of the pore fluid. The specimen was a sand of 0.7-mm mean grain diameter.

The shear wave speeds of the dry and saturated specimens may be inverted to give the frame shear modulus as a function of pressure. The shear wave equation[5] for a Biot medium is a simple equation which gives a unique solution for the shear modulus. The result is shown in Fig. 1. At high pressures, the shear modulus values for both the dry and saturated conditions are almost identical, but at low pressures, the difference is noticeable. The difference may be modeled as a negative pressure offset. It appears that saturation has reduced the value of the shear modulus by an amount equivalent to a constant reduction in pressure. (On a log–log plot, a constant offset produces a nonlinear shift that is more noticeable at the low end of the scale.) The cause is the capillary effect, which has a tendency to maximize the contact area between solid and fluid. In order to gain access to all unexposed solid surfaces, the fluid tends to force the solid grains apart, producing the equivalent of a negative pressure offset, giving a noticeable difference in shear modulus values at low pressures.

For the frame bulk modulus, the difference between the dry and saturated conditions is much bigger. Due to the surface roughness of sand grains, the grain–grain contact in the saturated condition may be characterized as a narrow fluid-filled gap bridged by a few solid contacts, as illustrated in Fig. 2. For a qualitative impression of sand grain surface roughness, the micrographs by Mack and Leistikow[6] are an excellent source. For modeling purposes, the grain–grain contact may be idealized as an equivalent fluid-filled gap of constant width, bridged by a single equivalent solid connection. In the dry condition, both compressional and shear stresses are channeled exclusively through the solid connection. In the saturated condition, the shear stress is still carried by the solid connection because the fluid is unable to support shear stress, but the compressional stress may be carried through both solid and fluid connections. It can be shown, as outlined in the Appendix, that there are two asymptotic conditions: a low-frequency regime in which the fluid in the gap can freely flow, and a high-frequency regime where the fluid is trapped and forced to support the compressional stress between the grains. The result is a stiffening of the grain–grain contact, analogous to the ''local flow'' model[7] for fluid saturated cracks in rocks. It is likely that the area of the fluid film is much greater than that of the solid connection and the result is a frame whose bulk modulus is very much larger than, and unconnected to, its shear modulus, and whose Poisson's ratio is very close to that of a liquid.

In support of the above argument, the compressional wave speeds shown in Fig. 1 of Ref. 1 were inverted to give the frame bulk modulus as a function of pressure, using Stoll's formulation[8] of Biot's theory and the previously obtained curve of frame shear modulus. The inversion was

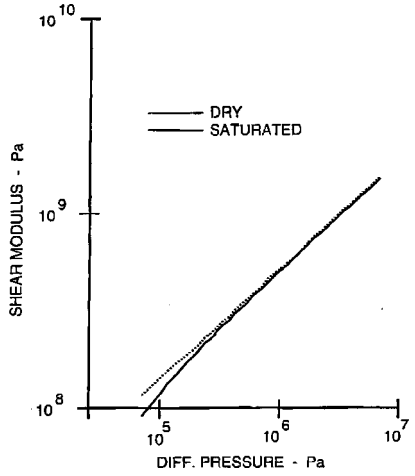


FIG. 1. Frame shear modulus as a function of differential pressure.

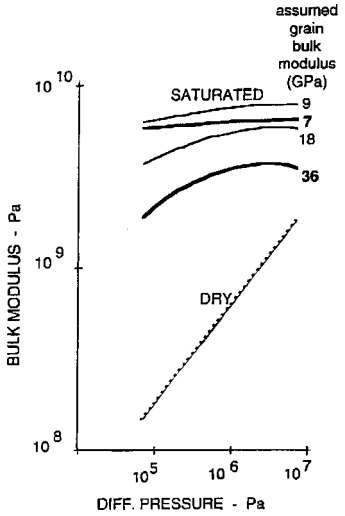


FIG. 3. Frame bulk modulus as a function of differential pressure.

done numerically using Newton's method. Although, theoretically, the equation gives several solutions for the frame bulk modulus, in practice only one solution is viable, and the convergence was rapid. The results are shown in Fig. 3. In the saturated case, since there is a strong dependence on the grain bulk modulus, the results for a number of assumed values of the grain bulk modulus are shown. It is noted that by using the ''Historical'' value of grain bulk modulus (36 GPa) the saturated and dry frame bulk moduli obtained are quite different, thus invalidating the assumption of equal moduli. The curve has the unusual behavior of increasing and then decreasing as pressure is increased; the bulk modulus of an elastic structure generally increases monotonically as a function of increasing stress.

When the ''Chotiros'' value of the grain bulk modulus (7 GPa) is used, the frame bulk modulus values are considerably greater than that of the dry condition, and increase monotonically with pressure. By applying these values to the grain–grain contact model illustrated in Fig. 2, the average fluid film thickness (a) was estimated to vary between 10 and 3 μ, and the area ratio between the solid and fluid connections (b^2/c^2) between 10 000 and 1000, over the indicated pressure range. This is an ongoing subject of research and a paper, with quantitative details, under preparation.

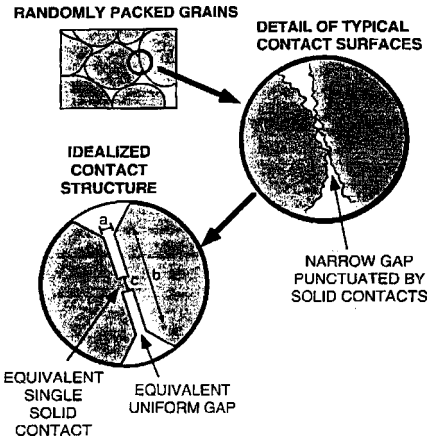


FIG. 2. Model of fluid-reinforced frame.

II. GRAIN BULK MODULUS

The second part of the comments in Ref. 1 concerns the grain bulk modulus. The ''Historical'' value of the grain bulk modulus is the one obtained from the sound speed in a solid quartz crystal. There are two main reasons why the operative value of grain bulk modulus is lower than the ''Historical'' value.

(a) Historically, it has been assumed that the bulk modulus of the materials that make up the frame is the same as that of the solid grains. In the foregoing, it was shown that a proportion of the pore fluid becomes an integral part of the frame, and hence of the grains as well. This has consequences on the operative value of grain bulk modulus. For example, at a porosity of 40%, if 10% of the pore fluid were counted as part of the frame, then the operative value of grain bulk modulus would go down from 36 to 18 GPa.

(b) So far, the discussion has been confined to a version of Biot's theory that was formulated by Stoll, but now it is necessary to go back to Biot's original equations. In Biot's original equations, the grain bulk modulus was not used directly. In its place were two other terms, the ''unjacketed bulk modulus'' and the ''coefficient of fluid content.'' The unjacketed bulk modulus is defined as the change in pore fluid pressure per unit change in the volume contained within the outer envelope of the frame, as measured in an unjacketed compressibility test.[9] The coefficient of fluid content is defined as the volume of fluid absorbed by a porous solid, per unit volume, per unit pore pressure increase, the inverse of which also has the dimensions of bulk modulus. Thus, there are three types of bulk moduli associated with the frame material. Biot and Willis[9] noted that if the porosity remains constant, independent of pore pressure, then the following relationship is applicable,

$$\gamma = f(c - \delta), \tag{1}$$

where γ is the coefficient of fluid content, f porosity, c compressibility of pore fluid, and δ compressibility of the solid material. In Stoll's formulation, this relationship is implicitly assumed, allowing the coefficient of fluid content to be eliminated entirely as an independent variable and the grain

bulk modulus to be substituted for the unjacketed bulk modulus.

The measurements by Molis and Chotiros[10] were also based on the above equation. The volume of fluid leaving a pressure chamber containing a mixture of sand and fluid was measured as a function of pore pressure, giving the coefficient of fluid content γ. Then, knowing the porosity f and fluid compressibility c, the grain bulk modulus δ was calculated using the above equation, and the values obtained were quite different from the ''Historical'' value.

It is quite possible that Eq. (1) may not be valid, as discussed by Biot and Willis. In this case, what was measured by Molis and Chotiros could no longer be equated to the grain bulk modulus. Deviations from Eq. (1) may take several forms, but the simplest is where an extra term δ_p is added. This term represents the tendency of the frame to absorb pore fluid as a function of pressure. The resulting equation is

$$\gamma = f(c - \delta - \delta_p). \tag{2}$$

A positive value of δ_p would indicate a situation where porosity increases with pore pressure. What was measured by Molis and Chotiros is no longer $1/\delta$, the grain bulk modulus, but the quantity $1/(\delta + \delta_p)$. Within the frame work of Stoll's formulation, the term δ_p is not permitted, but its effects may be simulated by a further reduction in the operative value of the grain bulk modulus.

It is very plausible that the contributions from both (a) and (b) caused the operative value of the grain bulk modulus to go from the ''Historical'' value to the ''Chotiros'' value, but more research is still needed for verification.

III. SUMMARY

The problem with the frame moduli values has been traced to the assumption that the frame moduli of the dry and saturated conditions must be the same. This assumption is not applicable to unconsolidated sand, as demonstrated by inversion of the data set in Fig. 1 of Ref. 1 for the frame moduli values; the frame moduli values obtained for the dry and the saturated conditions are quite different. A new submodel for the frame is under development, which recognizes the effect of restricted local fluid flow within each grain–grain contact in reinforcing the frame and producing an increased bulk modulus in the saturated condition. The shear modulus is also affected by the presence of the pore fluid but to a much smaller degree.

The problem with the grain bulk modulus value may be traced to two assumptions. The first is an implicit assumption that the bulk modulus of the frame material is equal to that of the solid grains. In reality, the frame is fluid-reinforced, which makes a small part of the pore fluid an integral part of the frame, resulting in a reduction in the net bulk modulus of the frame material relative to a purely solid frame, and hence an apparent reduction in the operative value of the grain bulk modulus. The second is an implicit assumption in Stoll's formulation of Biot's theory that porosity is independent of pore pressure. Biot's original equations used two parameters to represent the bulk properties of the frame material, one of which can allow for a change in porosity as a function of pore fluid pressure. This has two consequences: (a) what was measured by Molis and Chotiros can no longer be equated to the grain bulk modulus, as Hickey and Sabatier[11] correctly deduced; (b) postulating a porous solid in which porosity increases with pore fluid pressure, its effect, in the context of Stoll's formulation of Biot's theory, may be simulated by a further reduction in the operative value of the grain bulk modulus. These findings are consistent with those of Hickey and Sabatier.[11]

In summary, the ''Historical'' parameter set is incompatible with acoustic measurements, particularly the slow wave speed and attenuation and the normal incidence reflection loss. A new parameter set, that uniquely matches experimental measurements, referred to as the ''Chotiros'' parameter set, was found and published in the subject paper.[2] The new values of grain and frame moduli were controversial, as articulated in Ref. 1. Since then, progress has been made. Certain assumptions, that had been taken for granted in connection with Biot's theory, are found to be less than universal in their validity, but Biot's theory remains valid in its original form. Continuing within the context of Biot's theory, a new submodel for the frame, which includes a fluid-reinforced frame, and a revival of the coefficient of fluid content, is expected to resolve the controversy.

The underlying cause of the controversy lies in the definition of the boundary between frame and pore fluid. The division is not as simple as it seems. In Biot's equations, the pore fluid is regarded as a connected body of fluid that can move relative to the frame. Any part of the pore fluid that remains stationary relative to the frame and functions as an integral part of the frame must be counted as a component of the frame. Consequently, the operative grain and frame moduli are determined by the solid grains plus any part of the fluid that is held stationary relative to the frame. In some porous structures, all of the fluid is mobile, and this problem does not arise. In Gassmann's theory, this problem is unknown because no relative motion between frame and pore fluid is allowed.

ACKNOWLEDGMENTS

I thank Prof. Stoll for his interest in the problem and for the opportunity to discuss these very interesting issues. I thank C. J. Hickey and J. M. Sabatier, of the National Center for Physical Acoustics, for bringing the coefficient of fluid content to my attention, and for many useful discussions. This work is supported by the office of Naval Research, Code 321 OA.

APPENDIX: FLUID FILM BETWEEN TWO GRAINS

Consider a circular fluid film of thickness h, and radius r sandwiched in the contact area between two grains of sand. The passage of a sinusoidal compressional wave, at a frequency f, causes the distance between the grains, hence the film thickness h, to oscillate. A fraction x of the fluid will surge in and out of the contact region. In addition, the fluid may undergo a volumetric strain y. The lateral forces due to

the inertial reaction, u, and the viscous drag, v, associated with the moving part must balance the lateral compressional force w. Thus

$$|u|^2+|v|^2=|w|^2. \tag{A1}$$

Order of magnitude estimates of u, v, and w are approximately given by

$$|u|\approx x(2\pi f)^2\rho h\pi r^3, \tag{A2}$$

$$|v|\approx x(2\pi f)\eta\pi r^3/h, \tag{A3}$$

$$|w|\approx yK\pi rh, \tag{A4}$$

where ρ, η, and K, are density, viscosity, and bulk modulus of the fluid, respectively. Substituting from (A2), (A3) and (A4) into (A1), a ratio between x and y is obtained;

$$y^2/x^2\approx 4\pi^2r^4h^{-4}f^2K^{-2}(4\pi^2h^4\rho^2f^2+\eta^2). \tag{A5}$$

It is evident that as f tends to zero, y will be small compared to x, indicating that the fluid film will move freely with negligible volumetric strain. On the other hand, as f tends to infinity, x will be small compared to y, indicating negligible fluid flow and maximum volumetric strain, resulting in a compressional stress within the fluid that tends to resist changes in h, hence reducing the compressibility of the grain–grain contact.

[1] R. D. Stoll, ''Comments on 'Biot model of sound propagation in water-saturated sand' [J. Acoust. Soc. Am. **97**, 199–214 (1995)],'' J. Acoust. Soc. Am. **103**, 2723–2725 (1998).

[2] N. P. Chotiros, ''Biot model of sound propagation in water-saturated sand,'' J. Acoust. Soc. Am. **97**, 199–214 (1995).

[3] N. P. Chotiros, ''Inversion and sandy ocean sediments,'' in *Full Field Inversion Methods in Ocean and Seismic Acoustics* (Kluwer, Dordrecht, 1995), pp. 353–358.

[4] E. L. Hamilton, ''Elastic Properties of Marine Sediments,'' J. Geophys. Res. **76**, 579–605 (1971), see Fig. 8.

[5] R. D. Stoll and T. K. Kan, ''Reflection of Acoustic Waves at a Water-Sediment Interface,'' J. Acoust. Soc. Am. **70**, 149–156 (1981), Eq. (17).

[6] W. N. Mack and E. A. Leistikow, ''Sands of the World,'' Sci. Am. (Int. Ed.) **270**, 62–67 (August 1996).

[7] T. Mukerji and G. Mavko, ''Pore fluid effects on seismic velocity in anisotropic rocks,'' Geophysics **59**(2), 233–244 (1994).

[8] R. D. Stoll and T. K. Kan, ''Reflection of Acoustic Waves at a Water–Sediment Interface,'' J. Acoust. Soc. Am. **70**, 149–156 (1981), Eq. (15).

[9] M. A. Biot and D. G. Willis, ''The Elastic Coefficients of the Theory of Consolidation,'' J. Appl. Mech. **24**, 594–601 (1957), Eq. (24).

[10] J. C. Molis and N. P. Chotiros, ''A measurement of grain bulk modulus of sands,'' J. Acoust. Soc. Am. **91**, 2463(A) (1992).

[11] C. J. Hickey and J. M. Sabatier, ''Choosing Biot parameters for modeling water-saturated sand,'' J. Acoust. Soc. Am. **102**, 1480–1484 (1997).

Ultrasonic surface waves above rectangular-groove gratings

Luc Kelders
Laboratorium voor Akoestiek en Thermische Fysica, Katholieke Universiteit Leuven, Departement Natuurkunde, Celestijnenlaan 200D, 3001 Leuven, Belgium

Jean F. Allard
Institut d'Acoustique et de Mécanique de l'Université du Maine, UMR CNRS 6613, avenue Olivier Messiaen, 72085 Le Mans Cedex 9, France

Walter Lauriks
Laboratorium voor Akoestiek en Thermische Fysica, Katholieke Universiteit Leuven, Departement Natuurkunde, Celestijnenlaan 200D, 3001 Leuven, Belgium

(Received 10 July 1997; accepted for publication 29 January 1998)

A modal model, initially developed to modelize diffraction of electromagnetic waves by rectangular-groove gratings, is used to describe ultrasonic surface waves above the same structures in air. A simple analytical formulation which provides results comparable to the modal method is presented. Measurements performed on ultrasonic surface waves are compared to predictions obtained with the simplified formulation and the modal model. © *1998 Acoustical Society of America.* [S0001-4966(98)03105-1]

PACS numbers: 43.35.Pt [HEB]

INTRODUCTION

A rectangular-groove grating is represented in Fig. 1, with a surface wave propagating in direction x, parallel to the surface of the structure, in air. The wave is purely evanescent in the direction z perpendicular to the surface. The losses are neglected inside the slits and in the free air. Acoustic[1–3] and electromagnetic[4–8] waves above these structures have been studied for a long time; many more works have been performed in the electromagnetic domain. The reflection of acoustic waves, and the behavior of surface waves, can be described by the same formalism as for the transverse magnetic (TM) waves. This property was known by Ivanov-Shits and Rozhin,[1] who used the work by Vajnshtejn,[9] but these authors were not very explicit in the description of their modeling. After Ivanov-Shits and Rozhin, in the domain of audible acoustics, the models developed for electromagnetic waves were ignored, and oversimplified models were used.

The equations that describe the acoustic field and a TM field if the structure is perfectly conducting are the same, provided the right equivalent quantities are substituted. The magnetic field has a single component H_y different from zero, and is perpendicular to the incidence plane xz of Fig. 1. The electric field lies in the xz plane and is defined, with the time dependence $\exp(-j\omega t)$, by

$$E_x = \frac{-j}{Zk}\frac{\partial H_y}{\partial z}, \tag{1}$$

$$E_z = \frac{j}{Zk}\frac{\partial H_y}{\partial x}, \tag{2}$$

where the characteristic impedance Z and the wave number k are given by

$$Z = (\epsilon_0/\mu_0)^{1/2}, \tag{3}$$

$$k = \omega(\mu_0\epsilon_0)^{1/2}. \tag{4}$$

In these equations, μ_0 and ϵ_0 are the vacuum magnetic and electric permeability. The boundary conditions are $\partial H_y/\partial x = 0$ on the plates, and $\partial H_y/\partial z = 0$ at the bottom of the slits, the electric field being normal to the conducting surface.

The description can be transposed to acoustics by replacing H_y by the pressure p, and E_x, E_z by the velocity components v_z, and $-v_x$, respectively. Equations (3) and (4) must be replaced by

$$Z = (\rho K)^{1/2}, \tag{5}$$

$$k = \omega(\rho/K)^{1/2}, \tag{6}$$

where ρ is the density and K is the bulk modulus of free air ($K = \gamma P$ where P is the atmospheric static pressure and γ the ratio of the specific heats). With this transposition, Eqs. (1) and (2) become

$$v_z = -\frac{j}{\omega\rho}\frac{\partial p}{\partial z}, \tag{7}$$

$$v_x = -\frac{j}{\omega\rho}\frac{\partial p}{\partial x}, \tag{8}$$

and the boundary conditions become $v_x = 0$ on the plates, and $v_z = 0$ at the bottom of the slits, which are the appropriate equations for the acoustic problem.

Simple expressions for the wave number components k_x and k_z that characterize the surface wave were obtained from the models by Hurd[4] and Vajnshtejn,[5] and can be transposed to acoustics. If the distance a between the blades is sufficiently small, k_z can be written[10]

$$k_z = jk\,\tan k\left(h - \frac{\log 2}{\pi}\,a\right), \tag{9}$$

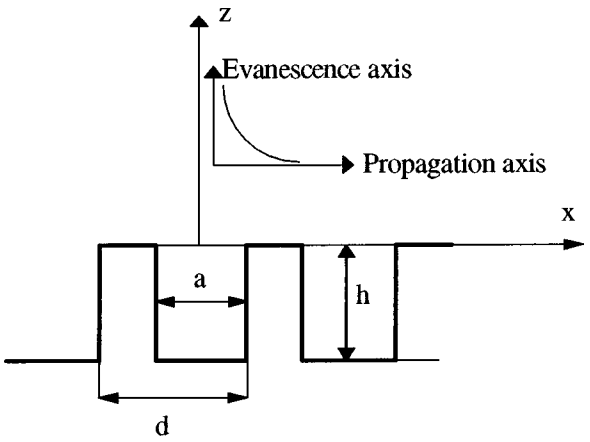


FIG. 1. A rectangular-groove grating. A surface wave that propagates in the x direction is represented over the structure.

$$k_x=(k^2-k_z^2)^{1/2}=k\left[1+\tan^2\ k\left(h-\frac{\log 2}{\pi}\,a\right)\right]^{1/2}. \tag{10}$$

This model is valid only for infinitely thin plates. A simple modal model, developed by Hessel *et al.*[6] in the context of electromagnetic waves, can be used for plates of finite thickness.

I. MODAL MODEL

The model was developed by Hessel *et al.*[6] to predict optical properties of metallic rectangular-groove gratings. Only a brief description of the model, transposed for acoustic waves, is presented. An incident acoustic wave that impinges upon the structure under an angle of incidence θ in the incidence plane xz (see Fig. 1) is related to a reflected wave and an infinite set of space harmonics scattered by the structure. Let $\hat{v}_i$, $\hat{v}_r$ and $\hat{v}_m$ the z components of the velocity field for these waves in the plane $z=0$. They can be written

$$\hat{v}_i=\hat{V}_i\hat{e}_i, \tag{11}$$

$$\hat{v}_r=\hat{V}_r\hat{e}_r, \tag{12}$$

$$\hat{v}_m=\hat{V}_m\hat{e}_m, \tag{13}$$

where $\hat{e}_i$, $\hat{e}_r$, and $\hat{e}_m$, are given by

$$\hat{e}_i=d^{-1/2}\exp(jkx\ \sin\ \theta), \tag{14}$$

$$\hat{e}_r=d^{-1/2}\exp(jkx\ \sin\ \theta), \tag{15}$$

$$\hat{e}_m=d^{-1/2}\exp\left[j\left(k\ \sin\ \theta+\frac{2\pi m}{d}\right)x\right]. \tag{16}$$

The pressure fields at $z=0$ for these waves are given by

$$\hat{p}_i=\hat{Z}_i\hat{v}_i, \tag{17}$$

$$\hat{p}_r=\hat{Z}_r\hat{v}_r, \tag{18}$$

$$\hat{p}_m=\hat{Z}_m\hat{v}_m, \tag{19}$$

where

$$\hat{Z}_i=\frac{jZ}{(\sin^2\ \theta-1)^{1/2}}, \tag{20}$$

$$\hat{Z}_r=-\hat{Z}_i, \tag{21}$$

$$\hat{Z}_m=\frac{Zk}{\hat{k}_m}. \tag{22}$$

The z wave number component $\hat{k}_m$ is given by

$$\hat{k}_m=j\left[\left(k\ \sin\ \theta+\frac{2\pi m}{d}\right)^2-k^2\right]^{1/2}. \tag{23}$$

In a slit, the acoustic field is a superposition of the different modes with z velocity field components at $z=0$ given by

$$v_0=V_0e_0, \tag{24}$$

$$v_n=V_ne_n \quad n\neq 0, \tag{25}$$

where

$$e_0=a^{-1/2}, \tag{26}$$

$$e_n=\left(\frac{2}{a}\right)^{1/2}\cos\left[\left(\frac{n\pi}{a}\right)\left(x+\frac{a}{2}\right)\right] \tag{27}$$

and the related pressure fields are given by

$$p_0=Z_0v_0, \tag{28}$$

$$p_n=Z_nv_n \quad n>0, \tag{29}$$

where

$$\begin{aligned} Z_0&=-jZ\ \cot(kh)\\ Z_n&=-jZ\,\frac{k}{k_n}\cot(k_nh) \quad n>0. \end{aligned} \tag{30}$$

The z wave number vector component k_n is given by

$$k_n=\pm j\left(\frac{n^2\pi^2}{a^2}-k^2\right)^{1/2} \quad n\neq 0. \tag{31}$$

In other slits, the field is identical, except for a phase shift $kd\ \sin\ \theta$ per cell. The continuity conditions in the $z=0$ plane for the z component of the velocity are

$$\begin{aligned} \hat{v}_i+\hat{v}_r+\sum_{m\neq 0}\hat{v}_m&=v_0+\sum_{n>0}v_n, \quad |x|<\frac{a}{2},\\ &=0, \quad \frac{d}{2}>|x|>\frac{a}{2}. \end{aligned} \tag{32}$$

For the pressure, within the aperture, the continuity equation is

$$\hat{p}_i+\hat{p}_r+\sum_{m\neq 0}\hat{p}_m=\sum_{n\geqslant 0}p_n, \quad |x|<\frac{a}{2}. \tag{33}$$

Equations (32)–(33) can be replaced by the set of linear equations

$$\begin{aligned} \sum_{n'\geqslant 0}V_{n'}\Big[&\hat{Z}_r\langle e_n,\hat{e}_r\rangle\langle e_{n'},\hat{e}_r^*\rangle\\ &+\sum_{m\neq 0}\hat{Z}_m\langle e_n,\hat{e}_m\rangle\langle e_{n'},\hat{e}_m^*\rangle-\delta_n^{n'}Z_n\Big]=-2\hat{Z}_i\hat{V}_i\langle e_n,\hat{e}_i\rangle \end{aligned} \tag{34}$$

where $\langle f,g\rangle$ is defined as the integral

$$\langle f,g\rangle = \int_{-a/2}^{a/2} f\cdot g\,dx. \tag{35}$$

Calculations have been performed with a truncated set of equations, by using 11 waveguide modes in the slit ($n \leqslant 10$) and 51 space harmonics ($-25 \leqslant m \leqslant 25$). The system of equations (34) can be rewritten

$$\sum_{n'} A_{nn'} V_{n'} = B_n, \quad n = 0,...,10. \tag{36}$$

Surface waves correspond to a singular system, because they are related to solutions with nonzero amplitude of Eqs. (36) without incident waves. The acoustic field above the structure consists in $\hat{v}_r$, and the modes $\hat{v}_m$. All these waves are evanescent in the z direction, but when $kd \ll 1$, the space harmonics are much more damped than $\hat{v}_r$ which is the dominant surface wave, with the x wave number component $k \sin\theta$ such as $\det[A]=0$, and $\sin\theta$ real >1 when losses are neglected. A simple program has been carried out to localize the zeros of $\det[A]$. It has been verified that with smaller bases for the space harmonics and the waveguide modes, the location of the zeros of $\det[A]$ was not noticeably modified for the different sets (a,b,h) considered.

II. A SIMPLE MODEL

If the contribution of the space harmonics for $m \neq 0$ and of the waveguide modes for $n > 0$ are not taken into account in Eqs. (34), the surface wave is related to

$$\hat{Z}_r \langle e_0, \hat{e}_r\rangle \langle e_0, \hat{e}_r^*\rangle + jZ\cot(kh) = 0, \tag{37}$$

which can be rewritten

$$-j\,\frac{Z}{(\sin^2\theta - 1)^{1/2}}\,\frac{4\sin^2\left(k\,\frac{a}{2}\sin\theta\right)}{ad(k\sin\theta)^2} + jZ\cot(kh) = 0. \tag{38}$$

For $ak \ll 1$, k_z and k_x are given by

$$k_z = jk\,\frac{a}{d}\tan(kh), \tag{39}$$

$$k_x = k\left(1 + \frac{a^2}{d^2}\tan^2 kh\right)^{1/2}. \tag{40}$$

When $a=d$, Eqs. (39) and (40) are very similar to Eqs. (9) and (10), h in Eqs. (9) and (10) being replaced by $h - a(\log 2)/\pi$. When porosity a/d is different from 1, it can be supposed that Eqs. (9) and (10) with an added dependence on porosity as in Eqs. (39) and (40) provide correct predictions for $k_x = k\cos\theta$ and $k_z = k\sin\theta$ from

$$k_z = jk\,\frac{a}{d}\tan k(h - a(\log 2)/\pi), \tag{41}$$

$$k_x = k\left(1 + \frac{a^2}{d^2}\tan^2 k(h - a(\log 2)/\pi)\right)^{1/2}. \tag{42}$$

In order to show the degree of validity of this hypothesis, $\sin\theta$ calculated with the modal model and with Eq. (42) is represented in Fig. 2 for a porosity a/d varying from 1 to 0.1, $d=1.1$ mm, $h=0.7$ mm, at 40 kHz. It appears that Eq. (42) provides a good order of magnitude for $\sin\theta$.

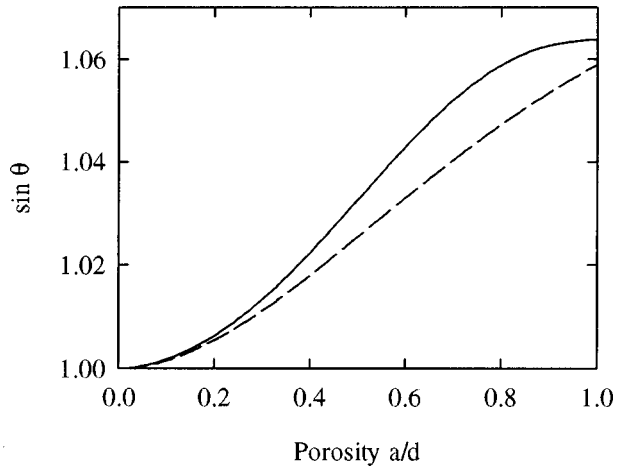


FIG. 2. Evaluation of $\sin\theta$ for $a=1.1$ mm, $h=0.7$ mm, at 40 kHz, from Eq. (42) — — —and from the modal model ———.

III. ILLUSTRATION

Measurements of the velocity of ultrasonic waves have been performed above a structure with $a=1$ mm, $d=1.1$ mm, $h=0.7$ mm from 20 kHz to 50 kHz. Velocity has been measured with ultrasonic bursts, and from interference patterns of standing waves, like in the pioneering work by Ivanov-Shits and Rozhin[1] at audible frequencies. More details on the experimental set up and the measurement technique can be found in Ref. 11. The velocity has been converted in equivalent $\sin\theta$, compared in Fig. 3 with the predictions obtained from the modal model and from Eq. (42). The agreement between experiment and predictions is good. Damping effects due to viscous forces and thermal

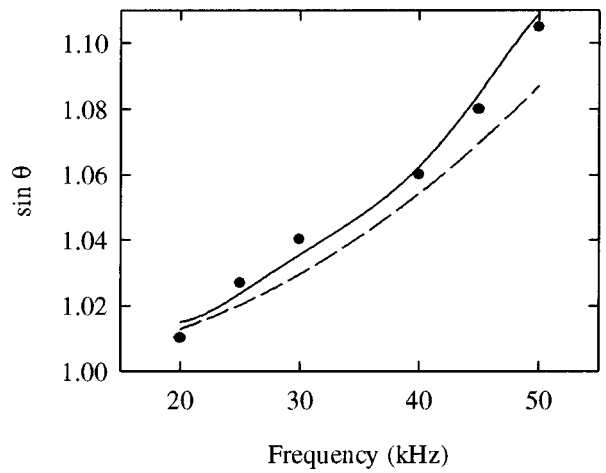


FIG. 3. Measured and predicted $\sin\theta$ for a $a=1.1$ mm, $d=1$ mm, $h=0.7$ mm. Measurement ●●●●●●●. Prediction from Eq. (42) — — —. Prediction from the modal model ———.

exchanges between air and the structure, which have not been taken into account in the calculations, can be neglected. It may be noticed that the space harmonics have not been observed, due to the strong damping in the z direction for these waves. For instance, for $m=1$, at 40 kHz, $k_z = j6.5 \times 10^3\ \mathrm{m}^{-1}$.

IV. CONCLUSION

A precise description of ultrasonic waves above rectangular-groove gratings can be obtained with a modal model initially developed in optics. The main characteristics of these waves can be obtained in a simpler way from a modified previous approximate result carried out for microwaves. The transposition of the previous works on electromagnetic surface waves to acoustics is easy and fruitful, and it is surprising that these works have generally been ignored by acousticians.

[1] K. M. Ivanov-Shits and F. V. Rozhin, ''Investigation of surface waves in air,'' Sov. Phys. Acoust. **5**, 510–512 (1960).
[2] R. J. Donato, ''Model experiments on surface waves,'' J. Acoust. Soc. Am. **63**, 700–703 (1978).
[3] G. A. Daigle, M. R. Stinson, and D. I. Havelock, ''Experiment on surface wave over a model impedance plane using acoustical pulses,'' J. Acoust. Soc. Am. **99**, 1993–2005 (1996).
[4] R. A. Hurd, ''The propagation of an electromagnetic wave along an infinite corrugated surface,'' Can. J. Phys. **32**, 727–734 (1954).
[5] L. A. Vajnshtejn, ''Electromagnetic surface waves on a comblike structure,'' Zh. Tekh. Fiz. **26**, 385–397 (1956).
[6] A. Hessel, J. Schmoys, and D. Y. Tseng, ''Bragg-angle blazing of diffraction gratings,'' J. Opt. Soc. Am. **65**, 380–384 (1975).
[7] J. R. Andrewartha, J. R. Fox, and I. J. Wilson, ''Resonance anomalies in the lamellar grating,'' Opt. Acta **26**, 69–89 (1979).
[8] J. R. Andrewartha, J. R. Fox, and I. J. Wilson, ''Further properties of lamellar grating resonance anomalies,'' Opt. Acta **26**, 197–209 (1979).
[9] L. A. Vajnshtejn, *Electromagnetic Waves* (Sovetskoe Radio, Moscow, 1957).
[10] R. E. Collin, *Field Theory of Guided Waves* (McGraw-Hill, New York, 1960).
[11] L. Kelders, W. Lauriks, and J. F. Allard, ''Surface waves above thin porous layers saturated by air at ultrasonic frequencies,'' J. Acoust. Soc. Am. (submitted).

Comments on "A re-examination of risk estimates from the NIOSH Occupational Noise and Hearing Survey" [J. Acoust. Soc. Am. 101, 950–963 (1997)]

Robert A. Dobie
Department of Otolaryngology-Head and Neck Surgery, The University of Texas Health Science Center at San Antonio, San Antonio, Texas 78284-7777

(Received 14 April 1997; accepted for publication 22 January 1998)

The recent paper by Prince *et al.* [J. Acoust. Soc. Am. **101**, 950–963 (1997)] uses as a primary outcome measure an estimate of hearing handicap based on a weighted average of pure tone thresholds at 1, 2, 3, and 4 kHz. This choice is unusual and inadequately justified. If four equally weighted frequencies are to be chosen, the ANSI standard referenced by Prince *et al.* would suggest 0.5, 1, 2, and 4 kHz. If 1, 2, 3, and 4 kHz are chosen, the weights from ANSI 3.5, 1969 would be very different from those used by Prince *et al.*
[S0001-4966(98)00805-4]

PACS numbers: 43.50.Qp, 43.64.Wn [GAD]

The primary outcome variable in the paper by Prince *et al.* (1997) is the excess risk of ''hearing handicap'' (HH), using a weighted average of hearing threshold levels (HTLs) at 1, 2, 3, and 4 kHz to estimate hearing handicap. This choice of outcome variable is unusual and, I believe, inappropriate. Most state and federal workers' compensation programs use either 0.5, 1, and 2 kHz or 0.5, 1, 2, and 3 kHz; in the most recent survey known to me, *none* use the 1, 2, 3, and 4 kHz frequency combination recommended by a Task Force of the American Speech and Hearing Association (ASHA), and this recommendation apparently never became official ASHA policy, although it is one of ten frequency combinations mentioned as ''commonly used or proposed'' in a recent American National Standard (ANSI S3.44, 1996).

It is important to conceptually distinguish between HH and noise-induced hearing loss (NIHL). HH is conventionally defined as ''the disadvantage imposed by an impairment sufficient to affect the individual's efficiency in the activities of daily living'' (AAO-HNS, 1988). The American Medical Association (AMA, 1993) and ASHA use essentially identical definitions, although the AMA substitutes ''binaural impairment'' for HH. It follows that HH might well be estimated using HTLs at frequencies important for speech understanding, and this is indeed why the AAO-HNS and AMA recommend the 0.5, 1, 2, 3 kHz pure-tone average.

NIHL, on the other hand, would probably be most efficiently detected using the frequencies most affected by NIHL. In the early years of exposure, one might choose 3, 4, and 6 kHz, while for some middle-aged workers late in a noise-exposed career, lower frequencies might be better. As a compromise, the Occupational Safety and Health Administration uses the 2, 3, 4 kHz pure-tone average in their definition of a ''standard threshold shift,'' intended to detect NIHL.

Some readers may confuse the concepts of hearing handicap and an ideal detector of NIHL when they read in this paper that:

> ''definitions [of hearing handicap] that exclude the higher frequencies tend to be less sensitive to noise damage...the most suitable definition [of hearing handicap] may depend on the population characteristics, such as age, exposure duration, and degree of hearing handicap already accrued, as well as whether one chooses to identify preclinical or later stages of hearing handicap.''

These comments would be very appropriate if the first two uses of ''HH'' were replaced by ''standard threshold shift'' and the third and fourth were replaced by ''NIHL.'' HH is defined to be independent of etiology. Thus if some cause of hearing loss produced initial threshold shifts at 8 kHz, we might wish to measure those shifts to get an early warning of more communicatively significant changes to come, but we would not call these changes ''HH,'' because isolated threshold shifts at 8 kHz do not affect everyday speech communication.

Having selected HTLs at 1, 2, 3, and 4 kHz to estimate HH, Prince *et al.* (1997) applied weights to each frequency: 0.24, 0.38, 0.34, and 0.24, respectively. These weights were obtained by taking the values of Table 5 of the American National Standards Institutes method for calculating the articulation index, (ANSI S3.5, 1969), then multiplying by 100 (Prince, personal communication). That table gives weights for 15 one-third octave bands, and the four bands including 1, 2, 3, and 4 kHz contribute only 36% of the total contributions to intelligibility of the 15 bands whose center frequencies range from 200 to 5 000 Hz. It would not be surprising if readers inferred from the weights used by Prince *et al.* (1997) that 3 and 4 kHz contribute about as much (0.34 +0.24) to speech intelligibility as 1 and 2 kHz (0.24+0.38), and that frequencies lower than 1 kHz make negligible contributions, but that is not at all what Table 5 of ANSI S3.5 says.

Table 5 shows that the spectral ''center of gravity'' is about 1600 Hz, i.e., frequencies below 1600 Hz contribute about as much as frequencies above 1600 Hz. If we wished to further divide the spectrum into *four* equally contributing

bands, they would be approximately 200–800, 800–1600, 1600–2700, and 2700–5600 Hz. The four audiometric frequencies best representing each of these bands would be 500 Hz, 1 kHz, 2 kHz, and 4 kHz, not the frequencies chosen by Prince *et al.* If for some reason we insist on using 1, 2, 3, and 4 kHz, it would be most reasonable to use the logarithmic midpoints of the intervals between these frequencies (approximately 1.4, 2.4, and 3.5 kHz) to divide the spectrum into four bands that do *not* contribute equally: frequencies from 200 to 1400 Hz (corresponding to the 1 kHz HTL) contribute 44% of overall importance, while the other three bands (related to HTLs at 2, 3, and 4 kHz) contribute about 26%, 17%, and 13%, respectively. It should be noted that all of these calculations from Table 5 of ANSI S3.5 assume that frequency-importance functions for nonsense syllables, as opposed to continuous discourse, are appropriate for selecting frequencies and weights for estimating HH; frequency-importance functions for more realistic speech materials show crossover frequencies of about 1300 Hz for monosyllable words (Studebaker and Sherbecoe, 1991) and 1100 Hz for continuous discourse (Studebaker *et al.*, 1987).

Prince *et al.* (1997) do present some results based on frequency combinations of 0.5, 1, and 2 kHz and 1, 2, and 3 kHz, but these are relegated to ''sensitivity analyses.'' As expected, choice of frequencies affects the results (excess risk of HH, in this case). Nevertheless, the authors appear to have based their primary conclusions, that there is excess risk of HH at time-weighted average exposure levels both above 85 dB and also between 80 and 84 dB, on their weighted average of 1, 2, 3, and 4 kHz. As discussed above, I believe this to be an inappropriate method of estimation of HH.

American Academy of Otolaryngology-Head and Neck Surgery-Subcommittee on the Medical Aspects of Noise of the Committee on Hearing and Equilibrium. (**1988**). *Guide for Conservation of Hearing in Noise* (American Academy of Otolaryngology-Head and Neck Surgery Foundation, Inc., Washington, D.C.), revised ed.

AMA (**1993**). *Guides to the Evaluation of Permanent Impairment* (American Medical Association, Chicago), 4th ed.

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

ANSI (**1996**). ANSI S3.44-1996, ''Determination of Occupational Noise Exposure and Estimation of Noise Induced Hearing Impairment'' (American National Standards Institute, New York).

Prince, M. M., Stayner, L. T., Smith, R. J., and Gilbert, S. J. (**1997**). ''A re-examination of risk estimates from the NIOSH Occupational Noise and Hearing Survey,'' J. Acoust. Soc. Am. **101**, 950–963.

Prince, M. (personal communication).

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., and Sherbecoe, R. L. (**1991**). ''Frequency-importance and transfer functions for recorded CID w-22 word lists,'' J. Speech Hear. Res. **34**, 427–438.

Response to "Comments on 'A re-examination of risk estimates from the NIOSH Occupational Noise and Hearing Survey'" [J. Acoust. Soc. Am. 103, 2734 (1998)]

Mary M. Prince,[a] Leslie T. Stayner,[b] Randall J. Smith,[b] and Stephen J. Gilbert[b]
National Institute for Occupational Safety and Health, 4676 Columbia Parkway, Cincinnati, Ohio 45226

(Received 29 October 1997; accepted for publication 22 January 1998)

Concern is raised by Dobie [J. Acoust. Soc. Am. **103**, 2734 (1998)] regarding a recent analysis [Prince *et al.*, J. Acoust. Soc. Am. **101**, 950–963 (1997)] of the NIOSH Occupational Noise and Hearing Survey data. Specifically, issues are raised concerning (1) definition of hearing handicap, (2) the use of frequency-specific articulation index (AI) weights applied to the binaural pure-tone average of 1, 2, 3, and 4 kHz, and (3) conclusions regarding significant excess risk based on this definition. We have reviewed the development of the definitions of hearing handicap and provide additional support for the use of a hearing handicap definition based on the binaural pure-tone average of 1, 2, 3, and 4 kHz and the weighting of specific frequencies. Furthermore, our definition of noise-induced hearing handicap is similar to one of several proposed by the International Standards Organization (ISO-1999, 1990) and the American National Standards Institute (ANSI S3.44, 1996). Additional analyses show that there is significant evidence of excess risk at daily exposure levels below 85 dB using any of the pure-tone average and/or weighting strategies we have examined. Hence we have provided additional support for our conclusions regarding exposure-response curves and we have reaffirmed that our methods are appropriate for the scope of our analysis. [S0001-4966(98)00905-9]

PACS numbers: 43.50.Qp, 43.64.Wn [GAD]

We are pleased to have the opportunity to clarify our rationale and methodology as it relates to three issues mentioned by Dobie (1998). These include (1) terminology and the use of hearing handicap, (2) use of articulation-based weights applied by frequency for the binaural pure tone average of 1, 2, 3, and 4 kHz, and (3) conclusions regarding significant excess risk based on the weighted and unweighted hearing handicap definition (Prince *et al.*, 1997).

I. DEFINITION OF HEARING HANDICAP

Dobie (1998) states that the choice of outcome variable (Prince *et al.*, 1997) is unusual and believes that it is inappropriate. The rationale for Dobie's statement appears to be that most state and federal workers' compensation programs use either a pure-tone average of 0.5, 1, 2 kHz or a pure-tone average of 0.5, 1, 2, 3 kHz and that ''none use the 1, 2, 3, 4 kHz frequency combination recommended by a Task Force of the American Speech and Hearing Association'' (ASHA, 1981).

We wish to clarify that this paper and its results were not to be applied nor were the analyses conducted for use by state and federal workers' compensation purposes. Our study was designed to evaluate damage-risk criteria for occupational noise exposure. The analysis focused on characterizing the dose-response relationship between cumulative noise dose and *noise-induced* hearing loss. Therefore its use to infer hearing handicap for compensation purposes is irrelevant to the scope of this analysis. In fact, our analysis specifically focused on ''noise-induced hearing handicap'' (NIHH) rather than ''hearing handicap'' which is independent of etiology.

Given this distinction, we believe it was appropriate to use the standard terminology for noise-induced hearing handicap used for this same purpose by international (ISO-1999, 1990) and American national standards (ANSI S3.44, 1996). The reference ISO-1999 and ANSI S3.44-1996 standards have defined hearing handicap as ''*a combination of hearing threshold levels at specified frequencies,*'' whereas ''*potential hearing impairment is directly assessed by the noise-induced permanent threshold shift (NIPTS) . . . for the exposure conditions and populations under consideration.*''

To better evaluate concerns regarding terminology, we revisited the literature relevant to the development of the 1959 American Academy of Ophthalmology and Otolaryngology (AAOO) definition which contained three frequencies (0.5, 1, and 2 kHz) and a 27 dB HL threshold for beginning handicap. Even before the adoption of the original AAOO (1958) rule, other rules for calculating percent hearing handicap had been employed, including one issued by the American Medical Association's (AMA) Council on Physical Medicine in 1947 (AMA, 1947). This method, called the Fowler–Sabine scale, used the four frequencies 512, 1024, 2048, and 4096 Hz but gave them weights of 15%, 30%, 40%, and 15%, respectively (Davis, 1978). In fact, Davis (1978) stated that this definition is a ''direct ancestor of the AAOO rules, which differ chiefly in their greater simplicity and their rejection of 4000 Hz from the calculation.'' According to Davis (1973), this method was revised in 1959 because ''it was essential to simplify the

[a]Industrywide Studies Branch, Division of Surveillance, Hazard Evaluations, and Field Studies.
[b]Risk Evaluation Branch, Education and Information Division.

more accurate but complicated Fowler-Sabine scale in order to gain acceptance. Otologists just wouldn't use the complicated table . . . .''

When the Occupational Safety and Health Administration (OSHA) promulgated its amendment for hearing conservation programs in 1981, the AAOO rule was criticized for including only three frequencies and for having a relatively high threshold for beginning handicap. The view at this time was that frequencies above 2 kHz clearly assisted the understanding of speech in noisy situations and this favored the inclusion of 4 kHz or at least 3 kHz in the calculations.

In light of mounting evidence that the higher frequencies were important for speech discrimination in nonideal conditions (distorted or with competing noise), the AAO (formerly AAOO) amended its formula to include 3 kHz (AAO, 1979). Later studies confirmed that hearing sensitivity above 2 kHz is critical for speech communication (Suter, 1985; Smoorenburg, 1992). A study by Phaneuf *et al.* (1985) confirmed that hearing sensitivity above 2 kHz was required to predict hearing disability and handicap. The study also reported that the highest precision in predicting ''judged hearing disability'' was obtained with a pure tone average of 1, 2, 3, and 4 kHz in the worse ear.

Terminology: Hearing impairment versus hearing handicap

In the original manuscript submitted to the journal, we used the term *hearing impairment* in a manner consistent with its use in the 1972 NIOSH criteria document (NIOSH, 1972) and with OSHA's definition of *material impairment of hearing* (46 Fed. Reg. 4082, 1981).[1] However, journal reviewers strongly recommended substituting *hearing handicap* to be consistent with the ISO-1999 (1990) and ANSI S3.44 (1996) standards. This seemed appropriate since additional analyses were included to compare the new NIOSH excess risk estimates to models used in these standards.

In our use of the term *hearing handicap*, the ears were given equal weighting and the frequency-specific pure-tone average was used to dichotomize the population into groups with and without hearing loss (as defined by OSHA and NIOSH). The proportions of workers with and without hearing loss were then compared to a reference value (based on ''control group'' pure-tone averages), a strategy consistent with recommendations of ISO-1999 (1990) and ANSI S3.44 (1996). Hence our definition of hearing handicap does not appear to be particularly unusual or inappropriate.

Our statements in the paper appear to be misinterpreted with regard to variation of excess risk estimates by definition of hearing handicap. To place the statements in their proper context, we discussed why the dose-response curves for NIHH may vary. We stated (Prince *et al.*, 1997) that:

''Hearing damage at 3 and 4 kHz is expected to occur sooner than loss at lower frequencies (0.5, 1, and 2 kHz). Definitions [of NIHH] that exclude the higher frequencies tend to be less sensitive to noise damage and may require longer durations of exposure to a given sound level to see significant excess risks in the population. Figures 4A and B suggest that the most suitable definition of hearing handicap may depend on the population characteristics, such as age, exposure duration, and degree of hearing handicap already accrued . . . . ''

There appears to be confusion regarding the concept of standard threshold shift and a criterion to prevent noise-induced hearing loss, for which OSHA and NIOSH ordinarily use the term ''material impairment of hearing.'' Dobie suggested that ''these comments would be very appropriate if *hearing handicap* was replaced by *standard threshold shift.*'' The purpose of *hearing handicap* (as meant here to represent material impairment of hearing), is quite different from that of *standard threshold shift.* While *hearing handicap* is used in the development of damage-risk criteria, a *standard threshold shift* refers to a change in hearing for an individual relative to a previous test and is used in the context of a hearing conservation program. The NIOSH Occupational Noise and Hearing Survey, a cross-sectional study, did not collect data on individual changes in hearing over time.

To further his arguments regarding our choice of frequencies, Dobie makes an analogy to the use of a single frequency (8 kHz). Dobie's comment that ''we wouldn't call these changes [at 8 kHz] 'HH' [hearing handicap] because isolated threshold shifts at 8 kHz do not affect everyday speech communication'' is specious. We never implied that 8 kHz or any single frequency is important for everyday speech communication. However, there is evidence that 3 kHz and 4 kHz are indeed important for speech recognition as are 1 kHz and 2 kHz. The fact that 3 kHz and 4 kHz are earliest and most severely affected by noise should not exclude them from our selection of a prevention criterion, but argue in their favor. Therefore, we consider it accurate to state that ''the addition of the most sensitive frequencies to a hearing handicap definition is a valid option if the goal is to have a measure that addresses both prevention and identification of noise-induced hearing handicap.''

II. USE OF ARTICULATION-BASED WEIGHTS FOR BINAURAL PURE-TONE AVERAGE

As mentioned previously, the idea of frequency-specific weighting is not a new concept (AMA, 1947). In our paper, the idea was conceived as an alternative that would balance the need to take into account both hearing sensitivity to noise-induced damage at 3 and 4 kHz and speech understanding. Since the frequencies do not contribute equally to speech recognition, as demonstrated by the research of French and Steinberg (1947) and others (Kalikow *et al.*, 1977; Suter, 1985; Quiggle *et al.*, 1959), a decision was made to weight the frequencies for the pure tone average with weights using the articulation index (ANSI S3.5-1969).

The articulation weights from Table 5 of ANSI S3.5-1969 were used as a relative guide to weighting the frequencies we chose to include in our definition of noise-induced hearing handicap. Given the importance of these four frequencies (1, 2, 3, and 4 kHz) in predicting noise-induced hearing handicap, we wanted to make some allowance for the relative contribution of each frequency. We were not using these data to represent 1/3 octave bands and do not have audiometric data corresponding to the 20 octave bands used in the calculation of the AI, or for each of the 15 1/3-

TABLE I. Excess risk percent by exposure level, age, duration exposed, and hearing handicap definition.

Daily exposure level (dB)	Hearing handicap definition: 1–4 kHz	1–4 kHz, AI weight[a]	1–4 kHz (alternate weighting scheme)[b]	1–3 kHz	0.5–2 kHz. (*AAOO definition*)	0.5–2 kHz. (*AAOO definition*)	0.5, 1, 2, 4 kHz (Fowler–Sabine weights)[c]
			Age 30, duration 2–4 years				
80	0.2	0.1	0.4	0.4	0.4	0.3	2.0
85	0.9	0.9	1.1	1.1	1.0	0.9	0.6
90	3.3	3.3	2.9	2.7	2.5	2.3	1.9
95	10.2	10.3	7.9	7.3	6.8	6.5	6.1
100	30.1	30.2	24.0	21.6	20.6	20.4	23.1
			Age 45, duration 5–10 years				
80	0.6	0.5	1.8	2.3	2.0	1.7	0.9
85	3.5	3.3	5.0	6.4	5.3	4.8	3.0
90	12.4	12.2	13.3	17.3	13.6	13.5	9.2
95	34.6	34.0	34.0	43.8	34.4	36.7	28.2
100	65.8	65.1	68.8	78.1	69.6	74.1	68.8
			Age 60, duration >10 years				
80	1.5	1.2	4.9	5.8	4.3	4.8	2.9
85	8.4	7.9	13.2	15.1	10.5	13.2	9.1
90	25.5	24.7	30.2	33.4	22.8	31.3	24.7
95	46.4	45.6	53.4	56.8	43.0	56.9	51.5
100	54.2	53.8	66.5	69.0	63.3	70.2	67.8

[a]AI-based were approximately 20%, 32%, 28%, and 20% for 1, 2, 3, and 4 kHz, respectively.
[b]Weights are 44%, 26%, 17%, and 13% for 1, 2, 3, and 4 kHz, respectively (Dobie, 1998).
[c]Weights are 15%, 30%, 40%, and 15% for 0.5, 1, 2, and 4 kHz, respectively (Davis, 1978).

octave bands. As calculated in Eq. (1) (Prince *et al.*, 1997), the hearing threshold level for the frequencies 1, 2, 3, and 4 kHz were given weights of approximately 20%, 32%, 28%, and 20%, respectively. In comparison to the unweighted binaural pure-tone average of 1–4 kHz, the AI-weighted version produced only minor changes in excess risks. These nominal changes in excess risk suggest that the unweighted binaural 1–4 kHz definition would also be adequate in describing the dose-response relationship. Most importantly, we have observed that nearly every other combination of frequencies and weightings examined results in the same or greater degree of excess risk.

III. CONCLUSIONS REGARDING SIGNIFICANT EXCESS RISK

Dobie appears to minimize the significance of our sensitivity analyses. Such analyses are important in exploring whether the inferences made from the data are consistent under different conditions or assumptions. The sensitivity analysis was a systematic method for evaluating modeling assumptions and decisions regarding the choice of definitions and use of frequency-specific weights.

Dobie states that in our analysis, excess risk estimates vary by choice of frequencies used, but that our primary conclusion, that ''there is excess risk of hearing handicap at time-weighted average exposure levels both above 85 dB and also between 80–84 dB,'' is based on the binaural AI-weighted pure-tone average of 1–4 AI definition, which he believes to be inappropriate. However, additional age-specific stratified analyses show that regardless of the definition of noise-induced hearing handicap (including the AAOO definition) used, our primary conclusions are supported. Table I shows additional analyses using several other definitions of noise-induced hearing handicap, all producing excess risk estimates that are higher than the binaural 1–4 kHz AI definition, particularly at daily exposure levels below 90 dB.

IV. SUMMARY

Based on the review of the literature and supporting data analyses, we conclude that Dobie's criticisms are tenuous. Further, we wish to reiterate that the results presented in our paper (Prince *et al.*, 1997) were not proposed as a basis of calculating hearing handicap for workers compensation purposes. Rather, our calculations were similar to those used by ANSI S3.44 (1996) and ISO-1999 (1990) in which the percentage of exposed workers with hearing handicap was based on a dichotomization of a binaural pure-tone average (i.e., average HTL$\geqslant$25 dB for selected frequencies) and compared to an appropriate control data base. These results may be used as a guide for developing occupational health standards for the protection of workers' hearing. The frequency-specific weighting was employed only as a relative guide to associate the pure-tone average with speech intelligibility and was not specifically advocated for use by state and federal compensation boards. Our analyses indicate that irrespective of the binaural pure-tone average definition or the frequency weighting used, excess risk estimates were similar or greater than the binaural 1–4 kHz AI definition of hearing handicap (Prince *et al.*, 1997).

ACKNOWLEDGMENTS

The authors gratefully acknowledge the advice of Dr. John R. Franks and Dr. Alice Suter in the preparation of this article.

[1]NIOSH (1972) used the term ''hearing impairment'' to define its criteria for maximum acceptable hearing loss, and OSHA later used a slightly modified term ''material impairment of hearing'' to define the same criteria (46 Fed. Reg. 4082, 1981). In this context, a worker was considered to have a material impairment of hearing when his or her binaural pure-tone average at the audiometric frequencies 1000, 2000, and 3000 Hz exceed 25 dB.

AAOO (**1958**). AAOO Committee on Conservation of Hearing (Subcommittee on Noise), ''Guide for the evaluation of hearing impairment,'' Trans. Am. Acad. Opthal. Otolaryng. **63**, 236–238.

AAO (**1979**). American Academy of Otolaryngology Committee on Hearing and Equilibrium, and the American Council of Otolaryngology Committee on the Medical Aspects of Noise, ''Guide for the evaluation of hearing impairment,'' AMA **241**, 2055–2059.

AMA (**1947**). AMA Council on Physical Medicine, ''Tenative standard procedure for evaluating the percentage loss of hearing in medicolegal cases,'' J. Am. Med. Assoc. **133**, 396–397.

ANSI (**1969**). ANSI S3.5-1969, ''American National Standard: Methods for the calculation of the articulation index'' (American National Standards Institute, New York).

ANSI (**1996**). S3.44-1996, ''American National Standard: Determination of occupational noise exposure and estimation of noise-induced hearing impairment'' (American National Standards Institute, New York).

ASHA (**1981**). American Speech-Language-Hearing Association Task Force on the Definition of Hearing Handicap, ''On the definition of hearing handicap,'' Asha **23**, 293–297.

Davis, H. (**1973**). ''Some comments on 'Impairment to hearing from exposure to noise' by K. D. Kryter,'' J. Acoust. Soc. Am. **53**, 1237–1239.

Davis, H. (**1978**). ''Hearing Handicap, Standards for Hearing and Medicolegal Rules,'' in *Hearing and Deafness, Fourth Edition*, edited by H. Davis and S. R. Silverman (Holt, Rinehart and Winston, New York), pp. 266–290.

Dobie, R. A. (**1997**). ''Comments on 'A re-examination of risk estimates from NIOSH Occupational Noise and Hearing Survey,''' J. Acoust. Soc. Am. **103**, 2734–2735.

46 Fed. Reg. 4,082,1981, ''46 Fed. Reg. 11, Occupational noise exposure; Hearing Conservation amendment, rule, and proposed rule,'' part III, pp. 4078–4179.

French, N. R., and Steinberg, J. C. (**1947**). ''Factors governing the intelligibility of speech sounds,'' J. Acoust. Soc. Am. **19**, 90–110.

ISO-1999 (**1990**). ''Acoustics—Determination of occupational noise exposure and estimation of noise-induced hearing impairment'' (International Organization for Standardization, Geneva).

Kalikow, D. N., Stevens, K. N., Elliot, L. L. (**1977**). ''Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability,'' J. Acoust. Soc. Am. **61**, 1337–1351.

NIOSH (**1972**). ''NIOSH criteria for a recommended standard: occupational exposure to noise,'' Cincinnati, OH: U.S. Department of Health, Education, and Welfare, Public Health Service, Center for Disease Control, National Institute for Occupational Safety and Health, DHSS(NIOSH) Publication No. HIM 73-11001.

Phaneuf, R., Hetu, R., and Hanley, J. A. (**1985**). ''A Bayesian approach for predicting judged hearing disability,'' Am. J. Indust. Med. **7**, 343–352.

Prince, M. M., Stayner, L. T., Smith, R. J., and Gilbert, S. J. (**1997**). ''A re-examination of risk estimates from the NIOSH Occupational Noise and Hearing Survey (ONHS),'' J. Acoust. Soc. Am. **101**, 950–963.

Quiggle, R. R., Glorig, A., Delk, J. H., and Summerfield, A. B. (**1959**). ''Predicting hearing loss for speech from pure tone audiograms,'' Laryngoscope **67**, 1–15.

Smoorenburg, G. F. (**1992**). ''Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiograms,'' J. Acoust. Soc. Am. **91**, 421–437.

Suter, A. H. (**1985**). ''Speech recognition in noise by individuals with mild hearing impairments,'' J. Acoust. Soc. Am. **78**, 887–900.

Species differences of distortion-product otoacoustic emissions: Comment on "Interpretation of distortion product otoacoustic emission measurements. I. Two stimulus tones" [J. Acoust. Soc. Am. 102, 413–429 (1997)]

M. L. Whitehead[a)]
University of Liverpool Veterinary Field Station, Leahurst, South Wirral L64 7TE, United Kingdom

(Received 14 October 1997; accepted for publication 20 January 1998)

In small mammals, there is convincing evidence for the existence of two discrete components of distortion-product otoacoustic emissions (DPOAEs), which have been called, variously, the active, low-level, or peak component, and the passive, high-level, or basal component. However, in humans, there is no convincing evidence for the existence of two analogous components of the DPOAE, at least for stimulus levels up to 85 dB SPL. These differences indicate that caution is required when generalizing from models of DPOAEs in small mammals to the human situation. [S0001-4966(98)00205-7]

PACS numbers: 43.64.Jb, 43.64Kc [BLM]

Mills (1997) presents a useful, simple model that explains several features observed in published measurements of distortion-product otoacoustic emissions (DPOAEs) in gerbils and rabbits (e.g., Norton and Rubel, 1990; Mills and Rubel, 1994; Whitehead *et al.*, 1992a, b, 1995). The model assumes that the total DPOAE signal consists of two components, one originating from the region of the peaks of the traveling waves in the cochlea, and the other originating basally in the cochlea. These ''peak'' and ''basal'' components correspond to what have previously been called ''active'' or ''low-level'' and ''passive'' or ''high-level'' DPOAE components, respectively.

An important feature consistently observed in rodents and rabbits is a notch in DPOAE growth functions. Specifically, for any one primary-frequency pair, the notch is seen at a constant L_1 of around 60–70 dB SPL, for any L_2 below about 50–60 dB SPL, as illustrated in the contour plots and 3-D plots DPOAE amplitude in Mills (1997; e.g., his Fig. 6) and other articles (e.g., Mills and Rubel, 1994; Whitehead *et al.*, 1992a, b, 1995). As noted by Mills, although the details of the notch vary slightly with stimulus-frequency parameters, the notch is seen at nearly all stimulus frequencies in almost all normal ears in small mammals. Phase measurements over the notch strongly suggest that it results from a phase cancellation of two discrete components (Whitehead *et al.*, 1992a, 1995; Mills, 1997). As such, the notch is perhaps the strongest evidence for the existence of two discrete DPOAE components in small mammals (Whitehead *et al.*, 1992a, b). In Mills' model the notch is reproduced by assuming that the peak and basal DPOAE components are 180° out of phase, resulting in a pronounced cancellation of the two components when they are of approximately equal amplitude.

In his Sec. IV D, Mills discusses the possible implications of his model for the assessment of hearing function, both in small mammals and humans. Mills' model may well be useful for interpreting DPOAE measurements in small mammals. However, the purpose of this letter is to point out that caution is required when extrapolating from models of small-mammal emissions to the human situation. This is because of substantial differences of the behavior of otoacoustic emissions between small mammals and humans. In particular, and with specific relevance to Mills' model, it is not clear from presently available data whether or not humans have two discrete DPOAE components that are equivalent to the active/low-level/peak and passive/high-level/basal components that have been demonstrated in rabbits and rodents.[1] In addition to the basic-science implications, this issue has relevance for the design and interpretation of clinical tests utilizing DPOAEs.

There are a number of pronounced differences in the behavior of DPOAEs, and of the other otoacoustic emission types, between humans and laboratory small mammals. In particular, in humans, at least for stimulus levels below 85 dB SPL, there is no highly consistent notch analogous to that observed at an L_1 of 60–70 dB SPL in small mammals. This can be seen, for example, in the contour plots of DPOAE amplitude in Whitehead *et al.* (1995; their Fig. 1). As discussed in Sec. III B of Whitehead *et al.* (1995), notches are sometimes seen in human DPOAE growth functions. However, they do not consistently occur within a narrow region of stimulus-level parameters for almost all stimulus–frequency pairs in all ears as in small mammals. In contrast, they are highly idiosyncratic, being seen only at some frequencies in any one ear, and varying greatly and more-or-less randomly (although mostly below 3 kHz) across ears in the frequencies and the stimulus levels at which they occur (e.g., He and Schmiedt, 1993). There is some association of these notches in DPOAE growth functions with the presence of spontaneous emissions and strong stimulus-frequency emissions around the frequency of the DPOAE (e.g., Wier *et al.*, 1988; Lonsbury-Martin *et al.*, 1990; Moulin *et al.*, 1992), suggesting that the notches observed in humans may

[a)]Address for correspondence: 46 Carlton Close, Parkgate, South Wirral L64 6RB, United Kingdom.

have a different origin to those in small mammals.

He and Schmiedt (1993) demonstrated that, in humans, notches in DPOAE growth functions are associated with pronounced idiosyncratic fine structure in the amplitude of the DPOAE as a function of frequency. The fine structure consists of ''ripples''—a series of peaks and valleys—that showed little or no saturation up to a stimulus level of $L_1 = L_2 = 65$ dB SPL. The pattern of fine structure tended to shift towards lower frequencies with increasing stimulus level. Thus, as stimulus level increases, a valley may shift down to a frequency at which there is a peak at a lower stimulus level. At such frequencies, there will be a notch in the DPOAE growth function if the DPOAE amplitude in the valley at a higher stimulus level is below the amplitude at the peak at a lower stimulus level.

The notches in DPOAE growth functions described in humans by He and Schmiedt (1993) show characteristics substantially different from those in rabbits and rodents. In rabbits and rodents, there is no idiosyncratic fine structure of DPOAE amplitude and, thus, no ripple of DPOAE amplitude as a function of frequency around the frequency of the notch (e.g., Whitehead *et al.*, 1992a). In humans, the specific valley in the fine structure associated with the growth-function notch is visible in the fine structure at stimulus levels lower and higher than that of the notch in the growth function. In rabbits and rodents, the growth-function notch, of course, is also seen as a local minimum in DPOAE amplitude as a function of frequency. However, in contrast to humans, this minimum is restricted to the narrow stimulus-level range of the notch, and is not visible at lower or higher $L_1 = L_2$ (e.g., Whitehead *et al.*, 1992a). Thus, whereas in humans the growth function notches are one reflection of a mechanism that operates to reduce DPOAE amplitudes in a narrow range of frequencies over a wide range of $L_1 = L_2$ in rabbits and rodents the notches only occur in that narrow range of $L_1 = L_2$ at which the two DPOAE components are $\sim 180°$ out of phase and approximately equal in amplitude. Mechanisms to explain the DPOAE fine-structure minima (and maxima) observed in humans include the mixing of DPOAE wavelets (e.g., Sun *et al.*, 1994), and the effects of reflection of cochlear-generated distortion energy from the DP-frequency place (e.g., Talmadge *et al.*, 1995; Brown *et al.*, 1996).

The species differences in the behavior of the DPOAE growth-function notches suggest that the notches in humans arise by a different mechanism from those in rabbits and rodents, and thus that the growth-function notches in humans are not necessarily evidence of discrete high-level and low-level DPOAE components.

Of the other significant differences between the otoacoustic emissions of humans and small mammals, the most pronounced are that spontaneous emissions are less prevalent in laboratory mammals than in humans, and that stimulus-frequency emissions are smaller whereas DPOAEs are larger in the small mammals, resulting in a 20–40-dB difference in the relative amplitudes of stimulus–frequency emissions and DPOAEs between humans and laboratory mammals (for review, see Whitehead *et al.*, 1996).

In humans, as in small mammals, the amplitudes of DPOAEs evoked by higher-level primaries are less reduced by trauma than are DPOAEs evoked by lower-level stimuli (for review, see Whitehead *et al.*, 1996). However, DPOAEs evoked by primaries of 75 dB SPL are reduced by various traumas in humans, and have been and are routinely and successfully used in clinical applications of DPOAEs by several groups. In the small mammals, DPOAEs evoked by 75 dB SPL primaries are regarded as dominated by the component variously called basal, high-level, or passive, and appear relatively unaffected by experimental traumas known to cause substantial hearing-threshold elevations (e.g., Mills and Rubel, 1994; Whitehead *et al.*, 1992b, 1995, 1996).

To summarize, whereas there is convincing evidence for two discrete DPOAE components in small mammals, there is no convincing evidence for discrete active/low-level/peak and passive/high-level/basal DPOAE components in humans, at least for those stimulus levels studied, i.e., up to 85 dB SPL. In the absence of a clear, consistent cancellation notch similar to that in rabbits and rodents, the overall picture of DPOAEs in humans is not that of two discrete components, but rather that of only one component which becomes gradually less vulnerable to trauma as stimulus levels increase (perhaps because of increasing saturation of the DPOAE generator). For this reason, models of DPOAE behavior in small mammals may not be directly applicable to the human situation.

What factors underlie these species differences of DPOAE behavior? I will briefly mention three possibilities. First, there may indeed be two discrete DPOAE components in humans, but which dominate the DPOAE over different stimulus-level ranges from those in small mammals (Whitehead *et al.*, 1995, 1996; Mills, personal communication). Thus, it may be the case that, in humans, the stimulus level at which the passive/high-level/basal component becomes dominant over the active/low-level/peak component is above 85 dB SPL. If so, then all of the available DPOAE data in humans would correspond to the active/low-level/peak component of small-mammal DPOAEs. According to this scenario, it is possible that, if human DPOAEs were measured at higher stimulus levels than previously used, they would show consistent notches at some L_1 above 85 dB SPL, similar to those seen at $L_1 \approx 65$ dB SPL in small mammals. In terms of Mills' model, such a finding could be interpreted as reflecting differences of cochlear-amplifier gain between species (Mills, 1997, personal communication).

A second possibility is that there is only one component to human DPOAEs. Mills (1997) makes the important point that the peak and basal DPOAE components observed in small mammals may actually be produced by exactly the same mechanism(s) within the cochlea. Thus, in his model, the two components effectively differ only in the shapes of the excitation patterns produced by the stimuli and, therefore, in the regions along the cochlea in which the DPOAEs are being produced. Further research is required to determine whether this is the case, or whether there are more fundamental, qualitative differences of generation of the two components seen in small-mammal DPOAEs. Nevertheless, in Mills' model, the distinction between peak and basal emissions lies in the precise vector summation, within various regions along the cochlea, of the contributions to the total

DPOAE from each point along the cochlear partition. Given the substantial differences in cochlear dimensions and the frequency range of hearing between humans and the laboratory small mammals, there is no *ad hoc* reason to suppose that this summation will be similar in humans to that in small mammals.

Finally, it is worth considering whether there may be more fundamental differences within the organ of Corti between mammalian orders. The outer hair cell, with its unique force-generation mechanism, appears to have evolved within the mammals—the other vertebrate classes do not have them. The evolutionary ancestors of the orders Lagomorpha and Rodentia split from those leading to the primates more than 100 million years ago. During this period, evolution could have resulted in subtle differences in the precise manner in which outer hair cells contribute to ''cochlear amplifier'' function and OAE generation. In this respect, it is worth noting the anatomical observation that humans and other primates have a more irregular arrangement of hair cells along the cochlea than do rabbits and rodents (for review, see Lonsbury-Martin *et al.*, 1988).

[1]There is, in fact, evidence for two components of DPOAEs in humans (see, e.g., Brown *et al.*, 1996). However, these do not appear to be analogous to the two components seen in small mammals. Rather, in humans, in addition to the ''main'' DPOAE component, there is a ''secondary'' component apparently originating from around the DPOAE-frequency place, and which may be thought of as an interaction of cochlear-generated distortion-products with the stimulus-frequency emission generator at the DP-frequency place. This secondary component shows different properties to both the active/low-level/peak and passive/high-level/basal components seen in small mammals.

Brown, A. M., Harris, F. P., and Beveridge, H. A. (**1996**). ''Two sources of acoustic distortion products from the human cochlea,'' J. Acoust. Soc. Am. **100**, 3260–3267.

He, N., and Schmiedt, R. A. (**1993**). ''Fine structure of the $2f_1-f_2$ acoustic distortion product: Changes with primary level,'' J. Acoust. Soc. Am. **94**, 2659–2669.

Lonsbury-Martin, B. L., Martin, G. K., Probst, R., and Coats, A. C. (**1988**). ''Spontaneous otoacoustic emissions in a nonhuman primate. II. Cochlear anatomy,'' Hearing Res. **33**, 69–94.

Lonsbury-Martin, B. L., Harris, F. P., Stagner, B. B., Hawkins, M. D., and Martin, G. K. (**1990**). ''Distortion-product emissions in humans: II. Relations to stimulated and spontaneous emissions and acoustic immittance in normally hearing subjects,'' Ann. Otol. Rhinol. Laryngol. **99** (Suppl. 147), 14–28.

Mills, D. M. (**1997**). ''Interpretation of distortion product otoacoustic emission measurements. I. Two stimulus tones,'' J. Acoust. Soc. Am. **102**, 413–429.

Mills, D. M., and Rubel, E. W. (**1994**). ''Variation of distortion product otoacoustic emissions with furosemide injection,'' Hearing Res. **77**, 183–199.

Moulin, A., Collet, L., and Morgon, A. (**1992**). ''Influence of spontaneous otoacoustic emissions (SOAE) on acoustic distortion product input/output functions: Does the medial efferent system act differently in the vicinity of an SOAE?,'' Acta Oto-Laryngol. **112**, 210–214.

Norton, S. J., and Rubel, E. W. (**1990**). ''Active and passive ADP components in mammalian and avian ears,'' in *Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer-Verlag, New York), pp. 219–226.

Sun, X.-M., Schmiedt, R. A., He, N., and Lam, C. F. (**1994**). ''Modelling the fine structure of the $2f_1-f_2$ acoustic distortion product. I. Model development,'' J. Acoust. Soc. Am. **96**, 2166–2174.

Talmadge, C., Tubis, A., Piskorski, P., and Long, G. (**1995**). ''Modeling distortion-product otoacoustic emission fine structure in humans,'' J. Acoust. Soc. Am. **97**, 3413.

Whitehead, M. L., Lonsbury-Martin, B. L., and Martin, G. K. (**1992a**). ''Evidence for two discrete sources of $2f_1-f_2$ distortion-product otoacoustic emission in rabbit: I. Differential dependence on stimulus parameters,'' J. Acoust. Soc. Am. **91**, 1587–1607.

Whitehead, M. L., Lonsbury-Martin, B. L., and Martin, G. K. (**1992b**). ''Evidence for two discrete sources of $2f_1-f_2$ distortion-product otoacoustic emission in rabbit: II. Differential physiological vulnerability,'' J. Acoust. Soc. Am. **92**, 2662–2682.

Whitehead, M. L., Lonsbury-Martin, B. L., Martin, G. K., and McCoy, M. J. (**1996**). ''Otoacoustic emissions: Animal models and clinical observations,'' in *Clinical Aspects of Hearing, Volume VI of the Springer Series in Auditory Research*, edited by T. R. Van De Water, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 199–257.

Whitehead, M. L., Stagner, B. B., McCoy, M. J., Lonsbury-Martin, B. L., and Martin, G. K. (**1995**). ''Dependence of distortion-product otoacoustic emissions on primary levels in normal and impaired ears. II. Asymmetry in L_1, L_2 space,'' J. Acoust. Soc. Am. **97**, 2359–2377.

Wier, C. C., Pasanen, E. G., and McFadden, D. (**1988**). ''Partial dissociation of spontaneous otoacoustic emissions and distortion products during aspirin use in humans,'' J. Acoust. Soc. Am. **84**, 230–237.

PROGRAM OF

The 16th International Congress on Acoustics and 135th Meeting of the Acoustical Society of America

The Sheraton Seattle Hotel and The Westin Seattle Hotel • 22–26 June 1998

Technical sessions will be held at two hotels—The Sheraton Seattle and The Westin Seattle. The hotel where each session will take place is noted with an (S) for Sheraton or (W) for Westin after the room name in each session heading.

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, 22 JUNE 1998 GRAND BALLROOM III (W), 8:00 TO 9:00 A.M.

Session 1aPLa

Plenary Lecture

James L. Flanagan, Chair
Research, Rutgers University, Busch Campus, Core Building 706, Piscataway, New Jersey 08855-1390

Chair's Introduction—8:00

Invited Paper

8:05

1aPLa1. Future directions in speech information processing. Sadaoki Furui (Dept. of Comput. Sci., Tokyo Inst. of Technol., 2-12-1, Ookayama, Meguro-ku, Tokyo, 152 Japan, furui@cs.titech.ac.jp)

Speech processing technologies, including speech recognition, synthesis, and coding are expected to play important roles in an advanced multimedia society with user-friendly human–machine interfaces. Speech recognition systems include not only those that recognize messages but also those that recognize the identity of the speaker. This paper predicts future directions in speech processing. It describes the most important research problems and tries to forecast where progress will be made in the near future and what applications will become commonplace as a result of the increased capabilities. The most promising application area is telecommunications. To solve various fundamental problems, a unified approach across speech recognition, synthesis, and coding is indispensable. Handling the common phenomenon of voice individuality from different aspects are: research on speaker adaptation in speech recognition, automatic speaker verification, voice conversion in speech synthesis, and the problems of very low-bit-rate speech coding quality variation from speaker to speaker. The methods that essentially solve such problems should be based on a common mathematical model. Research on the mechanism of speech information processing in our human brains, that is, how meanings of speech are understood and how speech is produced, is also crucial for epoch-making technological development in conversational speech understanding and natural speech synthesis.

Session 1aAA

Opera House Acoustics Workshop—CIARM, MCHA and Architectural Acoustics: Sound Field for the Audience, Part 1

Yoichi Ando, Cochair
Graduate School of Science and Technology, Kobe University, Rokkodai, Nada, Kobe, 657 Japan

Michael Barron, Cochair
Department of Architecture and Civil Engineering, University of Bath, Bath BA2 7AY, England

Chair's Introduction—9:15

Invited Papers

9:20

1aAA1. Spatial impression as measured in concert and opera auditoria. Michael Barron (Dept. of Architecture and Civil Eng., Univ. of Bath, Bath BA2 7AY, UK)

A coherent set of subjective experiments (published in 1973 and 1981) established the effect of spatial impression associated with early lateral reflections. The early lateral energy fraction was proposed as an objective measure. Some authors have preferred to use measures linked to cross correlation, presumably on the basis that this is probably closer to the mechanism used by the ear. The paper reviews the results of measurements of the early lateral energy fraction in 17 British concert spaces, as well as considering the measured results by Bradley in 13 North American concert halls. The discussion considers the relationship between early lateral energy fraction and hall width, the behavior of the measure close to the source and in both parallel-sided halls and fan-shaped halls. The design requirements for high spatial impression are also considered. The subjective effect of early lateral reflections is also influenced by sound level. The consequences for design of considering this additional variable in combination with the directional distribution will also be discussed. Spatial impression in concert halls and opera houses will be examined.

9:40

1aAA2. Effects of front/back energy ratios of early and late reflections on listener envelopment. Masayuki Morimoto (Environ. Acoust. Lab., Faculty of Eng., Kobe Univ., Nada, Kobe, 657 Japan, mrmt@kobe-u.ac.jp) and Kazuhiro Iida (Matsushita Comm. Ind., Yokohama, Japan)

A previous paper [M. Morimoto and K. Iida, J. Acoust. Soc. Am. **93**, 2282 (1993)] indicated that the listener envelopment (LEV) grows as the energy of reflections coming from the back of the listener increases. This paper investigates which is more effective for LEV, the F/B ratio of early or late reflections. In the experiments, a music motif is used as a source signal. The F/B ratios in the early and late reflections are changed independently, but the C value, the degree of interaural cross correlation, and the sound-pressure level of all sound fields are kept constant. The experiments were performed by using a paired comparison test. The task of the subjects was to judge which sound field had higher LEV. The experimental results indicate that the F/B ratio of both of the early and late reflections have an effect on LEV but that the latter is a little more effective for LEV than the former. This result suggests a relation between the perception of LEV and the law of the first wavefront.

10:00

1aAA3. Sound strength in concert halls: The role of the early sound field on objective and subjective measures. Jerald R. Hyde (Consultant on Acoust., Box 55, St. Helena, CA 94574)

A very important objective parameter in the acoustical understanding of concert halls is the total sound strength G which is essentially the acoustical ''gain'' of the space, normalized to source power. Measurement data are not well predicted by current theories and indicate that there must be other component factors of G in addition to the factors of distance (r) and room constant (A) of current theory. Measurements show that significant changes in G, independent of distance, are found to be directly proportional to the variation in the early reflected level G_{50} (refl.). To check this finding, a test was conducted by adding a reflector within the reverberant boundary of a hall. A significant increase in G was found in the seating area where the G_{50} (refl.) was increased. While this may seem obvious, current theory does not predict this. The factor G is therefore also found to be a function of early reflected energy where such energy acts in effect as an amplification of the source. The area receiving this early energy is energized by this amplification factor. Measurements show that specific design to provide the early reflected energy can increase G_{50} (refl.) by 2 to 3 dB and effectively amplify G by up to 2 dB. Subjective tests on the perception of loudness will be reported showing that these increases are quite significant and well above the subjective limit for loudness. A relationship with hall quality and improved ''intimacy'' is suggested with this design approach.

10:20

1aAA4. Subjective attributes for sound fields from a model of the auditory-brain system. Yoichi Ando, Hiroyuki Sakai, and Shinichi Sato (Grad. School of Sci. and Technol., Kobe Univ., Rokkodai, Nada, Kobe, 657 Japan)

Based on physiological responses, a model of the auditory-brain system may be proposed. The model consists of the autocorrelation mechanisms, the interaural cross-correlation mechanism between the two auditory pathways, and the specialization of human cerebral hemispheres for temporal and spatial factors of sound field. The specialization of the human cerebral hemisphere may relate to the highly independent contribution between the spatial and temporal criteria of sound fields on subjective preference judgments. In addition, based on the model, any other subjective attributes of sound fields can be described in terms of processes of the auditory pathways and the brain. For example, the phenomena of missing fundamental, loudness and coloration, the threshold of perception, and preferred delay of a single reflection are well described by the autocorrelation mechanism. The subjective diffuseness and ASW are described by the interaural crosscorrelation function. Even so-called ''cocktail party effects'' may well be explained by such specialization of the human brain, because speech is processed in the left hemisphere, and the directional spatial information is mainly processed in the right hemisphere. [Work was partially supported by the Ministry of Education, Grant-in-Aid for Scientific Research (C), 9838022, 1997.]

10:40

1aAA5. Reflected sound field by stage floors in the concert halls: Theoretical model analysis. Shinsuke Nakanishi, Kimihiro Sakagami, and Masayuki Morimoto (Environ. Acoust. Lab., Faculty of Eng., Kobe Univ., Rokko, Nada, Kobe, 657 Japan)

This paper theoretically analyzes a reflected sound field by the stage floor which is modeled as an infinite elastic plate with a point force excitation. Models of infinite elastic plates backed with either an air cavity or an absorptive cavity, which are often seen in stage floors of concert halls, are also considered. The far-field expression for the solution to the reflected field is derived by coupling the Helmholtz integrals for the sound fields surrounding the plate and the equation of motion for the plate. The main features of the reflected field are as follows: (1) A significant peak appears around the coincidence frequency. (2) The reflected field fluctuates due to interference between the reflection by the plate without the point force and the radiation from the plate excited by the point force. (3) The reflection characteristics show the panel absorption dip caused by a mass-spring resonance of the plate-cavity system and many small dips and peaks due to acoustic resonances of the back cavity. Effects of the parameters of the plate and cavity on the reflected sound field are discussed through numerical examples.

11:00

1aAA6. A fast room acoustical simulation algorithm based on the free path distribution. Michael Vorlaender (Inst. of Tech. Acoust., RWTH Aachen, 52056 Aachen, Germany, mvo@akustik.rwth-aachen.de)

A new algorithm is presented which provides estimates of impulse responses in arbitrary shaped rooms. It is applicable to any kind of room shape including nondiffuse spaces like workrooms or offices where, for instance, sound propagation curves are of interest. In the case of concert halls and opera houses it enables very fast predictions of room acoustical criteria like reverberation time, strength, or clarity. The algorithm is based on a conventional statistical ray tracing. However, a very low number of rays is used. During the ray tracing procedure, the distribution of free paths is recorded by considering the entire series of reflections per ray. Furthermore, the transition probabilities between free path classes and the absorption coefficients involved are stored. After the ray tracing, the free path distributions are evaluated by post-processing to create echograms. They include receiver-dependent direct sound and superposed exponential functions corresponding to classes of free paths. It is also possible to create reverberation tails without assuming *a-priori* diffuse sound fields and purely exponential decays.

Contributed Papers

11:20

1aAA7. The relationship between subjective preference for sound fields with vocal music and brain activity. Kiminori Mouri (Grad. School Sci. Technol., Kobe Univ., Rokkodai, Nada, Kobe, 657 Japan) and Yoichi Ando (Kobe Univ., Rokkodai, Nada, Kobe, 657 Japan)

The purpose of this study is to reconfirm the theory that the subjective preference for music sound fields depends on its autocorrelation function (ACF) of the source signal by means of analyzing continuous α-brain waves. Ando's theory describes that the most preferred delay time of early reflections and reverberation time are calculated by the effective duration (τ_e) of the ACF of music signals [Y. Ando, J. Acoust. Soc. Am. **74**, 873–887 (1983)]. On the other hand, it was demonstrated that the τ_e of the ACF of the α-brain wave range (8–13 Hz) from the left and right temporal lobes was changed by temporal factors relating to the subjective preference of sound fields [Y. Ando and C. Chen, J. Archit. Plann. Environ. Eng. AIJ **488**, 67–73 (1996); C. Chen and Y. Ando, *ibid.* **489**, 73–80 (1996)]. When a sound field has a constant delay time as a usual condition, it is assumed that continuous α-brain waves may be affected sequentially by the different running τ_e values of the ACF of the source signal.

11:35

1aAA8. New diagnostic system measuring orthogonal factors and calculating subjective attributes of sound fields in a real room. Masatsugu Sakurai (Yoshimasa Electron., Inc., Yoyogi, Shibuya, Tokyo, 151 Japan), Shinichi Aizawa (Yoshimasa Electron., Inc., Kobe, 657 Japan), and Yoichi Ando (Kobe Univ., Kobe, 657 Japan)

Based on the model of the auditory-brain system which consists of the autocorrelation mechanism, the interaural cross-correlation mechanism between the both auditory pathways, and the specialization of human cerebral hemispheres, a new diagnostic system was developed. The system works on a PC for Windows, thus there is no need for special additional devices. After obtaining the impulse response, four orthogonal factors including the SPL, the initial time delay gap between the direct sound and the first reflection, the subsequent reverberation time, and IACC are analyzed for the calculation of the scale values of global and individual subjective preference. In addition to the four factors, two more factors, τ_{IACC} and $\mathbf{W}_{\mathrm{IACC}}$, in the interaural cross-correlation function are figured out for evaluating the image shift of the sound source and the apparent source width (ASW), respectively. Also, the effective duration, τ_e, and fine structures of autocorrelation function of sound signals including the value of first maximum, ϕ_1, and its delay time τ_1, are analyzed. Keeping the

subjectively optimal condition, this system may be utilized for automatic control of sound fields by electroacoustic systems. The new system will be demonstrated at the presentation of this paper. [Work supported by the Ministry of Education, Grant-in-Aid for Scientific Research.]

11:50

1aAA9. Absorption by seating and audience—Whose figures, whose measurement method? Michael Barron and Steven Coleman (Dept. of Architecture and Civil Eng., Univ. of Bath, Bath BA2 7AY, UK)

Measuring the absorption coefficients of seating and audience is no easy matter because of the difference in size between auditoria and standard test chambers. An optimum test method remains still to be established. Two authors have published papers on this issue recently in the *Journal of the Acoustical Society of America*. The designer, though, is left confronted by contradicting figures and test proposals. Bradley (1992, 1996) has proposed a technique which involves measuring about five different configurations of seats to establish the absorption of infinite seat areas and absorption by the perimeter. Davies *et al.* (1994) deal with the perimeter absorption by testing with and without barriers placed over the front and sides of a test block of seating. Neither method is ideal; Bradley's depends on extrapolation to the smaller ratios of perimeter to area found in real auditoria, while Davies' method assumes that the barriers do not introduce any new problems of their own. Beranek in his new book (1996) has reanalyzed measured data from full-size concert halls to produce revised absorption figures. This paper seeks to find what common ground there is between the measurement methods and published data. It will draw on results of some tests on model seating in a model test chamber.

MONDAY MORNING, 22 JUNE 1998 GRAND BALLROOM A (S), 8:00 TO 10:30 A.M.

Session 1aAO

Acoustical Oceanography and Underwater Acoustics: Long-Range Propagation for Measurement of Ocean Processes I

Brian D. Dushaw, Chair

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

Chair's Introduction—8:00

Invited Papers

8:05

1aAO1. Acoustic thermometry of ocean climate: Comparison of acoustic, altimetric, and historical data. Robert C. Spindel (ATOC Group) (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105-6698, spindel@apl.washington.edu), B. D. Cornuelle, M. Dzieciuch, W. H. Munk, P. F. Worcester (Univ. California, La Jolla, CA 92093-0225), A. B. Baggeroer, D. Menemenlis, C. Wunsch (MIT, Cambridge, MA 02143), T. G. Birdsall, K. Metzger (Univ. Michigan, Ann Arbor, MI 48109-2122), C. Clark (Cornell Univ., Ithaca, NY 14850), J. A. Colosi (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), D. Costa (Univ. California, Santa Cruz, CA 95064), B. D. Dushaw, B. M. Howe, J. A. Mercer, R. C. Spindel (Univ. Washington, Seattle, WA 98105-6698), and A. M. G. Forbes (CSIRO, Marine Labs., Hobart, Tasmania 7000, Australia)

Approximately 15 months of data are now available from the 75-Hz ATOC transmissions from the Pioneer Seamount off Half Moon Bay, CA to horizontal arrays at 11 U.S. Navy horizontal arrays in the northeast Pacific, two vertical arrays, one near Hawaii and the other near Kiritimati (Christmas) Island. The data demonstrate that ocean basin temperatures can be determined with millidegree precision. Travel time variations at tidal frequencies agree well with predicted values. Low-frequency travel time variations provide range and depth averaged temperatures that are comparable to seasonal temperature variations derived from historical oceanography (Levitus94). They are also comparable (but not identical) to temperature variations derived from contemporaneous sea surface height variations measured by satellite altimetry. Preliminary results of combining the acoustic and altimetric data in a numerical general circulation model of the North Pacific are also available.

8:25

1aAO2. Computation of *T*-phase coda. Catherine D. de Groot-Hedlin (Scripps Inst. of Oceanogr., UCSD, La Jolla, CA 92122-0225)

T-phases result from the conversion of seismic to acoustic energy at the seafloor in the vicinity of an earthquake epicenter. The time-frequency characteristics of *T*-phase coda are observed to be strongly dependent upon the bathymetry of the ocean floor in the source region; thus earthquakes from the same geographic area generate similar *T*-phase coda. In this study, detailed bathymetry and sound speeds computed from the Levitus atlas are used to simulate the *T*-phase amplitudes which would be recorded at distances of several thousand kilometers from these natural sound sources. It is hypothesized that seismic/acoustic coupling occurs over a distributed region of the ocean floor, with seismic energy being scattered into low-order modes in shallow water. Comparison with observed coda indicates that *T*-phase onset times and durations are modeled quite accurately with this simple method. The feasibility of using this method to determine accurate source times and locations for small suboceanic earthquakes, given natural ocean variability, is examined in this study.

8:45

1aAO3. Coherence analysis of multimegameter range ocean acoustic signals. Matthew Dzieciuch and the ATOC Consortium[a] (Scripps Inst. of Oceanogr., Univ. of California, San Diego, 1GPP-0225, 9500 Gilman Dr., La Jolla, CA 92093)

The quality of acoustic travel time measurements depends on the coherence of the signal as well as the bandwidth. Ocean acoustic signals exhibit coherence properties in four ways: in time, in bandwidth, and in vertical and horizontal extent. (These are conjugate to Doppler broadening, pulse time spread, and rms vertical and horizontal arrival angle.) These properties have recently been measured and can be compared for several different ranges (80–5000 km), center frequencies (28–250 Hz), and geometries (source on-bottom versus off-bottom). For signals of less than perfect coherence, techniques from communications theory can be used to improve the quality of travel time measurements by relaxing the resolution requirement. For instance, if the coherence bandwidth is less than the signal bandwidth, it is possible to consider subbands of the signal as separate measurements. How to best combine these measurements then depends crucially on the SNR. While acoustic thermometry has relied on electronically controlled sources, by allowing for the coherence structure of stochastic sources (such as T-phases, explosives, or marine mammals) an improved signal processing strategy may result. [a]The ATOC Consortium: A. B. Baggeroer (MIT), T. G. Birdsall (U. Michigan), C. Clark (Cornell), J. A. Colosi (WHOI), B. D. Cornuelle (SIO), D. Costa (UCSC), B. D. Dushaw (U. Washington), M. Dzieciuch (SIO), A. M. G. Forbes (CSIRO, Hobart), C. Hill (MIT), B. M. Howe (U. Washington), J. Marshall (MIT), D. Menemenlis (MIT), J. A. Mercer (U. Washington), K. Metzger (U. Michigan), W. Munk (SIO), R. C. Spindel (U. Washington), D. Stammer (MIT), P. F. Worcester (SIO), and C. Wunsch (MIT).

9:00

1aAO4. Internal wave effects on single-frequency coherence, matched field processing, and adiabaticity. Kevin D. Heaney, Peter N. Mikhalevsky, Herb Freese (Sci. Applications Intl. Corp., 10260 Campus Pt. Dr., MS C4A, San Diego, CA 92121), and W. A. Kuperman (Scripps Inst. of Oceanogr., UCSD, La Jolla, CA 92093)

A set of PE calculations at various frequencies (10, 30, 75, and 150 Hz) through internal wave realizations was performed. To examine the range/frequency limits of coherent processing, the coherence against the range-independent PE was calculated every km out to 3000 km. The coherence between with and without internal waves drops to 0.8 (-2 dB) at ranges of 3000, 1800, 500, and 200 km for the 10, 30, 75, and 150 Hz, respectively. The sidelobe structure of matched-field processing results confirms these as range/frequency limits of matched-field processing. Vertical coherence as a function of range and frequency will be presented. The decomposition of the full field into normal mode amplitudes will be presented, as well as range limits for the adiabatic approximation. The validity of the first-order perturbation theory correction to the adiabatic approximation will be examined.

9:15–9:30 Break

9:30

1aAO5. Modal leakage in range-dependent oceans. Arthur B. Baggeroer (MIT, 77 Massachusetts Ave., Bldg. 5-204, Cambridge, MA 02139) and Edward K. Scheer (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

The validity of the adiabatic approximation for modal propagation at low frequencies and over long megameter ranges is evaluated. A dynamical ocean model for the sound speed profile perturbations is superimposed upon a canonical Munk profile to represent a range-dependent ocean. The channel is excited with a single mode and propagated with the Collin's FEPE code which was tuned to conserve the energy in each mode in a range-independent ocean. Modal amplitudes were determined by projecting the field on the local range-dependent modes. If one uses displacements which correspond to nominal experimental levels in the Atlantic, mode coherence is maintained only for ranges less than 1 megameter at 75 Hz; the energy appears to approach equipartitioning. Results starting with mode 1 and mode 10 were qualitatively the same. If the displacements are reduced by a factor of three, possibly more representative of Pacific levels, modal coherence is maintained for more than four megameters.

9:45

1aAO6. Coupled-mode sound propagation in a range-dependent, moving fluid. Oleg A. Godin (School of Earth and Ocean Sci., Univ. of Victoria, P.O. Box 1700, Victoria, BC V8W 2Y2, Canada, ogodin@uvic.ca)

Full-field acoustic methods for current velocity inversion require accurate and efficient mathematical models of sound propagation in a range-dependent waveguide with flow. In this paper, an exact coupled-mode representation of the acoustic field is derived for the corresponding 2-D problem. To account for the physics of the problem, normal modes in a corresponding range-independent waveguide are chosen as the local basis. Unlike the motionless case, vertical dependencies of acoustic pressure in individual normal modes are not orthogonal in the presence of currents. To overcome this difficulty, a five-dimensional state vector is introduced. Orthogonality of the state vectors corresponding to individual normal modes is established. In terms of the state vector, the set of linearized equations of the hydrodynamics becomes a single differential equation of the first order with respect to the range coordinate. By using the orthogonality, coupled equations are derived for range-dependent mode amplitudes. The resulting mode-coupling equations have the same form as those known for motionless case, but values of mode-coupling coefficients differ. The contributions to mode coupling from horizontal gradients in the sound speed, density, and flow velocity are isolated and analyzed. [Work suported by NSERC and ONR.]

10:00

1aAO7. Inversion of multimegameter-range acoustic data for ocean temperature. Brian D. Dushaw (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105-6698)

Ocean sound speed (temperature) derived from ray travel times obtained from multimegameter-range acoustic transmissions may be inaccurate when the reference ocean state is inadequate for linearized inversion. When the reference (e.g., the Levitus atlas) is significantly different than the ''true'' ocean, the ray paths defined by that reference inaccurately represent the true sampling. In addition, natural oceanic variations, such as the evolution of a summer mixed layer, can significantly change the ray sampling. The guiding principle for inversion is that ray travel times associated with the inverse solution must match the measured travel times. Inversion of travel times relative to an arbitrary mean can be inaccurate. Accurate inversions may be obtained by iterative methods or by a better choice of reference ocean, but the solution of this nonlinear problem is not always obvious. Inadequate depth resolution for the vertical structure in the inverse solution can make calculation of travel times difficult. A linearizing time-dependent reference ocean can reduce the solution uncertainties since the *a priori* model variances are less. Simulations employing the Levitus ocean atlas, so that the ''true'' ocean state is known, are used to define the problems of nonlinear inversion and to find accurate inversion methods.

10:15

1aAO8. Ray tracing in a turbulent, shallow-water channel. Christian Bjerrum-Niese, René Lützen, and Leif Bjørnø (Dept. of Industrial Acoust., Bldg. 425, Tech. Univ. of Denmark, 2800 Lyngby, Denmark, cbn@ipt.dtu.dk)

A ray tracing model can be used to simulate sound (10–100 kHz) transmitted through shallow water. The phase of the ray arrivals, primarily given by travel time, may be mutually independent in such a multipath

transmission. Consequently, the transmission loss in a receiving point is randomly valued due to the coherent interference of the multipath arrivals. This problem can be overcome by incoherent summation of the multipath arrivals. However, knowing that nature behaves coherently, this method is not preferred. Instead, the channel can be regarded as dynamic by allowing microfluctuations of the sound speed. Specifically, the channel may be modeled as given by horizontal layers, each assigned with an individual turbulent dissipation rate, and a translation of this is performed into diffraction and phase fluctuation parameters, following DiIorio and Farmer [J. Acoust. Soc. Am. **96**, 1056–1069 (1994) and Flatté *et al.*, *Sound transmission through a fluctuating ocean* (1979)]. Amplitude and phase fluctuations of every ray are realized as samples of a random process. The results for the dynamic channel show a smoother and easier interpretable transmission loss behavior than for the static channel. Furthermore, it is observed that amplitude fluctuations generally exhibit Ricean fading. [Work sponsored by the Danish Technical Research Council.]

MONDAY MORNING, 22 JUNE 1998 EAST BALLROOM A (S), 8:30 TO 10:45 A.M.

Session 1aBV

Bioresponse to Vibration/Biomedical Ultrasound: Novel Approaches to Ultrasound Imaging I

John M. Reid, Cochair
16711 254th Avenue, SE, Issaquah, Washington 98027

Shira L. Broschat, Cochair
School of Electrical Engineering and Computer Science, Washington State University, P.O. Box 642752, Pullman, Washington 99164-2752

Chair's Introduction—8:30

Invited Papers

8:35

1aBV1. Flow function method for the display of streamlines on ultrasonic color flow image. Shigeo Ohtsuki (Precision and Intelligence Lab., Tokyo Inst. of Technol., 4259 Nagatuta, Midori-ku, Yokohama, 226 Japan) and Motonao Tanaka (Tohoku Welfare Pension Hospital, Sendai, 983 Japan)

Ultrasonic color Doppler technique is useful for the observation of the blood flow in heart. The information on blood flow by this technique is Doppler velocity, which is the velocity component in the direction of the beam emitted from the ultrasonic probe. The ultrasonic color Doppler image displays the distribution of Doppler velocity on an observation plane using red and blue corresponding to toward flow and away flow, respectively. Streamlines overlapped on ultrasonic color Doppler image are considered to be suitable for the quantitative evaluation of the characteristics of the flow. In this investigation, a new concept of the flow function is proposed in order to obtain streamlines on an arbitrary section of three-dimensional flow for visualizing and evaluating the characteristics of the flow. Flow function at a point on a plane is defined as the flow rate along a path from a reference point. This function is a multivalued function. The section of the three-dimensional display of the function at a constant value shows streamlines. The streamlines separated by an arbitrary unit flow can be fixed with the approximated point sources of the unit flow. These point sources are named simple sources. These simple sources are start points or end points of streamlines.

8:55

1aBV2. Noninvasive assessment of shear rate (WSR) in humans by means of ultrasound. R. S. Reneman, L. Kornet, S. K. Samijo, P. J. Brands (Cardiovascular Res. Inst. Maastricht (CARIM), Dept. of Physiol., Maastricht Univ., P.O. Box 616, 6200 MD Maastricht, The Netherlands), and A. P. G. Hoeks (The Netherlands)

An ultrasound technique to assess velocity and shear distribution in arteries noninvasively was developed . Accurate assessment of velocities close to the wall, a prerequisite for reliable determination of WSR, is achieved by tracking of the arterial wall and adaptive filtering of the wall signal. Velocity (v) distribution in the vessel is obtained by RF cross correlation of the signals in between the walls, and determination of dv/dr throughout the cardiac cycle provides shear distribution. Maximum dv/dr near the wall is taken as peak WSR and average shear over the cardiac cycle as mean WSR. Wall shear stress (WSS) is estimated by multiplying WSR by whole blood viscosity. The system also allows assessment of arterial diameter, diameter changes, and intima-media thickness (IMT). In studies on humans it could be shown that at a younger age in the common carotid artery (CCA) MWSS is close to 1.5 Pa, a value in agreement with minimal expenditure of energy. MWSS, however, is not optimally regulated, because it decreases with age. In the CCA, WSR close to the bifurcations were lower than and associated with larger IMT than more upstream. Reflections from downstream are held responsible for these differences in WSR.

9:15

1aBV3. Two-dimensional measurement of acoustic properties of tissue using 3 to 50 MHz ultrasound. Hiroyuki Hachiya (Dept. of Information & Comput. Sci., Chiba Univ., 1-33 Yayoi-cho, Inage-ku, Chiba, 263 Japan, hachiya@ics.tj.chiba-u.ac.jp), Shigeo Ohtsuki (Tokyo Inst. of Technol., Yokohama, 227 Japan), and Motonao Tanaka (Tohoku Kousei-Nenkin Hospital, Sendai, Japan)

Application of ultrasound has been expanded in the medical field and used widely for diagnosis. Acoustic properties such as the sound speed of a bulk wave in biological tissues are important parameters for quantitative characterization of living tissues. But the relationship between ultrasonic characteristics and pathohistological changes in tissues is poorly understood. To investigate the relationship, we have proposed a high-precision noncontact measuring method of acoustic properties of tissues using 3- to 50-MHz ultrasound. In this method, the travel time of sound through the sample, and the difference in travel times from the transducer to the rigid reflector with and without the sample are estimated based on frequency-time analysis which does not require physical contact of ultrasonic probes to living or freshly excised tissue specimens. The components of multipath propagation in tissues are decomposed in a frequency complex plane. The travel time and the time difference are estimated using all waveform information; rms error of sound speed is less than 3 m/s. Two-dimensional measurement of acoustic properties in rat and human tissues is performed using the noncontact measurement method. The relation between the acoustic properties change and the physical structure change in tissues is getting clear.

9:35–9:45 Break

Contributed Papers

9:45

1aBV4. Three-dimensional Doppler ultrasound: A tool for the 21st century. Roy W. Martin (Anesthesiology and Bioengineering, Box 356540, Univ. of Washington, Seattle, WA 98195, rmartin@u.washington.edu), Dan Leotta, Xiang-Ning Li (Univ. of Washington, Seattle, WA 98195), and Trygve Hausken (Haukeland Univ. Hospital, Bergen, Norway)

Three-dimensional (3-D) ultrasound now offers a method to overcome former limitations of two-dimensional (2-D) Doppler. Important 3-D areas include the spatial delineation of the patency and pathways of blood vessels as well as the measurement of flow in them. In our laboratory a variety of 3-D uses has been investigated: detection and mapping of coronary arteries (3-D power Doppler, *in vitro*), mitral valve regurgitation (3-D color Doppler, *in vitro* and *in vivo*), and gastric flow (3-D color Doppler, *in vivo*). Several different 3-D approaches have been applied and results will be presented. However, in the future, it is believed that direct flow measurement can be achieved. Flow is equal to the integral of the velocities normal to a surface that cuts through a flow field. An arc of constant radius in a 2-D-sector scanner will inscribe an arc of a spherical surface, if rotated around the origin of the sector (e.g., multiplane transesophageal echocardiogram). The pulsed Doppler signals along the circumference of the inscribed arc represent the velocities normal to that surface. Integration of these velocities should provide direct measurement of the flow field cutting this inscribed surface. Algorithms for testing this strategy are under development.

10:00

1aBV5. Dynamic flow quantitation with spatial orientation guided digital color Doppler imaging: *In vitro* validation and initial *in vivo* experience. Xiang-Ning Li (Bioengineering, Box 356540, Univ. of Washington, Seattle, WA 98195, ning@u.washington.edu), Jing-Ming Jong, Trygve Hausken, Barbrina Dunmire, Bryan Goldman, Daniel F. Leotta, Kirk W. Beach, and Roy W. Martin (Univ. of Washington, Seattle, WA 98195)

The assumptions of geometric shape and uniform velocity profile prohibit the use of conventional Doppler methods for volume flow quantification with irregularly shaped nonvascular passages in space. A flow quantification method that integrates a three-dimensional (3-D) position and orientation sensing system with digital color Doppler imaging approach has been developed. Methods: The optimized ultrasound system settings were predetermined using a static, laminar flow model. The instantaneous flow rate was measured by the integration of the area-velocity product with angle correction. Initial *in vivo* studies quantified gastric emptying on five healthy volunteers 10 min after ingestion of a 500-ml meat soup, postprandially. Digital Doppler images were obtained continuously covering a transpyloric flow episode, immediately following the 3-D scan of the passage. The flow direction was derived from 3-D reconstruction using the centroids of two short-axis views. Results: *In vitro* validation showed the Doppler flow rates were within 10% for reference flow rates ranging from 6 to 30 ml/s. The average gastric emptying lasted 0.75 s with 3.0 ml in volume. Conclusions: With appropriate instrument settings, orientation determined digital Doppler can be used for the quantification of gastric emptying with adequate accuracy. This method minimized the angular ambiguity and geometric assumptions.

10:15

1aBV6. Assessment of vascular dialysis access with Doppler. Kirk W. Beach, Marla Paun, Suhail Ahmad, Mark Meissner (Univ. of Washington Medical Ctr., Box 356410, Seattle, WA 98195-6410), Robert Hickman, and Laurie Fauser (Children's Hospital, Seattle, WA)

Shunts between the arterial and venous systems are created to provide the blood flow required for hemodialysis. These shunts, created with interposed grafts or fistulas, often fail unexpectedly due to occluding thrombi. Such cases require surgical repair of the existing shunt or new shunt creation at an alternate site. Measuring pressure in the venous dialysis needle at a fixed flow rate is a standard method of surveillance. The method is not as effective as required because there is too much measurement variability. The variability is probably due to variability in placement of the dialysis needle. Neither ultrasonic duplex scanning nor x-ray angiography is viable for surveillance because the cost is too high. A simple ultrasonic Doppler method of measuring the change in flow between the dialysis needles as the dialysis pump is started and stopped can be used to compute the resistance of the graft during each dialysis. The cost is minimal, therefore making Doppler surveillance feasible. In addition, this method could be used to control the dialysis pump rate to an optimized value.

10:30

1aBV7. Using Doppler ultrasound to detect wall vibrations and flow velocity fluctuations in arteries. Melani I. Plett, Kirk W. Beach, and Marla Paun (Div. of Vascular Surgery, Dept. of Surgery, Univ. of Washington Medical Ctr., Box 356410, Seattle, WA 98195-6410)

Two arterial disorders, stenoses and perforations, result in fluctuations in both wall displacements and blood velocities that are not displayed by conventional ultrasound instruments. Wall vibrations with amplitudes on the order of a micron are too small to be displayed by M-mode. Narrowband velocity oscillations of less than 50-ms duration can easily be missed in viewing traditional Doppler spectral waveforms. Low-frequency velocity oscillations are attenuated by wall filters. Furthermore, the temporal relationship between flow velocities and wall displacements is not avail-

able in conventional systems, yet this relationship provides key information in studying stenoses and locating perforations. To simultaneously examine flow velocities and wall displacements in arteries, a customized ultrasonic, pulse-echo, multigate, quadrature phase demodulation system has been developed. The use of multigate, quadrature phase demodulation in the absence of a wall filter enables exploration of arterial behaviors temporally in the cardiac cycle with sufficient amplitude, frequency, and duration resolution to meet the rigorous demands of the analysis. Wall vibrations as high as 600 Hz with amplitudes less than a micron have been detected simultaneously with arterial jet velocities and flow oscillations. [This work was supported by a grant from DARPA, No. N00014-96-0630.]

MONDAY MORNING, 22 JUNE 1998 WEST BALLROOM B (S), 7:30 TO 10:45 A.M.

Session 1aEA

Engineering Acoustics: Underwater Acoustic Calibration, Methods, and Devices

Arnie L. Van Buren, Chair
Naval Undersea Warfare Center, Code 216, Building 1171, 1176 Howell Street, Newport, Rhode Island 02841

Invited Papers

7:30

1aEA1. Towards new UK underwater acoustical measurement standards in the 21st century. Roy C. Preston and Stephen R. Robinson (Ctr. for Mech. and Acoust. Metrology, Natl. Physical Lab., Teddington, Middlesex TW11 0LW, UK, rcp@npl.co.uk)

The importance of establishing common measurement standards to support underwater acoustical calibration and testing in the UK, and therefore the need to demonstrate calibration traceability, has progressively grown over the last few years. In the UK, the National Measurement System provides the means of achieving traceability through the establishment and dissemination of primary standards of measurement. This paper will outline the progress made so far in establishing the UK's underwater acoustical measurement standards, and look forward to plans to extend the capability to cover the whole frequency range of interest, over the ranges of temperature and hydrostatic pressure applicable to simulated ocean conditions. The validation of these measurement standards will also be discussed.

7:55

1aEA2. Acoustic projector calibration in reverberant environments: A survey of USRD measurement technology. S. E. Forsythe and P. L. Ainsleigh (Naval Undersea Warfare Ctr., Newport, RI 02841)

The use of pressure- and temperature-controlled tanks for calibration of Navy projectors is a cost-effective alternative to open-water or at-sea measurements. The Underwater Sound Reference Division (USRD) of the Naval Undersea Warfare Center maintains a large closed tank which is temperature and pressure controlled. This tank, like all closed vessels, suffers from a major drawback: it is highly reverberant, especially at low-frequency operation, which precludes the possibility of making steady-state projector response measurements (especially for highly resonant projectors with low-frequency resonances) as would be possible in a large body of open water. This presentation surveys the techniques developed at USRD to ensure that measurements made in the tank are traceable to those that would be made for the same projector in open water. Three broad categories of approaches will be presented: (1) estimating the steady-state response using a model that represents only the pre-echo portion of the received data (transient modeling), (2) estimating the echo-free steady state using a multiple-arrival model that accounts for echoes in the received data (multipath modeling), and (3) characterizing the wave field using multiple hydrophones to separate the direct wave field from the wave field due to echoes (spatial processing). [Work supported by ONR.]

8:20

1aEA3. Low-frequency calibration in water-filled pipes. L. D. Luker and J. F. Zalesak (Underwater Sound Reference Div., Naval Undersea Warfare Ctr. Div., Bldg. 1171B, Rm. 205, 1176 Howell St., Newport, RI 02841-1708)

For many years the Underwater Sound Reference Division has developed ever more powerful means for creating desired sound fields inside water-filled pipes in order to calibrate transducers. Early efforts resulted in manually controlled calibration chambers approximately 2 m long, creating standing-wave fields or traveling-wave fields propagating along the axis of the pipe at frequencies from 0.1 to 4000 Hz. These calibrators have been modified in recent years to measure also the acoustic properties of material samples and the wave fields are now computer controlled. These systems were followed by pipe-shaped calibrators that are submerged in troughs of water for checking towed-array sensor elements. The latest calibrator to be developed is the Long-Line Hydrophone Calibrator (LLHC). This calibrator has a chamber length of 57 m and is capable of calibrating hydrophone line arrays up to 49 m long at temperatures from 2 to 35°C and hydrostatic pressures up to 6.9 MPa. The LLHC uses transducers distributed along its length to provide wave fields equivalent to plane waves arriving from any angle of incidence at frequencies from 5 to 700 Hz. Research is currently underway to increase the frequency range of the LLHC to above 200 Hz. [Work supported by ONR.]

8:45

1aEA4. A comparison of hydrophone calibration by free-field reciprocity and by optical interferometry. Stephen P. Robinson and Roy C. Preston (Ctr. for Mech. and Acoust. Metrology, Natl. Physical Lab., Teddington, Middlesex TW11 0LW, UK, rcp@npl.co.uk)

As the national standards laboratory in the UK, the National Physical Laboratory has established two primary standard methods for the free-field calibration of hydrophones. In the frequency range 500 kHz to 20 MHz, the primary standard is realized using optical interferometry; in the range from 5 to 500 kHz, it is realized by the method of three-transducer spherical-wave reciprocity. Although the ranges of the primary standards are defined as above, both methods can be used in the frequency range 200 kHz to 1 MHz, enabling a comparison to be undertaken between these two independent methods. The comparison was made by calibrating reference hydrophones using both methods and comparing the results. In general, the results agree to within ± 0.4 dB or better, well within the combined uncertainties of the methods. In this paper, a comparison of results is presented and a discussion given of the major sources of uncertainty in the methods.

9:00

1aEA5. An intercomparison of hydrophone calibrations within Europe. Stephen P. Robinson, Roy C. Preston, and Geraint J. Green (Ctr. for Mech. and Acoust. Metrology, Natl. Phys. Lab., Teddington, Middlesex TW11 0LW, UK, rcp@npl.co.uk)

Intercomparisons between different laboratories are an important way of validating calibration methods and indicating where improvements are required. This paper describes a European intercomparison of free-field hydrophone calibrations in the frequency range 10–315 kHz. A total of 12 participants from seven European countries took part by calibrating three reference hydrophones, with project coordination provided by the National Physical Laboratory, UK. The agreement between the results was generally encouraging, with a majority of the results lying within ± 1 dB of the Grand Mean. However, some large variations were observed which give cause for concern, and the uncertainties in the calibrations were typically underestimated by the participants with the maximum differences from the Grand Mean almost invariably exceeding the quoted overall uncertainties. A discussion is given of possible reasons for the disagreement, and of useful further work identified as a result of undertaking the intercomparison.

9:15

1aEA6. A simple two-projector procedure for producing impedance-controlled wave fields for transducer calibration. S. E. Forsythe and A. L. Van Buren (Naval Undersea Warfare Ctr., Newport, RI 02841)

When calibrating hydrophones in an open body of water, it is sometimes desirable to establish a local impedance in the vicinity of the hydrophone that is accurately known. Three examples that are often used are (1) an impedance characteristic of a plane wave in open water (the so-called rho-c condition), (2) a pressure-only field with no local particle velocity (infinite impedance), and (3) a velocity-only field with no local pressure (zero impedance). This presentation describes a simple apparatus consisting of two spherical sources (or sources that appear spherical at the frequency of interest) and the description of the ratio of drive amplitudes required to produce the desired impedance condition. The useful ensonified volume, that region over which the desired impedance condition holds, is described. Examples of application to the calibration of pressure-sensitive and pressure-gradient hydrophones are given. [Work supported by ONR.]

9:30

1aEA7. Underwater facility for automated experimentation and measurement. Carlos Ranz and Pedro Cobo (Instituto de Acustica, CSIC, Serrano 144 28006, Madrid, Spain, iacrg32@fresno.csic.es)

The underwater tank at the Instituto de Acustica, CSIC ($7.5\times4.5\times4.5$ m), was built in 1970; it had facilities useful for the work expected at that time. The tank had two bridges. In both bridges only the vertical coordinate and the 360° turn were automated. In current experimental work, including calibration, the actual instrumentation and the sensors' motion control, together with signal acquisition and manipulation, showed the need for a drastic change in the tank structure. The mechanisms of motion of both bridges are now fully automated and can be controlled via software. This control can imply positioning at a given field point, for any or both sensors, or along a specified trajectory, including mapping, in the four coordinates X, Y, Z, and angle. The new installation includes the simultaneous signal acquisition, the specific design of ''calibration pages,'' files management, storing of full working sessions, and remote control, e.g., through the web, of the whole tank, including operation and signal acquisition and management. The tank is expected to be inserted in the European facilities network. The new installation, funded by the PN-CYTMAR and CSIC, is described in this presentation.

9:45

1aEA8. Calibration and testing of a neutrally buoyant *u*-*u* intensity probe. Kevin J. Bastyr and Gerald C. Lauchle (Penn State Univ., Grad. Prog. in Acoust., P.O. Box 30, State College, PA 16804, zabour@acs.psu.edu)

The theory of operation and the methodology for calibration of an underwater acoustic intensity probe are explained. The intensity probe, which is classified as a *u*-*u* probe, consists of two moving coil velocity sensors mounted separately in neutrally buoyant, collinearly oriented bodies. This two-velocity sensor probe is used to determine intensity in a manner analogous to that of pressure gradient hydrophones. The acoustic intensity is shown to be the imaginary part of the cross spectrum measured between the output voltages of the velocity sensors, scaled by a complex, frequency-dependent calibration constant. Due to the vector nature of intensity, these complex calibration constants contain magnitude and phase information. To establish the calibration constant, an experiment was conducted that employed the *u*-*u* intensity probe and a reference pressure transducer in a standing wave tube. The intensity in the tube is known, and measurements made of this field with the *u*-*u* probe are in agreement with the prediction. [Work supported by the Office of Naval Research, Multi University Research Initiative on Acoustic Transduction, Materials, and Devices.]

10:00

1aEA9. Calibration of a neutrally buoyant *p*-*u* intensity probe. James A. McConnell (Penn State Univ., Grad. Prog. in Acoust., P.O. Box 30, State College, PA 16804, univibe@sabine.acs.psu.edu) and Gerald C. Lauchle (Penn State Univ., State College, PA 16804)

Theory and test methods are presented for calibrating an underwater acoustic intensity probe containing two pressure hydrophones and one moving coil velocity sensor. The intensity probe is classified as a *p*-*u* probe and is neutrally buoyant in fresh water. In contrast to *p*-*p* and *u*-*u* probes, it determines intensity via direct measurement of the acoustic pressure and particle velocity. Specifically, intensity is computed by adjusting the time-averaged rms cross spectrum between the transducer output voltages by an appropriate calibration curve. Because intensity is a vector quantity, both a magnitude and phase calibration is required. This is accomplished by a straightforward technique involving a single reference hydrophone and *a priori* knowledge of the acoustic field where the calibration is performed. A simple experiment was conducted using a standard-wave tube to verify the effectiveness of the calibration technique. In this experiment, the pressure and velocity sensors were calibrated, then

an intensity measurement was made. Results show that the calibration technique is highly effective. [Work supported by the Office of Naval Research and the Naval Air Warfare Center Aircraft Division via the DoD Small Business Innovation Research Program.]

10:15

1aEA10. Test chamber for determining damage thresholds for high-amplitude underwater sound exposure in animal models. Thomas N. Lewis, Peter H. Rogers, James S. Martin, George S. McCall II (The George W. Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405), Joey G. Lloyd, Henry P. Cotten, and Gary W. Caille (Georgia Tech Res. Inst., Atlanta, GA 30332-0810)

In order to determine the effects of low-frequency underwater sound on small animal models it is desirable to expose them to well-characterized fields which closely simulate the (locally) plane-wave open ocean stimulus. It is also desirable to produce pure pressure and acceleration stimuli to isolate the effects of the individual components of the acoustic plane wave. It is necessary to produce sufficiently strong signals to enable damage thresholds to be determined. The range of frequencies over which such stimuli must be produced is quite broad since it must include the actual band of interest (100–500 Hz) to examine damage on a cellular and tissue level, as well as at scaled frequencies (~1–5 kHz) to examine the effects of organ structure. The problem is subtler than one might imagine because of the high compliance of the test animals due to the air in their lungs. The lungs will drastically alter the acoustic field in the chamber and change the loading on the sources and the chamber could alter the vibrational response of the lung. Careful consideration is required to ascertain what is required of the chamber to assure valid measurements and to determine the true effective sound pressure level.

10:30

1aEA11. A free-flooding rare-earth iron hexagonal transducer. Rong Rong Zhao (Shanghai Marine Electron. Equipment Res. Inst., 390 Jian Guo East Rd., Shanghai 200025, PROC, lpeng@stn.sh.cn)

This paper describes the development of a free-flooding rare-earth iron hexagonal transducer with an outer size of $\phi 160 \times 160$ and a weight of 4.9 kg. The vibrations of the rare-earth iron hexagonal transducer and the free-flooding water cavity in it are theoretically analyzed and discussed. The measurement result shows that the transducer may send out a source level of 185 dB, and the resonance frequency of the transducer drops down from 1.90 kHz in air to 0.85 kHz in water because of the cavity resonance. The free-flooding water cavity can therefore not only withstand the hydropressure, but also make the resonance frequency drop down considerably. The free-flooding rare-earth iron hexagonal transducer shows promise to be a lightweight, high-power sound source for operation at great depths in the ocean. [Work supported by CSSC.]

MONDAY MORNING, 22 JUNE 1998 GRAND CRESCENT (W), 9:15 A.M. TO 12:25 P.M.

Session 1aMU

Musical Acoustics: Signal Modeling in Music Synthesis and Processing

Naotoshi Osaka, Cochair
NTT Basic Research Laboratory, 3-1 Wakamiya, Morinosato, Atsugi-shi, Kanagawa, 243-01 Japan

Julius O. Smith III, Cochair
CCRMA, Music Department, Stanford University, Stanford, California 94305-8180

Chair's Introduction—9:15

Invited Papers

9:20

1aMU1. A flexible analysis/synthesis tool for transient signals. Tony S. Verma and Teresa H. Y. Meng (Elec. Eng. Dept., Ctr. for Integrated Systems, Stanford Univ., Stanford, CA 95305)

An overview of a flexible analysis/synthesis tool for transient signals that effectively extends the spectral modeling synthesis (SMS) parametrization of signals from sinusoids+noise to sinusoids+transients+noise is given. The extended model, by explicitly handling transients, provides a more realistic and robust signal analysis/transformation/synthesis tool. Although others have pointed out the need for partitioning SMS into sinusoids+transients+noise, they have not provided a flexible model for transients. The need for an explicit transient model arises because SMS analyzes a signal by first modeling the sinusoidal components of the signal. SMS then subtracts these sinusoidal components from the original signal leaving a residual signal that contains transients+noise. One drawback of SMS is it models this transient+noise residual solely as filtered noise. Our transient model, when used on this residual, first detects where possible transients occur, models the transients, then removes transients from the signal. This leaves a second residual that is slowly varying noise. Attacks are well preserved by the transient model and the noise model now works more effectively. Becuase the transient model is the frequency domain dual to sinusoidal modeling, it retains flexibility and allows for a wide range of signal transformations.

9:45

1aMU2. Lemur: A bandwidth-enhanced sinusoidal modeling system. Kelly Fitz and Lippold Haken (CERL Sound Group, Univ. of Illinois at Urbana—Champaign, Champaign, IL)

Lemur models sounds as a collection of partials having amplitude, bandwidth, and frequency envelopes. Bandwidth envelopes allow the neat representation of noisy parts of sounds which produce clutter in purely sinusoidal analyses. Lemur associates noise energy with sinusoidal components that are nearby in frequency. Bandwidth envelopes record the fraction of the partial energy that is attributable to noise energy, as opposed to sinusoidal energy. Amplitude envelopes are calculated to include the energy added by the

bandwidth association process. When partials are pruned from the model, in the analysis process or during editing, Lemur preserves the overall energy by distributing the removed partials energy among neighboring partials. For sounds with a good spread (in frequency) of partials, Lemur can represent noise energy in the appropriate spectral regions while retaining the manipulability and homogeneity of a purely sinusoidal model. This homogeneity is important for implementing certain kinds of model-domain behaviors and real-time timbral manipulation. Lemurs synthesis includes a spectral line widening technique to spread energy away from each partial's center frequency according to bandwidth envelope values.

10:10

1aMU3. Synthesis and control of synthesis using a generalized diphone method. Xavier Rodet (Responsable de l'Equipe Analyse/Synthese Head of the Analysis/Synthesis Team, IRCAM 1 place Stravinsky 75004 Paris, France)

Well known in speech synthesis, a diphone is a segment of sound extending from the middle of a phone to the middle of the next phone. More precisely, in order to allow for time and pitch modifications, a diphone is the set of synthesis parameters representing this segment in some given synthesis technique. A speech sentence is synthesized by concatenating diphones found in a dictionary. The concept of diphone has been extended into generalized diphone control in order to build musical phrases from dictionaries of sound segments by concatenating and articulating them. Sound segments stored in dictionaries can comprise any recorded sound as well as instrumental sounds. Any analysis–synthesis method such as source-filter synthesis, additive synthesis, or physical model can be controlled in terms of generalized diphones. Presented here are several important improvements to diphone control and synthesis implemented in a software named ''Diphone,'' such as hierarchical structure of sequences and subsequences of diphones, variable speed of transitions, various smoothing functions at diphone boundaries, etc. ''Diphone'' proved itself a very flexible and inspiring tool for composition and synthesis, which composers are using at IRCAM and outside of the institution.

10:35–10:45 Break

10:45

1aMU4. Sound representation and modification with multiresolution sinusoidal models. Michael Goodwin, Paolo Pradoni, and Martin Vetterli (SGI, 2011 N. Shoreline Blvd., Mountain View, CA, mmg@esd.sgi.com)

In the standard sinusoidal model, the analysis and synthesis are typically carried out at fixed frame rates. This can result in representation error for signals with dynamic behavior. To improve the modeling performance for such signals, the sinusoidal analysis-synthesis can be carried out in a multiresolution framework based either on a filter bank structure or on adaptive time segmentation. In the latter method, which will be of primary interest here, the adaptation is guided by a metric such as rate-distortion. Such an approach improves the model accuracy while preserving the ability to carry out the rich class of modifications afforded by the standard sinusoidal model.

11:10

1aMU5. Timbre morphing and interpolation based on a sinusoidal model. Naotoshi Osaka (NTT Basic Res. Labs., Information Sci. Res. Lab., 3-1, Wakamiya, Morinosato, Atsugi-shi, Kanagawa, 243-01 Japan, osaka@idea.brl.ntt.co.jp)

There have been several attempts to synthesize the timbre in between two sounds. Several synthesis methods are reported to work well, but are not automatic throughout the whole process or are not precisely described if one wishes to trace the algorithm. An automatic algorithm of synthesizing morphed sounds and interpolated timbre has been reported. Two original sounds are analyzed using the M&Q algorithm. Correspondent partials of each sound are found using Dynamic Programming. Two synthesis methods are introduced using a sinusoidal model. One is to find correspondent partials within correspondent frames of two sounds. Then an interpolated partial of a new frame is calculated. Peak-to-peak match of sinusoidal parameters is done for a new frame and synthesized sounds are acquired. The other method is the one which changes the procedure sequence of the first method. Correspondence is considered for sequent partials and not a framewise signal. These two methods are compared with another simple synthesis method, that is, wave-to-wave interpolation of one pitch period. Total performance of sounds are discussed using subjective test results.

11:35

1aMU6. Tracking the frequency components of musical tones based on global waveform fitting. Yinong Ding (DSPS RDC, Texas Instruments, Inc., 8330 LBJ Freeway, MS 8374, Dallas, TX 75243, ding@ti.com)

A novel approach to estimating sinusoidal parameters of musical tones is presented. The proposed algorithm uses a combined quadratic polynomial phase sinusoid and residual signal model, and a waveform fitting criterion, as opposed to the traditional energy fitting one as employed by short-time Fourier transform-based approaches. The proposed algorithm obtains sinusoidal parameters by minimizing the squared difference between the original and the synthesized musical tones. The signal synthesis is implemented implicitly and the synthesis errors are taken into account while the sinusoidal paremeters are estimated. A number of different phase models proposed for music analysis and synthesis are reviewed and their pros and cons are identified. This analysis justifies the use of the proposed signal model. Experiment results for the analysis and synthesis of a variety of musical instrument tones such as piano, trumpet, french horn, flute, and violin are given to demonstrate the accuracy and effectiveness of the proposed algorithm. A VLSI architecture for sinusoid synthesis using the parameters obtained from the proposed algorithm is also given to provide a complete solution to the parametric modeling-based music analysis and synthesis.

12:00

1aMU7. Efficient synthesis model for bowed strings using a nonlinear bow–string interaction model. Julius O. Smith III (Music Dept., Stanford Univ., Stanford, CA 94305)

A digital simulation method for bowed strings was developed which combines features of preexisting ''digital waveguide'' and ''commuted synthesis'' methods. In both methods, sampled acoustic traveling waves are explicitly modeled in the string, and string losses are *lumped* at one or two points along the string for computational efficiency. In the digital waveguide case, the bow–string interaction model is based on the theory of McIntyre, Schmacher, and Woodhouse [J. Acoust. Soc. Am. **74** (1983)]. The commuted synthesis model [J. Smith, Int. Computer Music Conf., pp. 64–71 (1993)] interchanges the order of the string and body (assumed linear and time invariant), and therefore cannot support a nonlinear bow–string interaction model. However, according to theory and measurement of bow–string interaction, disturbances sent out by the stick–slip process along the string are fundamentally *impulsive* in nature. The proposed method consists of a full nonlinear dynamic model for the bow and string, without body, followed by an ''impulse prioritizer,'' which detects the most important impulsive events reaching the bridge, followed by a commuted synthesis model which drives a second string model with body impulse-responses initiated at the time and amplitude of each high-priority pulse.

MONDAY MORNING, 22 JUNE 1998

CASCADE BALLROOM I, SECTION A (W), 9:15 A.M. TO 12:05 P.M.

Session 1aNSa

Noise: Non-Auditory Effects of Noise on Health

Lawrence S. Finegold, Cochair
AL/OEBN, Armstrong Laboratory, 2610 Seventh Street, Wright Patterson Air Force Base, Ohio 45459

Willy Passchier-Vermeer, Cochair
TNO, The Netherlands

Chair's Introduction—9:15

9:20–11:35
Panel Discussion

The session will consist of a panel discussion by a group of international scientists of the Non-Auditory Health Effects of Noise Exposure, plus presentation of two contributed papers on this topic. Both occupational and community noise exposure will be included in the discussion. Each of the invited panel members will initially provide a brief overview of their own work on the non-auditory health effects of noise. This will be followed by a discussion of issues related to the development of this area of research, including the advantages and disadvantages of various research methodologies and research design alternatives, the choice of specific health effects end points, and the transition of research data to governmental noise exposure policies.

Contributed Papers

11:35

1aNSa1. Study of noise levels in a neonatal intensive care unit. Chantal Laroche and Paula Fournier (Audiol. and Speech Lang. Pathol. Prog., Univ. of Ottawa, 545 King Edward, P.O. Box 450, Stn. A, Ottawa, ON K1N 6N5, Canada)

In the past few years, much interest has been focused on the integrated care approach, where noise is seen as an element which can compromise the normal development of newborns in neonatal intensive care units. The objective of this study was to develop a methodology evaluating the impact of training of the nurses, regarding noise levels existent within the neonatal intensive units of an Ottawa regional hospital. The method consisted of continuous noise level measurements performed by a computer (type 1 sound level meter), pre- and posttraining. $L_{\mathrm{Aeq,1sec}}$, as well as audio samples were recorded throughout eight work shifts. The average pretraining A-weighted noise levels were of 53, 61, and 65 dB, for the night, day, and evening shifts, respectively. These levels largely exceed the maximal A-weighted sound level of 45 dB, recommended by the World Health Organization to avoid negative effects on sleep. Throughout this study, it is likely that the training of nurses could contribute to the reduction of noise levels, but that interventions concerning noise control would also be necessary in order to ensure an acceptable sonorous ambiance for neonates requiring intensive care. [Work supported by University of Ottawa.]

11:50

1aNSa2. Trading level for number in the laboratory: Differential effects of aircraft noise on cardiovascular activation, annoyance, and assessment of quality of life. Iris B. Mauss (Univ. of Duesseldorf, Dept. of Cybernetical Psych., 40225 Duesseldorf, Germany, mauss@uni-duesseldorf.de), Joachim Vogt (Univ. Dortmund, FB 14, 44221 Dortmund, Germany), and Karl Th. Kalveram (Univ. of Duesseldorf, Dept. of Cybernetical Psych., 40225 Duesseldorf, Germany)

As aircraft become quieter, the question arises to what extent the decrease in noise is convertable into an increase in the number of flight movements without enhancing psychological and physiological noise effects. Conventional reasoning has it that it suffices to keep constant the

energy equivalent noise level. In a 3×3 design combining three number conditions (3/9/27 overflights during 30 min) with three noise level conditions (aircraft types emitting 71/76/81 dB(A) per overflight), this assumption was tested, using three groups of 12 subjects each for the number conditions, and repeated measurements for the level conditions. Comparisons between the equal-energy conditions (e.g., 3×76 dB vs 9×71 dB) did not show any significant differences regarding systolic and diastolic blood pressure and heart rate. Regarding ratings for annoyance in the lab situation, results were the same with a slight tendency for underenergetic number effects, that means, more but softer aircraft were preferred. However, if the subjects imagined to reside in a region continuously exposed to the experienced noise, quality of living was assessed as slightly more negative for residential areas with nine overflights, and clearly more negative for 27 overflights, though energy was fixed. This suggests overenergetic number effects for living quality beginning at about 18 overflights/h.

MONDAY MORNING, 22 JUNE 1998

CASCADE BALLROOM I, SECTION C (W), 9:15 A.M. TO 12:35 P.M.

Session 1aNSb

Noise and Archtiectural Acoustics: Novel Approaches to Noise Control Barriers I

David C. Hothersall, Chair
Department of Civil and Environmental Engineering, University of Bradford, Bradford BD7 1DP, United Kingdom

Chair's Introduction—9:15

Invited Papers

9:20

1aNSb1. Excess attenuation by reactive obstacles at a noise barrier edge. Kyoji Fujiwara, Chul-hwan Kim, and Tomonao Ohkubo (Dept. of Acoust. Design, Kyushu Inst. of Design, Shiobaru 4-9-1, Minami, Fukuoka, 815-8540 Japan)

The excess attenuation by small obstacles with a reactive surface installed at the noise barrier edge was studied. So far, the absorbing obstacle has been used to obtain large attenuation, keeping the barrier height constant. It was found that an obstacle with a soft surface at the barrier edge can suppress noise more effectively than the absorbing one does. To realize this soft surface an arrangement of one-quarter wavelength acoustic wells was used. Although these acoustic wells behaved as the soft surface in a certain frequency range, at a barely lower frequency range the sound pressure level behind the barrier was amplified. To overcome this negative effect a new arrangement of wells tuned to several different resonance frequencies is studied. As a result, the excess attenuation by a small round obstacle of 53 cm width and 33 cm height installed at the edge of a 3-m barrier is about 4.7 dB (OA) at the ground 25 m behind the barrier. When the barrier has a T shape and is of 2 m width and 5 m height, the excess attenuation is 7 dB (OA). Details of the cross sections of these barriers will be presented.

9:50

1aNSb2. Jagged-edge noise barriers. Penelope Menounou, Ilene J. Busch-Vishniac, and David T. Blackstock (Dept. of Mech. Eng., Univ. of Texas at Austin, TX 78712-1063)

Experimental, numerical, and analytical investigations have shown that noise barrier performance can be improved by making the top edge randomly jagged. The incident noise is a spherical N wave (produced by a spark in our model experiments). See Ho *et al.* [J. Acoust. Soc. Am. **101**, 2669–2676 (1997)], and oral papers by (i) Menounou *et al.*, (ii) Ohm *et al.*, and (iii) Rosenberg *et al.*, all in J. Acoust. Soc. Am. **101**, 3075(A) (1997). In many, but not all, cases the jagged edge improves insertion loss by up to 5 dB. The most promising theoretical approach is to model the edge as a directive, crooked line source [paper (i) above]. Besides yielding predictions that agree favorably with measurements [paper (iii) above], the line source model helps in understanding the following results: (1) the jagged edge smears out the diffracted signal and makes the tail ragged, (2) insertion loss varies significantly with receiver position, and (3) the portion of the edge on the most direct path between source and receiver plays a very important role. Because the model is easy to apply, it is a promising and powerful tool for further research on jagged barriers. [Work supported by Texas Advanced Technology Program.]

10:20

1aNSb3. Comparison of calculated and measured data for the attenuation of improved noise barrier. Kohei Yamamoto (Kobayasi Inst. of Physical Res., 3-20-41 Higashi-Motomachi, Kokubunji, Tokyo, 185 Japan)

The comparison between calculated and measured data was made to the additional noise attenuation of a barrier with an absorptive device on the top edge. The attenuation was determined at a highway with a noise barrier of 3 m height. By the use of the boundary-element method (2D), noise attenuation due to the mounting of the device was estimated. The results showed an agreement between the calculated and the measured value. However, the calculated values showed fluctuation when they were plotted against distance.

10:50–11:05 Break

11:05

1aNSb4. Active control of traffic noise around barriers. Denis Duhamel (ENPC-CERAM, 6 et 8 Ave. Blaise Pascal, Cité Descartes, Champs-sur-Marne, 77455 Marne-la-Vallée Cedex, France)

Active control was previously found efficient to reduce noise behind barriers for compact and immobile sources [Omoto and Fujiwara, J. Acoust. Soc. Am. **94**, 2173–2180 (1993)]. For real road traffic there are usually many sources at quite different positions, so the sound pressure is very complex. It is simulated by an incoherent line source in which different point sources placed along a straight line radiate independently. The structure of the sound pressure, mainly the coherence properties, are calculated around the noise barrier. This is done by a boundary element method to get precise results for low frequencies and near the barrier. A good coherence is necessary to obtain an efficient active control. Simulations of active control systems are done for different placements and numbers of secondary sources. A sufficient number of sources is determined according to the frequency to get a significant attenuation. Simulations of feedback control near the top of the barrier are investigated and compared to feedforward control. For a feedforward control a reference signal is necessary which is difficult to provide because of the low coherence of the sound pressure. A sufficient number of sensor microphones and their positions are also determined in that case.

11:35

1aNSb5. Active noise barrier based on the boundary surface control. Shiro Ise (Nara Inst. of Sci. and Technol., Takayama 8916-5, Ikoma, Nara, 630-01 Japan, ise@is.aist-nara.ac.jp) and Hideki Tachibana (Univ. of Tokyo, Tokyo, 106 Japan)

Based on the principle of the boundary surface control, it is possible to make a surface which reflects acoustic energy flow by controlling the pressure and its gradient on the surface. In order to realize this method, a system has to be made using multiple speakers and multiple microphones. The system should be adjusted so that the output signals of the multiple microphones which are located on the surface considered to be the location of the barrier are minimized. By changing the distance between the microphones, the best effect of the active noise barrier can be obtained. For instance, in a level lower to the ground, many microphones are needed to control higher frequencies. On the other hand, in a level higher from the ground, it is not necessary to use so many microphones to control lower frequencies. The merit of this method is that the apparent height of the barrier, depending on the frequency considered, can be changed. In this paper, the active noise barrier based on the boundary surface control is introduced and the effect of this method is calculated using BEM.

Contributed Papers

12:05

1aNSb6. Traffic noise barrier overlap gap analysis. Lloyd A. Herman (Dept. of Civil Eng., Ohio Univ., 141 Stocker Ctr., Athens, OH 45701) and Craig M. Clum (Development Consultants Group, Duluth, GA 30136)

Sound propagation through the gap produced by two parallel vertical barriers with overlapped ends was formulated for traffic noise sources. The method accounts for sound propagated from vehicle source positions for a maximum of ten roadway lanes. Six receiver regions were considered based on potential receiver locations with respect to a gap. The analysis identified both source and receiver regions according to the mechanisms that influence noise propagation in the vicinity of an overlap gap, which can result in: direct rays, diffracted rays from the top edge of one barrier, diffracted rays from the top edge of both barriers, rays reflected between the barriers, and rays that are both reflected and diffracted. The derived method could be implemented using various equations for sound propagation. The results of using equations approved by the United States Federal Highway Administration (FHWA) for traffic noise propagation are given. Field measurements for up to 30 receiver positions from four overlap gaps were compared with uncalibrated predictions. The equivalent continuous levels, A weighted, were overpredicted by 2–3 dB. When the analysis was based on an octave band characterization of the source, the mean overprediction was reduced to less than 1 dB.

12:20

1aNSb7. Rubber crumb as an absorbent acoustical material. Jaime Pfretzschner and Rosa M. Rodriguez (Instituto de Acústica, CSIC, Serrano 144, Madrid 28006, Spain)

In order to contribute to the solution of the environmental problem created by the tire wasteways, the possibilities of using rubber crumbs of waste tires as acoustic absorbent materials have been studied in a systematic way. Experimental absorption coefficients of rubber crumb granulates with mesh diameters between 1.4 and 7 mm have been tested in Kundt's tube and in a reverberation room. Absorption coefficient values so obtained are compared with computational values afforded through mathematical prediction models that use the intrinsical parameters of porous material such as porosity, tortuosity, and flow resistivity, which were measured on layers of the respective materials. The good broadband acoustic absorption of the material encourages its use as an alternative to the current absorbent acoustic screens, used for the protection against traffic noise.

Session 1aPAa

Physical Acoustics and Noise: Outdoor Sound Propagation I

Henry E. Bass, Cochair
National Center for Physical Acoustics, University of Mississippi, University, Mississippi 38677

Daniel Juve, Cochair
Ecole Centrale de Lyon, Acoustique, 36 Avenue Guy de Collongue, BP 163, 69131 Ecully, Cedex, France

Chair's Introduction—7:55

Invited Papers

8:00

1aPAa1. Propagation of sonic booms through the atmosphere. Richard Raspet and Henry E. Bass (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677, raspet@next1.ncpa.olemiss.edu)

During the past 20 years the authors have been involved in developing physically based prediction calculations for sonic booms and other nonlinear atmospheric waves. Early work developed a calculation method for the effects of atmospheric absorption and dispersion on sonic booms in a quiescent atmosphere. The validity of this method has been confirmed by recent collaborative work with the University of Texas [J. Acoust. Soc. Am. **100**, 3017–3027 (1996)]. More recently, the authors have been working on a method of predicting the effect of turbulence on sonic boom waveforms in the planetary boundary layer. A Monte Carlo approach has been employed so that histograms of measured variables which are needed for environmental assessment can be developed, rather than a single average value [J. Acoust. Soc. Am. **99**, 147–152 (1996)]. This prediction method can be combined with the earlier sonic boom propagation algorithm to predict the ensemble of received waveforms at the ground, given the waveform in the vicinity of the aircraft, based solely on physical considerations.

8:20

1aPAa2. Atmospheric sound propagation near the ground. Conny Larsson (Dept. of Meteorol., Uppsala Univ., P.O. Box 516, S-751 20 Uppsala, Sweden, conny@big.met.uu.se)

The most important meteorological effects on sound propagation are refraction-ground attenuation, atmospheric absorption, and scattering by turbulence. The wind and temperature gradients cause refraction of the sound rays, and hence influence the sound level. The curvature of a nearly horizontal sound ray can be calculated by using measurements of wind and temperature. Atmospheric absorption is important for outdoor sound propagation. The absorption varies for different parts of the world. Results from some stations in Scandinavia are given. Atmospheric absorption for the high frequencies changes during different times of the year. The absorption for low frequencies varies both over the day and the year. The size of the annual and diurnal variance is frequency dependent. An increase in the accuracy of calculations of outdoor sound levels can be made by considering the local climate. The values of atmospheric absorption need to be calculated for the local climate instead of using some global mean value. The time has come to use data from local meteorological stations. Turbulence scatters sound into sound shadow zones and causes fluctuations of the phase and the amplitude of the sound waves.

Contributed Papers

8:40

1aPAa3. Effect of waveform distortion on sonic boom noise penetration into a flat ocean. Tracie J. Ferguson and Victor W. Sparrow (Grad. Prog. in Acoust., Penn State Univ., 157 Hammond Bldg., University Park, PA 16802)

The penetrating underwater sound field due to an N-shaped sonic boom incident on a flat ocean surface is well described by existing theory. However, in realistic situations, most of the waves that hit the surface are not perfect N waves but are distorted by such processes as atmospheric turbulence. In the present research, N waves are created that contain the same amount of energy as experimentally measured sonic booms produced by aircraft in the U.S. Air Force. Each signature is input to an algorithm recently presented by Sparrow and Ferguson [AIAA Paper 97-0486 (1997)] to calculate the respective underwater sound fields. By comparing the underwater wave shapes and plots of peak pressure with respect to depth, initial results of the effect of waveform distortion on sonic boom noise penetration into a flat ocean are determined. Preliminary figures show that waveform distortions at the surface cause a slightly greater penetration depth underwater. However, the magnitude of the effect is dependent on the incident sonic boom. [Work supported by Armstrong Laboratory, Air Force Materiel Command, USAF, under Grant No. F41624-96-1-0003.]

8:55

1aPAa4. Sonic boom in the shadow zone. François Coulouvrat (LMM, Univ. Pierre et Marie Curie, CNRS (URA 229), case 162, 4 pl. Jussieu, F75252 Paris Cedex 05, France)

The classical geometrical acoustics theory predicts the amplitude of sonic booms only within the carpet. Inside the geometrical shadow, the sound field is known to emanate from rays shed by creeping waves propagating along the ground. A new, nonlinear, time-domain formulation of diffraction effects inside the shadow zone is given. The resulting equation is analogous to an unsteady, two-dimensional, transonic equation. In the linear case, it can be solved, so that the dispersion equation for creeping waves is recovered. If only the least attenuated one is taken into account,

an approximate equation of Burgers-type describing the nonlinear propagation of the creeping wave can be deduced. A dimensional study indicates that nonlinear effects are small, except for small sound-speed gradients. A numerical simulation shows that the amplitude decay of the signal compares favorably with Concorde measurements made by ISL at Biscarosse in 1973, at least sufficiently deep within the shadow zone. [Work supported by Aérospatiale Aéronautique–Toulouse.]

9:10

1aPAa5. Statistical data of excess attenuation for long-range noise propagation over sea. Kazuo Konishi (Ken-On Eng., Inc., 1-17-1 Shofudai, Tarumi, Kobe, 655 Japan, ken-on@ta2.so-net.or.jp) and Zyuniti Maekawa (Environ. Acoust. Labs., Osaka, 531 Japan)

At an environmental impact assessment of a new planned airport on an artificial island, many data of excess attenuation were obtained in long-range (5-km) sound propagation over sea continuously for 14 months versus meteorological conditions. Unfortunately, as presented at Inter-Noise 96 and 97, clear relations were not found between them using both multiregression analysis and multidimensional quantification analyses. Although it is known that the sound-speed profile is one of the most effective factors in long-range sound propagation, it seemed to be difficult to obtain the profile sufficient to predict the sound propagation over sea. Furthermore, it is very difficult to predict the excess attenuation, possibly because of the existence of unpredictable turbulence in the atmosphere. Therefore, using the measured data, attempts were made to estimate the excess attenuation related to aircraft noise at a receiving point. After statistical analysis, it has become clear that the excess attenuation is almost positive, and negative only at night. This fact is very useful for the environmental impact assessment of the planned airport.

9:25–9:35 Break

9:35

1aPAa6. Sound propagation in a city composed with hexagonal periodic buildings. Judicael Picaut (Lab. Acoustique, Univ. Le Maine, B.P. 535, 72017 Le Mans France) and Jean Hardy (Univ. Le Maine, B.P. 535, 72017 Le Mans France)

The noise propagation in central parts of towns is a problem of major importance. Because of multiple reflections upon the buildings the background noise obeys to a diffusion law in place of a free propagation law. Usually the diffusion law is obtained by means of a random disposition of the buildings. The originality of this work is to show that this hypothesis can be avoided since the example given is composed with hexagonal periodic buildings, with a semidiffusing law for the reflections. A hexagonal geometry is used, cumulating the advantages of simplicity for the refections laws and avoiding the avenue effects. A numerical simulation shows that a diffusion occurs for the rays and the dependence of the diffusion coefficient from the geometrical dimensions of the streets, the complexity of the building façades is investigated. The model suggests that for a realistic case two parameters suffice to describe the noise propagation in a town: the diffusion coefficient depending of the complexity of the façades and the streets width and an attenuation coefficient depending of the sizes of buildings and of the attenuation properties of the reflecting surfaces and of the air.

9:50

1aPAa7. Scale model experiments on the insertion loss of wide and double barriers. Glenn J. Wadsworth and James P. Chambers (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677)

The insertion loss of wide and double barriers is investigated through scale model experiments. Such configurations are common in outdoor sound propagation problems such as highway barriers and raised or depressed roadways. The Biot–Tolstoy–Medwin time domain wedge formulation for multiple diffraction [J. Acoust. Soc. Am. **72**, 1005–1013 (1982)] is used to predict the acoustic field directly in the time domain. Evaluating the insertion loss at discrete frequencies or in octave bands is accomplished via the fast Fourier transform (FFT). Comparisons to other frequency based models will also be presented.

10:05

1aPAa8. Incorporation of inhomogeneous, atmospheric turbulence into numerical propagation calculations. D. Keith Wilson (U.S. Army Res. Lab., 2800 Powder Mill Rd., Adelphi, MD 20783, dkwilson@arl.mil)

Homogeneous, atmospheric turbulence is often synthesized for use in numerical propagation calculations (the parabolic equation, for example) by multiplying Fourier modes by random phases and then inverse transforming. This method can be generalized to inhomogeneous turbulence by multiplying the empirical orthogonal functions (EOFs) by random phases; in homogeneous turbulence the EOFs become the Fourier modes. The EOF technique is applied in this paper to the case of turbulence in a vertically inhomogeneous, atmospheric shear layer, where the integral length scale is proportional to the height. The EOFs are determined numerically, and then the turbulence is synthesized and incorporated into calculations of scattering into a shadow zone. The results are compared to previous calculations involving homogeneous turbulence.

10:20

1aPAa9. Nonlinear aspects of outdoor propagation of acoustic pulses produced by weak explosions. Antonio Moreno, Carlos de la Colina, and Francisco Simón (Instituto de Acústica, del C.S.I.C., Serrano 144, Madrid 28006, Spain, iacma15@pinar2.csic.es)

This paper deals with some nonlinear characteristics of acoustic pulses generated by weak explosions with regard to establishing a probe signal for the experimental study of ground characteristics and outdoor sound propagation. These explosions cover the intensity range of signals commonly used on experimental studies, e.g., studies of the impulse response of rooms and the transmission loss and absorption characteristics of noise barriers. After a brief introduction of time history and spectra of such signals, compared to linear pulses and nonlinear pulses of explosions of a higher nonlinear range particular attention is paid to air absorption up to 20 kHz. Evolution of the pulse signature as it propagates as well as global pick attenuation and attenuation in $\frac{1}{3}$ octave frequency bands are discussed compared to linear processes and well established nonlinear theories, offering as a conclusion a simple and practical model that accounts well for the transformations undergone by the pulse when it propagates under reference atmospheric conditions. These characteristics are intended as a reference for comparison of experimental studies on the influence of the ground and obstacles of different acoustic nature involved in propagation processes. [Work supported by the Spanish National Research Program, DCICYT AMB 95-0101.]

Session 1aPAb

Physical Acoustics: General Topics in Nonlinear Acoustics

Andrés Larraza, Chair
Naval Postgraduate School, Department Physics, Code PH/LA, Monterey, California 93943

Chair's Introduction—8:10

Contributed Papers

8:15

1aPAb1. Steady streaming from a noncompact sphere. Ashok Gopinath (Dept. of Mech. Eng., Code ME/Gk, Naval Postgrad. School, Monterey, CA 93943, gopinath@nps.navy.mil) and Eugene H. Trinh (NASA-JPL, Pasadena, CA 91109)

An analytical study has been conducted of acoustic streaming around a rigid sphere in a plane standing acoustic field under the noncompact sphere approximation. Under these conditions the sphere size is allowed to be comparable to the (radian) wavelength of the acoustic field, thus making the usually ignored compressibility effects significant in the flow. However, the amplitude of fluid oscillations is assumed small compared to the sphere size, and their ratio serves as the appropriate perturbation parameter in the matched asymptotic technique being used. The solution calls for a Helmholtz type decomposition of the velocity field into the sum of the gradient of a potential function, and the curl of a suitably vectorized stream function. The role of the nonlinear interactions of both the potential function and the stream function is pointed out, and the resulting streaming flow patterns are contrasted with those obtained in the incompressible case. This work would find application in determining transport rates in containerless processing using acoustic levitation with *ultrasonic* frequencies. [Work supported by NASA.]

8:30

1aPAb2. An acoustic radiometer. Timothy G. Simmons, Bruce Denardo (Dept. of Phys. and Astron., and Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677), Andrés Larraza, and Robert Keolian (Naval Postgrad. School, Monterey, CA 93943)

An electromagnetic radiometer is a well-known physics demonstration and toy. When exposed to light of sufficient intensity, the vanes rotate in the *opposite* direction if radiation pressure were the mechanism. That is, the reflective white side leads the absorptive black side. This occurs due to thermal effects associated with the gas in the bulb, and can be overcome only with a very high vacuum. To demonstrate the effect of radiation pressure, we have built an *acoustic* radiometer consisting of an aluminum T-beam with a 10-cm-square pane attached to each end. One side of a pane is acoustically reflective (aluminum) and the other is acoustically absorbing (foam). The device is pivoted with a V-jewel bearing and a steel pin, and is placed in a transparent chamber with a uniform broadband noise field (2.5–15 kHz). A minimum sound-pressure level of roughly 125 dB (re: 10^{-12} W/m^2), which yields a theoretical net force corresponding to 2 mg on each pane, produces rotation in the ''correct'' sense with the absorptive side leading the reflective side. The apparatus will be demonstrated during the talk. [Work sponsored by the Office of Naval Research.]

8:45

1aPAb3. An acoustic Casimir effect. Andrés Larraza, Christopher D. Holmes, Robert T. Susbilla (Dept. of Phys., code PH/La, Naval Postgrad. School, Monterey, CA 93943), and Bruce Denardo (Univ. of Mississippi, University, MS 38677)

Theoretical and experimental results are presented for the force law between two rigid parallel plates due to the radiation pressure of broadband acoustic noise. The noise is in the band of 5 to 15 kHz and has an intensity of 133 dB (re: 10^{-12} W/m^2). Good quantitative agreement is shown between theory and experiment. These results constitute an acoustic analog for the Casimir effect, which is the force experienced by two closely spaced uncharged conducting plates due to the quantum electromagnetic zero point field. In contrast to this case, however, band limited noise can cause the force to be *attractive* or *repulsive* as a function of the distance of separation of the plates. [Work sponsored by the Office of Naval Research.]

9:00

1aPAb4. Research of finite-amplitude acoustic pulse propagation in a thin capillary filled with a viscous-elastic fluid. Igor B. Esipov (N. Andreyev Acoust. Inst., 4 Shvernik str., Moscow 117036, Russia) and Vladislav G. Mikhalevich (General Phys. Inst., Moscow 117942, Russia)

Acoustic wave propagation in a thin capillary filled with a viscous fluid is accompanied by additional attenuation because of sound energy outflow to the capillary wall due to viscous forces. It is known that a whole class of fluids exists, which change their viscous properties (shift modulus) under the applied stress. It is the so-called Bingham fluids. The results of experimental research of wide frequency band sound pulse propagation in the capillary filled with a raw oil are presented. A pulse was generated by laser radiation. The amplitude of the pulse reached up to 1 MPa and its time duration was about 20 ns. Acoustic signal reception was provided by a PVDF gage and registered by a digital oscilloscope in the frequency range of 100 MHz. Experiment shows a kind of induced ''transparency'' for the high-amplitude sound pulse propagated along a capillary of 0.4-mm radius. This effect of acoustic-induced transparency of viscous fluid shows a rapid growth of high-frequency components of the sound signal passed through the capillary with respect to the growth of laser pulse energy. Meanwhile, for the small-amplitude sound signal propagation the linear relation between laser pulse energy and sound pulse amplitude is valid. [Work is partially supported by the Russian Foundation of Basic Research (project 97-02-17789).]

9:15

1aPAb5. Stabilization of a capillary bridge far beyond the Rayleigh–Plateau limit using active feedback and acoustic radiation pressure. Mark J. Marr-Lyon, David B. Thiessen, and Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814)

A liquid bridge between two solid surfaces is known as a capillary bridge. Long bridges naturally become unstable to a symmetric mode by bulging near one end while the opposite end thins. Suppression of this mode is important for managing fluids in low gravity. For a cylindrical bridge in low gravity of radius R and length L, the slenderness $S=L/2R$ has a natural (Rayleigh–Plateau) limit of π beyond which the bridge breaks. Using acoustic radiation pressure to control the bridge's shape, and an optical sensor to detect the bridge's shape, an active feedback system in simulated low gravity was constructed. This system stabilized bridges significantly beyond the Rayleigh–Plateau limit, with S as large as 4.3 [Marr-Lyon *et al.*, J. Fluid Mech., accepted for publication]. To be useful in low gravity, this technique will have to be modified to work on liquid bridges in air. In preparation for an experiment to be performed on NASA's low gravity KC-135 aircraft, acoustic resonators in air that have the required

property that the sound amplitude can be spatially redistributed rapidly are being investigated and tested in normal gravity using soap-film bridges filled with sulfur hexafluoride. [Work supported by NASA.]

9:30

1aPAb6. Surfactant-bearing drop dynamics in microgravity. Xiaohui Chen and Robert E. Apfel (Dept. of Mech. Eng., Yale Univ., 9 Hillhouse Ave., New Haven, CT 06520-8286, xhchen@minerva.cis.yale.edu)

Experiments of large amplitude nonlinear oscillations of axially symmetric surfactant-bearing (triton or bovine serum albumin) water drops with different initial aspect ratios and drop volumes were performed during the second United States Microgravity Laboratory, USML-2 in November 1995. Through the numerical simulation of the oscillations of these surfactant-bearing water drops using the boundary integral method, the surface dilatational viscosity and the surface shear viscosity are found to play a major role in these oscillations. How they influence the damping coefficient and the drop oscillation frequency will be presented. The numerical and analytical results and a physical picture of the relation between the surface viscosities and the damping coefficient will be described. In addition, the relation between the initial drop aspect ratio, the surface viscosities, and the surface rotational velocity will also be presented. [Work supported by NASA through JPL, Contract No. 958722.]

9:45–10:00 Break

10:00

1aPAb7. The study of evaporation of multicomponent drops using the acoustoelectric levitator. Yibing Zheng, Robert Apfel, and Yuren Tian (Dept. of Mech. Eng., Yale Univ., New Haven, CT 06520, yibing.zheng@yale.edu)

The evaporation of single drops and drop arrays of multicomponent liquids is investigated using a novel acoustoelectric levitator. By simultaneous levitation and charging of the drops, 2D drop arrays with the drop diameters varying from 0.2 to 1 mm can be formed. Also by controlling both the acoustic and electric field intensities, arrays with different drop sizes and spaces between drops can be created. This allows us to study the interference effects among drops in an evaporating dense spray. Mixtures of different liquids such as alkanes (heptane, dodecane, and hexadecane) are studied. The evaporation rates of those mixtures with different molar ratios of components and different numbers of drops in an array are measured. The drop size and spacing are carefully controlled and varied. A video camera is used to record the drop size change with time, and the diameters of drops are calculated from the pictures. The results are compared with a theoretical analysis based on the diffusion-controlled evaporation model. [Work supported by NASA through JPL Grant No. 958722.]

10:15

1aPAb8. Particle separation by ultrasonic forces. Steven M. Woodside (Biotechnology Lab. and Dept. of Chemical and Bio-Resource Eng., Univ. of British Columbia, Vancouver, BC V6T 1Z3, Canada), Martin Groeschl, Ewald Benes (Institut fuer Allgemeine Physik, Technische Universitaet Wien, A-1040 Wien, Austria), Bruce D. Bowen, and James M. Piret (Univ. of British Columbia, Vancouver, BC V6T 1Z3, Canada)

A simple, effective device has been developed to separate mammalian cells from growth medium using the ultrasonic forces on suspended particles in a standing wave field. The ultrasonic force distribution in this acoustic separation device has been analyzed to provide a basis for further optimization and scale-up of the separation system. Measurements of the force distribution in the liquid using a microscope-based imaging system were compared to laser interferometric measurements of the velocity amplitude distribution on the transducer and reflector surfaces of the ultrasonic resonator. The axial force followed the surface velocity, being highest near the center, and approaching zero near the walls. The transverse force distribution and the local gridlike surface amplitude variation had characteristic wavelengths of 1.4 mm. These results suggest that the force distribution is determined predominantly by the response of the reflector and transducer to the standing wave. This understanding of the force distribution simplifies the modeling of particle retention and facilitates resonator design. By selection of appropriate resonator dimensions and construction materials, and modeling of the ultrasonic force on the particles, enhanced flow-through capacity and particle retention efficiency can be achieved.

10:30

1aPAb9. The two-dimensional acoustic standing wave and its application in coagulation of aerosols. Henryka Czyż (Inst. of Phys., Pedagogical Univ. of Rzeszów, 35-310 Rzeszów, Rejtana 16a, Poland, hczyz@atena.univ. rzeszow.pl)

The problem of motion of small solid or liquid particles suspended in gaseous medium (aerosols and smokes) has been studied intensively in connection with the technical application of acoustic coagulation for precipitation of gases. Along with its technical applications, the problem of motion of small particles in acoustic field (called the drift) is interesting by itself as one of the fundamental physical effects of sound. For theoretical and practical reasons, the two-dimensional standing-wave configuration seems to be of special interest. In the case when nodal and antinodal surfaces of the standing wave are convex, the particles move along some curves and gather in some particular points and not only in some planes. This means that one can obtain much greater concentration of particles in some definite regions than in the case of the plane standing wave. As a configuration of special theoretical interest it proposes the standing wave in a rigid elliptical cylinder with a strip-shaped source of uniform velocity distribution placed in the symmetry plane of the system. Some numerical examples will be presented during the conference [H. Czyż, Acustica **70**, 23–29 (1990)].

Session 1aPAc

Physical Acoustics: Cavities and Resonators

Robert M. Keolian, Chair
Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

Chair's Introduction—8:25

Contributed Papers

8:30

1aPAc1. Physical effects of macrosonic standing waves in oscillating cavities. Christopher C. Lawrenson, Thomas W. Van Doren, and Bart Lipkens (MacroSonix Corp., 1570 East Parham Rd., Richmond, VA 23228)

Previous work has demonstrated resonant macrosonic synthesis (RMS) for shaping pressure waves in oscillating closed cavities. Additional measurements of pressure waves are presented, with peak overpressures in excess of 400% of ambient pressure. Characteristic power consumption of various resonator shapes is presented. Resonator shape as well as pressure amplitude is shown to be important in determining power consumption. For a change in resonator geometry, one example shows a 765% increase in pressure amplitude for a fixed power delivery. The dependence on resonator geometry of hardening and softening resonance behavior is demonstrated by showing results for both different resonator shapes and different gases in the resonators. Physical effects of macrosonic pressures and particle velocities are demonstrated with pulverization of objects placed in the resonator cavity.

8:45

1aPAc2. Investigation of a particle velocity field in nonlinear standing waves. Yurii A. Ilinskii, Bart Lipkens, and Evgenia A. Zabolotskaya (MacroSonix, 1570 East Parham Rd., Richmond, VA 23228)

A one-dimensional model of nonlinear standing waves in an oscillating closed cavity has been developed at MacroSonix and was presented at the 1997 Acoustical Society meeting in San Diego. The present study focuses on the characteristics of the particle velocity wave field. Measurements of the particle velocity are very difficult to execute, but insight into the velocity field yields important information of the properties of an acoustical resonator. Waveforms and harmonic amplitudes of velocity as well as pressure waves along the resonator axis are calculated. Results are presented for three geometries: a cylinder, a cone, and a bulb-type resonator. The resonators are oscillated at the fundamental resonance frequency. Numerical calculations show that for dissonant resonators, in which harmonics of the fundamental frequency are not coincident with higher modal frequencies, Mach numbers on the order of 0.5 can be achieved, while for a cylindrical resonator Mach numbers are much more limited. For the resonators considered, the velocity waveforms have different characteristics: a shocked waveform for the cylinder, a sawtooth-type wave for the cone, and a disturbed sine wave for the bulb. An interesting result is that the velocity waveform is fundamentally different from the pressure waveform.

9:00

1aPAc3. Parametric excitation of a Helmholtz resonator. Wayne E. Prather, Bruce Denardo, and Richard Raspet (Dept. of Phys. and Astron., and National Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677, weprathe@olemiss.edu)

Parametric excitation of an oscillator results from the modulation of a parameter upon which the natural frequency of the oscillator depends. If the drive amplitude exceeds a threshold determined by dissipation, the response amplitude grows exponentially until it is saturated by a nonlinearity of the system. Hence, large amplitudes may be possible. An unsuccessful attempt to parametrically excite the fundamental mode of a Helmholtz resonator is reported. The apparatus consists of two equal-volume cavities connected by a U-shaped neck similar to a trombone slide, where the neck length is modulated at twice the Helmholtz frequency. In model systems, the effective quality factor diverges as the parametric drive amplitude approaches threshold from below. the extent to which a parametric drive amplitude is below threshold can thus be probed by performing a frequency sweep with a small-amplitude direct drive. However, the effective quality factor in our system is found to *decrease* as the parametric drive amplitude is increased, as a result of turbulence caused by the parametric drive. This indicates that parametric excitation by neck length modulation may not be possible. Current work on a Helmholtz resonator in which the cavity volume is modulated is also reported on. [Work sponsored by ONR.]

9:15

1aPAc4. Measurements of temperature and velocity oscillations in acoustic waves. Guadalupe Huelsz, Francisco Lopez-Alquicira, and Eduardo Ramos (Centro de Investigacion en Energia, AP 34 CP 62580, Temixco, Mor. Mexico, ghl@mazatl.cie.unam.mx)

Experimental techniques using cold and hot wire anemometry have been developed to measure acoustic temperature and velocity oscillations in gases, respectively, especially in the case where there is no imposed mean flow. The calibration of both instruments was made through the measurement of the pressure oscillation. Using these techniques the structures of thermal and viscous oscillatory boundary layers produced by an acoustic standing wave in air with 130-Hz frequency and pressure oscillation amplitude up to 200 Pa are determined. Experimental results are satisfactorily compared to theoretical results of a linear theory. The advantages and the limitations of these experimental techniques are discussed. [Partial support for this work has been provided by DGAPA-UNAM IN103197 and CONACYT G0044E projects.]

9:30–9:45 Break

9:45

1aPAc5. Acoustic methods for transport properties measurements in gases. Keith A. Gillis, James B. Mehl, and Michael R. Moldover (NIST, Bldg. 221, Rm. A111, Gaithersburg, MD 20899, keith.gillis@nist.gov)

Two novel acoustic resonators have been developed, one for measuring the viscosity and a second for measuring the Prandtl number of gases. For viscosity measurements, a double Helmholtz resonator, called the Greenspan viscometer, is used. This resonator consists of two chambers connected by a small duct. The Greenspan viscometer has a single, isolated, low-frequency mode that is fitted to determine the gas' viscous diffusivity. The Greenspan viscometer may be used for absolute measurements if the dimensions of the duct and chambers are accurately known. Otherwise, the viscometer may be calibrated with argon or helium. For Prandtl number measurements, a cylindrical resonator that has been modified by inserting an array of short, parallel plates halfway between the

ends is used. The lower longitudinal modes of this resonator are fitted to determine the Prandtl number. The odd order modes are attenuated primarily by the viscous diffusivity; the even order modes are attenuated primarily by the thermal diffusivity. Detailed acoustic models of both resonators have been developed and are necessary to realize the full precision and accuracy of these techniques. Transport property measurements of several gases will be presented. [This work is supported in part by the Office of Naval Research.]

10:00

1aPAc6. Optical techniques for sound generation and detection in acoustic resonators. Giuliana Benedetto, Roberto M. Gavioso, and Renato Spagnolo (IEN Galileo Ferraris-Acoust. Dept., Strada delle Cacce 91, 10135 Torino, Italy)

The most reliable method for sound-speed measurement in gases is based on the use of acoustical resonators. The measurement consists of the determination of the resonance frequencies of the cavity in the presence of a stationary acoustic field. Usually, the source and the receiver are condenser microphones. The main limit of these transducers is their inability to be used in a hostile environment, so there is an increasing interest to develop remote optical techniques for sound generation and detection. In this paper it is shown that it is possible to generate the acoustical field through a photoacoustic effect on the interior surface of the resonator, achieving a measurement precision comparable to that obtained when the source is a 1/4-in. condenser microphone. As a further improvement, it is also shown that sound can be generated by a silicon micromembrane facing the cavity, optically excited at the rear side. Measurements have been performed in a cylindrical stainless steel resonator of about 572 cm^3; the light source is the 514-nm line of an Ar^+ laser. An effort is being made to investigate the possibility of also substituting the receiving transducer with an optical detection scheme, based on laser beam deflection.

10:15

1aPAc7. Numerical study of a one-dimensional resonant gas oscillation in a closed tube. Takeru Yano (Mech. Sci. Dept., Hokkaido Univ., Sapporo, 060 Japan, yano@lebes-me.eng.hokudai.ac.jp)

The one-dimensional resonant gas oscillation in a closed tube is numerically studied by solving Navier–Stokes equations of gasdynamics. Accurate numerical solutions are obtained by using an eighth-order Padé type compact finite-difference scheme without any artificial viscosity and the second-order Runge–Kutta method. The solutions show that a shock wave of sufficiently steep front is formed even in the case of $M=O(10^{-5})$, where M is the acoustic Mach number at the sound source located at one end of the tube. The evolution process from the initial state of uniform and at rest to a quasisteady state including a shock wave is analyzed. The result is compared with earlier works based on the first-order weakly nonlinear theory.

10:30

1aPAc8. The use of a subgrid-scale quasi-stationary approximation in finite-difference time-domain simulations to calculate the absorption by Helmholtz resonators. John De Poorter and Dick Botteldooren (INTEC, Univ. of Gent, Sint Pietersnieuwstraat 41, B-9000 Gent, Belgium)

To calculate Helmholtz resonators using a finite-difference time-domain (FDTD) simulation technique, one has to deal with two important difficulties. First, the determining structures of a Helmholtz resonator are typically subwavelength geometries with dimensions that are smaller than the FDTD cell size normally used for the wavelengths of interest. Therefore, the FDTD cell size should be reduced, resulting in a very high computational cost. On the other hand, the viscous absorption in the resonator neck influences the resonator behavior a lot and has to be included in the calculation. The first problem was solved by using a coarse FDTD grid but applying detailed simulated quasistationary pressure distribution to modify the FDTD equations in the neighborhood of the resonator neck. For the second problem the analytical solution for viscous boundary layer absorption was introduced into the FDTD equations. However, this solution is only correct for the viscous absorption near an infinite plate. Therefore, the calculated quasi-stationary solutions were applied to correct for the nonuniform velocity distribution in the resonator neck. The results are compared with both analytical solutions and with experimental results.

MONDAY MORNING, 22 JUNE 1998 FIFTH AVENUE ROOM (W), 9:15 A.M. TO 12:15 P.M.

Session 1aPP

Psychological and Physiological Acoustics: Pitch—High and Inside, Low and Away

Adrianus J. Houtsma, Chair

IPO, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Contributed Papers

9:15

1aPP1. Absolute magnitude estimation of pitch. Andrzej Rakowski and Andrzej Miskiewicz (Chopin Acad. of Music, Okolnik 2, 00-368 Warsaw, Poland)

The pitch of 24 pure tones covering a frequency range of 63–12 500 Hz in 1/3-oct. steps was measured by absolute magnitude estimation. The tones were presented monaurally at a loudness level of 60 phons. Tone duration was 1 s. Sixteen music students, not possessing absolute pitch served as subjects. The experiment was carried out in individual listening sessions. The subjects were instructed to assign a number to the tone in such a way that their impression of how large the number was matched their impression of the pitch of the tone. The results show that the pitch function plotted in log–log coordinates against frequency is considerably steeper below than above 250 Hz. A similar shape of the pitch function was found in a previous study in which narrow-band noises were used [A. Rakowski, Proc. Fechner Day '97, pp. 91–96]. It is hypothesized that the change of the pitch function steepness at 250 Hz may also be observed in the perception of musical intervals. Intervals in low pitch registers may be perceived larger than those corresponding to the same frequency ratios in high-pitch registers.

9:30

1aPP2. Regular interval stimuli with more than one pitch. William A. Yost (Parmly Hearing Inst., Loyola Univ., 6525 N. Sheridan Rd., Chicago, IL 60626, wyost@luc.edu)

Iterated ripple noises (IRN) are an example of regular interval stimuli (RIS) that can have more than one pitch. IRN stimuli are generated by a cascade network of delay (d in ms), attenuate ($-1 \leqslant g \leqslant 1$), and add circuits. When $g = -1.0$ and the number of cascaded circuits (number of iterations) is less than four, the pitch of IRN stimuli has two values. The addition of two IRN stimuli, each generated with different values of d, can also have two pitches. Listeners were asked to discriminate between an IRN stimulus generated with $g = -1.0$ and the sum of two IRN stimuli with each IRN stimulus generated with a different value of d. Performance was measured as a function of the difference in the two values of d. The center value of d, g, and the number of iterations were also parameters of the experiments. For certain separations of d and for certain values of g, listeners had difficulty discriminating between the two stimuli. The discrimination procedure allows for an objective identification of the pitch of IRN stimuli when $g = -1.0$. The results will be discussed in terms of models of complex pitch. [Work supported by NIDCD and AFOSR.]

9:45

1aPP3. New experiments beyond the traditional Fourcin Pitch range. Johan Raatgever, Frans A. Bilsen, and Reshma Mungra (Appl. Phys. Dept., Delft Univ. of Technol., Delft, The Netherlands)

The Fourcin Pitch is a dichotic pitch phenomenon occurring when a white noise signal and its interaurally delayed replica (delay $T1$) are presented simultaneously with a second (uncorrelated) noise signal with interaural delay $T2$. Former experiments dealt with $T1$ between 2 and 10 ms, while $T2$ was kept at a value between 0 and 2 ms [Bilsen and Wesdorp, 8th ICA (1974)]. With proper polarities of the noises and the delays at opposite ears, the pitch would typically have a value corresponding to $1/(T1+T2)$ and the (lateralized) pitch image would correspond to the shortest delay ($T2$). In the present study, both delay times were chosen about equal and with a value above about 5 ms; again the delays were presented at opposite ears, but with equal polarities. It turned out that a pitch is perceived with a saliency estimated between the Fourcin Pitch and the dichotic repetition pitch (one noise source only), with a value corresponding to $2/(T1+T2)$, and with a (lateralized) pitch image corresponding to an interaural delay of $(T1-T2)/2$. This new result, also, can be explained by the central spectrum theory [Raatgever and Bilsen, J. Acoust. Soc. Am. **80** (1986)].

10:00

1aPP4. Comparing the accuracy of monotic and dichotic repetition pitch. Frans A. Bilsen, Han H. Dols, and Johan Raatgever (Appl. Phys. Dept., Delft Univ. of Technol., Delft, The Netherlands)

When white noise is added to its delayed replica, a pitch is evoked corresponding to the reciprocal value of the delay time (T). When the delayed replica is fed to the same ear, the pitch (monotic repetition pitch: MRP) is rather salient; when, instead, it is fed to the contralateral ear, the pitch (dichotic repetition pitch: DRP) is rather faint. MRP is readily explained by spectral information processing in the cochlea. For reasons of parsimony, DRP is assumed to be the result of central spectrum processing preceded by interaural (temporal) processing. The present paper reports measurements on the accuracy of pitch matching for both MRP and DRP. Because of the difference in saliency, a difference in accuracy might be expected, if pitch saliency and pitch accuracy would be positively correlated. It turned out that pitch accuracy, expressed as the relative standard deviation, is dependent on the delay time and the stimulus type. For T larger than about 20 ms (the infra pitch range) the accuracy is about 10% for both MRP and DRP. For T smaller than 20 ms (the pitch range), accuracy is about 0.5% for MRP and 3% for DRP.

10:15

1aPP5. Listeners' spectral weighting functions for pitch judgment. Huanping Dai (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131)

Two experiments were carried out to determine how listeners use spectral information in judging the pitch of harmonic complex tones. In particular, spectral weighting functions were obtained using a correlation technique in which random, independent perturbations were added to the frequencies of the spectral components. The correlation coefficient between the perturbations applied to a given frequency component and listeners' responses provides an estimate of the weight associated with that component. In the first experiment, all complex tones had the first 12 partials; the fundamental frequency was varied from 100 to 800 Hz. The obtained weighting function tended to show a peak for partials close to 600 Hz regardless of their harmonic numbers. In the second experiment, the fundamental was fixed at 200 Hz, and the number of partials was reduced by omitting the first two, four, or six partials. As these partials were excluded, the weights redistributed, with the maximum weight consistently going to the lowest partial present in the stimulus, i.e., at 600, 1000, or 1400 Hz. Overall, these spectral weighting functions reaffirmed the presence of a frequency dominance region, showing prominent influences on pitch judgment from low-ranked partials. [Work supported by NIH/NIDCD.]

10:30–10:45 Break

10:45

1aPP6. Pitch mechanisms for unresolved harmonics I: Duration effects. Louise J. White and Christopher J. Plack (Experimental Psych., Univ. of Sussex, Brighton, BN1 9QG, UK)

The experiment measured the effects of duration, fundamental frequency ($F0$), spectral region, and relative phase of harmonics on $F0$ discrimination for complex tones consisting of unresolved harmonics. The $F0$ discrimination was measured for complexes with durations between 20 and 160 ms, and with $F0$s of 62.5, 125 and 250 Hz. The complexes were bandpass filtered between either 2750 and 3750 Hz or between 5500 and 7500 Hz. Finally, either all the harmonics were in sine phase (SINE) or alternating harmonics had a cosine phase shift (ALT). The pitch of an unresolved ALT complex is one octave higher than that for an unresolved SINE complex with the same $F0$ [T. M. Shackleton and R. P. Carlyon, J. Acoust. Soc. Am. **95**, 3529–3540 (1994)]. For the SINE stimuli the improvement in $F0$ discrimination with duration was dependent on $F0$, with a larger effect of duration for the lower $F0$s. However, for the ALT stimuli the improvement with duration was the same as that for a SINE complex with twice the $F0$. Overall the results suggest that pitch, rather than $F0$ per se or spectral region, is the main determinate of the duration effect and also possibly of the integration time of the pitch mechanism.

11:00

1aPP7. Pitch mechanisms for unresolved harmonics II: Sensitivity to envelope phase. Christopher J. Plack and Louise J. White (Experimental Psych., Univ. of Sussex, Brighton, BN1 9QG, UK)

Introducing a phase change between two short bursts of a sinusoid can produce a pitch shift which corresponds to a shift in the peak of the long-term spectrum [I. V. Nabalek, Acustica **82**, 531–539 (1996)]. The present experiment examined whether a similar effect could be obtained for a complex tone consisting of unresolved harmonics where the phase change was applied to the envelope of the complex. For unresolved harmonics a pitch is heard corresponding to the fundamental frequency ($F0$) of the complex even though the $F0$ component is not physically present. It is likely that $F0$ is extracted from neural phase locking that provides a crude representation of the envelope of the complex (which repeats at $F0$). The complex consisted of the harmonics of a 250-Hz $F0$ filtered between 5500 and 7500 Hz. A phase advance of 0.75 $F0$ periods between two 20-ms bursts produced a downward pitch shift of almost 10 Hz when the

bursts were contiguous but no downward shift when the bursts were separated by a gap of 5 ms or more. The results are consistent with a long-term pitch mechanism sensitive to envelope phase that can be reset in response to discontinuities in the stimulus.

11:15

1aPP8. Cooperative behavior in the Fitzhugh–Nagumo neuron: Consequences for acoustic perception. Jonathan Tapson (Dept. of Elec. Eng., Univ. of Cape Town, UCT Private Bag, Rondebosch 7700, South Africa)

Recent studies have shown that a type of stochastic resonance, known as cooperative behavior, can occur in biological neurons. This behavior is characterized by phase locking between the pulse position modulation of the neuron's activation and weak sinusoidal input signals. The behavior is strongly influenced by the level of background noise in the system. Significantly, it is common for the phase of the neuron activation to lock on a subharmonic of any sinusoidal input stimulation or on subharmonics of a combination of tones. If the Fitzhugh–Nagumo neuron is simulated as an integrate-and-fire (Wiener) process, it can be shown that a single neuron demonstrates cooperative behavior modes which correspond to the known phenomena of acoustic perception (if the neural input is assumed to be a sonic stimulus). Most significantly, the perception of virtual pitch is apparent, as is the pitch shift effect. When presented with an amplitude modulated waveform, the neuron locks on peak features of the time representation of the waveform, thereby recreating the effects discovered by Schouten [Philips Tech. Rev. **5**, 286–301 (1940)]. It seems likely that this simple model explains some of the physiological bases for the nature of pitch perception.

11:30

1aPP9. The role of global processes in the perceptual cohesion of harmonic complex tones. Jeffrey M. Brunstrom and Brian Roberts (School of Psych., Univ. of Birmingham, Edgbaston, Birmingham B15 2TT, UK, J.M.Brunstrom@Bham.ac.uk)

Harmonic components typically fuse to form a single perceptual entity. Their spectral positions are consistent both globally, because of the regular pattern defined by the other components, and locally, because of their equidistance from adjacent components. This study explored the contribution of global regularity to harmonic fusion. On each trial, subjects heard a periodic complex tone, derived from a 14-harmonic complex, followed by an adjustable pure tone in a continuous loop. Either two, three, or four consecutive components (harmonics six and above) were removed from the original complex and replaced by a ''probe'' component, located at one of a set of regularly spaced positions spanning the gap created in the complex. Subjects were instructed to match the pitch of the pure tone to that of the probe. Generally, they were successful unless the probe coincided with a harmonic position in the complex, though to a lesser extent, this was true even when the probe matched a nonconsecutive harmonic position. This indicates that the cohesion established by a set of harmonics is not dependent upon spectral equidistance between adjacent components. Rather, the effects of this kind of spectral pattern can extend into a multicomponent gap in a harmonic series.

11:45

1aPP10. Accentuation of spectral pitch. Werner A. Deutsch (Acoust. Res. Dept. of the Austrian Acad. of Sci., Liebiggasse 5, A-1010 Vienna, Austria)

Lower harmonics of steady-state complex tones are resolved and produce individual spectral pitches when attention is focused to them. Attention to individual harmonics is enforced by sharp onset cues as simply switching partials on and off. Periodically rising and falling changes of the fundamental frequency of a complex tone (range: 1 oct, rate: 1–0.5 s) results in concurrent percepts. Initially accentuated spectral pitches of short duration are perceived followed by continuing integration of the partials into the complex tone according to the ''common fate'' rule. In comparison to the falling fundamental frequency phase higher partials accentuated during the rising phase are integrated faster. Concurrent spectral and fundamental frequency information exhibit a distinct pitch ambiguity transition immediately after the onset of the spectral pitch. The duration of the transition stage is assumed to reflect a kind of ''Gestalt'' generation time constant until the complex tone accepts its new incoming component. Raising the rate of fundamental frequency change to 4/s and more reduces the effect of the spectral pitch progressively. The pitch change of the complex tone dominates the spectral information. The concurrence is reduced furthermore when onsets are smoothed and consequently changes of timbre are perceived.

12:00

1aPP11. Perceptual segregation and pitch shifts of mistuned components in harmonic complexes and in regular inharmonic complexes. Brian Roberts and Jeffrey M. Brunstrom (School of Psych., Univ. of Birmingham, Edgbaston, Birmingham B15 2TT, UK)

It is unclear whether the perceptual segregation of a mistuned harmonic from a periodic complex tone depends specifically on harmonic relations between the other components. A procedure used previously for harmonic complexes [W. M. Hartmann, S. McAdams, and B. K. Smith, J. Acoust. Soc. Am. **88**, 1712–1724 (1990)] was adapted and extended to regular inharmonic complexes. On each trial, subjects heard a 12-component complex followed by a pure tone in a continuous loop. A mistuning of $\pm 4\%$ was applied to one of the components 2–11. Subjects adjusted the pure tone to match the pitch of the mistuned component. Near matches were taken to indicate segregation. The complexes were either harmonic, frequency shifted, or spectrally stretched. Performance in the inharmonic conditions was generally as accurate as in the harmonic condition. Small but consistent mismatches, pitch shifts, were found in all conditions. Their direction and size were similar to earlier findings for harmonic complexes. Subsequent tests using a range of mistunings showed that the perceptual segregation of components from regular inharmonic complexes could be sensitive to mistunings of 1.5% or less. These findings suggest that the effects of spectral regularity on auditory grouping are more general than those of harmonic relations.

Session 1aSA

Structural Acoustics and Vibration: Analytical Methods in Radiation and Scattering

Jerry H. Ginsberg, Chair
G. W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0405

Chair's Introduction—7:55

Contributed Papers

8:00

1aSA1. Algebraic aspects of acoustic scattering. Elisabeth Fournier, Yves Decanini, Antoine Folacci, and Paul Gabrielli (URA 2053 CNRS, SDEM, Université de Corse, Faculté des Sci., BP 52, 20250 Corte, France)

Acoustic scattering by an arbitrary number of identical cylinders is studied. The S-matrix associated with such a system is obtained by extending the KKR method developed by Kohn, Korringa, and Rostoker in condensed matter physics. Calculations are carried out for various boundary conditions (soft, rigid, and elastic cylinders). Scattering resonances are the poles of the S-matrix. They are determined in the complex plane of the reduced frequency. Scatterers with discrete symmetries, such as: C_{2v} : two cylinders, C_{3v} : three cylinders located at the vertices of an equilateral triangle, C_{4v} : four cylinders located at the vertices of a square, are more particularly studied. Symmetries are reflected in the S-matrix, which reduces to a block diagonal form where the blocks are labeled by the irreducible representations of the symmetry group. This work is a step toward a complete study of chaos (linked to statistical properties of the S-matrix) in the context of multiple scattering.

8:15

1aSA2. Free and forced vibration of a multisupported string. Benjamin A. Cray (Naval Undersea Warfare Ctr., Newport, RI 02841)

This paper examines a simple model: the vibration of an infinite string resting on sets of periodic supports, which are modeled as linear spring-mass-damper systems. The qth set of periodic supports may be offset by a distance xq. These offsets can thus be used to create spatially distributed gratings along the length of the string. Both the free and forced responses are examined, with the objective of gaining an understanding of the vibration response that governs more complicated arbitrarily stiffened elastic structures.

8:30

1aSA3. Eigenvalues gap optimization with sensitivity criteria in coupled acoustic-structural systems. Walter P. Casas[a] (Mech. Eng. Faculty, Natl. Ctr. Univ. of Peru, paucar@fem.unicamp.br) and Renato Pavanello (State Univ. of Campinas, Sao Paulo, Brazil, pava@fem.unicamp.br)

The design optimization can be made following a hard repetitive time consuming process, through the analysis of the design with a few modifications in its parameters, until the required system is encountered. To solve this problem, it is better to use the sensitivity analysis, letting one evaluate the model performance without realizing new analysis. The purpose of this work is to optimize the gap between two eigenvalues in coupled acoustic-structure systems in order to avoid the resonance phenomena in a certain frequency interval, taking the thickness of the structural elements as control variables subject to one constant structural volume. The finite element method is used for discretization of the system, and one nonsymmetrical formulation up with Hermissian matrices in displacement of the structure and fluid pressure describes the system, being necessary to use the left and right eigenvectors concept. For simplicity the nonzero and nonrepeated eigenvalue case is considered. The sequential quadratic programming algorithm is used for the constrained nonlinear optimization, and in the optimization process is considered the modal sensitivity analysis to choose the more sensible parameters. The validity of the study is verified by applying it to a coupled bidimensional system. [a] Currently at UNICAMP-FEM-DMC.

8:45

1aSA4. A modal approach for the acoustic and vibration response of an elastic cavity with absorption treatments and mechanical/acoustical excitation. Noureddine Atalla and Michel Tournour (G.A.U.S., Dept. Génie Mécanique, Univ. de Sherbrooke, Sherbrooke, QC J1K 2R1, Canada)

The calculation of the acoustic and vibration response of an elastic cavity is considered. The presented technique couples a modal decomposition for the structure *in vacuo* and the cavity with a coordinate transformation to allow for efficient determination of the coupled modes. Using these coupled modes, it rewrites the coupled equations accounting for the impedance condition at the treated elastic walls, internal acoustic sources, and mechanical excitations in terms of a compact forced system. The theory behind the method is presented together with validations and numerical examples.

9:00

1aSA5. The effect of large structural perturbations on the response of structural-acoustic systems. Steven Hahn and Aldo Ferri (Georgia Tech, MS 0405, Atlanta, GA 30332, steven.hahn@me.gatech.edu)

Due to high sensitivity, first-order perturbation analysis of structural-acoustic systems can be inaccurate for large perturbations. Using high-order derivatives to create a Taylor series approximation for the perturbed solution can result in slow convergence or even divergence. A rational polynomial or Pade approximation may overcome the poor convergence of the Taylor series by canceling out the pole causing the poor convergence. In this study, a finite element framework is used to describe the structural-acoustic system. External radiation and scattering problems are accommodated by truncating the infinite fluid region using exponential decay infinite elements. Changes to a nominal model are introduced through a structural perturbation that modifies the nominal structural stiffness matrix. An efficient method for calculating the solution derivatives with respect to the structural perturbation is presented. A Taylor series expansion is constructed using the derivative information and the convergence criteria for the series is examined. The local solution and derivatives are then used to construct a Pade approximation. The method is illustrated using several numerical examples. The approximation is shown to be quite accurate for large perturbations even when there are one or more nearby poles and the Taylor series fails to converge.

9:15

1aSA6. Synthesis of the radiated sound field from a vibrating plate-strip with free edges. Kai-Ulrich Machens (Inst. of Tech. Acoust., Univ. of Technol., Berlin, Germany)

A multipole expansion in elliptical coordinates was employed to calculate the directivity pattern of the radiation from free vibrations of a plate-strip with free edges. The Helmholtz equation can be separated in elliptical coordinates and an exact solution for a cylindrical geometry with elliptical cross section can be constructed by Mathieu functions. The radiation problem was solved by mapping the normal velocity of the rectangular cross section of the plate-strip onto an acoustically equivalent elliptical cross section and applying a multipole expansion. A comparison between this method and a standard wave-number transform showed good agreement in the case of a baffled plate, which proved the validity of this procedure. The directivity patterns of baffled and unbaffled plates are compared in order to estimate the influence of hydrodynamic short-circuit around free edges on the radiation. The compact representation of the sound field by the source-strengths is advantageous for the determination of the interference of superimposed eigenmodes for radiation. For applications in active noise control, a direct optimization criteria for secondary forces with respect to the radiated soundpower was obtained. It permits the study of efficiency and stability of the active measure with regard to number and location of the secondary forces.

9:30

1aSA7. A modal formulation for an arbitrary submerged structure in a heavy acoustic medium. Pei-Tai Chen (Dept. of Naval Architecture, Natl. Taiwan Ocean Univ., Keelung, Taiwan, ROC)

The paper studies radiational and dynamic characteristics for an arbitrary elastic structure submerged in a heavy acoustic medium. The equation of motion for the coupled structural/acoustical system is formulated. In the equation, the acoustic loading is expressed as a symmetric complex matrix derived by a surface acoustic reciprocal principle. Due to the symmetric formulation, the ratio of the complex input power exerted by an external force yields an expression for the ratio of quadratic representations of the vibration velocity. A consideration of the stationarity leads to a symmetric eigenvalue problem, in which the eigenvalue is related to the radiation efficiency for the vibration velocity corresponding to the eigenvector. Based on the eigensolution, the forced vibration of the coupled system is decomposed into a set of mutual orthogonal equations where each equation represents a one-degree-of-freedom mass/spring/dashpot system having an impedance determined by the eigensolution. A spherical shell and a slender spheroidal shell are chosen as numerical examples to demonstrate the present analysis.

9:45

1aSA8. Finite-difference studies of structural–acoustic interaction. Ulf Kristiansen (Norwegian Univ. of Sci. and Technol., NTNU Acoust. Group, Dept. of Telecommunications, N-7034 Trondheim, Norway)

The behavior of an acoustic fluid coupled to an elastic structure has been investigated by a numerical technique, i.e., the finite-difference method. The technique is simple to apply and convert the governing differential equations to a system of simultaneous algebraic equations in mesh-point unknowns. The chosen example is a 2D cavity enclosed by three hard walls and a flexible beam. The coupled modes are calculated. To incorporate an external field, the beam is assumed to be situated in an infinite baffle. Coupling to the external field is done either by extending the finite-difference model to this region, or by a boundary element method. Different examples of forcing functions will be discussed: mechanical, acoustical, and turbulent flow. The presentation will also discuss how the errors inherent in this type of numerical approximation might be reduced. Higher-order finite-difference operators and Richardson extrapolation are two such techniques.

10:00

1aSA9. Dispersion-free finite-element methods for the Helmholtz equation. Paul E. Barbone (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington Ave., Boston, MA 02215) and Isaac Harari (Tel Aviv Univ., Tel Aviv, Israel)

The problem of dispersion in finite-element approximations of solutions of the Helmholtz equation are considered. Dispersion errors can accumulate over large propagation distances. Such error is sometimes called the ''pollution error.'' Via a one-dimensional example, it is shown that the dispersion error is linked to the fact that standard finite-element approximations are not optimal (in the sense of H^1 best approximation) for the Helmholtz equation. Thus to eliminate dispersion error in finite-element solutions, the problem of finding the H^1 projection onto a finite-element space of a field satisfying the Helmholtz equation is examined. In one dimension, exact optimality of the finite-element approximation is achieved, and thus exactly zero dispersion is obtained. This leads to a finite-element formulation that is nodally exact in one dimension on *any* linear mesh. In higher dimensions, or formulation is approximately optimal and nonsymmetric. The formulation itself yields an expression for the ''distance'' from optimality, related to the lack of symmetry. Examples are shown that demonstrate the performance of the method in one and two dimensions. [Supported by ONR.]

10:15

1aSA10. A wave-based prediction technique for vibroacoustics: Comparison with the finite-element technique and experimental validation. Paul Sas, Wim Desmet, and Dirk Vandepitte (Dept. of Mech. Eng., Div. PMA, Katholieke Universiteit Leuven, Celestijnenlaan 300B, B-3001, Heverlee, Belgium, wim.desmet@mech.kuleuven.ac.be)

In a coupled vibroacoustic analysis, a structural and acoustic finite-element model must be solved simultaneously. This results in a very large, nonsymmetrical coupled model, which requires a computationally expensive solving procedure. To reduce the size of a coupled model, a wave-based prediction technique has been developed. Instead of dividing the structural and acoustic domain into small elements and solving the dynamic equations within each element using simple approximating shape functions, the entire structural and acoustic domains are described by plane-wave functions, which are exact solutions of the structural and acoustic homogeneous wave equation. To these wavefunctions, particular solutions of the inhomogeneous wave equations are added, so that the governing dynamic equations are exactly satisfied. The contributions of the wavefunctions to the coupled vibroacoustic response are determined by applying the boundary conditions in a weighted residual formulation. The wave technique has been used for the prediction of the acoustic pressure in the cavity of a car breadboard model by applying the measured structural velocity distribution as boundary conditions to the car cavity. The prediction results of the wave model are compared with the finite-element predictions and the measured cavity pressure field.

Session 1aSC

Speech Communication: Speech Technology and Speech Reception (Poster Session)

Michael D. Hall, Chair
Department of Speech and Hearing Sciences, University of Washington, Box 357920, Seattle, Washington 98195-7920

Contributed Papers

All posters will be on display from 9:15 a.m. to 12:15 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 9:15 a.m. to 10:45 a.m. and contributors of even-numbered papers will be at their posters from 10:45 a.m. to 12:15 p.m. To allow for extended viewing time, posters will remain on display until 8:00 a.m. on Tuesday, 23 June.

1aSC1. Factors influencing the detection and intelligibility of speech in competing message environments. Mark A. Ericson (AL/CFBA, 2610 Seventh St., Wright-Patterson AFB, OH 45433-7901), Robert S. Bolia (Systems Res. Labs., Dayton, OH 45440), and W. Todd Nelson (Wright-Patterson AFB, OH 45433-7901)

While monitoring multiple conversations, a listener often needs to know not only what was said but who said what to whom. An exploratory investigation into these questions involved eight talkers saying similar, simultaneous messages. Four male and four female talkers speaking call sign phrases and logical phrases were digitally recorded. All phrases were adjusted in overall gain to have the same average power and onset time. Between one and seven competing messages were played back over loudspeakers in an anechoic chamber. Thereby, angular separation between the speech signal and speech babble was also a controlled factor. Physical attributes of the eight talkers' speech were analyzed using the Multi-Dimensional Voice Program from Kay Elemetrics. Nineteen acoustic attributes of the talkers' voice samples were quantified. Voice samples were analyzed during sustained phonation and continuous speech conditions. A correlational analysis was used to investigate which aspects of the talkers' speech facilitate detection and intelligibility against a cacophony of competing voices. The salient features of the eight talkers' voice samples will be discussed.

1aSC2. The enhancing effect of subsequent context in the perception of the sentence-initial word. Takayuki Kawashima (Dept. of Psych., Div. of Humanities and Sociology, Univ. of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113 Japan), Makio Kashino (NTT Basic Res. Labs, 3-1 Morinosato Wakamiya, Atsugi, Kanagawa, 243-01 Japan), and Takao Sato (Univ. of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113 Japan)

The authors examined whether the perception of the initial word in a Japanese sentence is enhanced by semantically related subsequent context. Phonemic restoration effects under periodic noise replacement were compared for two types of sentences. In the first condition, the initial word of a sentence was semantically predictable from the subsequent context (high-predictability condition). In the second condition, the initial word was not directly predictable from the subsequent context (low-predictability condition). The intelligibility of the sentence-initial words was higher in the high-predictability condition than in the low-predictability condition (58% vs 10%). On the other hand, the intelligibility of the subsequent portions was not significantly different between the two conditions (71.1% vs 66.3%). These results show that word perception can be enhanced by subsequent semantic context in Japanese.

1aSC3. Effects of context-sensitive phonetic variation and lexical structure on the uniqueness of words. Edward T. Auer, Jr. and Lynne E. Bernstein (Spoken Lang. Processes Lab., House Ear Inst., 2100 West Third St., Los Angeles, CA 90057)

Relatively little is known about the optical phonetic speech characteristics to which speechreaders are attuned. However, it is known that phonetic context does affect visual confusability of phonemes. In study 1, behavioral experiments were performed to examine effects of context-sensitive phonetic variation on the visual confusability of consonants and vowels. Hearing and deaf participants identified consonants or vowels presented in carefully constructed sets of nonsense syllables. Vowels were presented in four different consonant–vowel–consonant environments. Consonants and frequently occurring consonant clusters were presented in 14 different disyllabic nonsense words environments. These disyllabic environments varied target consonant position (initial, medial, and final) and target consonant adjacent vowel. Vowel identification was relatively accurate and insensitive to surrounding consonantal context. Consonant identification significantly varied as a function of both position and vowel context. In study 2, computational experiments were performed to assess the importance of patterns of context-sensitive visual confusability on the uniqueness of words in the language. The computational experiments in study 2 provide evidence that choice of a segmental intelligibility model does matter for modeling the uniqueness of words. However, the distribution of words in English substantially preserves lexical uniqueness even when phonetic variability is taken into account. [Work supported by NIH/NIDCD Grant No. DC02107.]

1aSC4. Amplitude modulation cues for perceptual voicing distinctions in noise. Brian P. Strope and Abeer Alwan (Dept. of Elec. Eng., UCLA, 66-147E Eng. IV, 405 Hilgard Ave., Los Angeles, CA 90095)

Fricatives are produced by forming a sufficiently narrow constriction in the vocal tract to generate a turbulent noiselike source. With voiced fricatives, vibrating vocal folds introduce low-frequency energy at the first few harmonics of the fundamental frequency. Measurements using the relative level of the first harmonic have been shown to be good indicators for voiced/unvoiced fricative distinctions [e.g., Pirello *et al.*, J. Acoust. Soc. Am. **101**, 3754–3765 (1997)]. In the current study, fricatives [s z] with the vowels [i a u] were recorded as CV syllables from four talkers. Perceptual thresholds for syllable detection and voicing distinction were measured in noise. Subjects were able to make reliable voicing distinctions at low SNRs after the fricatives were manually isolated from the vowels and spectral components below 3 kHz were removed. Temporal analyses indicate that during voiced fricatives the vibrating vocal folds can modulate the pressure source behind the constriction sufficiently to generate an acoustic pitch-rate amplitude-modulation cue in high-frequency regions. Predictions of the perceptual experiment's thresholds are evalu-

ated for psychoacoustic models of amplitude-modulation detection which use the standard deviation of the temporal envelope, modulation filter banks, and summary autocorrelograms. [Work supported by NIDCD Grant No. DC02033 and NSF.]

1aSC5. Effects of context on time-compressed speech intelligibility in older adults. Nancy Vaughan and Ronda Kasper (Washington State Univ., 601 W. First Ave., Spokane, WA 99201-0399)

The purpose of this study was to assess the effects of context on sentence intelligibility at high compression rates in older adults. A group of older participants listened to time-compressed low-predictability sentences from the SPIN-R test and high-context sentences from the Connected Speech Test (CST). Many studies have shown that speech intelligibility declines significantly at a 60% compression rate especially for older adults. An algorithm called the Synchronized Overlap Add (SOLA) has been shown to produce intelligibility scores above 80% on the CST for adults aged 65–75 years at 60% compression. This method uses an autocorrelation function to combine common elements in overlapping sequential frames of the speech signal, producing highly intelligible speech at high rates of compression. An apparently similar algorithm used in another study of older adults produced intelligibility scores around 50% for the low-predictability sentences of the SPIN-R test at a 60% compression rate. The question of whether the varying results of time-compressed speech intelligibility was due to differences in algorithms or to differences in context of the speech materials remains unanswered. The results of the current study are discussed in terms of speaker gender, context, and methods of time-compression.

1aSC6. Dichotic presentation of speech signal using a critical filter bank for bilateral sensorineural hearing impairment. D. S. Chaudhari and P. C. Pandey (Indian Inst. of Technol., Bombay, India)

Sensorineural hearing-impaired listeners face a particular problem in view of a decrease in the frequency resolving capacity of an ear due to the spread of masking of frequency components by adjacent frequency components. Filtering a speech signal by a bank of critical band filters and adding signals from alternate bands for presenting to the two ears is likely to reduce the effect of the spread of masking along the cochlear partition, and thus may help in improving the speech intelligibility. This processing scheme has been implemented with 18 critical bands for experimental evaluation through listening tests. The scheme was found to be helpful in improving recognition scores and transmission of speech features in listening tests with normal hearing subjects with simulated sensorineural hearing loss. Listening tests were carried out using 12 vowel–consonants–vowel and consonant–vowel nonsense syllables presented in quiet (no masking noise), on ten subjects with mild to very severe sensorineural hearing loss. The stimulus response confusion matrices were analyzed using relative improvement in recognition scores and information transmission analysis. The relative improvement in recognition scores is up to 25%. Information transmission analysis indicated that the overall improvement is contributed to by improvements in transmission of the speech feature of voicing, place, and manner. The mean response time has also been found to decrease significantly.

1aSC7. A new method of sequential information analysis for consonant perception. Chin Tuan Tan, Yit Chow Tong, and Joseph S. Chang (School of Elec. & Electron. Eng., Nanyang Technolog. Univ., Singapore)

A new method of information analysis which considers all entries in the confusion matrix is proposed for identifying important articulatory features for consonant perception. For a system of N features, the important features are selected sequentially in a series of iterations. In the first iteration, the feature whose transmitted information is maximum is selected. In the second iteration, the first feature is combined with each of the remaining $N-1$ features to form $N-1$ sets of two joint features. The second feature is selected by choosing the set of joint features which has the maximum transmitted information. The transmitted information for a set is determined by summing its unconditional transmitted information and a correction factor to account for redundancy. The third and subsequent important feature are selected in the same manner. The selection process terminates when the transmitted information for the selected set of joint features in the last iteration is equal to the maximum possible transmitted information for the confusion matrix. The proposed method of analysis was tested with several confusion matrices published in past research papers, and the results showed that it was able to select a common set of important features consistently across the confusion matrices.

1aSC8. Aural method of speech intelligibility enhancement. Wojciech J. Majewski (Inst. of Telecommunications and Acoust., Wroclaw Univ. of Technol., Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland, majewski@aipsa.ita.pwr.wroc.pl)

Among a variety of speech enhancement techniques, one subjective technique, called the Hollien binaural filtered (HBF) approach [Hollien, *The Acoustics of Crime* (Plenum, New York, 1990)], was proposed. It utilizes a special binaural hearing technique: The modified (usually filtered) speech signal is fed to the listener's dominant ear at a relatively high level and a low-energy unprocessed speech signal is simultaneously sent to the contralateral ear. Unfortunately, there were no investigations related to the effectiveness of this technique of speech enhancement. In the present paper the results of the experiments performed in order to quantitatively evaluate this method are presented. The speech material consisted of nonsense word lists. For each measuring condition, one list containing 100 nonsense words was presented to five experienced listeners. Due to possible applications of the HBF technique in forensic phonetics, poor quality recordings were examined. The major difficulty in the experiments as to obtain the improvement of speech intelligibility by means of conventional speech enhancement techniques. When, however, such technique yielded an increase in speech intelligibility score (the largest increase was 5%), then the HBF technique provided further speech enhancement, the largest increase was 8% and 12%, respectively, in relation to modified and unprocessed speech.

1aSC9. A training system to improve the usefulness of hearing aids. Takashi Ikeda (Kurume Natl. College of Technol., 1232, Komorino-machi, Kurume, 830 Japan), Akira Watanabe, Mitsuaki Hino, and Yuichi Ueda (Faculty of Eng., Kumamoto Univ., 2-39-1, Kurokami, Kumamoto, 860 Japan)

An audiovisual system was proposed previously. It can make the hearing-impaired notice phonemic cues by using vision together with hearing. In the preliminary experiments by a hearing impaired person with a prototype hearing aid, it was noted that the information of visualized speech could contribute to combine sound images and their understanding as phonemes. First, to generalize this method, a new type of training system that consists of a personal computer, audio equipment, training programs and data is presented. The system provides output sounds of hearing aid and visualized speech together automatically. Next, the results of experiments, which are conducted with the groups of hearing-impaired persons by sound of the single resonant analysis type hearing aid, are presented. The contribution of visual information in the training will be examined. Finally, it will be discussed that the design to realize a practical single resonant analysis type hearing aid is small enough to use in the field test.

1aSC10. Voice transmission through vibration pickups. Paolo E. Giua (CNR Istituto di Acustica, via Cassia 1216, Roma, Italy, I-00189)

Devices that pick up vibrations induced in the body by the speaker voice have been used in speech communication for many years in order to enhance intelligibility in noisy environments. This and other interesting applications seem to justify a study of this ''accelerometric'' speech in order to obtain useful correlations with the ''microphonic'' speech. Many factors seem to influence the voice spectrum of the accelerometric speech. A constant general factor is the low-pass filter effect of the acoustical bone conduction. But the main one seems to be the coupling of the voice to the bony transmission channel. This coupling is phoneme dependent. For instance, voiced stop consonants are more amplified than voiced fricatives, while vowels are modified according to their qualities. Moreover speaker features and accelerometer features are to be taken into account. The present results refer to an accelerometer inserted in the ear canal. Main features of this type of accelerometric speech are shown and intelligibility measurements as a function of speakers and their vocal effort are reported. An attempt to recover microphonic speech corrupted by noise by means of a simultaneous accelerometric speech is described.

1aSC11. Enhancing second formant identification in speech by listeners with significant hearing loss. Peggy B. Nelson and Sally G. Revoile (Univ. of Maryland School of Medicine, 16 S. Eutaw St., Ste. 500 Baltimore, MD 21201)

Listeners with significant hearing loss have consonant recognition abilities that appear limited by their inability to use spectral peak information in the higher frequencies, specifically in the highly informative second formant ($F2$) frequency region. For them, ensuring the audibility of speech may not result in full recognition of consonants and vowels that differ by $F2$ characteristics. In the current study, listeners with moderate to severe hearing loss identified naturally spoken consonant-vowel-consonant stimuli (CVCs) with a wide range of $F2/F1$ formant frequency characteristics at levels that ensured the audibility of the $F2$ region. CVCs were presented in three conditions: (a) unmodified, (b) filtered to enhance the $F1$ and $F2$ regions, and (c) frequency-modulated to transpose the $F2$ peak to a lower, more detectable frequency region. The amount of modulation for condition (c) was determined by measuring individual spectral peak detection thresholds for each listener using the rippled-noise method described by Summers and Leek (1994). Closed-set identification of the CVC stimuli was measured and consonant confusion matrices generated. The amount of improvement due to enhancing the spectral contrast and to transposing the $F2$ region will be reported. Correlations between listeners characteristics and enhancement due to contrast sharpening and/or $F2$ transposition will be reported. [Research supported by NIDCD K08-000124 and by the University of Maryland School of Medicine.]

1aSC12. Subjective versus objective intelligibility of sentences in listeners with hearing loss. Kathleen M. Cienkowski, Charles Speaks, and Amy S. McKee (Dept. of Commun. Disord., Univ. of Minnesota, 115 Shevlin Hall, Minneapolis, MN 55455)

The relation between estimated intelligibility and correct key-word repetitions of sentences was examined in listeners with normal hearing and hearing loss. The ten lists of CID sentences [S. R. Silverman and I. J. Hirsh, Ann. Otol. Rhinol. Laryngol. **64**, 1234–1245 (1995)] were arranged in two randomly ordered groups. Listeners were asked to complete two tasks: (1) rate the intelligibility of 50 sentences on a scale of 0% to 100%, (2) repeat each of the 50 sentences, which were scored as the number of key words repeated correctly. Sentences were presented at five signal-to-noise ratios, +6, +4, +2, 0, and −2 dB. The 0 dB S/N point was established by the method of adjustment using the revised speech intelligibility rating passages to reach a criterion of 50% intelligibility. Correlations of rated intelligibility and correct repetition of key words were 0.85 for listeners with normal hearing and 0.87 for listeners with hearing loss. These findings confirm earlier work [Speaks *et al.*, J. Speech Hear. Res. **15**, 590–602 (1972)] with listeners with normal hearing and demonstrate that listeners with hearing loss preserve the ability to rate intelligibility of speech.

1aSC13. Effects of stimulation mode on speech recognition by cochlear implant users. Qian-Jie Fu and Robert V. Shannon (Dept. of Auditory Implants and Percept., House Ear Inst., 2100 West 3rd St., Los Angeles, CA 90057)

This study investigated the effects of both mode of stimulation and frequency allocation on speech recognition in three subjects implanted with Nucleus-22 multichannel cochlear implant. A four-channel processor with continuous interleaved sampling speech processing strategy was implemented through a custom interface. The temporal envelopes from four broadbands were used to modulate 500 pps, 100 ms/phase pulse trains and then delivered to four electrodes 16, 11, 6, 1 with three stimulation modes: BP+1, BP+3, BP+5. Ten different frequency allocations were used and each frequency allocation mapped different insertion depths of these four electrodes based on Greenwood's place-to-frequency function. Results showed that the recognition score was highly dependent on the frequency allocation in all stimulation modes, indicating that it may be indispensable to match the frequency allocation to the tonotopic location of activated electrodes. Results also showed that a speech processor with BP+5 stimulation mode was more robust to the mismatching between the frequency allocation and tonotopic location of electrodes, indicating that a speech processor with more broad stimulation, such as a monopolar stimulation mode, may be advantageous for most cochlear implant users, especially for those who have either shallower insertion depth or uneven nerve survival patterns. [Work supported by NIDCD.]

1aSC14. Continuous assessment and modeling of speech transmission quality. Martin Hansen and Birger Kollmeier (AG Medizinische Physik, Universität Oldenburg, D-26111 Oldenburg, Germany, martin@medi.physik.uni-oldenburg.de)

This paper addresses the question of whether subjects are able to assess the perceived time-varying quality of speech material continuously. A method inspired by Hamberg and de Ridder [J. Opt. Soc. Am. A **12**, 2573–2577 (1995)] is presented by which subjects continuously rate the perceived speech quality by moving a slider along a graphical scale. In this way, temporal variations in quality can be monitored quantitatively and differences between, for example, alternative speech transmission systems, can be analyzed in an informative way. The usability of this method is illustrated with an experiment in which different sequences of sentences with a total duration of 40 s were degraded in quality with a modulated noise reference unit. The modulation depth was varied with time and the subjects task was to assess the perceived quality. The results show that subjects can monitor speech quality variations closely with a delay of approximately 1 s. The subjective results are discussed in view of a psychoacoustically motivated objective speech quality measure [Hansen and Kollmeier, ICASSP '97, 1387–1390 (1997)].

1aSC15. Coherence-based sub-band decomposition for efficient reverberation and noise removal in enclosed sound fields. Joaquin Gonzalez-Rodriguez and Javier Ortega-Garcia (Dept. of Audiovisual and Commun. Eng., E.U.I.T. Telecomunicacion, Univ. Politecnica de Madrid, Spain, jgonzalz@diac.upm.es)

The issue of speech enhancement in reverberant rooms has been extensively addressed. The use of sub-band decomposition through nested linear arrays has been shown to be very efficient for this task. Recently, the coherence/noncoherence nature of the sound field has been used to separate the wideband speech signal in two sub-bands, obtained from the sinc function that characterizes the diffuse field prediction for a given separation between microphones when many independent sources are present and the reverberation time is not too short. The processing scheme

proposed here takes advantage of both ideas. After a sub-band decomposition from a linear log-spaced 15 microphone array, a further subdivision is made in each processing band, dividing in a coherent sub-band, where a Generalized–Sidelobe–Canceller is applied, and a noncoherent sub-band, where a Wiener postfilter is used, resulting in a total of six analysis subbands. Promising results have being obtained with room simulations (image method), and further results are being obtained with the Carnegie Mellon University (CMU) multichannel-recorded database. The authors thank Dr. Finn Jacobsen (Technical Univ. Denmark) for his fruitful discussions about spatial coherence, and Dr. Richard Stern (CMU, PA, USA) for the multichannel database. [Work supported by CICYT under Project TIC97-1001-C02-01.]

1aSC16. Continuous digit speech recognition using subword unit HMM juncture. Jeh-Seon Youn and Kwang-Seok Hong (Dept. of Electron., Sung Kyun Kwan Univ., Chun-Chun Dong 300, Chang-An Gu, Kyung-Gi Do, Suwon 440-746, Korea)

In this paper, a new recognition system is described for digit strings of various length in Korean under fluent speech conditions using subword HMM juncture. A robust automatic recognition system for connected digit strings under spontaneous speech conditions is crucial for a number of applications such as voice dialing and credit card or personal identification numbers. It is thought that the recognition of connected digits is a relatively easy task, but the output of the recognizer must be a very exact transcription of the input speech. The process of recognition is to construct the database of digit string models by concatenating subword clusters not trained on target application through the typing task, avoiding the recording of training data each time the digit changes. Thus, the acoustic segments divide into N subword clusters, and each subword cluster generates the HMM model. The goal is to give users a flexible number recognition that can add new connected digits without collecting a new corpus or training new models, namely that can be easily modified for different applications with minimum amount of training and modification effort.

1aSC17. Stochastic word models for articulatorily constrained speech recognition and synthesis. John Hogden, David Nix (Los Alamos Natl. Lab., Los Alamos, NM 87545), Vincent Gracco, and Philip Rubin (Haskins Labs., New Haven, CT 06511)

Stochastic word models based on hidden Markov modeling (HMM) techniques are commonly used for speech recognition and are increasingly being used for speech synthesis. An alternate technique for generating stochastic word models using Maximum Likelihood Continuity Mapping (MALCOM) will be presented. MALCOM generates a stochastic model of speech assuming (1) that speech sounds are periodically emitted as a point moves smoothly through a low-dimensional space called a continuity map (CM), and (2) that the sound emitted at time t is a probabilistic function of the position of the point at time t. The assumptions underlying MALCOM are intended to mimic speech production in that (1) speech sounds are produced as the articulators move slowly through a low-dimensional articulator space, and (2) the speech sound produced at time t is a function of the articulator positions at time t. A method for finding CM trajectories (analogous to articulator trajectories) corresponding to a known sequence of phonemes will be shown, and how to use these trajectories as word models for speech recognition and synthesis will be shown. A preliminary test of the theory will be presented in which articulator trajectories estimate given phoneme sequences and phoneme temporal positions are compared to measured articulator trajectories.

1aSC18. Elliptical wave radiation from a curved waveguide. Rolphe F. Fehlmann (Swiss Software Eng. College, Morgartenstr. 2, CH-3012 Berne, Switzerland)

This paper deals with the investigation of the wave radiation properties of the resonance modes in 3-D curved waveguides. With the finite element method (FEM), a curved waveguide has been modeled and the computer simulations have shown that the wave radiation changes from circular waves in two dimensions to spherical waves in three dimensions, respectively, to beam formation. Interestingly enough, however, for some certain modes there exist elliptical and ellipsoidal waves, respectively. A visualization of these waves emanating from such nonaxisymmetric waveguide geometries is presented as pressure fields as well as phase fronts. This so far unknown elliptical field radiation is a phenomenon which future acousticians of the 21st century will have to cope with.

1aSC19. An efficient Korean speech recognition for name and digit on a communication network. Hyeon-Gu Lee and Kwang-Seok Hong (Dept. of Electron., Sung Kyun Kwan Univ., Chun-Chun Dong 300, Chang-An Gu, Kyung-Gi Do, Suwon 440-746, Korea)

The necessity of service using speech recognition goes on increasing day by day, but for the difficulty of speech recognition, the efficient application service is not activated yet. For the security of users and the convenient use of a disabled person, a desire for using speech about the user's name, ID number, or an account which is mainly used in service is increasing accordingly as the information and banking service on communication network increases. This paper presents the efficient Korean speech recognition method for name and digit string on a communication network. Because most of the traditional methods recognize the full name at once, there are some problems, in that they must have all patterns of the name or require very accurate segmentation of the phoneme, etc. So the unit of Korean spelling pronunciation that is contained in the name is recognized, and then a connection is made to the full name recognition. In the recognition of the digit, semisyllable and the combination of a connected semisyllable are used. That is, 20 semisyllables from 0 to 9 are used as the pattern of the start semisyllable or the end semisyllable of the digit string in recognition, and 100 connected semisyllables from 00 to 99 are used as the pattern of the connection part of each digit string. In comparison with the existing method, all names and digits, regardless of their kind and figure can be recognized.

1aSC20. A multilevel training procedure for speech recognition in noise. Chin Tuan Tan, Yit Chow Tong, Joseph S. Chang (School of Elec. & Electron. Eng., Nanyang Technolog. Univ., Singapore), and Lai Peng Chan (Audiological Services Clinic, Singapore)

A multilevel training procedure is proposed for automatic speech recognition in noise. In the multilevel training procedure, six levels of Gaussian noise, corresponding to six signal-to-noise ratios (SNR), were added to the training samples. The procedure was tested in three speech recognizers implemented, respectively, on the basis of the following three speech recognition techniques: dynamic time warping (DTW), discrete hidden Markov modeling (DHMM), and neural network (NN). The tests on the three recognizers were conducted using a common database for isolated word recognition of the ten digits. The results of the first series of the tests showed that when the recognizers were trained using clean training samples without noise and tested with noisy test samples with additive Gaussian noise, the performance of the three recognizers degraded significantly as the SNR decreased. In the second series of tests, the multilevel training procedure was applied. The results showed that, when the recognizers were trained with noisy training samples and tested with noisy test samples, the three recognizers were, in contrast to the first series, much better able to maintain their performance as the SNR decreased. The multilevel procedure is therefore effective in improving the performance of recognizers for automatic recognition in noise.

1aSC21. An implementation of the recognition system for Windows 95 commander. Seong-Kwon Lee and Soon-Hyob Kim (Dept. of Comput. Eng., Kwangwoon Univ., Wolgy-dong, Nowon-gu, Seoul, Korea)

This paper is a study on the window commander recognition system. The applied recognition algorithm is a one-stage/DP with DMS(dynamic multisection) model which generates a reference pattern and recognizes 60 commanders in Windows 95. From the results of experiments, the optimal recognition considering a tradeoff of real time is obtained by applying a DMS model with 20 section and 13 order PLP coefficients for feature parameter. The whole system is constructed by a soundblast card within an IBM-PC/586. After the commander is recognized, the commander is processed automatically. The references are made by five male speakers in a noisy laboratory environment. The recognition experiment is performed for 60 different commanders and the recognition rate was 96.2% in multispeaker independent.

1aSC22. Fractal modeling of a glottal waveform for high-quality speech synthesis. Naofumi Aoki and Tohru Ifukube (Res. Inst. for Electron. Sci., Hokkaido Univ., Sapporo, Japan, aoki@sense.hokudai.ac.jp)

From several psychoacoustic experiments, it is well known that the oversimplified glottal waveform degrades the quality of synthesized speech. For the enhancement of the quality of synthesized speech, it is desired to generate a glottal waveform in detail using an appropriate model which requires a few control parameters. In this study it is supposed that the differential glottal waveform, which is estimated by linear prediction filter, is modeled as a random fractal signal. The differential glottal waveform is regarded as the first integral of white noise whose frequency characteristic is approximated as -6 dB/oct. This signal can be interpreted as Brownian motion in terms of fractal theory. In order to generate random fractal signal, a kind of wavelet synthesis method called the midpoint displacement method is employed. Psychoacoustic experiments were conducted to examine whether or not the proposed method is effective. Experimental results indicate that the quality of original speech is fully preserved even if more than 90% of wavelet coefficients are replaced with a random fractal signal generated by the proposed method. The wavelet coefficients that cannot be replaced are to be the control parameters in this model.

1aSC23. Interactive Korean speech recognition for a mobile robot. Jae-Young Lee, Seok-Hyun Yoon, Kwang-Woo Chung, and Kwang-Seok Hong (Dept. of Electron., Sung Kyun Kwan Univ., Chun-Chun Dong 300, Chang-An Gu, Kyung-Gi Do, Suwon 440-746, Korea, kshong@yurim.skku.ac.kr)

This paper is a study on the implementation of a speech recognition mobile robot to which the interactive Korean speech recognition techniques are applied. The speech command uttered the sentential connected word and is asserted through the wireless mic system. This speech signal transferred LPC-Cepstrum and short-time energy which are computed from the received signal on the DSP board to a notebook PC. In the notebook PC, the DP matching technique is used for the recognizer and the recognition results are transferred to the motor control unit which outputs pulse signals corresponding to the recognized command and drives the stepping motor. It has been implicitly assumed that, for connected word recognition, each word in the string can be followed by any other word in the string. However, for some connected word recognition tasks, there is an explicit set of rules governing which words can logically follow other words to form valid sentences in the language. In this paper, a grammar network was applied to reduce the recognition speed of the recognizer, so that real time recognition is realized. The misrecognized command is revised by interface revision through conversation with the mobile robot. The recognition system asks the user for help whenever there are two or more recognition candidates whose likelihood scores are all high and it is difficult to resolve small differences in the string. The recognizer asks the user to verify the first choice decision; if it is not verified, the recognizer asks the user to verify the second choice, etc. Alternatively, the recognition system can list the confusing candidates and ask the user to choose the correct string based on the string index number in the list.

1aSC24. Synthesis of fricative sounds using an aeroacoustic noise generation model. Daniel J. Sinder, Michael H. Krane, and James L. Flanagan (CAIP Ctr., Rutgers Univ., Piscataway, NJ 08855)

Results from speech synthesis with an aeroacoustic fricative source model are presented. The source model, which is implemented in the time domain, is based on Howe's reformulation of the acoustic analogy [Howe, J. Fluid Mech. **71**, 625–673 (1975)]. It specifies the source as a function of flow conditions and the geometry of the vocal tract downstream from a constriction where a turbulent jet is formed. This is accomplished by combining elements of aeroacoustic theory and a model for jet behavior with a traditional model for wave propagation in the vocal tract. Using this formulation, the source strength, impedance, and spectrum can be determined through post-processing. The fricative model is implemented in an articulatory speech synthesizer. Evaluation of the resulting synthesis is emphasized. Comparisons are made with physical experiments in static configurations. Additionally, speech fricatives are compared with natural speech as well as synthetic speech from other models. These comparisons are made in the frequency domain and through listening tests in which fricatives in V-C-V sequences are judged for naturalness and intelligibility. [Research supported by NSF/ARPA IRI-9314946 and ARPA DAST 63-93-C-0064.]

1aSC25. Glottal source parameter estimation by comparison of measured signals with simulated signals. David G. Druker (Dept. of Speech Pathol. and Audiol., Univ. of Iowa, 361 East Sherman Ave., Salt Lake City, UT 84115, david.druker@hsc.utah.edu), Ingo R. Titze (Univ. of Iowa, Iowa City, IA 52242), and Brad H. Story (Wilbur James Gould Voice Res. Ctr., Denver, CO 80204)

Oral flow, electroglottographic (EGG), and photoglottographic (PGG) signals were simultaneously digitized from a single subject phonating sustained vowels. Analogous simulated signals—oral flow, vocal fold contact area, and glottal area, respectively—were produced by a driven, kinematic glottal source model coupled to a wave-reflection-type vocal tract. The vocal tract area function used in the simulation was estimated from magnetic resonance imaging (MRI) data previously published for the subject. This essentially removed the vocal tract as a confounding factor in the analysis. The measured and simulated signals were compared and an attempt was made to adjust model parameters so that the least-squared differences between the signals were minimized. The model parameters, which included subglottal pressure, vocal fold adduction, vocal fold convergence, and fundamental frequency were observed to be estimates for the glottal configuration and control parameters used by the subject.

1aSC26. SGP analysis of excised larynges in anechoic chamber. Paul F. Castellanos, Sang Hung, Craig Skinner (Div. of Otolaryngol., Head and Neck Surgery, Univ. of Maryland School of Medicine, Baltimore, MD 21201), and Samuel A. Elder (Dept. of Phys., U.S. Naval Acad., Annapolis, MD 21402)

Single glottic pulse analysis will be presented of experimental data obtained on excised larynges in an anechoic chamber. Attention is directed to understanding the nature of the acoustic source, and comparison will be made with data taken from live subjects. *In vivo* measurements include those taken in the chamber as well as those recorded by Sondhi tube.

1aSC27. Perceptual effects of spectral envelope and *F*0 manipulations using the STRAIGHT method. Hideki Kawahara (ATR HIP Labs., Wakayama Univ., Wakayama, 640 Japan, kawahara@hip.atr.co.jp) and Reiko Akahane-Yamada (ATR HIP Labs., Kyoto, 619-02 Japan)

A versatile speech manipulation method, STRAIGHT (Speech Transformation and Representation using Adaptive Interpolation of weiGHTed spectrogram) [Kawahara, ICASSP (1997)], was applied to natural speech in order to test the effects of speech parameter manipulations. STRAIGHT consists of three procedures called STRAIGHT-core, TEMPO (Temporal domain Excitation extraction using a Minimum Perturbation Operator) and SPIKES (Synthetic Phase Impulse for Keeping Equivalent Sound). STRAIGHT-core is a method to extract a smoothed time-frequency representation, which is free from interferences due to the source periodicity. TEMPO is used to extract *F*0 and other source-related information. SPIKES provides artificial ''naturalness'' to the synthetic speech. As a base line, resynthesized speech, using the STRAIGHT method, was found to provide equivalent ''naturalness'' compared to the original speech when no parametric modification was introduced. Simultaneous manipulation of the spectral envelope and *F*0 illustrated that stretching the frequency axis of the spectral envelope, using the quantity of the cubic root of *F*0 stretching, preserved ''naturalness'' the best. Perceptual effects of group delay used in the SPIKES procedure were also investigated. [Work supported by CREST (Core Research for Evolutional Science and Technology) of Japan Science and Technology Corporation (JST).]

1aSC28. Modeling and synthesis of the lateral consonant /l/. Adrienne M. Prahler (Res. Lab. of Electron. and Dept. of EECS, MIT, 50 Vassar St., Cambridge, MA 02139, aprahler@mit.edu)

The lateral consonant in English is generally produced with a backed tongue body, a midline closure of the tongue blade at the alveolar ridge, and a path around one or both of the lateral edges of the tongue blade. In prevocalic lateral consonants, the release of the closure causes a discontinuity in the spectral characteristics of the sound. Past attempts to synthesize syllable-initial lateral consonants using formant changes alone to represent the discontinuity have not been entirely satisfactory. Data from prior research has shown rapid changes in not only the formant frequencies but also in the glottal source amplitude and spectrum and in the amplitudes of the formant peaks at the consonant release. Further measurements have been made on additional utterances, guided by models of lateral production. Based on these data, new synthesis attempts have incorporated changes in source amplitudes, formant bandwidths, and the location of a pole-zero pair. Including these additional parameters improves the naturalness of the synthesized lateral-vowel syllables in initial perception tests. Results of listening tests indicating the relative importance of the various factors to the naturalness of the synthesized syllables will be reported. [Work supported in part by a LeBel Fellowship and by NIH Grant No. DC00075.]

1aSC29. End corrections in an acoustic model of the vocal tract. J. Dang (ATR HIP, 2-2 Hikaridai Seikacho Soraku-gun, Kyoto, 619-02 Japan)

Open end corrections are required in an acoustic model to compensate for the effects caused by abrupt area changes in the vocal tract, such as the open ends of the side branches and the opening between the teeth. The end corrections were studied by means of a number of mechanical models which imitated vocal tract geometries. The results showed that the end correction greatly depends on the local geometries in the vicinity of the open ends. On the basis of these measurements, two empirical formulas have been proposed to account for the acoustic effects of the geometries [Dang *et al.*, TR-H-219 (1997)]. In this study, these formulas were applied to morphological data of the vocal tract. The range of the open end correction coefficient (OECC) for estimating an accurate vocal tract transfer function was evaluated. The OECC of the nasopharyngeal port was found to be about 0.8 for the oral tract side and 0.3–0.6 for the nasal tract side. For the ostia of the paranasal cavities, the value was about 0.8 for the cavity side and 0.28 for the nasal tract side. An OECC of about 0.4 was sufficient to compensate for end radiation of the laryngeal tube outlet for front vowels.

1aSC30. Modeling the transitory behavior of speech using a time-varying transmission line. Amit S. Rane, Derrick C. Wei, Lisa E. Falkson, and Abeer A. Alwan (Speech Processing and Auditory Percept. Lab., Dept. of Elec. Eng., UCLA, 66-147E Eng. IV, 405 Hilgard Ave., Los Angeles, CA 90095, alwan@icsl.ucla.edu)

In this study, a transmission-line model of speech production [Rael *et al.*, J. Acoust. Soc. Am. **97**, S1(5) (1995)] is modified so that changes in the vocal-tract (VT) area functions, reflecting changes in the VT shape from one sound to another, are properly modeled. Each VT shape is divided into uniform cylindrical sections and each section is modeled by its equivalent transmission line comprising resistors (R), capacitors (C), and inductors (L). A radiation impedance terminates the transmission-line sections. For voiced sounds, the KLSYN88 glottal waveform [D. H. Klatt and L. C. Klatt, J. Acoust. Soc. Am. **87**, 820–857 (1990)] is used as input. The model is simulated in time and frequency by using the analog-circuit simulator HSPICE. To simulate the transitory characteristics of speech, different interpolations such as linear, cubic, and Fourier transform are used for area transitions. The interpolated values result in a ''sampled'' area function and the RLC elements of each section are varied according to these time-sampled values. These variations are accomplished by utilizing time-varying independent voltage sources. Examples using American English dipthongs will be presented. Formant frequencies (both steady state and transitions) of the synthesized and natural dipthongs are in close agreement. [Work supported by NSF Award IRI-95-03089.]

1aSC31. Pitch-synchronous decomposition of mixed-source speech signals. Philip J. B. Jackson and Christine H. Shadle (Dept. Electron. and Comput. Sci., Univ. of Southampton, Southampton SO17 1BJ, UK)

As part of a study of turbulence-noise sources in speech production, a method has been developed for decomposing an acoustic signal into harmonic (voiced) and anharmonic (unvoiced) components, based on a hoarseness metric [Muta *et al.*, J. Acoust. Soc. Am. **84**, 1292–1301 (1988)]. Their pitch-synchronous harmonic filter (PSHF) has been extended (to EPSHF) to yield time histories of both harmonic and anharmonic components. These corpus include many examples of turbulence noise, including aspiration, voiced and unvoiced fricatives, and a variety of voice qualities (e.g., breathy, whispered). The EPSHF algorithm plausibly decomposed breathy vowels, but the harmonic component of voiced fricatives still contained significant noise, similar in shape to (though weaker than) the ensemble-averaged anharmonic spectrum. In general, the algorithm performed best on sustained sounds. Tracking errors at rapid transitions due to jitter and shimmer were spuriously attributed to the anharmonic component. However, the extracted anharmonic component clearly exhibited modulation in voiced fricatives. While such modulation has been previously reported (and also in hoarse voice), it was verified by tests on synthetic signals, where constant and modulated noise signals were extracted successfully. The results suggest that the EPSHF will continue to enable exploration of the interaction of phonation and turbulence noise.

1aSC32. Automatic creation of CV templates for formant type speech synthesis based on HMM-based segmentation and syllable boundary detection. Takahiro Ohtsuka, Chang-Sheng Yang, and Hideki Kasuya (Faculty of Eng., Utsunomiya Univ., Utsunomiya, 321 Japan)

A formant type speech synthesis method has an enormous advantage in that it allows one to generate speech with various voice quality variations and talker individualities. But it has suffered from unnatural speech quality, not because of theoretical limitations but because of an incomplete set of rules for the synthesis. Any insufficient approximation to the acous-

tics causes degradation of the perceived quality of synthetic speech. A novel formant type speech synthesizer in Japanese based on concatenation of CV (consonant-vowel) formant-source templates obtained from natural utterances has been investigated, in which multiple sets of formant and voice source parameter values are used for each of the CV syllables. This paper describes an automatic method to create the CV formant-source templates from speech corpus. The ARX (autoregressive with exogenous input) analysis method is first used to automatically extract formant and voice source parameters and then an HMM based segmentation is performed to locate the CV segments. The segments are further analyzed to detect a starting point of the syllable. A distance measure is used to decide the number of templates needed for each of the syllables. The method is proved to be useful in creating CV templates by experiments performed on 503 Japanese sentences.

1aSC33. A study on natural-sounding Japanese phonetic word synthesis based on the pitch waveform concatenation. Yasuhiko Arai, Ryo Mochizuki (AVC Development Lab., CED, Matsushita Commun. Ind. Co., Ltd., 600 Saedo-cho, Tsuzuki-ku, Yokohama, 224 Japan, arai@adl.mci.mei.co.jp), and Takashi Honda (Meiji Univ., Kawasaki, 214 Japan)

A speech output system in which messages are composed by embedding some phonetic words in the prerecorded sentence patterns requires very natural-sounding phonetic words to be embedded. Especially in the case where a very large or an unlimited vocabulary is required, synthesis of very natural-sounding phonetic words would be desired. Therefore, a novel Japanese VCV-concatenation synthesis method based on the pitch waveform concatenation has been studied. The VCV waveforms are extracted from the VCV-balanced spoken word database and pitch waveforms are extracted by means of the excitation synchronous pitch waveform extraction method [Y. Arai *et al.*, ICSLP **96**, 1437–1440 (1996)]. The wavelength of each pitch waveform is modified depending on the pitch modification rate, and the global pitch contour of each VCV waveform is modified along with the target pitch pattern, while the microprosody and the small perturbation in pitch frequency inherent in that VCV waveform are retained. The auditory tests confirmed the superiority of the novel synthesis method. Finally, the advantage of the spoken word database, which consists of Japanese VCV-balanced word sets uttered forcibly by two pitch accent types (so-called type -0 and type $+1$ in Japanese) differently from the normal pitch accent, is discussed.

1aSC34. The Hablarte text-to-speech system. Helen E. Karn (Office of the Dean, Georgetown College, Georgetown Univ., Box 571003, Washington, DC 20057-1003, karnh@gunet.georgetown.edu)

This paper presents an overview of the architecture of Hablarte, an unrestricted text-to-speech system for Spanish which runs on DOS-based personal computers. The input to Hablarte is an ASCII text with conventional Spanish orthography and punctuation. The front-end text processing routines convert the input text's graphemes to phonemes, determine phonological phrases, syllable structure, and the location of primary stress, and assign prosodic values to the phonetic representation [H. E. Karn, ''Aspects of phonetics and phonology in a text-to-speech system for Spanish,'' Doctoral dissertation, Georgetown University, 1996]. The acoustic signal is generated by Mbrola v2.02, using the Spanish male diphone database and utilities of the MBROLA project [http://tcts.fpms.ac.be/synthesis/mbrola.html]. Evaluations of Hablarte's text processing components and sample audio output are given.

1aSC35. Pitch-controlled variable bitrate CELP speech coding. Robert Oberhofer and Frank Owens (Univ. of Ulster, Jordanstown, Co. Antrim, BT37 0QB, Northern Ireland, r.oberhofer@ulst.ac.uk)

Most research in efficient speech coding concentrated for many years on algorithms which produced a fixed bit rate. Fixed bit rates are, however, not a requirement for modern packet-based telecommunications and computer networks as well as voice storage applications. This makes the implementation of speech coders feasible, which adapt to the changing properties of speech for improved efficiency or quality. Variable bit-rate coding has therefore become the focus of considerable research activity in recent years [A. Gersho and E. Paksoy, ''Variable Bit Rate Coding,'' Signal Processing VII, Theories and Applications, EUSIPCO 1994, pp. 1169–1173; L. Zhang *et al.*, ''A CELP Variable Rate Speech Codec with Low Average Rate,'' ICASSP 1997, pp. 735–738; B. C. Xydeas, ''Source Driven Variable Bit Rate Prototype Interpolation Coding,'' ICASSP 1996, p. 220; B. Shen *et al.*, ''A Robust Variable-Rate Speech Coder,'' ICASSP 1995, p. 249]. While all of these coders focus on classification procedures for efficient bit allocation, the analysis frame size remains mainly static. The coder featured in this paper adapts the analysis frame size and bit allocation according to the pitch of the signal. A pitch frame extractor and classifier at the front end feeds the detected frame into the speech coding back end, which uses traditional CELP-based techniques [B. S. Schroeder *et al.*, ''CodeExcited Linear Prediction (CELP): High Quality Speech at Very Low Bit Rates,'' ICASSP 1985, pp. 937–940]. This approach allows the reduction of the bit rate from a constant 4800 bps (bits per second) with original CELP to typically 3200–4000 bps, depending on the speech situation. This figure is even reduced to typically 2600–3200 with the inclusion of Voice Activity Detection. The speech quality equals the quality of the original CELP coder.

1aSC36. Assessment of the intelligibility and perceived quality of speech produced by text-to-speech engines. Michael D. Hall, Erica B. Stevens, Richard Eyraud, Denise M. Padden, and Patricia K. Kuhl (Dept. of Speech and Hearing Sci., Univ. of Washington, Seattle, WA 98195, mdhall@u.washington.edu)

The intelligibility and quality of speech produced by current text-to-speech (TTS) engines was assessed using well-established tests of speech intelligibility. TTS engines included Microsoft's Whistler, AT&T's Watson, and Elan Informatique's ProVerbe. The speech intelligibility tests varied in complexity and in the degree to which syntax and lexical knowledge could aid perception. Stimuli included CVC words, semantically meaningful sentences, semantically anomalous sentences, and paragraphs representing different styles of discourse. Intelligibility was evaluated using transcription accuracy for isolated word and sentence items, as well as ratings of the clarity of articulation for word and paragraph items. Perceptual errors on word items also were analyzed by phoneme and word position (initial, medial, or final). Additional ratings were used to determine the perceived ''naturalness'' of meaningful sentences and paragraph items. Across levels of speech complexity, clear and consistent differences between TTS engines were obtained in terms of both intelligibility and naturalness. Possible explanations of the results will be discussed, as will potential areas where text-to-speech synthesis can generally be improved.

1aSC37. Acoustic-phonetic features for the automatic recognition of stop consonants. Ahmed M. Abdelatty Ali, Jan Van der Spiegel (Dept. of Elec. Eng., Univ. of Pennsylvania, 200 South 33rd St., Philadelphia, PA 19104-6390), and Paul Mueller (Corticon, Inc., King of Prussia, PA 19406)

Despite the recent successes in the field of automatic speech recognition, more research is still needed in order to understand the variability of speech and the acoustic characteristics of speech sounds in different contexts and for different speakers. In this paper, the acoustic-phonetic characteristics and the automatic recognition of the American English stop consonants are investigated. The acoustic features that exist in the literature are evaluated and new features are proposed. To test the value of the

extracted features, a knowledge-based acoustic-phonetic system for the automatic recognition of stops, in speaker independent continuous speech, is proposed. The system uses an auditory-based front-end processing and incorporates new algorithms for the extraction and manipulation of the acoustic-phonetic features that proved to be rich in their information content. Several features, which describe the burst frequency, formant transitions, relative amplitude, spectral shape, and duration, are combined in the recognition process. Recognition accuracy of 95% for voicing detection and 90% for place of articulation detection are obtained for TIMIT database continuous speech of multiple speakers from different dialect regions. The obtained results are analyzed and compared to previous work. [Work was supported by Catalyst Foundation.]

1aSC38. An 800 bps VQ-based LPC voice coder. Xianglin Wang and C.-C. Jay Kuo (Elec. Eng. Dept., USC, 3740 McClintock Ave., Los Angeles, CA 90089-2564)

In this paper, it is investigated how to efficiently apply VQ on LPC so as to realize a very low bitrate speech coder with reasonable speech coding quality. The coding bitrate has been pushed to 800 bps and the result is compared with that of a traditional LPC 2.4 kbps speech coder. In LPC10 standard(FS1015), scalar quantization is adopted to quantize the LPC parameters. However, when vector quantization is considered, the speech coding rate can be further reduced without obvious degradation of speech quality. In our coding scheme, instead of direct encoding of LPC parameters, LSF coefficients are encoded which are obtained from a transformation of LPC. In addition, in order to utilize the correlation among LPC parameters, three frames of speech are processed together so that some kind of prediction could be utilized. A weighted version of MSE is selected as the distortion measure in building the VQ codebooks and encoding. Also, due to their different characteristics, separate codebooks are built and used for voiced and unvoiced speech to better match their statistical properties. The result of our experiments shows that our 800 bps speech coding scheme can achieve similar quality to that of the LPC 2.4 kbps speech coder.

1aSC39. A study on the speech analysis and recognition for Korean children. Eunjung Park, Jieun Kam, Inchan Paek, Y. Kwon (Hanyang Univ., Ansan, Kyonggi-Do 425-791, Korea), K. S. Lee, and Sung-il Yang (Hanyang Univ., Ansan, Kyonggi-Do 425-791, Korea)

This paper describes a speaker-independent Korean ten-digit recognizer for males and females from 10 to 13 years old, who have a serious variation of speech characteristics. DHMMs with time normalized ten-dimensional LPC, CEPSTRUM, LSF coefficients, and ten-band energies in frequency domain as the feature vector are used. The pitch and formant informations are extracted to analyze speech characteristics. Through experiments, DHMM recognizer shows the performance of recognition rates, 67.1% (LPC vectors), 80.8% (CEPSTRUM vectors), 66.7% (band energies in frequency domain), and 83.8% (LSF vectors).

MONDAY MORNING, 22 JUNE 1998 WEST BALLROOM A (S), 8:00 TO 10:15 A.M.

Session 1aSP

Signal Processing in Acoustics and Speech Communication: Modeling and System Identification

Karl M. Reichard, Chair
Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804-0030

Contributed Papers

8:00

1aSP1. Signal models for broadband model-based array processors. Edmund J. Sullivan (Code 82101, Naval Undersea Warfare Ctr., 1176 Howell St., Newport, RI 02841) and James V. Candy (Lawrence Livermore Natl. Lab., L-156, Livermore, CA 94551)

The recently developed theory for broadband model-based acoustic array processing is based on a Kalman-type estimation scheme. Here two sets of equations are required; the state equations and the measurement equations. In the narrow-band case, the measurement equation is based on either a plane wave or curved wavefront monochromatic signal model. In the broadband case the signal model, i.e., the measurement equation, is represented as an autoregressive moving average (ARMA) process. In this approach the parameters of interest, which are the bearing and range, are combined with the ARMA parameters in a single state vector. This provides a scheme wherein these parameters are adaptively estimated in a recursive manner. However, the order of the ARMA process depends upon the spectral characteristics of the signal. In this study, various realistic broadband signal spectra are examined in terms of the required ARMA representations. Examples will be shown for several cases of representative surface ship radiation spectra.

8:15

1aSP2. Simulation of the effects of thermal turbulence on beamforming and time reversal techniques using Gaussian beam summation and Fourier modes superposition techniques. Christian Lhuillier, David Fiorina, Jean-Luc Berton (Commissariat a l'Energie Atomique, Cadarache, DER/SSAE/LSMR, F-13108 Saint-Paul-lez-Durance cedex, France, ssaelsmr@macadam.cea.fr), and Daniel Juve (Laboratoire de Mecanique des Fluides et d'Acoustique, UMR CNRS 5509, F-69131 Ecully, France)

In-service inspection of Fast Breeder Reactors will require submerged ultrasonic imaging devices. For this purpose, the CEA (French atomic commission) is developing an orthogonal imaging system, based on beamforming applied to two linear arrays. In order to evaluate the performance of this technique for realistic physical conditions, one must take into account the possible effects of weak thermal turbulence on the propagation and focusing of ultrasound. To achieve this study, a numerical model has been implemented. It is based on the Fourier modes superposition method to generate independent realizations of turbulent fields and on the Gaussian beam summation method to calculate the associated acoustic pressure fields induced by point sources. First, the model has been used to show the aberrations that are created in the focusing of a multielement transducer.

Second, the model has been used to verify the efficiency of the time reversal technique for the correction of these aberrations, in the case of slow temporal evolutions of the turbulence.

8:30

1aSP3. Advanced system identification techniques for acoustics enhancement. Yuchang Cao and Mark Poletti (Industrial Res. Ltd., Gracefield Rd., P. O. Box 31-310, Lower Hutt, New Zealand)

This paper proposes a novel scheme of adaptive system identification for acoustics enhancement applications. The proposed scheme incorporates a variable exponentially weighted-step-size LMS algorithm, called the VES algorithm, the time-average, and the group measurement methods. A key contribution of this study is the development of the VES algorithm based on the statistics of the overall varying characteristics of the estimate of room impulse responses. The modeling error of the VES algorithm can be reduced to -20 dB after only $2N$ iterations (N is the length of adaptive filter), about 4 times faster than the normalized LMS algorithm under the test conditions. The increase of the computation load of the VES algorithm is small. When the algorithm is employed in a multichannel system identification, the increase of the computation load is only about 10%. Since the identification scheme has to operate under strong interference in most cases, the time average and group measurement approaches have also been investigated and are suggested to be employed in the scheme to further improve the identification accuracy. The proposed techniques can be used to eliminate undesired feedback, equalize the frequency response, and allow the design of more advanced multichannel assisted reverberation systems, which can provide larger loudness and reverberant gains than conventional systems.

8:45

1aSP4. Convergence characteristics of frequency-domain LMS adaptive filters. Karl M. Reichard and David C. Swanson (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804, kmr5@psu.edu)

The frequency-domain implementation of the adaptive LMS algorithm offers several advantages over the time-domain LMS algorithm. The primary advantage is that the orthogonality of the individual bins in the signal spectra permit the implementation of a single complex LMS filter for each FFT bin independently, each with a separate convergence parameter. The result is a uniform rate of convergence for all FFT bins as opposed to the time-domain algorithm which converges only as fast as the dominant eigenvalue of the input data. An additional parameter which affects frequency-domain LMS convergence is the overlap between successive FFT buffers. It is shown that the maximum advisable overlap between FFT buffers is 50%. Greater overlap between FFT buffers can lead to oscillations (and even instability) in the adaptive algorithm. Experimental results are presented which illustrate the convergence characteristics of a frequency-domain adaptive active noise control system for several different types of reference inputs.

9:00

1aSP5. A simple genetic algorithm for active noise control. Antonio Minguez and Manuel Recuero (Inst. de Investigacion del Automovil (INSIA), Univ. Politecnica de Madrid, Ctra. Valencia km. 7, 28031 Madrid, Spain)

Genetic algorithms in active noise control (ANC) eliminates the need to know the cancellation path, necessary for a conventional gradient descent-type algorithm (such as filtered-x and filtered-U) to remain stable. However, genetic algorithms show some disadvantages: slow convergence, high residual error noise, and high computational complexity. This paper presents a simple genetic algorithm that improves these quality factors, simplifying the basic stages of the genetic algorithm: population size, selection, reproduction, and mutation. To reduce the algorithm computation the population size is low, 6 chromosomes (filter coefficients) maximum, and only the best chromosome is selected so the reproduction stage disappears. Next generation of chromosomes is generated mutating the selected chromosome with a variable random number which amplitude is directly proportional to the mean-squared error in the current generation. This variable mutation is the most important parameter because it determines the stability, the convergence, and the final residual level. Good results have been obtained for narrow-band signals in single (SISO) and multiple (MIMO) ANC configurations and with FIR or IIR filters. These results shows fast convergence, only four or five generations are necessary with very reduced populations (6 chromosomes in SISO systems and 2 to 3 chromosomes in MIMO systems) and an attenuation better than 20 dB.

9:15

1aSP6. Reducing sibilants in recorded speech using psychoacoustic models. Markus Sapp, Martin Wolters, and Joerg Becker (Institut für Elektrische Nachrichtentechnik, RWTH Aachen, 52056 Aachen, Germany)

Sibilants are a known problem in speech recording. Even though they can often be decreased by a different placement of the microphones, there still is the necessity for methods that reduce these artifacts. Therefore new investigations on sibilants in German, English, Spanish, and French have been made to describe their properties in time and frequency domain. To find an adaptive algorithm which considers these properties a reliable method is needed for detection and classification of the different kinds of sibilants. It is shown that the psychoacoustic unit ''sharpness'' is well correlated with the appearance of these sounds and that specific loudness, an intermediate step in calculating sharpness, can be utilized to get information about the spectral properties of each sibilant. An algorithm is presented which employs sharpness to detect sibilants and reduces them using variable filters.

9:30

1aSP7. Machine recognition of Hindi consonants and distinctive features using vector quantization. Pradip K. Das (Human Resource Development Group, Rm. No. 505, CSIR Complex, Dr. K. S. Krishnan Marg, New Delhi-110 012, India) and S. S. Agrawal (Central Electrons Eng. Res. Inst., CSIR Complex, Dr. K. S. Krishnan Marg, New Delhi-110 012, India)

This paper describes the results of experiments conducted to test the efficiency of the VQ algorithms to recognize and distinguish minute differences among consonant features of Hindi. The database was created by combining all the frequently occurring 29 Hindi consonants with vowel /a/ to form CV-type syllables and recorded by 20 male Hindi speakers. The acoustic features in the form of cepstral vectors were derived for each syllable file in the training set. For each syllable, the set of 20 codebooks for 20 speakers were first clubbed to form the universe for that syllable. The LBG algorithm was used to generate the codebook of size 64. Using the codebooks as reference templates the recognition rate was found to be 73.28% out of the total of 580 (29*20) test samples. Further analysis was carried out to see how different consonants were recognized when they were clubbed into different groups according to the manner and place of their articulation. The average recognition rate for the consonants corresponding to the same place of articulation and the same manner of articulation was found to be 83.38% and 76.25%, respectively. Results of detailed distinctive feature analysis carried out to find the differences in the rate of recognition are also discussed.

9:45

1aSP8. Synthesis of Hindi consonants with natural sounding quality using the PC-based Klatt synthesizer. Rajesh Verma and Shyam S. Agrawal (Central Electrons Eng. Res. Inst. Ctr., CSIR Complex, NPL Campus, Hill Side Rd., New Delhi 110 012, India)

This paper describes experiments conducted to synthesize Hindi consonants with natural sounding quality using the PC-based Klatt synthesizer. All the most frequently used 29 Hindi consonants have been synthesized in the form of CV and VC syllables. For this purpose, a detailed spectral analysis was carried out to determine minute acoustic differences among the various categories of consonants such as unaspirated/aspirated, voiced/unvoiced, dental/retroflex, liquid/trill, etc.. The synthesis was done using a modified version of the Klatt Synthesizer simulated on a PC. It uses about 60 parameters for the source and vocal tract variations in 5-ms duration. Variations in important parameters were made until a satisfactory quality of consonant was perceived and a clear distinction occurred with its cognate sound. For example, the distinction among dental plosives (t, th, d, dh) and retroflex plosives (t*, th*, d*, dh*) have been achieved by changing formant frequencies at the burst position, intensity of burst, and formant transitions from burst to the target vowel. Similarly the sound /r/ in Hindi takes various forms such as trill or fricative depending upon the context. This is a very special characteristic of /r/ in Hindi. In this same manner the special features of all the 29 consonants are discussed with references to their synthesis in five different vowel contexts.

10:00

1aSP9. A rejection method for the Korean isolated word recognition system. Dong-Hwa Kim (Dept. of Information and Commun., Miryang Natl. Univ., Nei 2 Dong, Miryang, Gyung Nam, 627-130, Korea, dhkim@arang.miryang.ac.kr), Hyung-Soon Kim, and Young-Ho Kim (Pusan Natl. Univ., Jang Jun Dong, Gyem Jung Gu, Pusan, 609-735, Korea)

The small vocabulary isolated word recognition systems as well as large vocabulary continuous speech recognition systems can be applicable in many areas. For the isolated word recognition systems to be deployable in actual applications, the ability to reject the out-of-vocabulary is required. This paper presents a rejection method which uses the clustered phoneme modeling combined with postprocessing by likelihood ratio scoring. Our baseline speech recognition was based on the whole-word continuous HMM. Six clustered phoneme models were generated using the statistical method, monophone clustering algorithm, from the 45 context independent Korean phoneme models, which were trained using the phonetically balanced Korean speech database. The performance of this method was assessed in terms of the out-of-vocabulary rejection rate and the accuracy on the pre-defined-vocabulary. The performance test for speaker independent isolated words recognition task on the 22 section names shows that this method is superior to the conventional postprocessing method, which is performing rejection according to the likelihood difference between the first and second candidates. Furthermore, these clusters phoneme models do not require retraining for the other isolated word recognition systems with different vocabulary sets.

MONDAY MORNING, 22 JUNE 1998 GRAND BALLROOM B (S), 8:00 TO 10:48 A.M

Session 1aUW

Underwater Acoustics: Underwater Acoustic Communications

Marshall H. Orr, Chair
Naval Research Laboratory, Code 7120, 4555 Overlook Avenue, SW, Washington, DC 20375-5350

Chair's Introduction—8:00

Invited Paper

8:05

1aUW1. A review of progress in acoustic telemetry and the challenges of the underwater channel. Arthur B. Baggeroer (MIT, Rm. 5-204, Cambridge, MA 02139)

High data rate acoustic telemetry in the ocean is limited by both bandwidth and reverberation. Attenuation limits the available bandwidth and most acoustic environments are dominated by reverberation which is spread in both delay and Doppler. The availability of low power VLSI electronics has had an enormous impact upon acoustic telemetry because the algorithms of digital communication systems could be implemented in oceanographic sensors which have severe constraints upon available energy and space. The first demonstration was the DATS (Digital Acoustic Telemetry System) for reverberant acoustic channels. It used MFSK with an (8,4) Hamming code, frequency hopping, and explicit synchronization of time and Doppler. Since then, (i) coherent signaling, (ii) source and channel coding, (iii) adaptive equalizers, and (iv) frequency and spatial diversity and networked configurations all have been used. Bit rates of 5–20 kbs at short ranges and 0.5–1 kbs at long ranges have been demonstrated. Higher rates can probably be obtained with forward error control methods; however, there the relatively slow speed of sound leads to some unique latency issues. Nevertheless, the availability of two-way links leads to acoustic telemetry networks which can be addressed just as any other node on a net. Some other interesting issues in acoustic telemetry concern increasing the data rate within the same bandwidth, low probability of intercept covert communications, methods for high Doppler situations, and network topologies which are robust to node outages. This presentation will provide a brief overview of acoustic telemetry and the evolving research areas, and present some of the challenges of the underwater channel.

8:35

1aUW2. Environmental impact on phase coherent underwater acoustic communications: A physics based approach. T. C. Yang and A. Al-Kurd (Naval Res. Lab., Washington, DC 20375)

Phase coherent acoustic communication systems, which employ joint adaptive channel equalization and phase-locking synchronization, have been shown to yield a much higher symbol rate (e.g., 2 kb/s at 3.5 kHz) by successfully tracking and updating time varying phase fluctuations in signals dominated by multipath components. This communication technique holds high promise for many applications. Its performance, however, can strongly influence (and could be limited by) the signal and noise fluctuation characteristics of the propagation channel. A statistical underwater acoustic communication model which incorporates random media propagation physics is needed to predict the performance of phase coherent acoustic communication algorithms. In this paper, the ocean acoustic propagation physics and signal processing issues which must be addressed to develop a physics-based performance prediction model are outlined. Acoustic communication simulation results for a dynamic underwater acoustic channel with time varying internal waves, fine and microfine structures, and surface waves are presented. Initial performance assessments (predictions) are given based on previously measured signal fluctuation statistics as characterized by the signal amplitude fading and phase spectra. [Work supported by the Office of Naval Research.]

8:47

1aUW3. Observations of coherence frequency and amplitude fluctuations in a shallow-water acoustic communications channel. Richard Shockley, Joseph Rice, Vincent McDonald (Space and Naval Warfare Systems Ctr., San Diego, Code D881, San Diego, CA 92109), Dale Green (Torrey Sci. Corp.), John Proakis (Northeastern Univ.), and John Newton (Polar Assoc., Inc.)

This work examines the statistics of a fading, dispersive shallow-water channel, specifically the distribution of amplitudes and the correlation between fluctuations in amplitude at different frequencies in the 8- to 12-kHz band. Signals considered include frequency-hopped, 50-ms duration, CW pulses and 500-ms linear FM chirps sent over a 15-m deep, 1-km range channel. Physical origins and implications for communications receivers of the observed non-Gaussian or non-Rayleigh behavior are briefly discussed. [Work supported by the SSC San Diego Independent Research Program, sponsored by the Office of Naval Research.]

8:59

1aUW4. Experimental performance of concatenated coding in shallow-water channels. Daniel Kilfoyle (Woods Hole Oceanograph. Inst., Woods Hole, MA 02543, dkilfoyle@whoi.edu)

In a series of phase coherent underwater acoustic communication experiments near the New England continental shelfbreak, the coding gain realized by a concatenated outer block code and trellis coded modulation (TCM) inner code was measured and compared to ideal asymptotic limits. Viterbi decoding of the inner code was fully integrated into the adaptive decision feedback equalizer structure. Coding gain is typically not realized until after decision delays of several code constraint lengths. Delayed decisions for the tap update algorithm, however, degrades equalization performance. Two specific techniques were employed to mitigate this conflict, including linear extrapolation of the carrier phase correction and a unique evolution of the tap weights for each state of the Viterbi trellis. Coding gains were realized for both eight-level phase shift keyed (8 coding states) and 16 quadrature amplitude shift keyed (128 coding states) waveforms. Bit error rate as a function of output signal-to-noise ratio was computed to demonstrate practical, as opposed to asymptotic, coding gains. The extent to which coding increased the system performance envelope will then be discussed.

9:11

1aUW5. Comparative performance of nonbinary Viterbi and sequential decoding in adverse underwater acoustic channels. Dale Green, Richard Justice (Torrey Sci. Corp.), and Joseph Rice (Space and Naval Warfare Systems Ctr., San Diego, CA 92109)

The application of higher-order alphabets in convolutional forward error correction coding is discussed, specifically the use of Viterbi decoding to obtain maximum likelihood performance with short-constraint-length codes, and the use of Fano sequential decoding with long-constraint-length codes. It is demonstrated that both techniques perform very well even near the respective computational cutoff (Rcomp) region. An 8-ary, constraint-length-3 coder combined with Viterbi decoding provides probability of bit error in a Rayleigh faded channel on the order of 10^{-4} at Eb/No=11 dB. With an 8-ary, constraint-length-10 coder combined with Fano sequential decoding operating also at Eb/No=11 dB, the probability of undetected error may approach 10^{-12}. The probability of erasure is less than 10^{-4} with a reasonable amount of computer resources. Viterbi decoding with short-constraint-length codes produces occasional bit errors which may be corrected with the concatenation of a simple block code applied to the input message. Performance comparisons are drawn from a side-by-side analysis of the two using data obtained during telesonar experiments in San Diego Bay. [Work supported by the ONR 321 Sensors, Sources and Systems Program and SSC San Diego Contract No. N66001-96-C-6003.]

9:23

1aUW6. Signal fluctuation statistics and performance of phase coherent underwater acoustic communications: An experimental evaluation. A. Al-Kurd and T. C. Yang (Naval Res. Lab., Washington, DC 20375)

Underwater acoustic communications data using phase modulated signals were collected during the Littoral Warfare Advanced Demonstration 98-1 experiments. The signals were projected from both a drifting modem and from a towed frequency source. Data of several different baud rate were transmitted and used to (1) study the temporal variation of the acoustic impulse response of the ocean and (2) evaluate the performance of the phase coherent communication algorithm. The amplitude and phase variations are analyzed on a fast scale (within the transmission packet of 5 s) and a low scale (minutes to hours). The performance of the communication algorithm is correlated with the phase rate and amplitude fluctuation statistics, the signal-to-noise ratio, etc. The data are compared with simulation results and used to guide performance predictions. [Work supported by the Office of Naval Research.]

9:35

1aUW7. Adaptive array processing for acoustic communications through time-varying ocean channels. James C. Preisig (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanograph. Inst., Woods Hole, MA 02543)

In many underwater environments, the acoustic communications channel is reasonably modeled using a micro/macro multipath model. While this model was developed by Flatte [S. M. Flatte, *Sound Transmission Through a Fluctuating Ocean* (Cambridge U.P., Cambridge, 1979)] and others to describe the effect which random channel fluctuations caused by linear internal waves have on propagating narrow-band acoustic signals, it provides a useful paradigm for modeling the effects of a wider range of environmental fluctuations on wideband communications signals. With this paradigm as a starting point, an analytic two-level model for the effect of random channel fluctuations is presented and compared with experimental data. Using the two-level model, a two-stage adaptive array processing algorithm is developed for processing acoustic communications signals. The algorithm exploits the space/time correlation structure of the received signals which is implied by the channel fluctuation model. The

performance of this algorithm in terms of both bit error rate and computational efficiency is analyzed and compared with that of other multichannel communications algorithms. [Work supported by the Office of Naval Research under Grant No. N00014-95-1-1316.]

9:47

1aUW8. Phase-encoded frequency-hopped signaling for underwater acoustic communications. Geoffrey Edelson (Sanders, A Lockheed Martin Co., Adv. Systems Directorate, P.O. Box 868, Nashua, NH 03061-0868) and Bradford Gillespie (Univ. of Washington, Seattle, WA 98195-2500)

Phase-coherent digital communications systems require complicated receivers to overcome the intersymbol interference caused by the time-varying multipath propagation of shallow-water acoustic channels. However, many of these channels feature ''sparse'' impulse responses in that the received signal arrives in distinct clusters separated by time periods in which little signal energy arrives at the receiver. If the phase-coherent message is divided into multiple ''minipackets'' such that the minipacket duration is shorter than the time span between the distinct multipath clusters, then the receiver configuration can be simplified considerably. The short packets may be frequency hopped so that a given frequency is vacated until subsequent multipath arrivals have died out. Each minipacket is decoded using a maximum-likelihood decoder. If each chip is not phase-encoded, the system reduces to a noncoherent frequency-hopped modulation technique. Experimental results are included to demonstrate the potential performance of this approach.

9:59

1aUW9. Doppler spread in an undersea acoustic transmission channel. Joseph Rice, Paul Baxley, Homer Bucker, Richard Shockley (Space and Naval Warfare Systems Ctr., San Diego, Code D881, San Diego, CA 92109), Dale Green (Torrey Sci. Corp.), John Proakis (Northeastern Univ.), and John Newton (Polar Assoc., Inc.)

This paper examines spectral dispersion of narrowband sound energy transmitted between two stationary, horizontally displaced, undersea, electroacoustic transducers. Any frequency smearing arises kinetically and is imparted by the aggregate motions of scatterers influencing the forward-propagated, physical channel. From this time-variant cause, hypotheses are drawn about the statistics of the spectrally dispersive effect. The hypotheses are tested by analyzing representative acoustic data in the 8–12-kHz band obtained from ocean experiments. To further study the phenomena, a numerical ray model is constructed with 3-D Gaussian beams, evolving sea–surface boundary, and dynamic volume inhomogeneities. The model represents the sea surface as a nonstationary mosaic of polygonal tiles, each reflecting energy with an appropriate Doppler shift. Evolution of the faceted sea surface is governed by assumed oceanographic processes. Spectral density is computed by accumulating the energy contributions at a specified receiver location. In conclusion, the implications of spectral dispersion on undersea acoustic signaling are summarized. [Work supported by the SSC San Diego Independent Research Program, sponsored by the Office of Naval Research.]

10:11

1aUW10. Doppler tolerant link. Anthony D. Matthews (Coastal Systems Station Dahlgren Div., Naval Surface Warfare Ctr., Code R21, 6703 W. Hwy. 98, Panama City, FL 32407, matthews_tony@ccmail.ncsc.navy.mil)

A specialized probe signal is used for environmental adaptation, frequency scaling, and alertment in an underwater acoustic digital communications link. The link is designed to permit one-way message traffic to be sent from a static node to a moving node. The moving node creates an unknown range rate within known range rate limits. Link design goals include power conservation, a bit error rate of 1/10 000, and tolerance of a 25 dBs noise field. Design parameters, simulation experiment results, and field experiment results (if available) are presented. [Sponsored by ONR.]

10:23–10:48

Panel Discussion

MONDAY MORNING, 22 JUNE 1998 GRAND BALLROOM A&B (S), 11:00 A.M. TO 12:00 NOON

Session 1aPLb

Plenary Lecture

Ralph R. Goodman, Chair

Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804-0030

Chair's Introduction—11:00

Invited Paper

11:05

1aPLb1. Acoustical studies of the upper ocean boundary layer. David Farmer (Inst. of Ocean Sci., 9860 West Saanich Rd., Sidney, BC V8L 4B2, Canada)

Ocean surface processes control the vertical transfer of heat, mass, and momentum, affecting water properties and influencing climate. At higher sea states the upper ocean can be one of the most energetic environments on the surface of the planet, presenting special observational challenges to the oceanographer. Acoustical methods provide new approaches to upper ocean measurement and have already deepened our understanding of wave breaking, Langmuir circulation, and bubble distributions. High-frequency sonar images of bubble clouds delineate the patterns of near surface circulation while acoustic Doppler reveals the corresponding flow field; high-resolution coherent Doppler can resolve the microscale turbulent velocity fluctuations. Interpretation of ambient noise allows us

to sense wave breaking and presents an intriguing inverse problem requiring separation of sound generation and propagation effects. *In situ* acoustical sensors can now provide reliable determinations of bubble size distributions. Our experience with many of these techniques is limited, but significant advances can be expected through the combined application of diverse approaches to the study of this climatically sensitive boundary.

MONDAY AFTERNOON, 22 JUNE 1998 GRAND BALLROOM III (W), 1:00 TO 2:00 P.M.

Session 1pPLa

Plenary Lecture

Leo L. Beranek, Chair
975 Memorial Drive, Suite 804, Cambridge, Massachusetts 02138-5755

Chair's Introduction—1:00

Invited Paper

1:05

1pPLa1. Sound transmission through suspended ceilings. F. P. Mechel (Landhaustrasse 12, D-71120 Grafenau, Germany)

Suspended ceilings passing over a partition wall between adjacent rooms are a path of sound transmission from room to room for which, until recently, no analytical and satisfactory theory existed, although suspended ceilings are a common system of construction. The principal features of an analysis are outlined for the sound fields in the adjacent rooms, in the plates and absorber layers of the suspended ceiling and in the common plenum above the ceiling. The theory is based on a modal analysis. It not only shows the mechanisms of the sound transmission, which in contrast to common ideas, are quite different on the emission and on the receiving sides, but it also gives numerical results which agree (within the precision of measurements) with experimental data. The influence of geometrical parameters (e.g., plenum height, room width) and of material parameters (e.g., absorber layer thickness) can be demonstrated quantitatively.

MONDAY AFTERNOON, 22 JUNE 1998 CASCADE BALLROOM II (W), 2:15 TO 3:40 P.M.

Session 1pAAa

Opera House Acoustics Workshop—CIARM, MCHA and Architectural Acoustics: Opera House Acoustics: State of the Art

Yoichi Ando, Cochair
Graduate School of Science and Technology, Kobe University, Rokkodai, Nada, Kobe, 657 Japan

Robeto Pompoli, Cochair
Via Comacchio 199, 44100 Ferrara, Italy

Chair's Introduction—2:15

Invited Papers

2:20

1pAAa1. The development of room acoustical demands on operas within the last 50 years. Gerhard Müller and Helmut A. Müller (Müller-BBM GmbH, Robert-Koch-Str. 11, D-82152 Planegg, Germany)

After World War II, many destroyed operas in Europe required reconstruction. Since the period of the royal operas was past, architects tried to develop new shapes based upon the new social philosophy. For the acoustical planning, this has led to the following goals: (1) equal acoustical quality in all areas of the auditoria, (2) high speech intelligibility, and (3) uniform sound transmission from the pit and the stage to the audience. The deducted measures were the following: (1) application of an arenalike seating arrangement of the public to expose everybody to a strong direct sound, (2) low reverberation, and (3) reflection of sound energy to compensate the increasing distance of the listeners from the sound sources. These principles have led to the realization of slightly fan-shaped

rooms with a relatively small volume or additional sound-absorbing claddings. Surprisingly, these opera halls have been quoted worse than the old ones, by the musicians and the audience. The reason for the lower reputation of the operas, designed applying the rules of the 1950s, will be explained with the help of two examples in which improvements have been realized afterwards. The measured relevant room acoustical data will be presented and the differences discussed.

2:40

1pAAa2. Alternate venues for opera performances. Wagner started it . . . should we retreat now? Christopher Jaffe (Jaffe Holden Scarbrough Acoust., Inc., 114A Washington St., Norwalk, CT 06854, dcoppola@jhsacoustics.com)

Although there has been a resurgence of interest in opera over the last 15 years in America the current outlook for continued financial support for this form of musical theater as well as classic concerts is bleak. At one time opera was enjoyed by the unwashed as well as sophisticates and a number of today's creative opera directors are trying to appeal to a broader segment of the market. This paper will discuss the break with tradition that Richard Wagner and Adolph Appia succeeded in achieving at Bayreuth in terms of architectural acoustic design and how these changes were integrated with new approaches to scenic design and stage lighting. In addition, examples of contemporary opera companies that are creating a more intimate and emotionally powerful experience for their audiences by breaking out of the picture-frame productions of proscenium theater are given. Among these companies are the Colorado Opera and the Opera Theatre of Saint Louis.

3:00

1pAAa3. Orchestra pit acoustics: From Bayreuth to Broadway. Mark Holden (Jaffe Holden Scarbrough Acoust., Inc., 114A Washington St., Norwalk, CT 06854, mholden@jhsacoustics.com)

The design of orchestra pits for opera, ballet, and multiuse halls is examined from an historical perspective. The earliest orchestra pit designs for early European opera houses were open or ''bathtub'' pits holding few musicians on a single level. Pits remained similar until 1872 when Wagner's Bayreuth pit changed the musician/vocalist orientation by extending the pit below the stage edge in stepped tiers. Modern pit designs have varied from large ''bathtub'' pits (the Metropolitan Opera, San Francisco) to ''modified Bayreuth'' pits (Houston, Seattle, etc.) with varying degrees of success. Issues concerning musicians' comfort, loudness levels, and hearing loss are discussed along with ways to improve acoustics in the overhung area. Finishes and materials for walls, ceilings, and floors are discussed as well as adjustable acoustic panels, drapes, and walls that can be used to tune the pit for varying programs. Finally, solutions to the dilemma of Bayreuth versus Bathtub pits will be discussed.

3:20

1pAAa4. Why do traditional opera houses work so well for opera? Nicholas Edwards and David Kahn (Acoustic Dimensions, 24 Styvechale Ave., Coventry CV5 6DX, UK)

Most computer models of room acoustics assume geometric acoustics (as if sound behaves like light). This has assisted our understanding of how room shape (fan shape versus rectangular, for example) can affect the acoustics of concert halls. When one applies geometric acoustics modeling to the traditional opera house, typically all one finds is a single ceiling reflection. The geometric approach holds little promise for understanding the magic of opera house acoustics. One of the acoustical attributes ignored by the geometric model is edge diffraction of sound. Our research has shown edge diffraction to be essential in predicting the superior acoustics of the traditional opera house. We have developed an alternative to the geometric acoustics model based on edge diffraction. In a two-dimensional representation, the locus of points with equal delay time is an ellipse; in three dimensions, the locus is the volume of revolution of an ellipse. By studying where this ellipsoidal volume intersects with the balcony fronts, we can locate where, at any given time, the edge-diffracted sound is coming from. This model is applied to LaScala, and a ''movie'' is developed to show how the balcony front edges contribute to the sound that is heard.

Session 1pAAb

Opera House Acoustics Workshop—CIARM, MCHA, Architectural Acoustics, and Musical Acoustics: Sound Field for Singer and Orchestra, Part I

Yoichi Ando, Cochair
Graduate School of Science and Technology, Kobe University, Rokkodai, Nada, Kobe, 657 Japan

Roberto Pompoli, Cochair
Via Comacchio 199, 44100 Ferrara, Italy

Chair's Introduction—3:50

Invited Papers

3:55

1pAAb1. Measurement of objective criteria for performers in the ''Teatro di San Carlo'' opera house. Carmine Ianniello, Gino Iannace, Luigi Maffei, and Rosario Romano (DETEC–Univ. of Naples, Federico II, Piazzale Tecchio 80, 80125 Naples, Italy, maffei@unina.it)

In the last two decades some attention has been paid to the acoustics for performers in concert halls. The findings of these studies have shed some light on the structure of the sound field preferred for the ensemble. With respect to a concert hall, the performing condition is quite different in an opera house. Usually, the problems are related to: (a) the balance between singer(s) and the orchestra in the pit; (b) the ease of singing; (c) the ensemble of singers; (d) the ensemble of instrumentalists. To the knowledge of the authors no objective criterion has been suggested yet for (a). Some criteria are reported for the ensemble of instrumentalists (d). Useful information for (b) and (c) is reported in [A. H. Marshall and J. Meyer, Acustica **58**, 130–140 (1985)]. Although well aware of the lack of accepted objective criteria, which should be well correlated with subjective criteria for the above-mentioned topics, these authors report the results of measurements carried out in the ''Teatro di San Carlo'' in the aim of quantifying objectively some of the items (a)–(d). In particular, a criterion is suggested that is related to the comfort of singers on the basis of the study of Marshall and Meyer.

4:15

1pAAb2. Sound fields in orchestra pits. Juergen Meyer (formerly, Physikalisch-Technische Bundesanstalt, D-38116 Braunschweig, Germany, juergen.meyer@ptb.de)

The main problem for musicians sitting in opera house orchestra pits is the general loudness of the orchestral sound. The sound field occurring at the musicians' ears is not only affected by direct sound, early reflections, and reverberation, but—particularly for seats under the stage—by standing waves. Whereas direct sound and diffuse reflections appear to come from comparatively large distances, the sound due to standing waves is localized just around the head. Therefore, it is extremely disturbing for the players, impairing the tonal control of their own instruments as well as the ease of ensemble. Disturbing direct sound may be reduced (5–10 dB) by special screens directly behind the musician's head. Standing waves between the floor and cover of the pit are critical if the musicians' ears are in the neighborhood of the antinodes; this does not occur with the lowest mode, but for modes having two or three half-wavelengths, assuming the ears at a height of 1.15 m above the floor. Therefore, a distance between floor and cover of 2.0–2.75 m (or below 1.80 m) may be recommended, as in these cases only one disturbing mode occurs which can be damped by well-tuned absorption.

4:35

1pAAb3. Listening in the pit. Christopher N. Blair (ARTEC Consultants, Inc., 114 W. 26th St., New York, NY 10001)

Temporal masking effects are a controlling factor in the ability of performers to hear each other well. Difficulties in ensemble communication and perceived weakness of the acoustic response of the auditorium can be exacerbated by high early energy levels often found in orchestra pits. These problems become particularly acute in orchestra seating positions under stage overhangs. Drawing on experiences as an opera conductor as well as an acoustic designer, the author: (1) provides an overview of the psychoacoustic challenges facing the opera house designer relative to orchestra satisfaction; (2) explores, using architectural examples, the reasons why some pits are highly regarded by orchestras and conductors while others, even those similarly configured, are not; and (3) proposes protoptypical design solutions to improve the listening environment both inside the pit and elsewhere.

4:55

1pAAb4. Small-scale modeling of stage to pit balance: A pilot study. John P. O'Keefe (Aercoustics Engineering Ltd., 50 Ronson Dr., Ste. 127, Toronto M9W 1B3, Canada)

Measurements in full-scale auditoria have demonstrated that the acoustical balance between stage and orchestra pit sound sources is influenced by the seat dip phenomenon. In general, low-frequency balance favors the orchestra pit for listeners on the orchestra level and is about even for listeners on the balcony. Full scale measurements suggest that the rake of the floor and the presence of local (early) reflectors increase balance to the benefit of the singers on stage. A simple 1:24 scale model has been built to test these and other phenomena in a controlled fashion. Measurements have also been performed in small-scale models of existing theaters. Early findings confirm the influence of seat dip on stage to pit balance.

5:15

1pAAb5. Compromises in orchestra pit design: A ten-year trench war in The Royal Theatre, Copenhagen. Anders Chr. Gade (Dept. of Acoust. Technol., Tech. Univ. of Denmark, Bldg. 352, DK-2800 Lyngby, Denmark, gade@dat.dtu.dk) and Bo Mortensen (Royal Acad. of Fine Arts, Copenhagen K, Denmark)

The ''old stage'' of The Royal Theatre in Copenhagen—a classical horseshoe theatre with an almost flat audience floor and four balcony levels—is the primary opera and ballet theatre in Denmark. In the early 1980s the orchestra pit was enlarged and its size made flexible. However, in the following years this new flexibility caused a formal battle between the orchestra, the singers, and management, because each group had different interests as to how the pit should be configured for each new production. In the 1990s, increased concern about the musicians' hearing and the increase in the popularity of opera finally opened the way for a new renovation of the pit, which solved most of the problems. The paper describes the many lessons learned during the ten-year period in which the authors were involved with the case as acoustic consultants. Questions of how pit configuration influences mutual hearing and exposure levels in the pit, balance between singers and orchestra as judged by the audience, communication between stage and pit, as well as the influence of pit floor construction on orchestra timbre will all be illustrated by results of numerous objective measurements, subjective surveys, and computer simulations.

Contributed Papers

5:35

1pAAb6. Study on experimental equations for on-stage acoustics design. Shun-ichi Nakamura (Tokyo Natl. Univ. of Fine Arts and Music, Tokyo, Japan, ezn0255@niftyserve.or.jp)

The relationship between the configuration of an orchestra shell and the acoustics inside have been studied based on the results obtained from the impulse response analysis using scale models of a series of 27 different conditions. Three points on a time profile curve were chosen as the acoustic factors in order to represent the onset process, the stationary state, and the decay process of the reflecting sound field, respectively. The averaged values adopted for discussion have been calculated using the data of three 1/1 octave bands around 1 kHz measured at three different positions in the main part of a stage floor. A remarkable difference in acoustics is found in the onset process between the cases with and without sound diffusers on the inner surface of a stage enclosure, and in the decay process as well. A faster rising in the onset process introduced by sound diffusers is followed by a faster decaying. They have a close connection. Finally, it is found to be possible to derive the experimental equations which relate these factors to the size and the shape of an orchestra shell, considering also the effect of reflections from the audience side of a hall.

5:50

1pAAb7. Measurements of orchestra acoustics. How reflecting surfaces influence ''Klangfarbe'' and the possibility of playing in time. Tor Halmrast (Statsbygg, pb. 8106 Dep. N-0032, Oslo, Norway)

When investigating room acoustics, and also stage acoustics, most concern is often based on ''how much energy arrives in the time interval between time x and,'' and not on how these reflections are distributed within that time interval and how that affects the Klangfarbe and rhythmic conditions. The stage acoustics of the Oslo Concert Hall was measured (MLS) with and without the orchestra on the stage, and with different settings of reflecting surfaces to achieve a better ''ensemble'' feeling (TOR measurements: Through Orchestra Response). Reflecting surfaces above, behind, and to the side of the orchestra were examined, with a questionnaire for each musician and a number of tonemeisters in the audience area. Similar measurements of the stage acoustics were done also in several European halls on the tour of the Oslo Philharmonic Orchestra/Mariss Jansons including Gasteig/Munich and Musikverein/Vienna. For further analysis of the ability of playing in time, measurements were done in the Hall for the Norwegian Radio Orchestra, with analysis of both orchestra tutti and multichannel recording situations. The paper analyzes the aspects of playing together in an orchestra, and especially how different placements of reflecting and diffusing elements affect the Klangfarbe and rhythmic conditions, both for the orchestra and the audience.

Session 1pAAc

Architectural Acoustics: Room Acoustic Measurement Techniques, Absorption and Diffusion (Poster Session)

Daniel C. Bruck, Chair
Bruck Richards and Chaudiere, 105 NE 56th, Seattle, Washington 98105

Contributed Papers

All posters will be on display from 2:30 p.m. to 6:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 2:30 p.m. to 4:15 p.m. and contributors of even-numbered papers will be at their posters from 4:15 p.m. to 6:00 p.m.

1pAAc1. Measurements and modeling of commonly used resilient channels in lightweight floor structures. Jonas Brunskog and Per Hammer (Eng. Acoust., LTH Lund Univ., P.O. Box 118, SE-221 00 Lund, Sweden)

A common part in lightweight floor structures is resilient channels between the joists and the ceiling. The purpose of the channels is to lower impact sound transmission in a broad frequency band. The channels are normally complicated to describe accurately from a structural acoustic point of view. The intention in this paper is to describe the channels by means of a simple structural acoustical model. The channels are divided into two types of elements. The first element is a beam. The second constitutes the mounting of the channels to the joist. The two elements are treated and analyzed both separately and connected to each other. The measurements are done by mainly measuring the accelerations on stiff bodies on the top and bottom of the test sample. The procedure is then repeated with a new set of stiff bodies. From these measured values (four different) one can determine the global two-port description. As a check there is also the reciprocity. As a comparison a method is used where the actual point force acting on the spring is measured and used. The model can be used as a base for analyzing the acoustical behavior of the resilient channels.

1pAAc2. Evaluation of the flanking transmission in a scale model of two adjoining rooms. Alessandro Cocchi, Giovanni Semprini, and Simone Secchi (DIENCA, Univ. of Bologna, via Risorgimentoz, 40136 Bologna, Italy)

Flanking transmission is the sound propagation through the lateral structures of two adjoining rooms. Evaluation of this transmission is possible with a statistical model of two rooms that considers mainly the reduction in vibration transmission between the partition and the lateral structures. With this model, which is the object of a project of international standard (Pr EN 12354), it is possible to evaluate the transmission through each path, considered independent of each other. The research carried out at the Acoustics Laboratory of the University of Bologna has validated the CEN model by measuring the vibration reduction index and the sound transmission in a scale model. The vibration transmission index has been measured both with transient and steady excitation of the structures, the first by means of a hammer and the second of an electromagnetic shaker. The sound reduction index of each path of transmission has been calculated measuring the mean pressure level in the source room and the mean vibration level of each structure in the receiving room. The paper presents the validation of the model with a comparison of the different kinds of measurement techniques.

1pAAc3. Classical analysis of multilayered walls transmission loss in comparison with the F.E.M. and S.E.A. approaches. Marco Fringuellino and Claudio Guglielmone (Acoust. Dept., I.E.N.G. Ferraris, Strada delle Cacce 91, 10135 Torino, Italy)

Prediction models of the acoustical behavior of multilayered soundproofing walls are of increasing practical interest in the field of building trade, particularly for designing applications. This paper reports a classical approach to the problem, whose aim is the theoretical calculation of the transmission loss of acoustical insulating multilayered walls. The model is based on the straight compilation of the ''progressive impedance'' encountered by the acoustical wave in its propagation through the wall. The pressures ratios at the ending surfaces of each layer are then computable as a function of the progressive impedances obtained before. The same is true for the total ratio of the incident and transmitted pressure. The characteristic impedance and the propagation constant of each layer are needed as input parameters. The results of the classical analysis are compared with a set of measurements performed on multilayered soundproofing walls. At present an effort is being made to develop a Finite Element Method and a Statistical Energy Analysis approach of the problem, to compare these theoretical results with the classical ones on the same set of walls. The final aim is to show the accuracy, the advantages, and the main limits of each of these three methods.

1pAAc4. Optimizing techniques for field testing of buildings. George Dodd (Acoust. Res. Ctr., Univ. of Auckland, Private Bag 92019, Auckland, New Zealand, g.dodd@auckland.ac.nz)

As national building codes and acoustic comfort standards become more common as tools for quantifying and promoting acoustical quality in dwellings there is a greater need for field testing of buildings. The intent of this project is the promotion of a greater use of field testing by providing more convenient measurement methods. Primarily the focus has been on identifying the ideal sound source. Explosive sources using charges produced for use in powder powered tools have emerged as satisfying the majority of our needs. Continuing study is aimed at optimizing such a source. In particular, it needs to produce adequate power for measuring at the low frequencies required for implementing the new spectrum adaptation terms for the traditional single-figure ratings of building insulation.

1pAAc5. Loss factor at the boundary of a single-leaf wall under a vibrational field of diffused bending waves. Soichiro Kuroki (Faculty of Eng., Kagoshima Univ., 1-21-40 Korimoto, Kagoshima, 890 Japan, kuroki@aae.kagoshima-u.ac.jp) and Masahito Yasuoka (Univ. of Tokyo, 1-3 Kagurazaka, Shinjuku, Tokyo, 162 Japan)

The internal loss factor (eta_{int}) is an important parameter in the case of calculating the transmission loss of a panel by using SEA (statistical energy analysis). Though energy loss at the edge is fairly large in practical walls, eta_{int} have been ordinarily treated including the edge loss. So it must be necessary to treat quantitatively the energy loss at the boundary of the wall. In this paper, assuming that a panel in excitation is a two-dimensional vibration field of diffused bending waves, a formula of the decay time of a panel is deduced. The subject for experimental study is an aluminum panel mounted in the opening between transmission suite. The panel is excited by a piezoelectric ceramic diaphragm. Vibrational decay time is measured by the impulse response method. The calculation method of the energy absorption coefficient and loss factor at perimeter is indicated by using measured values of decay time and sound radiation.

1pAAc6. Transmission loss measurement of multilayered partitions in time domain. Andre L. Cherman and Roberto A. Tenenbaum (Acoust. and Vib. Lab, Federal Univ. of Rio de Janeiro, Rio de Janeiro, Brazil)

Transmission loss through single-layered partitions depends on the superficial specific mass *m* of the material. A multilayered partition has, at each interface, an impedance jump that improves its transmission loss. Due to these characteristics, multilayered partitions can provide better sound insulation than a single-layered partition with the same thickness, being much lighter. In this work an experimental setup was built to study the transmission loss of a sample inside a duct. The partition is inserted between two sections of the duct and a pulse is generated at one side. The incident and the transmitted signals are captured by free-field microphones, and the time domain series, acquired with a two channel digital oscilloscope, provides the transmission loss. The impedance of the materials tested was found experimentally and compared with tables, when available. The experiment was performed in order to evaluate transmission loss in single-, double- and three-layered partitions. The experimental results were compared with numerical simulation using the software LAYER, developed in our laboratory, showing good agreement.

1pAAc7. A safe method of truncating an echogram calculated by a sound ray technique. Jean J. Embrechts (Natl. Fund Sci. Res., Dept. of Acoust., Univ. Liege, Sart Tilman, B28, B4000 Liege 1, Belgium, jjembrechts@ulg.ac.be)

A recent international test of room acoustics computer simulations (presented by Prof. Vorlander at the 1995 ICA Congress in Trondheim) has shown the deviations that can occur between several programs when they are calculating the same acoustical parameters. Beyond data evaluation, the method of deriving these parameters from a computed echogram may be of significant importance. Therefore, the different steps in the calculation of the echogram by a sound ray technique are analyzed. In particular is shown the influence of the time discretization step (dt) and of the late part of this diagram (the tail) on the parameters T30 and EDT. An algorithm is proposed to reduce their influence. It consists in the truncation of the echogram at a well-chosen instant, independent of dt. Indeed, the tail of the echogram is greatly affected by statistical errors, and it is replaced by a linear decay, with a precisely defined slope, which is also independent of dt. The acoustical parameters are finally derived from this modified echogram. This algorithm is compared with other crude truncation procedures. It is shown that it significantly enhances the stability of the calculated results. Its performance is limited by the definition of T30 and EDT themselves.

1pAAc8. Sound sources and simulation software in room acoustics. Catherine Semidor and Leonie Couthon (ERIAC, Ecole d'Architecture et de paysage de Bordeaux, Domaine de Raba, F 33405 Talence Cedex, France, catherine.semidor@bordeaux.archi.fr)

There is a lack of data in directivity, power, and phase, and most of the time software users have no data at their disposal, so they use an omnidirectional sound source, or they calibrate the simulation with the measurements they operated *in situ*. But the use of the room is correlated to the kind of source that emits inside of it, so it is fundamental to know the statistical characteristics of emission of real sound sources. In this article, a survey about the way sound sources are taken into account in simulation software is proposed. Then an analysis is presented of a three-dimensional recording of real sound sources (musical instruments), to access to data in power and directivity, which tries to respond to the questions raised. And finally, the results obtained with different simulation software are presented. In conclusion, the need for standardization of sound source data is discussed.

1pAAc9. Applying the equivalent source method to the sound radiation from musical instruments in rooms. U. Peter Svensson, Mayumi Nakano, Kimihiro Sakagami, and Masayuki Morimoto (Environ. Acoust. Lab., Faculty of Eng., Kobe Univ., Rokko, Nada, Kobe, 657 Japan)

Musical instruments often have rather complicated radiation characteristics which have been documented for the case of free-field radiation. This paper studies the effect of the source location relative to the room boundaries on the sound radiation from musical instruments in rooms. The musical instrument is modeled by the equivalent source method, or source substitution method, so that the real source is replaced by a set of simpler sources which approximately reproduces the vibrations on the instrument's surface. A single corner with rigid walls is used to model the effects of the room boundaries on the sound radiation. As an example, a tympani is studied and theoretical expressions for its surface vibrations are used as input data. The directivity, and total power output, are then studied for various source locations. Possibilities for simplified source directivity modeling for room acoustical simulation are discussed.

1pAAc10. A study of the reverberant field in rooms with various absorptions, including prediction of spatial variation. Linda P. Franzoni (Ctr. for Sound & Vib., Dept. of Mech. & Aerosp. Eng., N. C. State Univ., P.O. Box 7910, Raleigh, NC 27695-7910)

The sound fields of two rectangular rooms of different size and proportion have been studied using a computer simulation. Point sound sources and absorbing walls are mathematically modeled and the wave equation is solved, the direct-field is removed, and the result is an exact calculation of the reverberant field in the room at any spatial location. Both the overall volume-averaged sound-pressure level and the spatially varying cross-sectional averaged levels are computed and compared to theoretical results. A new theoretical result for the spatial variation is developed using an energy balance applied to cross-sectional slices of the room. Good agreement is shown between this result and the calculated values. The same theory gives overall levels by integrating cross-sectional averages, this is compared to classical results (Sabine, etc.) and to the computer generated data. An improvement can be made to the classical theory by reducing the source power to account for first reflections. Some experimental results will also be presented for comparison. [Work sponsored by NSF.]

1pAAc11. Studying room acoustics using a monopole-dipole microphone array. Durand R. Begault and Jonathan S. Abel (San Jose State Univ., MS 262-2, NASA Ames Res. Ctr., Moffett Field, CA 94035)

The use of a sound field microphone for examining the directional nature of a room impulse response was reported recently [R. Essert, J. Acoust. Soc. Am. **100**, 2837(A) (1996)]. By cross-correlating monopole and colocated dipole microphone signals aligned with left–right, up–down, and front–back axes, a sense of signal direction of arrival is revealed. The current study is concerned with the arrays ability to detect individual reflections and directions of arrival, as a function of the cross-correlation window duration. If this window is too long, weak reflections are overlooked; if too short, spurious detections result. Guidelines are presented for setting the window width according to perceptual criteria. Formulas are presented describing the accuracy with which direction of arrival can be estimated as a function of room specifics and measurement noise. The direction of arrival of early reflections is more accurately determined than that of later reflections which are quieter and more numerous. The transition from a fairly directional sound field at the beginning of the room impulse response to a unidirectional diffuse field is examined. Finally, it is shown that measurements from additional dipole orientations can significantly improve the ability to detect reflections and estimate their directions of arrival. [Work supported by NASA.]

1pAAc12. Audibility of edge diffraction in auralization of a stage house. Rendell R. Torres and Mendel Kleiner (Chalmers Rm. Acoust. Group, Dept. of Appl. Acoust., Chalmers Univ. of Tech., S-41296 Gothenburg, Sweden, rendell@ta.chalmers.se)

A recent model for edge diffraction based on physical acoustics was applied to the auralization of a source in a stage house. In this paper the results of two methods of comparison (computation of frequency responses and evaluation of binaural listening tests) are used to judge the audibility of diffraction effects associated with such geometries. The computed impulse responses are also compared via binaural listening tests with measured impulse responses from a physical scale model. Current results indicate that, from an auralization standpoint, first-order diffraction is significantly more influential than second-order diffraction, which can lie more than 45 dB below the direct or specularly reflected arrivals. Thus, higher-order (and computationally more prohibitive) edge diffraction calculations may possibly be simplified or even neglected in auralization programs. [Work supported by Axel & Margaret Axson Johnsons Stiftelse, Sweden.]

MONDAY AFTERNOON, 22 JUNE 1998 GRAND BALLROOM A (S), 1:30 TO 4:15 P.M.

Session 1pAO

Acoustical Oceanography and Underwater Acoustics: Long-Range Propagation for Measurement of Ocean Processes II

John A. Colosi, Chair
Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institute, MS 11, Woods Hole, Massachusetts 02543

Contributed Papers

1:30

1pAO1. Constraints on ocean internal wave spectra from long-range acoustic transmission data. John Viechnicki, Michael G. Brown (Appl. Marine Phys., Rosenstiel School of Marine and Atmospheric Sci., Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149), and the ATOC Group[a)]

Both full wave and ray-based predictions of acoustic wavefields at long range in deep ocean environments are compared to their measured counterparts, including the SLICE89 and AET data sets. The environment for these experiments was characterized by weak climatological range-dependent sound-speed structure, weak mesoscale structure, and moderately strong (typical) internal-wave-scale variability. The acoustic measurements are used to provide constraints on the statistics of the internal wave field, using the Garret–Munk spectrum as a starting point. Our attention is focused on the following parameters to which our simulations are sensitive: the strength parameter E, the mode number cutoff $j_{\max}$, and the horizontal wave-number cutoff values $k_{\min}$ and $k_{\max}$. [Work supported by ONR.] [a)]The ATOC Group: A. B. Baggeroer, T. G. Birdsall, C. Clark, J. A. Colosi, B. D. Cornuelle, D. Costa, B. D. Dushaw, M. Dzieciuch, A. M. G. Forbes, B. M. Howe, D. Menemenlis, J. A. Mercer, K. Metzger, W. H. Munk, R. C. Spindel, P. F. Worcester, and C. Wunsch.

1:45

1pAO2. Measurements of the travel-time bias caused by internal waves. Matthew Dzieciuch and the AST Group[a)] (Scripps Inst. of Oceanogr., Univ. of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093)

Theoretical work and computer simulations predict that the travel-time of acoustic signals is biased by the ocean internal wave field. The path integral theory for scattering predicts that the size of the travel time bias depends on the logarithm of the center frequency of the acoustic signal. A recent experiment, in which two acoustic signals with different center frequencies were transmitted simultaneously, makes it possible to measure the bias experimentally. Preliminary results suggest that the bias is about 50 ms at 3500-km range, which is roughly as large as expected. [a)]The AST group is T. G. Birdsall (U. Michigan), B. M. Howe (U. Washington), J. A. Mercer (U. Washington), K. Metzger (U. Michigan), R. C. Spindel (U. Washington), and P. F. Worcester (SIO).

2:00

1pAO3. Basin-scale internal-wave tomography using a large-aperture vertical array. John A. Colosi (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), Bruce D. Cornuelle, Matthew A. Dzieciuch, Walter H. Munk, and Peter F. Worcester (Scripps Inst. of Oceanogr., La Jolla, CA 92093)

Basin-scale acoustic transmissions which are received at a large-aperture vertical line array (VLA), like those deployed for the Acoustic Thermometry of Ocean Climate (ATOC) program, provide several acoustic fluctuation observables that can be used for internal-wave tomography. A single receiver provides observations of wavefront travel-time variance, pulse time spread, and temporal signal coherence. A VLA adds the observations of vertical wavefront coherence and modal time spread and time coherence. The combination of wavefront and modal observations provides tomographic vertical resolution of average internal-wave energies as a function of depth. In addition, vertical and temporal wavefront coherences are complementary observations, since spatial and temporal internal-wave variabilities are related through the dispersion relation. Also, vertical and temporal coherences help test the assumption of a separable internal-wave spectrum in terms of frequency and vertical wave number. First efforts to combine several of these acoustic fluctuation observations from the ATOC North Pacific VLAs in an internal-wave tomographic analysis will be presented. Directions for future research will be discussed.

2:15

1pAO4. A statistical theory of low-frequency sound propagation through internal waves. Alexander G. Voronovich (NOAA/ERL/ETL and Univ. of Colorado/CIRES, 325 Broadway, Boulder, CO 80303, agv@etl.noaa.gov)

A statistical theory of propagation of low-frequency sound waves through stochastic internal wave (IW) field is suggested. This theory is based solely on the weakness of sound scattering over distances of the order of correlation radius of IW field. Estimates show that a corresponding small parameter is of the order of a few percent for sound frequencies up to 50–100 Hz. Hence the theory is quite adequate for practically interesting cases. The governing equations include both average field and cross correlators of acoustic modes and are generally 3D. Those equations are finite difference rather then differential. This allows one to avoid making rather restrictive assumptions which can be easily violated. However, in the appropriate limit those equations reduce to standard ''couple power equations'' by Dozier and Tappert. Decrement of attenuation of average mode amplitudes is linearly related to the spacial spectrum of IW. Thus, by measuring propagation of tonal low-power sound signals over distances of a few hundred kilometers with the help of mode-resolving line vertical arrays, it is possible to infer in detail spacial spectrum of different IW modes (a task which would be extremely expensive to solve with the help of traditional contact measurements).

2:30

1pAO5. Partially saturated transmission scintillation and the bias of the sonar equation. Nicholas C. Makris (MIT, Rm. 5-204, 77 Massachusetts Ave., Cambridge, MA 02139, makris@keel.mit.edu)

The statistical properties of partially saturated ocean-acoustic intensity measurements are analyzed. This extends and generalizes previous analysis for completely saturated measurements [J. Acoust. Soc. Am. **100**, 769–783 (1996)]. While the random component of the instantaneous field is a circular complex Gaussian random process in both saturated and partially saturated cases, a purely deterministic component is added in the latter but not the former. In both cases, the random component may contain a fluctuating signal as well as noise. The present analysis, therefore, is relevant not only to transmission scintillation but also to the general analysis of deterministic signals in additive noise. Analytic expressions are derived for the probability distributions and moments of finite-time averaged intensity and log-transformed intensity as a function of the time-bandwidth product of the measurement. From these, general formulas are obtained for the bias of the sonar equation and the standard deviation of the logarithmic measures commonly used in ocean acoustics. Asymptotic expressions are also derived for the various distributions examined and their moments.

2:45

1pAO6. 1000-km square scale propagation from ocean acoustic tomography experiment in the Kuroshio Extension region. Hidetoshi Fujimori, Iwao Nakano, Toshiaki Nakamura, Guan Yuan (Japan Marine Sci. and Technol. Ctr., 2-15 Natsushima-cho, Yokosuka, Kanagawa, 237 Japan), Takashi Kamoshida, and Akio Kaya (Oki Electric Industry Co. Ltd., 4-10-3 Shibaura, Minato-ku, Tokyo, 108 Japan)

An ocean acoustic tomography observation was carried out at JAMSTEC (Japan Marine Science and Technology Center) in the Kuroshio Extension region, east of Japan, from July to September 1997. In this observation, five transceivers (consisting of a sound source and hydrophone array) were used and arranged 28°−36° N, 143°−154° E area. The longest distance between transceivers was about a 1000-km range and moored about a 1100-m depth. These transceivers were communicated to eight times (in August) or four times (in July and September) a day of each other and a time series of 1000-km square propagation data in this observation area were recorded and analyzed.

3:00–3:15 Break

3:15

1pAO7. Preliminary results from a large-scale 3-D tomography experiment in the Kuroshio Extension region. Gang Yuan, Hidetoshi Fujimori, Toshiaki Nakamura, Iwao Nakano (Japan Marine Sci. and Technol. Ctr., 2-15 Natsushima-Cho, Yokosuka, 237 Japan), Takashi Kamoshida, Tatsuo Zaima, and Akio Kaya (Oki Electric Industry Co. Ltd., Shibaura 4-10-3 Minato-Ku, Tokyo, 108 Japan)

A large-scale three-dimensional acoustic tomography experiment has been conducted in the Kuroshio Extension region from July to September 1997. Five 200-Hz transceivers were moored in the 1000-km-diam region approximately centered at 33° N, 149° E. Larger-scale environmental surveys (XCTD, CTD, XBT, and ship-boarded ADCP) were conducted along several acoustic transmission sections during the deployment and recovery cruises. The purpose of this study is to show the results of the preliminary analysis of this data. The stochastic inversions provide the time evolution of the large-scale three-dimensional temperature fields. The 4-D temperature fields dynamically show the variability of Kuroshio Extension with high time resolution. An active and resolvable mesoscale cold eddy is present with variability typically dominated by motions with spatial scales of 100–200 km. The 2-D temperature field along a 1000-km section derived from acoustic data agreed well with that determined from the XCTD measurements during the deployment cruise. The preliminary results of the data analysis suggest that it is possible to monitor 4-D structures of the large-scale ocean (1000×1000 km) by tomographic measurements (five transceivers giving ten paths), which is very difficult and almost impossible to do by traditional point measurements.

3:30

1pAO8. Monitoring temperature variability along the California coast using acoustic tomography. Thierry Morvillez, Ching-Sang Chiu, and Curtis A. Collins (Dept. of Oceanogr., Naval Postgrad. School, Monterey, CA 93943, chiu@nps.navy.mil)

The electronic emissions of a low-frequency sound source placed by the Acoustic Thermometry of Ocean Climate (ATOC) project on Pioneer Seamount were monitored by a bottom-lying receiver on Sur Ridge from April 1996 to February 1997. The processed signals show a stable arrival pattern that was repeated in all the transmissions during the 11 months. Using the processed data, a tomographic analysis to study the coastal ocean variability along this California transmission path was conducted.

Systematically, the analysis involved forward acoustic modeling of the arrival structure using ray theory, associating the observed arrivals with the modeled arrivals, extracting the travel times of the arrivals, inverting the travel times for temporal and spatial temperature changes, and interpreting the observed temperature variations. In particular, the tomographic estimate was compared to the temperature and wind measurements from an *in situ* mooring deployed by the Monterey Bay Aquarium Research Institution (MBARI). The comparison shows that the tomographic estimate is of high quality and that the observed temperature variations were linked to coastal upwelling and downwelling events. The data, methods, and results, demonstrating fully the feasibility of using tomography to study coastal temperature variability in central California on a long-term basis, are presented.

3:45

1pAO9. Ocean acoustic tomography using explosive charges to observe the submerged mesoscale oceanic variability in the East Sea of Korea. Jung Yul Na, Sang Kyu Han (Dept. of Earth and Marine Sci., Hanyang Univ., 1271 Ansan, 425-791, Korea), and Taebo Shim (Agency for Defense Development, Chinhae, 645-600, Korea)

A tomography system consisting of aircraft-deployed shots as sources and a vertical line array moored by a receiver ship was used to survey a large area (120×120 km) where a mesoscale warm eddy appears frequently. The position of the shots was measured by the GPS on the aircraft and the detonation time (or depth) was monitored by sonobuoys which had been deployed close to the shot position. Results show that individual multipaths between a shot and the receiver contain substantial information for a vertical slice of the ocean with a warm eddy. With high accuracy, GPS positioning of shot locations, and receiver position, this rapidly deployable system can be used for monitoring the submerged structure of the warm eddy during the period of diffusive SST in summer in the East Sea of Korea. [Work supported by KOSEF.]

4:00

1pAO10. Three-dimensional classical dynamics of acoustic rays in a quasirealistic ocean. Michael A. Wolfson (Earth Systems Sci. Ctr., Penn State Univ., University Park, PA 16802) and Steven Tomsovic (Washington State Univ., Pullman, WA 99164-2814)

Simulations on multimegameter range scales are presented for the dynamics of classical acoustic rays in a three-dimensional ocean waveguide which contains sound-speed inhomogeneities derived from an ocean general circulation model. The inhomogeneities include scales equal to and larger than mesoscales (>20 km). Both horizontal and vertical two-dimensional deterministic ray propagation through mesoscale environments are known to exhibit chaotic behavior in the sense of exponential divergence of neighboring trajectories that are initially infinitesimally separated. The extra degree of freedom provided by an acoustic phase speed that is fully three dimensional allows a richer, more complex structure to be revealed. We use a Hamiltonian prescription for the ray dynamics to reveal this structure geometrically via the phase space evolution of the associated Lagrangian manifold. The linear transformation matrices for this system which describe how elements of phase space volume distort under the action of the dynamics are examined for a large subset of trajectories. It is shown that their associated exponential rates of divergence vary orders of magnitude. The relative importance of horizontal versus vertical folding in the trajectory manifolds is also discussed.

MONDAY AFTERNOON, 22 JUNE 1998 EAST BALLROOM A (S), 1:40 TO 4:45 P.M.

Session 1pBV

Bioresponse to Vibration/Biomedical Ultrasound: Novel Approaches to Ultrasound Imaging II

Shira L. Broschat, Cochair
School of Electrical Engineering and Computer Science, Washington State University, P.O. Box 642752, Pullman, Washington 99164-2752

John M. Reid, Cochair
16711 254th Avenue, SE, Issaquah, Washington 98027

Invited Papers

1:40

1pBV1. Acoustic microscopy in pathological specimens. Mike Halliwell, Joe Eavis, Peter NT Wells, and Jem Berry (Bristol Univ., Bristol, Medical Physics, Bristol General Hospital, Bristol BS1 6SY, UK)

The application of acoustic microscopy to biological targets is extremely limited. Devices operating at gigahertz frequencies have been proposed as replacements for conventional light microscopes with the advantage of utilizing unstained specimens. However, there remains the inconvenience of sectioning and very slow scanning speeds. Applications to unsectioned, thick specimens of unstained tissues with a view to reconstructing three-dimensional images remain uninvestigated. A 60-MHz scanning acoustic microscope was constructed with two coaxial transducers separated by about 10 mm. Linear and compound scans were performed throughout a specimen volume of about 5 mm^3. Standard pulse-echo slices, velocity maps, highly compounded images, and three-dimensional representations of the specimen were produced. Water-filled airways, in fetal lung, of 40 μm diameter were visualized, and compounding was found to yield this resolution in both axial and lateral directions. Three-dimensional images demonstrated airway morphology. Acoustic scanning microscopy has been shown to be feasible in 5 mm^3 soft tissue specimens. The images demonstrate pathological features in fresh unstained material. Potential for three-dimensional reconstruction of pathological morphology remains to be explored. [Work was supported by the Foundation for the Study of Infant Death.]

2:00

1pBV2. Tissue harmonic imaging in cardiology and radiology applications. Michalakis A. Averkiou, James R. Jago, David N. Roundhill, and Jeffry E. Powers (ATL Ultrasound, P.O. Box 30003, Bothell, WA 98041)

Harmonic imaging, the formation of ultrasound images from the harmonic components of the sonification signal, has been used in diagnostic ultrasound applications that utilize microbubble contrast agents. In the presence of microbubbles the harmonic signals are mainly due to nonlinear bubble oscillations. Recently, harmonic imaging has been applied without contrast agents, the harmonic signals arising from nonlinear propagation. Near-field artifacts and aberrations are reduced, and tissue borders are enhanced. The harmonic signals are at a lower amplitude and suffer more attenuation than the fundamental. However, the wide dynamic range, digital architecture, and signal processing capabilities of modern diagnostic ultrasound systems make it possible to utilize this tissue-generated harmonic energy for image formation. The nonlinear parabolic wave equation is used to model tissue propagation. A time domain numerical solution for a nonaxisymmetric source function is presented. Theory and experiments demonstrate harmonic generation in tissues. *In vitro* animal tissue experiments and *in vivo* collected radio frequency data are shown. A novel broadband pulsing scheme that uses two pulses with a phase difference of 180° is introduced, which results in cancellation of the propagated fundamental component and doubling of the second harmonic. Cardiac and abdominal harmonic images are shown.

2:20

1pBV3. Optical Sonography™, a new method of diagnostic imaging. George F. Garlick (Adv. Diagnostics, Inc., 2400 Stevens Dr., Ste. C, Richland, WA 99352)

The characterization of Optical Sonography™ has been measured and is reported. This new soft tissue imaging method, which utilizes the capabilites of acoustical holography, passes a plane wave of ultrasound through the tissue which imposes phase and amplitude information on the wave. This wave then carries information of the three-dimensional structure of the soft tissue being imaged. When combined with an ultrasonic reference wave, an acoustical hologram is formed which is reconstructed with coherent light (hence the name Optical Sonography™) to provide a real time image of the soft tissue. With Optical Sonography™ the image resolution is determined by the wavelength of sound interacting with points in the tissue rather than the focus size of the ultrasound as in the reflective mode. Thus theoretical limits of resolution are achievable; comparative resolution images will be presented. In addition, since over 99% of the ultrasound intensity passes through soft tissue interfaces, greater sensitivity is achieved. Imaging results using Optical Sonography™ will be presented.

2:40

1pBV4. Ultrasonic focusing and steering through the skull: Toward brain imaging. Mickael Tanter, Jean-Louis Thomas, and Mathias Fink (Laboratoire Ondes et Acoustique, E.S.P.C.I., Universit C5 Paris VII, U.R.A. C.N.R.S. 1503, 10 rue Vauquelin, 75005 Paris, France, michael@clara.loa.espci.fr)

The time-reversal method has been experimentally applied to the focusing through the skull bone. This technique needs the knowledge of the Green's function of a dominant scatterer available in the medium. Moreover, it is shown that the skull induces severe attenuation of ultrasound and reduces the efficiency of the time-reversal approach. Nevertheless, an improvement of time-reversal processing for ultrasonic waves focusing through an aberrating and absorbing layer can be achieved. When the attenuating layer is located close to the array, this method consists of inversing the amplitude modulation before the time-reversal operation. Conversely, when the attenuating layer is located some distance away from the transducer array, numerical backpropagation is used to discriminate between the amplitude modulations resulting, respectively, from interference and absorption. Moreover, although time reversal is only self-adaptative to the artificial source implanted inside the imaging volume, this method is generalized to steer the ultrasonic beam on other points surrounding this acoustic source. This process has already been applied to hyperthermia through the skull bone and it can be applied to noninvasive brain imaging.

3:00–3:15 Break

Contributed Papers

3:15

1pBV5. High resolution ultrasound imaging system for soft tissues. S. G. Ye, X. Z. Liu, X. F. Gong, W. Y. Zhang (Inst. Acoust. and Lab. Modern Acoust., Nanjing Univ., Nanjing 210093, PROC), Y. Q. Jia, and X. Y. Tan (Jiangsu Cancer Hospital, Nanjing 210009, PROC)

The great interest in high resolution (higher than 0.1 mm axially and 0.3 mm laterally) ultrasound imaging systems for application to the skin, eye, intestinal wall, etc., gives rise to the development of a high resolution ultrasound imaging system. In this study, the high resolution ultrasound imaging system with a frequency range from 20 to 40 MHz was developed. In the system described here, the different resolutions were obtained by means of the focused transducers with different center frequencies in the frequency range mentioned above. In designing and fabricating the PVDF transducer, the KLM equivalent circuit including internal loss was developed and the influence of the PVDF film and backing materials on the transducer characteristics were analyzed as well. The experimental results were compared with the analyzed results. High resolution ultrasound imaging of some soft tissue samples was obtained. Besides the ultrasound parameters such as speed of sound, the attenuation coefficient and backscattering coefficient were measured in the frequency range from 20 to 40 MHz. [Work supported by NSF of Jiangsu Province, China.]

3:30

1pBV6. Inverse imaging of the breast using a conjugate gradient-neural network technique. Xiaodong Zhang, Shira L. Broschat, and Patrick J. Flynn (School of Elec. Eng. and Comput. Sci., Washington State Univ., P.O. Box 642752, Pullman, WA 99164-2752)

Early detection of a breast tumor greatly improves the chances of survival. Thus it is important to develop high-resolution imaging techniques capable of detecting small tumors. In this paper, an ultrasound imaging technique is presented that is based on solving the two-dimensional inverse scattering problem. An object is reconstructed using an iterative algorithm that combines the conjugate gradient method and a

neural network (CGNN) approach. The neural network technique exploits prior knowledge of the statistical characteristics of the different tissues inside the breast to enhance the performance of the conjugate gradient (CG) method. The CGNN algorithm has been tested on two small models simulating a breast tumor surrounded by glandular tissue. One model presents a tumor 3 mm in size and the other a tumor 1 mm in size. Results show that the CGNN algorithm is more accurate than the CG method. In addition, convergence occurs more rapidly; with the CGNN algorithm approximately 40% fewer iterations are needed to obtain the inverse solution. The CGNN method also appears to be robust to noise. Results are still reasonably good for a signal-to-noise ratio of 35 dB. Numerical results will be presented and discussed. [Work supported by NSF and the Carl M. Hansen Foundation.]

3:45

1pBV7. A filter-based pulse-echo coded-excitation system for real-time cardiac imaging. Emad S. Ebbini, Jian Shen, and Idris Elbakri (Dept. of EECS, Univ. of Michigan, Ann Arbor, MI 48109-2122)

A new filter-based approach to pulse-echo imaging capable of parallel processing of multiple image lines has been developed. A major application of this approach is in true real-time 3D imaging, e.g., in cardiac imaging. The new approach utilizes a set of coded waveforms on transmit and conventional beamforming on receive. The transmitted codes produce a coded wavefront in the region of interest (ROI) and the receive beam is designed to be wide enough to allow echoes from the ROI to be received while rejecting echoes from outside the ROI. Image formation is achieved using a filter bank structure (post-beamforming) with every filter designed to extract echoes from a given direction in the ROI. The filter bank is derived from a regularized pseudoinverse operator of the (discretized) propagation operator from the ROI to the receiving array. Image reconstructions from speckle-generating phantoms using the new approach demonstrate its robustness and its potential in medical imaging applications. A complete description of the algorithm and its fundamental spatial and contrast resolution limits will be presented with illustrative experimental data. System architectures for robust real-time implementation will be discussed. [Work funded by NIH Grant HL57167.]

4:00

1pBV8. Main error sources in ultrasound transmission tomography imaging. Krzysztof J. Opielinski and Tadeusz Gudra (Inst. of Telecommunication and Acoust., Wroclaw Univ. of Technol., Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland)

Ultrasound transmission tomography (UTT) consists in the usage of information included in ultrasonic pulses penetrating through an object in order to reconstruct sectional images of its inner structure. The contrast in the images corresponds to the distribution of local values of the acoustic parameter measured in projections. The measurement errors usually distort the tomographic image of the inner object structure. How different measurement errors are manifested and how big the changes they bring about in the reconstructed tomographic image can be fundamental for the determination of the ultrasound transmission tomography application range. The main sources of errors in ultrasound transmission tomography are the measurement accuracy and geometry, the ultrasonic refraction, the changing of measurement conditions, the measurement method, the reconstruction algorithm, and the disappearance of the receiving signal. Some different tomographic measurement errors (caused by the above-mentioned sources) were computer simulated and compared with the real ones. The influence of these errors on the ultrasonic tomographic imaging was shown. All of these errors refer to the run-time measurements in the parallel ray projection geometry using the convolution and backprojection algorithm for the image reconstructions.

4:15

1pBV9. Aggressive region growing filtering for speckle reduction in ultrasound images. Yan Chen (Siemens Med. Systems), Ruming Yin, Patrick Flynn, and Shira Broschat (Washington State Univ., Pullman, WA 99164-2752)

Speckle noise is present in all conventional ultrasound images, and complicates the process of diagnosis as it can mask small but diagnostically significant features. Speckle reduction techniques can be applied prior to image formation [e.g., Chen *et al.*, J. Ultrasound Imag. **18**(2), 122–139 (1996)] or after image formation. Post-formation techniques can be implemented in software, whereas pre-formation techniques generally require hardware modifications. The approach described here is applied after image formation, and consists of two steps. In the first step, small ''seed'' regions are defined around each pixel in the image, and a homogeneity criterion is used to guide a procedure which identifies the maximal-size homogeneous region containing the pixel of interest. Once determined, this region is smoothed using a truncated arithmetic mean filter. Results on synthetic and real images demonstrate that this method preserves the sharpness of image edges yet reduces speckle noise, and compares well with weighted median filtering [Loupas *et al.*, IEEE Trans. Circuits Syst. **236**, 129–135 (1989)] and homogeneous region growing median filtering [Koo and Park, J. Ultrasound Imag. **13**, 211–237 (1991)]. [Work supported by NSF and the Carl M. Hansen Foundation.]

4:30

1pBV10. Image-based calculation of elevational B-scan separation. Jochen Krücker, Theresa A. Tuthill, Gerald L. LeCarpentier, J. Brian Fowlkes, and Paul L. Carson (Dept. of Radiology, Univ. of Michigan Medical Ctr., Kresge III, 200 Zina Pitcher Pl., Ann Arbor, MI 48109-0553, jokr@umich.edu)

A speckle decorrelation technique was extended to high-frequency imaging to calculate the elevational (out of plane) separation of B-scan images. Pure speckle images were obtained with a 13-MHz, 1.5-D array in an elevational sweep over a phantom in 25-μm steps. For each pixel the plane-to-plane autocorrelation function was computed for a number of frames and fit to a Gaussian to find the depth-dependent scaling factor between the rate of speckle decorrelation and the actual frame spacing. From this the frame separation in sets of images containing speckle and structure was determined. To avoid using specular signals the mean intensity over standard deviation was employed to automatically detect areas of speckle. Based on the speckle decorrelation in these regions and using a 200*300 pixel subimage the frame spacing (100 μm) was calculated with 2% accuracy. The error increased with decreasing number of pixels and increasing separation of the frames. The technique can now be combined with high frame rate techniques to determine the elevational transducer motion during complex-motion 3-D scanning of the breast. [Work partially supported by the U.S. Army, Medical Contract DAMD17-96-C-6061 and the German Academic Exchange Service.]

Session 1pEA

Engineering Acoustics: Nondestructive Evaluation

Bernhard R. Tittmann, Chair
Department of Engineering Science and Mechanics, Pennsylvania State University, State College, Pennsylvania 16802

Contributed Papers

1:00

1pEA1. Acoustic detection of a non-metallic object embedded in a ground material. Hasson M. Tavossi and Bernhard R. Tittmann (Penn State Univ., Dept. of Eng. Sci. and Mech., State College, PA 16802)

The preliminary results of acoustic detection of a nonmetallic object embedded in a ground material model at a shallow depth is investigated. The frequency range of the acoustic wave is optimized to the composition of the ground material model to obtain minimum attenuation of the received signal. The investigation considers both theoretical and experiment aspects of the problem to provide a better understanding of the interaction of the acoustic waves with the ground model and the embedded nonmetallic object. One part of the theory is aimed at finding the optimal acoustic signal with maximum ground penetration as a function of frequency, band width, signal energy, and other wave characteristics. The other part of the theory considers ground material characteristics such as grain size and contact areas, density, rigidity, and porosity. In the experimental investigation both through-transmission and pulse-echo methods are used. The results include the relationship between the optimal frequency for maximum amplitude of the received signal, wave velocity, band-width, incident direction, signal energy, attenuation, ground material, and the embedded object characteristics.

1:15

1pEA2. High-frequency imaging of thickness degradation in steel containment vessels and liners. Joseph E. Bondaryk (Eng. Technol. Ctr., 200 Boston Ave., Ste. 2500, Medford, MA 02155, bj@arctic.mit.edu)

An area of concern in some nuclear power plants is thickness degradation of embedded steel containments just below the interface where the steel enters concrete. This is an area which cannot be examined by traditional UT testing, due to the inaccessibility imposed by the concrete. In this numerical feasibility study, high-frequency vibrational sources placed above the interface are used to excite elastic waves in the steel, which propagate into the embedded area. The waves that reflect and scatter from the surface roughness caused by thickness degradation are detected and used to determine and map the degradation of the steel. A range-dependent, stratified layer, wave-number integration based, numerical model (OASES) is used to calculate the field in the scenario. Decomposition of the field into longitudinal, shear, and surface waves is used to determine the optimal transducer orientations for transmit and receive. [Work supported by the Nuclear Regulatory Commision.]

1:30

1pEA3. The piezoelectric implant method compared to classical ultrasonic spectroscopy: Application to the monitoring of polymer curing. Agnes Raquois, J. Mohamed Tabellout, Jacques Emery (Gr. Ultrasons. Lab. de Chim. et Phys. des Materiaux Polymeres, UMR CNRS 6515, Univ. du Maine, 72035, Le Mans, France, ultrason@aviion.univ-lemans.fr), and Yves Jayet (GEMPPM-INSA, 69621 Villeurbanne Cedex, France)

Previous studies have described an original method enabling the *in situ* real time monitoring of a composite polymer structure by means of an inserted piezoelectric sensor. An analytical one-dimensional model has demonstrated that the electrical impedance of such an inserted element depends in particular on the acoustical parameters of the surrounding media. Then, by solving the inverse problem by means of a nonlinear optimization algorithm, it is possible to determine these acoustical parameters characteristic of the materials in contact with the inserted piezoelectric plate. This technique has been applied to the monitoring of the cure of gel-forming and glass-forming polymers such as polyurethanes and epoxies and compared to classical ultrasonic spectroscopy method. Thus it is possible to detect chemical evolution, transitions, and damage phenomena in such systems, provided that the material is homogeneous in the ultrasonic field or in the area surrounding the piezoelectric element. Particular attention has been focused on a polyimide (PMR-15) which is used as a matrix for composites in high-temperature applications. [Work supported by SNECMA/HISPANO-SUIZA.]

1:45

1pEA4. Interface wave source concepts for a seismic sonar concept. Frederick E. Gaghan, Steven R. Baker, and Thomas G. Muir (Phys. Dept., Naval Postgrad. School, Monterey, CA 93943-5000)

A seismic sonar concept is being developed which employs guided interface (Rayleigh or Scholte) waves to detect objects buried in sea floor and beach sediments. These waves have the desirable characteristics for this application that (1) they are localized near the interface and (2) the propagation from a compact source is cylindrical. As a result, the wave intensity on targets of interest is enhanced and the reverberation suppressed, compared to a bulk wave. To selectively excite Rayleigh or Scholte waves, a harmonic source employed at the interface must excite elliptical particle motion in the vertical and propagation, or radial, directions. Several such source concepts will be described. Results to date of field tests employing these sources in narrow-band pulse tests will be described. [Work supported by ONR Code 321OA.]

2:00

1pEA5. Approximate computational model for sound generation due to unsteady flows in pipes. Michael Krane, Daniel Sinder, and James Flanagan (CAIP Ctr., Rutgers Univ., Piscataway, NJ 08855)

A model for the generation of sound due to unsteady vortical flow in pipes of varying cross-sectional area is described. The model is based upon Howe's acoustic analogy and an approximate model for the vorticity field. This formulation allows the sound generation to be specified in terms of properties of the unsteady behavior of the vortical flow and the potential flow solution for the pipe that would exist in the absence of vortical inhomogeneities. The approximate model for vorticity formation and evolution is presented, and the potential flow solution is obtained from the axial duct shape. Computed estimates of the time evolution of both a jet passing through a pipe constriction and the sound this flow produces are presented and compared to experimental data. [Research supported by NSF/ARPA IRI-9314946 and ARPA DAST 63-93-C-0064.]

2:15

1pEA6. Acoustic-wave detection by optical beam deflection. James N. Caron, James B. Mehl (Dept. of Phys., Univ. of Delaware, Newark, DE 19716-2570, jcaron@udel.edu), and Karl V. Steiner (Univ. of Delaware, Newark, DE 19716-3144)

Audio-frequency acoustic signals in gases have been detected by measuring the bending of an optical beam directed through the gas to a custom-built position-sensitive photodetector. The beam deflection is caused by a transverse acoustic density gradient, and can be calculated when the acoustic field is known. The spectrum of a cylindrical resonator was recorded using both a calibrated microphone embedded in one end of the cylinder, and deflection of a laser beam passing through the center of the cylinder along a diameter. Beam deflection is expected for the first plane-wave mode of the resonator and other modes having nonzero average transverse density gradients along the beam path. The observed beam deflections were in qualitative agreement with calculations for all modes which could be resolved. For the first plane-wave mode, the measured and calculated deflections agreed within a few percent. Refinement of this technique could be applied to the absolute calibration of microphones. The beam-deflection technique was found to be sufficiently sensitive for the recording of voice and musical instruments, both in digital form on laboratory instruments and on audio tape. [Work supported by the Army Research Office.]

2:30–2:45 Break

2:45

1pEA7. Laser-based ultrasonic NDE using gas-coupled laser acoustic detection. James N. Caron, Yuqiao Yang, James B. Mehl (Dept. of Phys., Univ. of Delaware, Newark, DE 19716-2570, jcaron@udel.edu), and Karl V. Steiner (Univ. of Delaware, Newark, DE 19716-3144)

An optical beam-deflection technique has been applied to the ultrasonic inspection of materials. The technique is similar to air-coupled detection, except the airborne wave is detected through beam deflection instead of with an electro-acoustic transducer. It is a noncontact, laser-based, detection technique that does not depend on reflection of the detection laser beam from the surface. It is thus independent of the surface optical properties of the material under test, and is particularly suitable for testing polymer/graphite composites, which have rough surfaces of poor optical reflectivity. The beam is directed parallel to the surface of a solid material under investigation and is incident on a custom-built position-sensitive photodetector. A wave is radiated from the surface into the surrounding gas during reflection of elastic waves in the solid. The transverse acoustic density gradients cause a beam deflection which is clearly observable with the photodetector. Applications of gas-coupled laser acoustic detection (GCLAD) with laser ultrasound generation will be illustrated with (1) waveforms showing numerous, well-resolved echoes, (2) C-scans showing subsurface defects, and (3) surface acoustic waves. The samples tested include several types of polymer/graphite composites. [Work supported by the Army Research Office.]

3:00

1pEA8. Direct detection of acoustic waves by laser light diffraction and proposals of the optophone. Yoshito Sonoda (Dept. of Elec. Eng., Kyushu Tokai Univ., 9-1-1 Toroku, Kumamoto 862, Japan, ysonoda@ktmail.ktokai-u.ac.jp)

The conventional light diffraction method cannot be applied to the measurement of the acoustic waves with a long wavelength, where the diffracted optical waves propagate in the penetrating optical beam and cannot be measured because of the disturbance by the latter. In this study, the Fraunhofer diffraction method, which was developed as a new means to detect the electromagnetic radiation scattered within the penetrating light beam in the plasma-nuclear fusion research, is applied to the measurement of acoustic waves. It is experimentally demonstrated that the acoustic wave with a wavelength longer than 8.5 mm can be detected by a visible laser of 5 mW, and the amplitude of the electrical output signal from a photodiode is linearly proportional to the sound pressure and the probe laser power, as predicted by the theory. The frequency property and directivity of the optical sound-antenna are roughly estimated by the preliminary experiment. Furthermore, the Optophone (the optical microphone) of various types are proposed as a new acoustic detection method.

3:15

1pEA9. Nondestructive SAW technique of acoustic transient spectroscopy to study deep centers in semiconductor heterostructures. Peter Bury and Igor Jamnický (Dept. of Phys., Žilina Univ., 010 26 Žilina, Slovak Republic, bury@fel.utc.sk)

A new technique of acoustic deep level transient spectroscopy (A-DLTS) using a high-frequency transverse acoustoelectric signal produced by a surface acoustic wave (SAW) electric field was used to study deep centers in semiconductor heterostructures. The basic idea of the presented A-DLTS technique is connected with the fact that the time development of the acoustoelectric response signal arising from the interaction between the SAW electric field and free carriers at heterostructure interface after an injection bias voltage pulse has been applied reflects relaxation processes associated with the thermal recombination of excited carriers moving toward their equilibrium state. The technique of computer evaluation of isothermal acoustoelectric transients applying the date compression algorithms [P. Bury and I. Jamnický, Acta Phys. Slovaca **46**, 693–700 (1996)] was used to calculate deep center parameters as activation energies and capture cross sections. The technique has been applied for both Si MIS and GaAs/AlGaAs heterostructures. [Work supported by Grant of Slovak Ministry of Education No. 1/4261/97.]

3:30

1pEA10. Acoustic spectroscopy of deep centers in high resistivity GaAs using a light beam generated interface. Peter Bury (Dept. of Phys., Žilina Univ., 010 26 Žilina, Slovak Republic, bury@fel.utc.sk)

An acoustoelectric effect on a light beam generated interface of high resistivity Cr-doped GaAs is used as a probing tool to study bulk deep centers using acoustic deep-level transient spectroscopy (A-DLTS). The space charge inhomogeneity, produced by the illumination of part of the photoresistive semiconductor by weakly absorbed light, can be generated by applying a high-frequency electric field in the proper conditions for the fast mode shear acoustic wave. The amplitude of such a generated acoustic wave depends on the concentration of free carriers excited from the deep centers at the light–dark interface. When the illumination of the semiconductor is turned off, the photoexcited carriers recombine thereby establishing thermal equilibrium and temperature dependence of the time develop-

ment of the amplitude. The acoustoelectric response signal is then investigated to determine the deep centers' parameters. The A-DLTS technique, based on the computer evaluation of acoustoelectric isothermal transients registered after the excitation illumination [P. Bury and I. Jamnický, Acta Phys. Slovaca **46**, 693–700 (1996)], was used to calculate activation energies and capture cross-sections of deep centers in Cr-doped GaAs. [Work supported by Grant of Slovak Ministry of Education No. 1/4261/97.]

MONDAY AFTERNOON, 22 JUNE 1998 METROPOLITAN BALLROOM (S), 1:00 TO 3:00 P.M.

Session 1pED

Education in Acoustics: Take Fives—Sharing Ideas for Teaching Acoustics

Thomas D. Rossing, Chair
Physics Department, Northern Illinois University, DeKalb, Illinois 60115

Do you have a novel demonstration, a new laboratory experiment, a favorite video, a recorded sound example, or a new idea for teaching acoustics that you would be willing to share with your colleagues? At this session, there will be a sign-up board for scheduling the presentations. No abstracts are printed, although some presenters may have handouts to distribute to their colleagues. You may sign up for more than one presentation (but not consecutively); presentations are strictly limited to 5 minutes. Keep them short! Keep them fun!

MONDAY AFTERNOON, 22 JUNE 1998 FIFTH AVENUE ROOM (W), 2:15 TO 6:00 P.M.

Session 1pMU

Musical Acoustics and Psychological and Physiological Acoustics: The Purposeful Use of Nonlinear Distortion in Musical Performance: The Eric Dolphy/Jimi Hendrix Celebratory Session

Edward M. Burns, Chair
University of Washington, JG-15, 1417 NE 42nd Street, Seattle, Washington 98105-6248

Chair's Introduction—2:15

Invited Papers

2:20

1pMU1. Some nonlinear aspects of the ear as a sound detector and a sound emitter: Comparisons with properties of musical instruments. Arnold Tubis, Carrick L. Talmadge (Dept. of Phys., Purdue Univ., West Lafayette, IN 47907), and Glenis R. Long (Purdue Univ., West Lafayette, IN 47907)

A compressive nonlinearity in the mechanics of the cochlea is responsible for (1) the dynamic detection range of about 120 dB; (2) the level-dependent modification of the ability to detect one tonal component in the presence of another (masking); (3) the stabilization of cochlear self-oscillations, which are associated with sounds broadcast from the ear (spontaneous otoacoustic emissions); and (4) the psychoacoustic detection and detection in the ear canal of frequency components (intermodulation distortion products) not present in the original acoustic signal. The spontaneous otoacoustic emission spectra are characterized by quasi-regular frequency spacings, with some ears having over 30 narrow-band spectral peaks. The intermodulation distortion level is higher than that claimed for mediocre stereo systems and varies with frequency. Several interesting analogies between the modeling of otoacoustic emissions and the sound spectra of musical wind instruments are discussed. [Work supported by NIDCD.]

2:45

1pMU2. Deliberate use of distortion in singing. Ingo Titze (Dept. of Speech Pathol. and Audiol. & The Natl. Ctr. for Voice & Speech, Univ. of Iowa, Iowa City, IA 52242)

Singers are able to create aperiodic vibration of the vocal folds by either desynchronizing the primary oscillators (the vocal folds) or by engaging secondary sound sources (the false vocal folds, the aryepiglottic folds, or the velum). The desynchronization of the primary oscillators can be accomplished by creating a slight left-right asymmetry and then using the vocal tract (an acoustic oscillator) to entrain one vocal fold, leaving the other vocal fold to oscillate independently. Other aperiodicities can be created by driving the

vocal folds at large amplitudes (with the critical parameter being A/L, the amplitude to length ratio of the vocal folds). In this mode of distortion, the highly nonlinear stress-strain characteristics of the vocal fold tissues are exploited. Infant cries, shouts, and groans are characteristic of this type of nonlinearity. Some videotaped examples and some computer simulations will be given to illustrate these phenomena.

3:10

1pMU3. An acoustical appreciation of the extended performance aspects of Eric Dolphy's musical craft. Douglas H. Keefe (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131)

The late 1950s and early 1960s was an era of rapid innovation in jazz in which musicians transformed the compositional and improvisional form of the music. While instrumentalists have incorporated extended performance techniques in a unique manner from the beginning of jazz, an important aspect of the avant garde movement was the breadth of innovation in woodwind instrument performance. Eric Dolphy was a pioneer in this respect, a virtuoso performer of the alto saxophone, bass clarinet, and flute, who helped bring the latter two instruments into prominence within jazz as solo instruments for improvisation. Literal acoustic distortion is present in woodwind multiphonics used by Dolphy and others, produced not merely by using unconventional fingerings and changes in embouchure, but also by alterations in the resonances of the player's vocal tract [Wilson (1996)]. Aesthetic distortions include the sheer temporal complexity of melodic lines, which cannot be separately resolved by many listeners, and the inclusion of speechlike and bird-song elements within extended musical forms lacking a conventional harmonically related musical structure. Such factors contribute to the overall artistic impact of this music.

Contributed Paper

3:35

1pMU4. Eric Dolphy's playing style in the Vienna Art Orchestra's performance. Alexandra Hettergott (Inst. of Wiener Klangstil, Univ. of Music, Singerstrasse 26a, A-1010 Vienna, Austria)

Aiming on the very musical expression intended, ''tone colorist'' Eric Dolphy in his developed avantgarde jazz playing style on alto sax and bass clarinet as well as on flute makes use of several extraordinary playing techniques (e.g., multiphonics). In musically portraying him [*Nine immortal non-evergreens for Eric Dolphy*, Amadeo 537 096-2], the now 20-year-old Vienna Art Orchestra in a postmodern way renders Dolphy's compositions in their very performance style, adopts and adorns his melodies typically performed in a highly expressive, noise-inducing way with their own instruments' rich means of expressions, allowing also for brass instruments. Consequently, the bass clarinet in *Hat and Beard* is given in solos of trumpet, trombone, and alto sax, the flute in *Gazzelloni* reflected in both trumpet and f- horn solos, or the alto sax in *Straight Up and Down* expressed in solos of either trumpet, trombone, and alto sax, respectively. In sound analyses along with sound examples the paper compares several of Dolphy's playing techniques in his original late recordings (1964) with the recent performances by the Vienna Art Orchestra (1995), focusing on the wind instruments' (unusual) sounds as part of the personal aesthetics of expression in both model Dolphy and the Vienna Art Orchestra.

Invited Papers

3:50

1pMU5. When bad amplification is good: Distortion as an artistic tool for guitar players. Tilmann Zwicker (Feldstraße 5, D-82057 Icking, Germany) and Søren Buus (Northeastern Univ., Boston, MA 02115)

This presentation surveys the development of and the approaches to distortion for electric guitars, which plays an extremely important role in modern popular music, particularly in Rock, Blues, Jazz, and Pop. The signal processing for guitar amplification has always included nonlinear effects that are usually heard as distortion, but also lends other important characteristics to the guitar tone. Originally, distortion occurred as musicians tried to coax more volume out of their amplifiers than they were designed to deliver. However, this ''design flaw'' soon became a tremendous asset. Guitarists discovered that, if handled properly, distortion gave their instruments an expressive, harmonically rich, and highly controllable voice rivaling that of horn players. Today, distortion for electric guitars has been cultivated to the extent that it is deliberately designed into every guitar amplifier. Characteristics of electric-guitar signals and distortion devices will be discussed and illustrated with examples. The reasons why distortion is a desirable artistic tool and related psychoacoustic aspects also will be discussed.

4:15

1pMU6. Jimi Hendrix' use of distortion to extend the performance vocabulary of the electric guitar. James J. Fricke (Experience Music Project, 110-110th Ave. NE, Ste. 550, Bellevue, WA 98004)

Experience Music Project is an interactive music museum scheduled to open in Seattle in 1999. Sound Lab, the hands-on portion of the museum, will use a variety of exhibit techniques to lead the visitor to explore music composition and performance, the technology of popular music making, and the physical principles involved in sound production and reproduction. The role of acoustics and signal processing in the development and realization of these exhibits is discussed. Jimi Hendrix, who pioneered the use of electronic effects in extending the expressive vocabulary of the electric guitar, is central to one planned exhibit. Hendrix' ''Drivin' South'' guitar solo from 1967 is presented as an example of the extended expressive possibilities afforded by a highly overdriven, amplified electric guitar. This use of distortion for artistic effect is analyzed, and the performance is compared to an undistorted performance of the same piece. The electronic means for achieving the distorted guitar sound is presented using artifacts from EMP's Jimi Hendrix collection, and the resulting signal modification is demonstrated and analyzed.

4:40–4:55 Break

4:55–6:00 Dolphy/Hendrix Memorial Jam

Following the formal session there will be a jam session in which several of the presenters will participate, along with local musicians expert in the techniques pioneered by Dolphy, Hendrix, and others.

MONDAY AFTERNOON, 22 JUNE 1998

CASCADE BALLROOM I, SECTION A (W), 2:15 TO 5:30 P.M.

Session 1pNSa

Noise: Impulse Noise Assessment

Paul D. Schomer, Chair
U. S. Army Construction Engineering Research Laboratory, P.O. Box 9005, Champaign, Illinois 61826-9005

Chair's Introduction—2:15

Invited Papers

2:20

1pNSa1. Impulse noise and startle. Ian H. Flindell (Inst. of Sound and Vib. Res., Univ. of Southampton, Southampton, UK)

It is well known that impulsive noise caused by gunfire, explosive devices, and high speed overflight by aircraft at low altitudes can cause an acoustic startle effect (ASE). The ASE is one of many possible startle effects which are normally associated with unexpected stimuli. The ASE always consists of a sequence of largely defensive reflex actions with physical, physiological, and psychological components. It is generally accepted that the ASE can be a direct contributor to accidents in inherently risky situations, but what is more controversial is the possibility of some, presently undefined, linkage to longer term health effects. Based on a critical review of the literature, this paper reports some general findings on the possible theoretical bases for any such linkage and makes a number of suggestions for future research. The available evidence suggests that if there is such a linkage, it must be mediated either through individual differences in reactivity and habituation, or through the less well-specified ''general stress hypothesis.'' The fact that many of the direct physiological responses observed in the ASE are not generally considered harmful in other contexts, such as exercise, is likely to be an important consideration.

2:40

1pNSa2. Comparison of procedures for the quantification of the impulsivity of environmental sounds. Giovanni Brambilla (CNR Istituto di Acustica, via Cassia 1216, I-00189 Roma, Italy, gianni@idac.rm.cnr.it)

Any procedure for the quantification of the impulsivity of environmental sounds should meet the following requirements: (i) sufficiently adequate in describing the human perception of impulsivity; (ii) easy to perform in the field by means of basic instrumentation or in the laboratory without requiring expensive apparatus and too much time. Among the several procedures proposed in the last decade, those more frequently used or promising have been considered in this paper and compared to one another. For this purpose data collected in earlier European studies have been reanalyzed. The data considered refer to 26 environmental sound stimuli heard in laboratory at 55 L_{Aeq} level for 20 s by a number of people varying between 48 and 112 subjects. The subjective data on clear perception of impulsivity have been related to the selected descriptors, such as LAI-LAS, increment, etc. The potential of these objective parameters in describing the perception of impulsivity, as well as the determination of the threshold value for the onset of impulsivity, has been investigated by means of their correlation with the subjective data and determination of sensitivity and peculiarity of each descriptor.

3:00

1pNSa3. On fitting functions to data for the percent of a community highly annoyed by noise such as impulsive noise. Paul Schomer (U.S. Army Construction Eng. Res. Labs., P.O. Box 9005, Champaign, IL 61826, schomer@uiuc.edu) and Karl Hirsch (Inst. for Noise Protection, Duesseldorf, Germany)

Survey respondents ''highly annoyed'' (HA) typically are categorized as those responding in the top 30% on an anchored numerical scale or in the top two categories on an adjectival scale. The percent HA typically ranges from 0% to 40%. At higher percents HA (e.g., $>15\%$), the noise stimulus is clear and dominant and the reliability to the percentage values is high. At low percentages, the reliability is low. Some respondents may be annoyed any time they hear a sound. Depending on the number of such respondents, the percent HA can vary greatly on a percentage basis (4% HA versus 2% HA is a change of 100%). Therefore, alternatives to using small percentages HA (e.g., $<10\%$) are considered when fitting transition functions to HA data. One method is

to approximate the transition function by three asymptotes. The transition (sloped) asymptote would be fit by the 50% point and the slope when the percentages HA were large. Another way to avoid small percentages HA is to consider the group ''little annoyed'' or ''happy.'' This category might include the bottom two categories on a five-point adjectival scale or the bottom 30% on a numerical scale. This paper explores these issues.

3:20

1pNSa4. On the rating of high-energy impulsive sounds: Optimization of a level-dependent conversion procedure with the help of new field survey results. Joos Vos (TNO Human Factors Res. Inst., P.O. Box 23, 3769 ZG Soesterberg, The Netherlands, vos@tm.tno.nl) and Edmund Buchta (Inst. für Lärmschutz, Düsseldorf, Germany)

A few years ago, Schomer [Noise Control Eng. J. **42**(5), 179–191 (1994)] developed a new rating procedure for high-energy impulsive sounds, in which a level-dependent conversion of the *C*-weighted sound exposure levels (CSELs) of the impulses into *A*-weighted sound exposure levels (ASELs) of equally annoying vehicle sounds was applied. In a recent reanalysis, Vos [Noise Control Eng. J. **45**(3), 123–131 (1997)] showed that at least for artillery sounds, the optimal parameter values to be used in this level-dependent conversion procedure were different from those suggested by Schomer. The experimental results from a new field survey on the community response to artillery and road-traffic sounds (Buchta and Vos, these proceedings) provide a new opportunity to optimize the conversion procedure. The present effort is especially interesting because here, in contrast to previous analyses, large sets of different CSEL values per noise zone are available (for most of the zones the impulses originating from tens of different firearm/shooting range combinations were considered), and the annoyance ratings for the impulsive and road-traffic sounds had been determined with similar questions and for the same group of respondents, facilitating a comprehensive estimate of the adequate parameter values.

3:40–3:55 Break

3:55

1pNSa5. Community response to artillery and road-traffic sounds—A new field survey. Edmund Buchta (Inst. für Lärmschutz, Arnheimerstrasse 107, D-40489 Düsseldorf, Germany, e.buchta@t-online.de) and Joos Vos (TNO Human Factors Res. Inst., Soesterberg, The Netherlands)

With the help of a questionnaire, subjective reactions to artillery sounds (muzzle bangs mainly from 120- and 155-mm caliber weapons, detonating high-explosive shells, and other explosions) were determined for over 400 respondents divided among 17 different residential areas. Also, for the same respondents, the subjective effects of (local) road-traffic sounds were determined, enabling a comprehensive comparison of dose–response relations for the, typically, high-energy impulsive sounds and the road-traffic sounds. For the sake of comparison with other field surveys, the noise dose for the shooting sounds was expressed as the yearly average *C*-weighted day–night level (L_{Cdn}), and that for the road-traffic sounds was expressed as the *A*-weighted day–night level (L_{Adn}). Similarly, for both sound types the community response in each residential area was expressed as the percentage of respondents being ''highly annoyed.'' From the comparison of the two dose–response relationships it could be concluded that for numerically equal L_{dn} values, the artillery sounds were more annoying than the road-traffic sounds. Overall, the difference is equivalent to the change in annoyance produced by a 5-dB shift in either L_{Cdn} or L_{Adn} of the sounds.

Contributed Papers

4:15

1pNSa6. Effect of low-frequency content on the rate of growth of annoyance of impulsive sounds. Sanford Fidell, Karl Pearsons, and Laura Silvati (BBN Technol., 21128 Vanowen St., Canoga Park, CA 91303)

Since useful field studies of the annoyance of high energy impulsive sounds are costly and difficult to conduct, comparatively little is known empirically about community response to high energy impulsive noise exposure. Several attempts have therefore been made to generalize the results of controlled exposure studies of the relative annoyance of impulsive and nonimpulsive noises to community-level analyses. Although such studies have yielded a range of level-dependent growth functions for the annoyance of impulsive sounds, it is commonly observed that the annoyance of impulsive sounds increases more rapidly with level than that of nonimpulsive signals. The findings of a recent set of paired comparison judgments, in which the very low-frequency content of sonic booms was accurately reproduced, suggest that the steep slope of human hearing sensitivity at low frequencies may provide a good account for the common observation, at least for the case of sonic booms. If this is so, then *ad hoc* ''adjustments'' to the annoyance of impulsive sounds based on rise time, crest factor, and similar waveform characteristics may be superfluous. [Study supported by U.S. Air Force Air Combat Command.]

4:30

1pNSa7. Prediction of the difference between CSEL and ASEL of blast sound for correlation purposes with annoyance. Karl W. Hirsch (Inst. für Lärmschutz, D-40489 Duesseldorf, Arnheimer Str. 107, Germany, ifl.acoustics@t-online.de)

ISO-TC43/SC1/WG45 plans to have three categories of impulsive sound sources for purposes of noise ratings. Each category, ''ordinary impulsive'' (adjustment 5 dB), ''highly impulsive'' (12 dB), and ''high-energy impulsive'' (significantly more than 12 dB to ASEL) sound, will receive the indicated adjustment to its measured SEL. But these adjustments based on categories of sources do not take into account the change of ''impulsiveness'' with propagation. This omission raises uncertainties—especially for high-energy blast sounds. Recently published proposals by Buchta and Vos use the differences between CSEL and ASEL at the receiver to determine impulse adjustments that correlate with annoyance. However, this level difference is not reliably obtained by measurement because the spectrum is influenced by local properties of the ground due to superposition, and because the A-weighted levels are close to the background noise level at typical receiver distances. For blasts from explosive sources (e.g., demolitions, weapon firing), this paper presents a theoretical approach to predict CSEL-ASEL using an energy concept

forthe sound spectrum. This approach yields an expected value for the dif-ference between C- and A-levels that depends on the acoustical energy of the sound at the source and on the frequency-dependent attenuation along the propagation path.

4:45

1pNSa8. A further test of the relevance of ASEL and CSEL in the determination of the rating sound level for shooting sounds. Joos Vos (TNO Human Factors Res. Inst., P.O. Box 23, 3769 ZG Soesterberg, The Netherlands, vos@tm.tno.nl)

In a previous study on the annoyance caused by shooting sounds [Vos, Proc. Internoise '96, **5**, 2231–2236], it was shown that an almost perfect prediction of the annoyance, as rated indoors with the windows closed, was obtained on the basis of the weighted sum of the outdoor A-weighted and C-weighted sound exposure levels [ASEL (L_{AE}) and CSEL (L_{CE})] of the impulse sounds. Basically, the annoyance (y) was given by $y=c+\alpha L_{AE}+\beta\cdot\Delta L\cdot L_{AE}$, in which $\Delta L=L_{CE}-L_{AE}$. In that study, just one type of facade attenuation (FA) was simulated. In the present study, the annoyance was determined for four FA types that varied from low (windows slightly opened) to very high (20 dB at 16 Hz to 50 dB at 8 kHz). Again, the sounds were produced by firearms ranging in caliber from 7.62 to 155 mm, and were presented at different levels. In all four FA conditions, the annoyance could be very well predicted by the equation given above, multiple r^2 varied from 95% to 97%. Relative to the prediction on the basis of ASEL only, the increase in r^2 varied from 7% in the slightly opened windows condition to 55% in one of the high FA conditions. It was concluded that for the determination of the rating sound level, the acoustic parameters ASEL and CSEL are very powerful.

5:00

1pNSa9. Impulse noise metrics and their application to noise from low-flying military jet aircraft. G. Kerry, P. D. Wheeler, T. I. Hempstock (Telford Inst. of Acoust., Univ. of Salford, Salford M5 4WT, UK), and D. J. James (HQ PTC, RAF Innsworth, Gloucester GL3 1EZ, UK)

It is recognized that noise from fast low-flying military jet aircraft is a major factor in the public concern over military low-flying training operations. While much research has been carried out worldwide, potential health risks from noise and environmental noise nuisance remain poorly understood. Part of the difficulties arise in quantifying the exposure and in selecting appropriate risk or tolerability criteria. Much depends on the nature of the noise. This review investigates the definitions and characteristics of impulsive noise and the various metrics in current use for assessing both hearing hazards and environmental annoyance and considers their application to noise from low-flying military jets.

5:15

1pNSa10. Wideband noise signatures from low-altitude military jet overflights. C. Lomax, G. Kerry (Telford Inst. of Acoust., Univ. of Salford, Salford M5 4WT, UK), and D. J. James (HQ PTC, RAF Innsworth, Gloucester GL3 1EZ, UK)

Public concern about the noise from low-altitude military jet overflights and its alleged effects on the health and annoyance of those overflown has resulted in some restrictions on flying activities at low level, even bans within some countries, being imposed. Much past research into aircraft noise emissions has been based on band-limited noise measurements. Full pressure signatures are required if all possible noise effects are to be fully understood. In 1995 a series of trials was completed in the UK under controlled conditions, and typical wideband noise signatures obtained for a variety of speeds and heights of a Tornado aircraft. Samples of these signatures are presented and the various pertinent characteristics highlighted. The question as to whether these signatures should be considered impulsive is discussed.

MONDAY AFTERNOON, 22 JUNE 1998

CASCADE BALLROOM I, SECTION C (W), 2:15 TO 5:05 P.M.

Session 1pNSb

Noise and Architectural Acoustics: Novel Approaches to Noise Control Barriers II

Michael R. Stinson, Chair
Institute for Microstructural Sciences, National Research Council, Building M-36, Montreal Road, Ottawa, Ontario K1A 0R6, Canada

Chair's Introduction—2:15

Invited Papers

2:20

1pNSb1. Scale modeling of railway noise barriers. David C. Hothersall, Kirill V. Horoshenkov, and Philip A. Morgan (Dept. of Civil and Environ. Eng., Univ. of Bradford, Bradford, West Yorkshire BD7 1DP, England, dchothersall@bradford.ac.uk)

Noise barriers are an efficient way of reducing the noise from railways, particularly the noise produced in the region of the rails and wheels. Several designs of noise barrier have been suggested in the literature which aim to improve the insertion loss over that obtainable from a plane screen of the same height. The positioning of these barriers is governed by the limitations imposed by the structure gauge of the railway. Scale modeling has been carried out in a humidity-controlled anechoic chamber at a scale of 1:20. Appropriate materials have been selected to model the impedance of the track ballast, absorbing barrier surfaces and grassland at the reduced scale. To model the interaction between the side of the train and the barrier realistically, models appropriate for a high-speed train cross section have been used. Experimental results for the insertion loss of shaped and multiple-edge barriers are compared with the results from numerical modeling. The efficiency of each of the designs is assessed. The generation of time histories for the passage of trains and the calculation of noise indices are considered.

2:50

1pNSb2. Specific problems on the use of noise barriers on viaducts. Jean-Pierre Clairbois, Peter Houtave, and Vincent Trefois (Acoustical Technologies Ave., Brugmann, 215 B-1050 Brussels, Belgium)

Whatever the location of highways or high-speed roads is (urban, suburban, country), we frequently encounter viaducts or bridges. Building noise barriers on viaducts is not a simple matter, mainly for 3 important reasons: aesthetic, safety and stability, and the problem of viaduct joints. Aesthetically, the wish is to reduce the height of the barriers. One also very often wishes to use transparent barriers: for that case, these reflecting materials can degrade the performance of the barrier, either due to the effect of parallel reflecting walls, or due to the reflections which occur with high sided vehicles passing close to these. One should then achieve a specific study of the shape and materials that have to be used in such an application, not forgetting that each case is different. Another important problem occurring with the barriers on viaducts is that, once the barrier is effective and reduces the noise originally coming from the traffic on a viaduct, one could listen back to the noise due to bad noisy joints, noise coming from below viaducts, and not protected with the barriers. This paper will present some measurements and design studies made in such situations.

3:20

1pNSb3. Source type effect on the efficiency of noise barriers. Philippe Jean (Ctr. Scientifique et Technique du Bâtiment, 24, Rue Joseph Fourier, 38400 Saint-Martin-d'Hères, France)

The use of numerical techniques such as the Boundary Element Method (BEM) is well suited for the study of noise barriers of complex shapes. Full studies, up to 4000 Hz, can only be done with 2D approximations, where the geometry of the problem is infinite, which is a realistic approximation for noise barriers. It also implies that the sources are infinite coherent line sources, which is a gross approximation for traffic noise. Recent transforms have been presented in the literature to overcome this last restriction. Simple integrals of 2D results now permit the rapid consideration of any combination of point sources or infinite incoherent line sources. Such transformations have been tested with a BEM code based on a variational formulation. The efficiency of different screen modifications is analyzed in the light of these more realistic source representations, and the importance of such developments is assessed by comparing the results with the classical coherent source approximation.

3:50–4:05 Break

4:05

1pNSb4. A new *in situ* method for the acoustic performance of road traffic noise reducing devices. Jean-Pierre Clairbois (Acoustical Technologies Ave., Brugmann, 215 B-1050 Bruxelles, Belgium), Jacques Beaumont (ENTPE/LASH F-6951 Vaux-en-Velin Cedex, France), Massimo Garai (DIENCA / Universita di Bologna, Bologna, Italy), and Gerold Schupp (IBP fhG, Stuttgart, Germany)

When they drew the European standards for ''road traffic noise reducing devices'' (NRD), the members of CEN/TC226 (road equipments)/WG6 (antinoise devices) found problems in using ISO 354 and ISO140 for characterizing sound absorption and airborne sound insulation of NRD. The problems are due to the fact that these standards have been drawn for building products, used inside buildings. The measurement of the sound absorption under a reverberant field does not reflect the real absorption coefficient found with road traffic. In addition, many road administrations wish to get a field test method for testing installed products. A European funded research (''ADRIENNE'') has been started in January 1995 and finished in December 1997 with the objectives to be able to measure the *in-situ* sound absorption and airborne sound insulation of NRD for the widest range of existing products (i.e., flat and nonflat ones). This paper shall present the results of the research, which is based on the combination of MLS signal and signal analysis, a substraction technique to enlarge the lowest reliable frequency limit, and a specific window. The fundamental method has been developed with three laboratories, and initial validation with eight laboratories shows promising results for a future standard.

Contributed Papers

4:35

1pNSb5. Full-scale model investigations into insertion loss of highway noise barriers under different atmospheric conditions. Scott D. Hansen, Courtney B. Burroughs, and Kenneth E. Gilbert (Grad. Prog. in Acoust. and Appl. Res. Lab., P.O. Box 30, University Park, PA 16802)

The insertion loss of highway noise barriers can change significantly with atmospheric conditions. The insertion loss of a 14×80-ft tongue and groove wooden barrier constructed at the Pennsylvania Transportation Institute test track near State College, PA was measured under different atmospheric conditions. Point sources consisted of pure tones and random noise. Both sources and receivers were held at fixed distances from the barrier. Wind and temperature profiles were estimated from measurements using similarity theory. Measurements were taken as winds varied in both direction and strength for different temperature profiles. Results from parabolic equation model calculations of the propagation over the barrier are compared to the measured data. A statistical analysis of the changes in the insertion loss measurements is performed. This study demonstrates the effects of atmospheric conditions on the insertion loss of the barrier.

4:50

1pNSb6. An innovative approach to predict sound field in a complex outdoor environment. Shahram Taherzadeh, Kai M. Li, and Keith Attenborough (Eng. Mech. Discipline, The Open Univ., Milton Keynes MK7 6AA, UK)

It is well known that the fast field program (FFP) is an accurate numerical scheme for calculating the sound field above a flat ground and in a range-independent environment. On the other hand, the boundary ele-

ment method (BEM) is most superior for a situation with complicated boundary conditions, such as the presence of noise barriers. However, the successful application of the BEM requires an accurate Green's function. This requirement somewhat limits the general applicability of the BEM only to the condition of homogeneous medium or other media with a simple sound-speed profile where analytical Green's function exists. In order to predict the propagation of sound in a complex outdoor environment, a hybrid method is developed that combines the advantages of the BEM and the FFP. In this numerical scheme, the sound field for an inhomogeneous medium is first computed by the FFP. The numerical result is then applied as the input Green's function for use in the BEM. The new hybrid scheme, fast field boundary element method (FFBEM), is used to investigate the performance of sound barriers and screens in upwind and downwind conditions. [Work supported by EPSRC of UK.]

MONDAY AFTERNOON, 22 JUNE 1998 ASPEN ROOM (S), 1:15 TO 4:45 P.M.

Session 1pPAa

Physical Acoustics and Noise: Outdoor Sound Propagation II

Daniel Juve, Cochair
Ecole Centrale de Lyon, LMFA, UMR, CNRS, BP 163, Ecully 69131, France

Henry E. Bass, Cochair
National Center for Physical Acoustics, University of Mississippi, University, Mississippi 38677

Chair's Introduction—1:15

Invited Papers

1:20

1pPAa1. Surface waves over a convex impedance ground. Kai Ming Li and Qiang Wang (Eng. Mech. Discipline, Faculty of Technol., The Open Univ., Milton Keynes, UK)

Laboratory experiments are reported for the study of surface waves above a convex impedance surface. Monopole, horizontal dipole, and vertical dipole sources are used in these experiments. The model ground surface is an open-meshed rubber mat with a hole size of 22-mm square and 12-mm depth. The rubber mat is then laid on a rigid masonite curved surface. There are generally good agreements between measurements and predictions based on the residue series theory. In an earlier theoretical study [see J. Acoust. Soc. Am. **89**, 107–114 (1991)], Raspet *et al.* have identified the contribution due to the surface wave poles. However, the present experimental results have suggested that the surface wave pole contributions not only contain the surface waves but also include the creeping waves. The experimental results also show that there are two groups of waves, which are well separated, arriving at different times. The earlier group of waves may be attributed to the creeping waves, while the later group of waves is identified as the surface waves. [Work supported by EPSRC and an Open University Studentship.]

1:40

1pPAa2. Modeling the effects of refraction and diffraction on wave propagation in a 3D, turbulent atmosphere medium. Kenneth E. Gilbert and Xiao Di (Appl. Res. Lab. and Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804)

In ocean acoustic modeling, it has been established that many nonturbulent, deep-water propagation problems can be treated by the ''$N\times$2-D'' approximation, where 2-D ''slices'' are computed along N radials. In the atmosphere, the validity of the $N\times$2-D approximation is questionable because of ever-present, multiscale turbulence. In this paper, a new approach that allows the inclusion of multiscale turbulence in practical calculations is presented. The approximations that are physically motivated by the relative contributions of refraction and diffraction in multiscale turbulence are discussed. It is shown, for example, that for horizontal atmospheric structures larger than 10's of wavelengths, horizontal diffraction is negligible, and the $N\times$2-D approximation suffices. For smaller turbulence structures, horizontal diffraction is a dominant effect and must be accounted for. The theoretical/numerical method presented allows efficient calculations for both large- and small-scale 3-D atmospheric structure. Numerical examples are given to illustrate the phenomena involved. [Work supported by DARPA and the Pennsylvania State University Applied Research Laboratory.]

2:00

1pPAa3. Scattering of acoustic impulse waveforms by turbules. Charles G. Don and Ian D. McLeod (Dept. of Phys., Monash Univ., VIC 3168, Australia)

When a short duration acoustic pulse propagates through the atmosphere its waveform can be markedly altered. Some are broadened, others have increased peak heights while occasionally ''inverted'' regions occur in the tail. In an effort to explain such changes, scattering by a turbuloid characterized by a Gaussian profile was considered. The effective size of the turbuloid, s, can be estimated from the correlation length, L, of the wind speed trace obtained as the pulse was propagating. The strength of the turbuloid, q, can be positive or negative, depending on whether the wind speed decreases or increases, while its magnitude depends on L as well as the number of turbuloids and the standard deviation of the refractive index variation. The received pulse waveform is sensitive to the values of s, q, and the position of the turbuloid relative to the source. By appropriate adjustment of the parameters, many of the observed waveforms can be reproduced; the choice of parameters giving some insight into the conditions which must have existed in the atmosphere during propagation.

2:15

1pPAa4. Sound propagation through time-dependent random media. Ph. Blanc-Benon, K. Wert, and D. Juvé (Ecole Centrale de Lyon, LMFA UMR CNRS 5509, BP 163 69131 Ecully Cedex, France, acous@selene.mecaflu.ec-lyon.fr)

In recent years, several authors have studied the impact of turbulence on sound propagation through numerical simulations. Currently, these simulations model the turbulence using a frozen turbulence hypothesis and a random Fourier modes (RFM) technique, such that the turbulent fluctuation at any point in the medium (either scalar or vectorial in nature) is calculated from the sum of a limited number of time-independent random Fourier modes. The RFM technique is quite easy to implement, and has been used to make accurate predictions of ensemble-averaged sound-pressure level for atmospheric propagation. In this paper, the extension of the RFM method to include simple time evolution of the fluctuations is presented. However, this method produces fields that do not satisfy the Navier–Stokes equations and thus cannot exhibit either the dynamics or the structures found in atmospheric boundary layer (ABL) turbulence. The desire to include ABL turbulence characteristics in simulations of atmospheric propagation has led to consideration of large eddy simulations (LES) for turbulence modeling. Here, numerical simulations obtained using both turbulence models in conjunction with a parabolic equation solver to calculate time-evolving acoustic fields, which are subsequently ensemble averaged to determine statistical characteristics, will be presented. [Work supported by CNRS/NSF and IDRIS under project 960742.]

2:30

1pPAa5. Acoustic measurements in an airport environment and the identification of the sound radiated by aircraft wake vortices. Neal Fine (Eng. Technol. Ctr., 240 Oral School Rd., Mystic, CT 06355-1208), Frank Rees, and William Von Winkle (Flight Safety Technol., Inc., New London, CT 06320)

Results are presented of an effort to measure the sound generated by aircraft wake vortices. The objective of the research is to estimate the spectral shape and source levels characteristic of the sound generated by aircraft wakes during approach and landing. Measurements were made at JFK airport in August 1997, for various aircraft, using three Bruel & Kjaer outdoor microphones. The data were recorded to tape, and spectrum analysis was performed using a B&K portable spectrum analyzer. The three microphones were placed near the beginning of runways 13L and 22L, arrayed in several configurations parallel to and perpendicular to the runway centerline. Initial spectrum analysis of the data revealed a prominant broadband ''hump'' in the one-third octave sound-pressure level spectra, centered at approximately 50–60 Hz for a B-747 landing (with a bandwidth of roughly 20–30 Hz). The timing of the hump is consistent with that expected for wake vortex sound and it is hypothesized to be generated by the wake vortices. Cross correlations between microphones arrayed parallel to the runway will test this hypothesis. Additional data, which are to be recorded in January 1998, will also aid in identifying the wake vortex sound characteristics.

2:45

1pPAa6. Spectral element analysis of sound propagation in a muffler. Wen-Hwang Lin and Munir M. Sindir (Rocketdyne Div., The Boeing Co., P.O. Box 7922, 6633 Canoga Ave., Canoga Park, CA 91309-7922)

This paper presents a novel numerical method based on the least-squares spectral element formulation to analyze sound propagation in an expansion-chamber muffler with and without a mean flow. The method solves the linearized acoustic field equations derived from the full Navier–Stokes equations. Effects of the mean flow on the acoustic wave propagation in the muffler were taken into consideration. The method solves the first-order acoustic field equations in a finite number of elements which represent the chamber muffler. Within each element the method first approximates the solutions to these equations by a series of unknown coefficients with known basis functions, forms the residual of the approximation, and then minimizes the integral of the squares of the residual with respect to the unknown coefficients. The resultant system equations were discretized by the spectral element method for spatial derivatives and by the dual time stepping for the temporal derivative. Finally, the algebraic equations were solved by a Jacobin preconditioned conjugate gradient method. Numerical results were presented for pressure contours and transmission loss of sound waves propagating at various frequencies in the muffler. These results were compared with the analytical values available, and the comparison showed very good agreement.

3:00–3:15 Break

3:15

1pPAa7. The effects of turbulent intermittency on scattering for varying degrees of saturation. David E. Norris (Grad. Prog. in Acoust., Penn State Univ., University Park, PA 16802, norris@essc.psu.edu), D. Keith Wilson (U.S. Army Res. Lab., Adelphi, MD 20783-1197), and Dennis W. Thomson (Penn State Univ., University Park, PA 16802)

Atmospheric turbulence is inherently inhomogeneous and intermittent. Periods of high activity are embedded in periods of relative calm. The local spatial and temporal changes in sound speed associated with intermittency increase the likelihood of measuring large values of scattered acoustic signals. Previous work successfully predicted the probability density functions (pdf) of fully saturated, scattered signals measured within an acoustic shadow zone [Wilson *et al.*, J. Acoust. Soc. Am. **99**, 3393–3400 (1996)]. The more general case, in which the scattered signal is partially saturated, was considered in this paper. A theory was developed through the use of the Rice–Nagakami pdf. There are two free parameters: one to describe the degree of intermittency, and a second to describe the degree of saturation. For validation purposes, outdoor propagation measurements were made over a flat, hard ground at a range of 283 m, and at frequencies from 90 to 500 Hz. Calculated and measured intensity pdf's were compared and found to be in excellent agreement. Both the intermittency and saturation increased with frequency, with the scattering at 500 Hz being almost fully saturated.

3:30

1pPAa8. The effects of atmospheric turbulence on the cross correlation between wind and travel time fluctuations. David E. Norris (Grad. Prog. in Acoust., Penn State Univ., University Park, PA 16802, norris@essc.psu.edu), Leif Kristensen, Jakob Mann (Risø Natl. Lab., Roskilde, Denmark), Dennis W. Thomson, and David C. Swanson (Penn State Univ., University Park, PA 16802)

Previous measurements over flat, hard ground showed strong coherence in received acoustic phases across frequencies from 90 to 660 Hz [D. E. Norris and D. W. Thomson, J. Acoust. Soc. Am. **101**, 3102(A) (1996)]. A recent test program was conducted at the same site in which concurrent acoustic and wind data were recorded. Similar coherences in the phase and travel time fluctuations were observed, and cross correlations between wind and travel time fluctuations were calculated. A theory was developed to predict the cross correlations by considering the effects of the two-dimensional horizontal wind spectrum on the direct source–receiver path. The theoretical parameters include the horizontal separation between the acoustic and wind sensor positions, and the bearing offset between the wind and propagation directions. Calculated and predicted cross correlations were compared and found to be in good agreement. The bearing offset was found to significantly influence both the magnitude and shape of the cross correlations.

3:45

1pPAa9. Acoustic determination of scale and structure constants in fluctuating fluid. Volker Mellert and Bernhard Schwarz-Roehr (Dept. of Phys., Univ. of Oldenburg, D-26111 Oldenburg, Germany)

The power spectrum of tubulence (e.g., fluctuating wind, **v**, and temperature T) is characterized by structure constants, which are related to scale and strength of fluctuation. The von Kármán spectrum is a common model for atmospheric turbulence, and is determined by intensity and scale parameters c_v^2 c_T^2 and L_0. The quantities are usually measured by recording time series, e.g., of **v** and T, at one location. Required is a frozen turbulent structure, being carried along the sensor by an average flow. This measuring method fails if the average movement of the fluid is insufficient. However, sound waves provide an easy and fast tool for determining the desired parameters. The complex amplitude $A(\mathbf{r})$ of the wave is changed by the refractive index fluctuations. Successive wavefronts ''see'' different spatial representations of the turbulence. Evaluating the statistics of this ensemble yields the desired parameters, provided a theory of wave propagation through turbulence is available. Based on the parabolic equation method, analytical results have been derived allowing for the calculation of the structure constants (Ostashev *et al.*, accepted for publication in J. Acoust. Soc. Am.). Spatial correlation $\langle A(\mathbf{r}_1)\cdot A(\mathbf{r}_2)\rangle$ yields c_v^2, c_T^2, and L_0 acoustically. Results are compared with ''classical'' sensor measurements in the atmospheric boundary layer.

4:00

1pPAa10. Model experiments and predictions of sound field in a downward refracting medium. Qiang Wang and Kai Ming Li (Eng. Mech. Discipline, The Open Univ., Milton Keynes, UK)

Laboratory experiments are conducted to investigate the diffraction of sound by concave surfaces for both monopole and dipole sources. The cylindrical concave surface in an otherwise homogeneous medium is used to simulate an outdoor situation corresponding to a downward refracting medium. A normal mode solution is developed to predict the sound field diffracted by the curved surface. The analysis is based on a conformal transformation in which a stricter analogy is established. Analytical expressions for dipole sources are deduced directly from that for a monopole source. Furthermore, the normal mode solutions are validated by comparison with that due to the boundary element method. Finally, the experimental results are compared with the normal mode predictions for an exponential sound-speed profile as well as a bilinear sound-speed profile. [Work supported by EPSRC and an Open University Studentship.]

4:15

1pPAa11. Inclined acoustic sounding of the atmospheric boundary layer. Igor P. Chunchuzov (Inst. of Atmospheric Phys., 3 Pyzhevskii per., Moscow 109017 Russia)

The method of inclined acoustic sounding of the atmospheric boundary layer has been proposed and realized experimentally. The method is based on the phenomena of the waveducted sound propagation along ground surface under the presence of the positive wind shear. In the experiment the wind velocity profiles and their large-scale time fluctuations were continuously controlled with the SODAR up to the height of 300 m and with the anemometers placed at the mast of 56 m height. The fluctuations of the acoustic pulse travel time with the periods in the range of 3–30 min were measured between the special pulse generator and the three receivers placed on the ground surface 2.5 km apart from each other, forming a horizontal triangle antenna. The obtained cross coherence and phase spectra between the pulse travel time and the wind velocity fluctuations show a peak of coherence with the maximum of about 0.8 to 0.9 for the period of 8 min. Such low-frequency wind velocity fluctuations near the ground surface are thought to be caused by the gravity wave modes propagating in the stable stratified troposphere with the periods close to the Brunt–Vaisalla period. The inclined acoustic sounding is proposed to be used for the acoustic tomographic monitoring of the three-dimensional wind field in the atmospheric boundary layer.

4:30

1pPAa12. Calculation of sound reduction by a screen in a turbulent atmosphere using the parabolic equation method. Jens Forssén (Dept. of Appl. Acoust., Chalmers Univ. of Technol., Gothenburg, Sweden)

Results from applying a Crank–Nicholson parabolic equation method (CNPE) are presented in situations with a thin screen on a hard ground in a turbulent atmosphere, and with the acoustic source at ground level. The results are evaluated by comparison with Daigle's model, which uses the sound scattering cross section by Tatarskii together with diffraction theory. The results show a fairly good agreement for situations where the receiver is above ground, thus indicating that both methods are applicable to the problem. When the receiver is at ground level the two methods lead to significant differences in insertion loss since only the PE method predicts that turbulence causes an increased sound level in the case without a screen. For the situations considered in the paper a turbulent atmosphere shows a significant influence on sound reduction by screens. [This work is financially supported by the Swedish Environmental Protection Agency (SNV) and the Swedish Transport and Communications Research Board (KFB).]

Session 1pPAb

Physical Acoustics: Scattering From Fluid-Loaded Objects

Raymond Lim, Chair
Coastal Systems Station, Naval Surface Warfare Center, Code R22, 6703 West Highway 98, Panama City, Florida 32407-7001

Chair's Introduction—12:55

Contributed Papers

1:00

1pPAb1. Scattering by partially buried shells. Raymond Lim (Coastal Systems Station Dahlgren Div./Naval Surface Warfare Ctr., Code R22, 6703 W. Hwy. 98, Panama City, FL 32407-7001)

An extension of a previously derived transition-matrix solution for acoustic scattering by a three-dimensional elastic obstacle partially buried at a water/sediment interface [Lim, J. Acoust. Soc. Am. **99**, 2498 (1996)] was carried out and numerically executed to allow for the obstacle to be layered. The formulation adapts an approach previously used to improve convergence for high-aspect-ratio obstacles in a homogeneous fluid host, where spherical harmonics were employed to expand the exterior surface fields of the object [A. Boström, J. Acoust. Soc. Am. **76**, 588–593 (1984)]. Although the analytical solution is expected to be exact, numerical stability is best if the obstacle layering is smooth and obeys certain convergence restrictions. In particular, for surface fields expanded in basis functions of spherical geometry, the origin-centered sphere that circumscribes the inner boundary of each homogeneous obstacle layer should never overlap the origin-centered sphere that inscribes its outer boundary. While this restriction becomes a problem for elongated shells, the acoustic response of spherical shells or shells with near-spherical cores can be calculated. Results are presented to examine the dynamics of partially buried shelled obstacles.

1:15

1pPAb2. Visualization of the energy flux in an ensonified fluid-loaded elastic sphere. Cleon E. Dean and James P. Braselton (Phys. Dept., P.O. Box 8031, Georgia Southern Univ., Statesboro, GA 30460-8031)

The energy flux in an ensonified fluid-loaded elastic sphere is visualized by its Poynting vector field in both its instantaneous real and time-averaged complex forms. The instantaneous Poynting vector field is shown in both freeze-frame and animated versions. The instantaneous Poynting vector field can be used to follow the propagation of elastodynamic waves in the elastic sphere. The complex Poynting vector field, which is a time-averaged representation of the energy flow, is shown in terms of separate real and imaginary parts. The real part of the complex Poynting vector field shows the traveling wave component of the elastodynamic wave in the ensonified sphere. The imaginary part of the complex Poynting vector field shows the standing wave part of the elastodynamic wave in the ensonified sphere. The patterns formed by the Poynting vector field, in both its instantaneous and the real and imaginary parts of the complex form, can be identified with the various modes of the ensonified sphere and used to explain the physical processes going on inside the ensonified sphere.

1:30

1pPAb3. Experimental measurements and theoretical modeling of acoustic scattering by a cube. Victor F. Humphrey, Jingdong Zhang (Dept. of Phys., Univ. of Bath, Claverton Down, Bath BA2 7AY, UK), and Paul A. Chinnery (Winfrith Technol. Ctr., Winfrith, Newburgh, Dorchester, Dorset DT2 8X5, UK)

Acoustic backscattering can be used to assess quantitatively the size and concentration of suspended sediments above the seabed. This technique depends on a knowledge of the backscattering form function for the scattering particles, which for simplicity are normally assumed to be spheres. In order to investigate the scattering from other shapes a series of measurements have been made of the acoustic backscattering behavior of both nonmetallic (glass) and metallic (steel) cubes over a wide frequency range. A laboratory system using a parametric array was used to make measurements between 20 and 200 kHz on cubes of side 20 mm, corresponding to ka values up to 10. Results are presented for a range of incidence angles involving incidence on a face, an edge, and a corner. At higher frequencies elastic resonances are observed, while at lower frequencies the average form function shows characteristics similar to those due to creeping waves on a sphere. The results are compared with theoretical predictions obtained with a boundary element model and a combined finite element/boundary element model. The agreement is observed to be very good. [Work partially supported by the MAST program of the EC, contract MAS3-CT95-0017 (TRIDISMA).]

1:45

1pPAb4. Direct and inverse obstacle scattering by shape deformation. Dilip N. Ghosh Roy, Jeremy A. Warner (SFA, Inc., 1401 McCormick Dr., Landover, MD 20785, dilip@acoustics.nrl.navy.mil), and Luise Couchman (Naval Res. Lab., Washington, DC 20735-5350)

The results of direct and inverse scattering of plane acoustic waves from penetrable and impenetrable objects are reported here. It is assumed that the scatterer boundary is a superposition of an arbitrary deformation on an underlying simple geometry. The direct problem is solved via the Padé extrapolation of the boundary variations. This results in solving only certain algebraic recursion relations and requires neither Green's function nor integral representations. The inverse problem of recovering the obstacle's shape and material parameters from the far-field scattering data is solved by Gauss–Newton minimization. The calculation of the scattered field and its Jacobian involves no more than solving a series of Helmholtz scattering problems in the same domain, namely, exterior to the simple shape instead of the iteratively updated deformed surfaces, leading thereby to substantial computational simplifications. Moreover, since no integral equation is used, the ambiguity in the solution because of the interior eigenvalues does not arise. Several two-dimensional obstacles of various shapes are inverted for their boundaries as well as their material parameters of mass density and wave number. Finally, the procedures described can be extended straightforwardly to transmission and three-dimensional problems.

2:00

1pPAb5. An alternative derivation of Howe's (1975) formula for the scattering of hydrodynamic flow into sound by solid bodies. Dan Lin (KT35, M. W. Kellogg Co., 601 Jefferson Ave., Houston, TX 77210) and Alan Powell (Univ. of Houston, Houston, TX 77204-4792)

An alternative way to derive the Green's function in the presence of a solid body is presented and applied to the case of scattering of hydrodynamic flow into sound. This method utilizes an imaging system of a point source outside a solid sphere. The ''incident potential'' and the ''scattered potential'' are obtained in terms of velocity potentials from the original source and the image sources. The total potential and the Green's function can then be easily derived. The resultant Green's function $G=\delta(t-\tau-|\mathbf{x}-\mathbf{Y}|/c)/(4\pi|\mathbf{x}-\mathbf{Y}|)$ is exactly the same as that derived by Howe [J. Fluid Mech. **67**, 597–610 (1975)] while the physical meaning is much more clear and less mathematics is involved. Here, **x** is the vector from the origin to the observation point, and **Y** is the incompressible flow vector potential in the presence of the solid body. For the case of scattering of hydrodynamic flow into sound, the above Green's function is substituted in Lighthill's solution $p(x,t)=\rho_0\int(\partial^2 u_i u_j/\partial y_i \partial y_j)G\,dV\,d\tau$. The far-field scattered acoustic pressure is then obtained in terms of flow vorticity ζ and the vector stream function Ψ_x for a unit potential flow around the solid body in the x direction: $p(x,t)=(\rho_0/4\pi xc)\partial^2/\partial t^2\int\Psi_x\cdot\zeta(t-|x|/c)dV$, as given by Dowling.

2:15

1pPAb6. Interaction of an acoustic plane wave with a circular elastic plate in water. Yannick Eudeline, Hugues Duflo, Jean-Louis Izbicki, and Jean Duclos (L.A.U.E. Pl. R. Schuman, 76610 Le Havre, France)

An experimental investigation of the interaction of an acoustic plane wave with a circular elastic plate is presented. The experimental setup consists of a duraluminum 2-mm-thick disk with different values of the radius (between 10 and 50 mm). Monostatic and bistatic experiments are performed. In the first case a transducer turns around the disk (in the perpendicular plane), sends the incident wave, and receives the diffracted signal in each observation direction. In the second experiment we use two transducers, the emitter and the receiver which turns around the circular plate. In each observation direction, we have calculated the FFT of the diffracted signal for frequencies lower than 3 MHz. We have represented the results in a two-dimensional image in which appears the spectral response of the plate. Two kinds of results are obtained: (i) Lamb waves propagation on the plate which are detected by their angular reemission of energy in water, (ii) particular frequencies are also detected close to eigenfrequencies of the circular plate in vacuum.

2:30

1pPAb7. Scattering by an elastic cylinder buried in a sedimentary fluid media. Arnaud Coatanhay and Jean-Marc Conoir (L.A.U.E., Université du Havre et CNRS (UPRESA 6068), Pl. R. Schuman, BP 4006, 76610 Le Havre, France, coatanha@iut.univ-lehavre.fr)

A new method for predicting the acoustic field scattered by an infinitely long elastic cylinder buried in a sedimentary fluid media is developed. The cylinder is parallel to the interface and the incident wave propagation direction is perpendicular to the cylinder axis. The method describes the interaction between the cylinder and the interface in modal theory. The approach is based on a generalized method of images which is originally restricted to rigid and soft plane interfaces. In this way, the acoustic field in sediment can be expressed as two cylindrical waves series whose emitting centers are symmetric with respect to the interface. The acoustic field transmitted by the plane interface is then calculated with the use of the steepest descent method. The goal is to detect buried objects by analyzing the resonances. At quite a high frequency, a good agreement is observed with regard to experiments.

2:45–3:00 Break

3:00

1pPAb8. Transient diffraction of a plane step pressure pulse by a hard sphere. H. Huang (Consultant, Bowie, MD 20715) and G. C. Gaunaurd (Naval Surface Warfare Ctr., West Bethesda, MD 20817-5700)

Transient diffraction of sound pulses by a sphere is a classical problem extensively studied in acoustics (and in EM) for nearly a century. In spite of this volume of work, a satisfactory treatment of the sound field in the neighborhood of the shadow boundary is not available, and it is still one of the unsolved problems of diffraction theory [see comments in F. Friedlander, *Sound Pulses* (1958), p. 148]. To our knowledge, no treatment of this case has yet appeared. The modal series solution for this situation is presented here. The time histories of a sufficiently large number of terms of the modal series are accurately computed; a task that is possible today due to the existence of better computers and of more sophisticated computational algorithms. The Cesàro summation method [E. Whittaker and G. Watson, *Modern Analysis* (1965), p. 155] is used to completely eradicate the Gibbs' phenomenon effect. This approach yields the diffracted field in the entire space-time domain. In particular, a complete treatment of the diffracted sound field is developed for a hard sphere insonified by a step pressure pulse, including the neighborhood of the shadow boundary where, heretofore, no analytical solution exists.

3:15

1pPAb9. Bistatic sonar cross sections and time-frequency signatures of simple targets insonified by an acoustic pulse. Guillermo C. Gaunaurd (Naval Surface Warfare Ctr., Carderock Div.—Code 684, West Bethesda, MD 20817-5700) and Hans C. Strifors (Defense Res. Establishment (FOA 6), S-17290 Stockholm, Sweden)

The area of acoustics is concurrently trying to develop the concept of impulse sonar. Using a time-frequency approach the very broadband echoes generated by a target when such devices are used can show the time evolution of the resonance features that identify the target. This will enhance the target recognition capability of the traditional sonar cross section (SCS). To this end, the bistatic scattering interaction of acoustic pulses of short duration with targets of simple shapes are studied. The targets are solid spheres and spherical shells. To model the output of an impulse sonar, the waveform incident on the target is a short, ultrawideband pulse that consists of a few sinusoidal cycles within a Blackman window. For each target the bistatic SCS is displayed as a function of frequency and bistatic angle. Analyzing the scattered waveform when the designed short pulse is incident on each target for a few selected bistatic angles gives information of how and how much scattered pulses are affected by the scattering direction, or the relative location of the transmitter and receiver. The methodology developed by us could be used to assess the performance, including the possibility of target recognition, of any impulse sonar system.

3:30

1pPAb10. New results for diffraction from an impedance wedge. Andrew N. Norris (Rutgers Univ., 58 Brett Rd., Piscataway, NJ 08854-8058, norris@jove.rutgers.edu) and Andrey V. Osipov (Inst. of Radiophysics, St. Petersburg State Univ., Uljanovskaja 1-1, Pedrodvorets 198904, Russia)

Acoustic diffraction from the tip of a wedge with impedance boundary conditions is a well covered topic, with extensive work reported in the 1950s. One of the main quantities of interest is the diffraction coefficient, D, for plane-wave incidence. Some new, simple, and useful results for the diffraction coefficient are reported. It will be demonstrated that D factors into the product of three terms: (1) the amplitude at the tip for a unit plane-wave incident from the direction of interest, (2) the same amplitude for a plane wave coming from the observation direction, and (3) an expression involving only trigonometric functions. Furthermore, the magnitudes of the tip, or edge, amplitudes (1) and (2) are very simple if the faces of the wedge are purely reactive (imaginary impedances). The simplicity

of the formula for D allows us to answer questions such as: Can the diffraction vanish if the impedances of the wedge faces are properly chosen? The answer is generally in the affirmative, at least for selected incidence and observation directions. These findings will be illustrated by several examples and applications will be discussed.

3:45

1pPAb11. Physical nature of the first circumferential wave of a Lucite cylinder immersed in water. Farid Chati, Fernand Léon, and Gérard Maze (LAUE, UPRESA CNRS 6068, Université du Havre, Pl. R. Schuman, 76610 Le Havre, France)

Acoustic scattering from a Lucite cylinder of infinite length immersed in water and excited by a normal incident plane acoustic wave is investigated at low frequencies. Numerical calculations and experiments are performed for a cylinder with a radius $a=0.01$ m. The experimental results are obtained with the method MIIR. The resonances of the first circumferential wave are easily isolated in the frequency range 100–300 kHz. A good agreement is obtained between experimental and theoretical results. First, in order to analyze the physical nature of vibration modes associated with this wave, the evolution of the resonance location and that of their corresponding widths are investigated theoretically when the fluid density is perturbated. Contrary to aluminum or steel cylinders, the computed calculus in vacuum are not in good agreement with the experimental results related to the first wave. The detection of this wave depends mainly on the outer fluid. Second, the dispersion curves of the phase velocity for this wave are considered. A noticeable difference of the phase velocity toward the high frequency is observed when the fluid density is modified. The phase velocities toward the high frequency, in the presence and in the absence of water, enable us to characterize this wave.

4:00

1pPAb12. A new model to predict echoes of large, smooth targets, with finite dimension transducers, at a low cpu cost. Jean-Luc Berton, Cecile Riffard, Christian Dominjon (CEA, CE Cadarache, DER/SSAE/LSMR, 13108 St. Paul lez Durance Cedex, France, berton@brandy.cad.cea.fr), and Dominique Dedieu (Ecole Centrale de Lille, 59651 Villeneuve d'Ascq Cedex, France)

A model called SYTRAC.L, for the prediction of echoes from large, hard reflectors, facing finite dimensions transducers, is proposed. For great reflectors, the introduction of the acoustic ray theory allows low cpu costs and more reasonable memory needs, with comparison with other methods where a numerical integration is performed on the reflector, since this integration is eliminated. In a first stage, this model does not take into account reflector edge diffraction effects, but these are being introduced with GTD approaches. The validation of the formulation and of the numerical implementation was done by comparison with analytical results in two canonical configurations. Experimental comparisons for plane and cylindrical reflectors have also been performed. This model is used to estimate the accuracy of telemetries on large structures, at distances less or greater than the far-field limit. As an example, results of a parametric study of echo characteristics from cylindrical and plane targets are described according to their distance and tilting.

4:15

1pPAb13. Prediction of echoes from small targets: Modelization and application to a telemetry problem. Cecile Riffard, Jean-Luc Berton, and Claude Estienne (CEA, CE Cadarache, DER/SSAE/LSMR, 13108 St. Paul lez Durance, Cedex, France, berton@brandy.cad.cea.fr)

A model for the prediction of echoes from finite dimension transducers and targets was developed on the basis of a formulation published by other authors. The numerical implementation in the context of telemetries on small axisymmetrical targets was validated by comparison with simulation results available in the literature. In the context of telemetry in liquid sodium, the calibration of ultrasonic celerity may be done by measurement of the difference of time of flight between two small targets at given distances. Diffraction effects exist, due to the superposition of delayed contributions of the centers and edges of both the transducer and the target. Thus, the shapes of the echoes of both targets may be different from each other. The study of the resultant error on the difference of time of flight, according to the geometry of the system, constitutes the second part of the paper. The so-called ''diffraction error'' remains acceptable for a 40-mm transducer illuminating a 30-mm-diam target at normal incidence at 4.5 MHz, even if they are separated by only 300 mm, which is clearly nearer than the far-field limit.

4:30

1pPAb14. Matrix theory of elastic resonance scattering. Myoung S. Choi and Yong M. Cheong (Non-Destructive Evaluation Team, Korea Atomic Energy Res. Inst., P.O. Box 105, Yusong, Taejon 305-600, Korea, mschoi2@kaeri.re.kr)

A fundamental matrix theory is developed for the exact isolation (in both magnitude and phase) of resonance amplitudes contained in the elastic field scattered by a penetrable target. Similarities and differences between this matrix theory and the classical resonance scattering theory [L. Flax, L. R. Dragonette, and H. Uberall, J. Acoust. Soc. Am. **63**, 723–731 (1978)] are discussed. The matrix theory is then applied to cylindrical and spherical fluid-filled cavities, and analytical expressions for the resonance amplitudes are derived. In the vicinity of resonance frequencies, the resonance amplitude expressions have forms similar to the Breit–Wigner formula. In order to check the validity of the matrix theory, numerical calculations are performed for a cylindrical water-filled cavity in an aluminum medium. Resonance spectra of the cavity are presented for some partial waves and their characteristics are also discussed.

Session 1pPAc

Physical Acoustics: Nonlinear Wave Propagation in Fluids

Steven G. Kargl, Cochair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

Patrick L. Edson, Cochair
Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02115

Chair's Introduction—1:55

Contributed Papers

2:00

1pPAc1. New symmetries and conservation laws for lossless KZK equation. Oleg A. Sapozhnikov (Dept. of Acoust., Phys. Faculty, Moscow State Univ., Moscow 119899, Russia, olegs@na.phys.msu.su) and Andrei G. Kudryavtsev (Moscow State Univ., Moscow 119899, Russia)

The Khokhlov–Zabolotskaya–Kuznetsov (KZK) equation is widely used to describe intense acoustic beams. No general analytical solution has been obtained for this nonlinear equation, so numerical methods are usually employed. Known analytical approaches are associated with various approximations. An additional powerful mathematical tool is group analysis that enables the finding of general symmetrical properties of differential equations. These symmetries help to generalize known analytical and numerical solutions, to derive new solutions, and to obtain conservation laws. In this work, results of group classification of the KZK equation with arbitrary nonlinear term are presented. It is shown that the KZK equation can be written in Euler–Lagrange form. All the classical Lie symmetries (geometric symmetries) are found and some classes of Lie–Backlund symmetries are considered. The largest number of the symmetries corresponds to the cases of quadratic and cubic nonlinearities. Besides scaling and translation groups of symmetries, there exist additional transformation groups which are not evident from the physical point of view. Examples of obtaining new solutions and deriving reduced equations are presented. New conservation laws for intense acoustic beams are obtained using the Noether theorem. [Work supported by RFBR and CRDF.]

2:15

1pPAc2. Nonlinear wave propagation from a discrete annular array: Theory. Steven G. Kargl (Appl. Phys. Lab., Univ. of Washington, 1013 N. E. 40th St., Seattle, WA 98105), Ronald A. Roy, and Patrick Edson (Boston Univ., Boston, MA 02215, ronroy@bu.edu)

Theoretical investigations of finite amplitude wave propagation from bounded beam sources has been considered by Lee and Hamilton [J. Acoust. Soc. Am. **97**, 906–917 (1995)] as well as others. These researchers have restricted their analyses to a source geometry of either an unfocused circular piston or a weakly, spherically focused circular piston. When the peak pressure does not exceed approximately 150 MPa, the Khokhlov–Zabolotskaya–Kuznetsov (KZK) nonlinear parabolic wave equation yields numerical solutions in agreement with experiments. However, little research has been reported on finite amplitude wave propagation from an array of discrete sources. Time domain simulations, based on the KZK equation, will be discussed for an eight element discrete annular array immersed in fresh and sea water. The source signal is typically a short sine wave tone burst with a center frequency of 500 kHz. The transitions from linear acoustics to nonlinear acoustics will be discussed, and comparisons between a spherically focused circular source and a properly phased array will be presented. While some data from recent experiments may be shown, the scope of this talk will be restricted to the theoretical treatment and the numerical simulations for the discrete annular array. [Work supported by ONR.]

2:30

1pPAc3. Nonlinear wave propagation from a discrete annular array: Experiments. Patrick Edson, Ronald A. Roy (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, ronroy@bu.edu), and Steven G. Kargl (Univ. of Washington, Seattle, WA 98105)

To date, efforts to model nonlinear wave propagation from a bounded beam source have chiefly been concerned with unfocused and weakly focused circular piston sources [Lee and Hamilton, J. Acoust. Soc. Am. **97**, 906–917 (1995)]. When the peak pressure does not exceed approximately 150 MPa, the Khokhlov–Zabolotskaya–Kuznetsov (KZK) nonlinear parabolic wave equation yields numerical solutions in agreement with experiments. However, little research has been reported on nonlinear wave propagation from an array of discrete sources. Theoretical and experimental results for an eight-element annular array will be presented. The source signal is typically a ten-cycle sine wave tone burst with a center frequency of 500 kHz. The outer diameter of the array is 0.14 m, and the surface areas of the elements are equal. While theoretical predictions will be shown, the experimental aspects of this research will be emphasized. [Work supported by ONR.]

2:45

1pPAc4. On the asymmetric waveform of shock waves in high-amplitude focused sound fields. Halvor Hobaek (Dept. of Phys., Univ. of Bergen, Allegt. 55, N-5007 Bergen, Norway)

The asymmetric shape of shock waves generated in high-amplitude sound fields, where the peak positive pressure tends to be much higher in amplitude than the peak negative pressure, is well known. Numerical simulations based on the KZK equation show excellent agreement with experimental observations of the field from a focusing circular sound source with a pronounced asymmetry. This equation accounts for diffraction, absorption, and nonlinear effects. Thus the asymmetry must be due to diffraction, because the plane wave limit of the KZK equation (Burgers' equation), accounting for nonlinearity and absorption only, predicts symmetric waves (if the initial sound is sinusoidal). However, experimental

results indicate that the shock front is strongly connected with the contribution from the central part of the source, while diffraction gives rise to the edge wave which propagates independently of the shock, these two signals can easily be distinguished. Spectral analysis of the measured waveforms on the axis show that each harmonic component has an almost constant phase offset (a symmetric *N* wave has none). The edge wave can have negligible influence on this phase shift. Thus the question is, how does diffraction cause asymmetry? Results on the analysis of this problem will be presented.

3:00

1pPAc5. The parametric efficiency of a bifrequency focusing Gaussian nonlinear source. M. Serhat Özyar (Dept. of Elec. and Electron. Eng., Middle East Tech. Univ., Ankara, 06531, Turkey, ozyar@rorqual.cc.metu.edu.tr)

The field distribution from a bifrequency focusing nonlinear source was calculated. The simple formulation of Novikov *et al.* [''Nonlinear underwater acoustics,'' ASA, 86–90 (1987)] for the Gaussian wavefront was adopted for simulations. Radius of aperture or beamwidth (a), radius of curvature (D), mean primary frequency ($f_m=(f_1+f_2)/2$), and difference-frequency ($f_d=f_1-f_2$) are essential parameters describing the performance of a parametric focusing source. The geometrical ratio D/a and the down–shift were varied between 5 and 200 by changing f_d and the resulting difference–frequency pressure wave and primary intensity distributions were plotted against the propagation axis for $D/a=2$. A similar calculation was repeated for $f_d=50$ kHz by varying f_m. Separate effects of D and a were also investigated where the down–shift ratio was fixed to 20. Results were plotted for D/a: 1–4. Finally, an efficiency was defined as $\eta=$(parametric pressure)/(primary intensity). It was presented that η made a peak in the prefocal region and reached to a minimum value at the geometrical focus. The variation of η in the postfocal region was also discussed.

3:15

1pPAc6. Generation of harmonics in the near field of finite amplitude sources. Eugeniusz Kozaczka and Grayna Grelowska (Naval Acad., ul. Smidowicza 71, 81-919 Gdynia, Poland, ekoz@amw.gdynia.pl)

It is well-known that the fine structure of the near field brings difficulties to both the theoretical and the experimental investigations. However, in many applications the phenomena occurring in the near field are the most important ones. It is caused by such a measurement setup arrangement and the choice of the parameters of used transducers that the measured or diagnosed objects are situated in the near field. The paper presents the results of the experimental investigations of the finite amplitude wave radiated by mono- and biharmonic sources of different radiating apertures. The measurements were carried out in the near field of the transducer using the high-precision device which controlled the movement of the receiver. The pressure was measured using a 1-mm-diam PVdF needle hydrophone. The results of investigations will be presented as the 3-D distribution of a primary wave and its harmonics. The distributions of chosen harmonic will be shown in planes perpendicular to the acoustic axis of the transducer or along this axis. The experimental results were compared with the numerical ones.

3:30–3:45 Break

3:45

1pPAc7. Raman–Nath diffraction caused at the focal point of a concave type transducer. Y. Watanabe, H. Miyaki, and K. Fujita (Dept. of Electron., Doshisha Univ., Kyotanabe 610-0321, Kyoto, Japan, kwatanab@mail.doshisha.ac.jp)

To estimate the intensity of sound pressure in the vicinity of the focal point of a concave type transducer, Raman–Nath diffraction caused at the focal point of a concave type transducer is theoretically and experimentally discussed. Sound field, near the focal point, including the acoustic nonlinearity, is theoretically calculated by using the KZK equation. Light intensity of zeroth order Raman–Nath diffraction is also calculated as a function of the sound pressure at the focal point. In the high intensive sound region, it is found that the acoustic nonlinearity makes the light intensity curve disappear at the local peaks. The He–Ne laser whose beam spot diameter is 1.2 mm is supplied as a light source. Light intensities of the zeroth-order Raman–Nath diffraction caused at the focal point are observed by changing the driving voltage of the transducer. The experimentally obtained graph shows the relatively same pattern compared with the theoretically calculated one. The sound pressure at the focal point is estimated by fitting those two results. It is confirmed that the sound pressure estimated by the optical method shows good agreements with the estimated one by the hydrophone measurements.

4:00

1pPAc8. Parameter sensitivity in nonlinear and dissipative time-reversed acoustics. Claes M. Hedberg (Mech. Eng., Univ. of Karlskrona/Ronneby, Karlskrona, Sweden, claes.hedberg@hk-r.se)

Time-reversed nonlinear acoustics has shown that a propagated signal retains information about the original signal after shock formation and even after merging of shocks. This information is naturally very dependent on the accuracy of the measured or calculated data. Presented here is a more detailed analysis of the estimation of the dimensionless parameters in the problem: the original amplitude, the distance traveled, and the dissipation over nonlinearity ratio. The robustness of alternative methods and their sensitivity to noise and incomplete data are investigated.

4:15

1pPAc9. ''Induced transparency'' of an acoustical waveguide containing a liquid with nonlinear viscosity. Oleg M. Zozulya (N. Andreyev Acoust. Inst., 4 Shvernik str., Moscow 117036, Russia)

Sound propagation in an acoustical waveguide containing a liquid that depends on a shear rate $\dot{\gamma}$ viscosity is considered. A number of rheological models adequately representing such dependence in a wide range of $\dot{\gamma}$ variation are well known (power law, Carreau model for stuctural viscosity, etc.). In most of the models viscosity decreases with the increase of $\dot{\gamma}$. It can result from the destruction of some internal microstructures in the liquid and their self-reconstruction with a reduction of load (as, e.g., in a sol–gel transition). For such liquids the effect of ''induced transparency'' for intense sound waves is expected. In this work the effect is determined for harmonic waves, propagating in an acoustical waveguide with cross–section dimensions that are small in comparison with the sound wavelength. The dependence of viscosity on $\dot{\gamma}$ is taken into account by retention of the first nonlinear term in the expansion of shear tension σ_τ in shear rate $\dot{\gamma}$: $\sigma_\tau=\eta\dot{\gamma}-\xi\dot{\gamma}^3$, where η and ξ are linear and nonlinear parameters of viscosity, respectively. Sound attenuation in acoustical waveguides is due to the partial transmission of acoustical energy to the viscous waves being generated in the vicinity of waveguide walls. It is shown that the nonlinear viscosity gives rise to the effective viscous wave self-interaction, which results in the reduction of the energy dissipation. The application of the effect to the measurements of the nonlinear viscosity parameter is considered and the importance of this parameter for the evaluation of the oil quality is discussed. Work was partially supported by the Russian Foundation of Basic Research (project 97-08-17789).]

4:30

1pPAc10. Application of the three-wave interaction to the problem of acoustical monitoring of complicated media. Samuil A. Rybak and Oleg M. Zozulya (N. N. Andreyev Acoust. Inst., 4 Shvernik str., Moscow 117036, Russia)

In a nonlinear medium with resonant dispersion the three-wave interaction is considered. The medium consists of oscillators distributed in eigenfrequency range. The quadratic nonlinearity is taken into account in surrounding compressible medium. The three-wave interaction in such medium is sensitive to the changing of the medium parameters. Space and time scales of energy transfer among the interacting waves are determined. The eigenfrequency distribution leads to the effective Landau attenuation of wave amplitudes and results in finite time interaction among waves. On the other hand, this finite time permits the same efficiency of interaction not only for waves exactly satisfying the resonant conditions but also for wave packets. In the particular case of Lorentz distribution the influence of the distribution parameters on the energy transfer is determined. Some exact solutions expressed in special functions describing the three-wave interaction are obtained. [Work was partially supported by Russian Foundation of Basic Research.]

4:45

1pPAc11. Application of nonlinear interaction of sound waves in sound reproduction. Weiguo Dong, Qunli Wu, and Shih Fu Ling (School of Mech. and Production Eng., Nanyang Technolog. Univ., Nanyang Ave., Singapore 639798)

Audio sound reproduction using the nonlinear interaction of sound waves is investigated. Special emphasis is given to the sum and difference frequencies generated by the nonlinear interaction of two sound waves in a duct. The numerical and experimental results for scattered sum and difference frequency sound are presented and compared. The relationship between the difference frequency and the two primary sound waves are studied. It is found that the flat frequency response of the difference frequency can be achieved by adjusting the amplitudes of two primary waves. Finally, the sound reproduction using a single bi-frequency source is also investigated. The results for sum and difference frequency generation of this case are compared with those of the dual sources.

MONDAY AFTERNOON, 22 JUNE 1998 GRAND CRESCENT (W), 2:15 TO 5:30 P.M.

Session 1pPP

Psychological and Physiological Acoustics: Peripheral Processes, Models, and Loudness

Mary A. Cheatham, Cochair
Audiology and Hearing Science 2-240, Northwestern University, 2299 North Campus Drive, Evanston, Illinois 60208

C. Craig Formby, Cochair
Division of Otolaryngology-HNS, University of Maryland School of Medicine, 419 West Redwood Street, Suite 360, Baltimore, Maryland 21201

Contributed Papers

2:15

1pPP1. Frequency-dependent saturation in inner hair cell responses. M. A. Cheatham and P. Dallos (Audiol. and Hearing Sci., The Hugh Knowles Ctr., Frances Searle Bldg., Northwestern Univ., Evanston, IL 60208)

Inner hair cell (IHC) recordings are made in guinea pig from cochlear turns 2, 3, and 4 where best frequencies (BF) are approximately 4000, 1000, and 250 Hz, respectively. Input–output functions for the ac receptor potential are measured for inputs at, below, and above the BF of each recording location. Although all responses saturate at high intensities, the response magnitude at which saturation occurs depends on stimulus frequency. For example, in turn 2 more linear responses are observed below BF while in turn 4 a different pattern emerges. At this apical recording site, which responds preferentially to very low frequencies, IHCs exhibit more linear behavior above BF. In other words, response magnitude at saturation exceeds that measured for inputs at and below the BF of the cell. The reasons for this dependence on BF are evaluated by comparing input–output functions for ac responses recorded intracellularly from IHCs and extracellularly from the organ of Corti fluid space. [Work supported by Grant No. 5 R01 DC00089, National Institute on Deafness and Other Communication Disorders, National Institutes of Health.]

2:30

1pPP2. Linear and nonlinear modeling techniques for cochlear mechanics. Karl Grosh (Dept. of Mech. Eng. and Appl. Mech., Univ. of Michigan, Ann Arbor, MI 48109-2125), Anand Parthasrathi (Univ. of Michigan, Ann Arbor, MI 48109), and Alfred Nuttall (Oregon Health Sci. Univ., Portland, OR)

New experimental techniques have made possible the measurement of the cochlear response to electrical, mechanical, and acoustical stimuli as never before possible. By controlling the various inputs to the cochlea, these measurements hold the promise for delivering sufficient information for developing morphologically based cochlear mechanics models capable of representing the amplification and nonlinear dynamics seen in the data. Nonlinear dynamics takes the form, for example, of harmonic generation and compression. Activity may take the form of outer hair cell mediated force generation. These new data in part motivate the need for new modeling techniques to accurately and efficiently simulate the response arising from the morphology and possible activity of the cochlea. In this paper, hybrid methods are presented for three-dimensional cochlear models which generalize to include activity and nonlinearity in a straightforward fashion. These methods are based on modal–finite element and boundary element–finite element hybrid methods. This approach enables the inclusion of as much structural inhomogeneity as needed for the accurate simu-

lation of the underlying response. The framework for the inclusion of activity, nonlinearity, and inhomogeneity (e.g., radial variation in morphology) is given. Results for time harmonic forcing (e.g., stapes motion) are presented.

2:45

1pPP3. Rejecting a cochlear-partition model with constant mass. Timothy A. Wilson (Dept. of Elec. Eng., Univ. of Memphis, Memphis, TN 38152)

The question of whether a simple low-order, linear, passive model of cochlear-partition dynamics can account for ''global'' (overall stiffness change, range of best frequencies and best places) and ''local'' (peakiness of frequency- and place-selectivity, amount of phase-angle change around the best place or frequency) attributes of cochlear response to tones in Békésy's observations in cadavers remains unresolved. With answering that question the goal, a simple stiffness-damping-mass model of cochlear-partition dynamics with constant mass and exponentially varying stiffness and damping, was investigated computationally using a box-cochlea model of cochlear hydrodynamics. Constrained only by Békésy's reported change in stiffness by about two orders of magnitude, the model's global and local response attributes were determined over a reasonably exhaustive set of mechanical parameter values. It was found that the only combinations of mechanical parameters resulting in reasonable cochlear maps had magnitude-response shapes neither qualitatively resembling nor quantitatively consistent with observations by Békésy or anyone else, postmortem or *in vivo*.

3:00

1pPP4. Determination of roots of eikonal equation for WKB solution in cochlear models. Anand A. Parthasarathi (Dept. of Biomed. Engr., Univ. of Michigan, Ann Arbor, MI 48109, anand@engin.umich.edu), Karl Grosh (Univ. of Michigan, Ann Arbor, MI 48109-2125), and Alfred L. Nuttall (Oregon Health Sci. Univ., Portland, OR 97201-3098)

The WKB-theory-based asymptotic methods have been used to obtain solutions to analytical cochlear models [E. de Boer, Phys. Rep. **105**(3), 141–226]. The local wave number for this solution is obtained as the root of the eikonal equation or the dispersion equation. For a three-dimensional rectangular model, multiple complex roots exist, which could give rise to additional waves on the basilar membrane. However, previous results were based on a single root locus determined using initial estimates obtained from the long-wave approximation. In this study, the winding-number integral technique [P. R. Brazier-Smith and J. F. M. Scott, J. Sound Vib. **145**(3), 503–510] is used to obtain a definitive identification of the multiple roots and their dependence on the number of acoustic cross modes. *A priori* information about the characteristics of the dispersion relationship was used to obtain an initial estimate of the search region. Thereafter the region was refined and the roots located using moments of the winding integral and secant iterations. It was found that higher-order roots are strongly evanescent away from the resonance location. Also, as the number of acoustic cross modes of the duct increases, more roots appear. The influence of these roots on the solution will be discussed.

3:15

1pPP5. A signal detection theoretic analysis of the Carney model of auditory processing. Leslie M. Collins and Lisa C. Gresham (Dept. of Elec. and Comput. Eng., Box 90291, Duke Univ., Durham, NC 27708-0291, lcollins@ee.duke.edu)

Computational models of human auditory processing provide a platform on which to apply signal detection theory (SDT) directly to the ''output'' of various peripheral processes. Theoretical performance predicted by standard SDT analyses which do not incorporate such models often substantially exceeds measured performance. These discrepancies are commonly compensated for by adding a physiologically unspecified source of ''internal noise.'' By applying SDT to auditory models it is possible to evaluate whether the performance predicted at various stages of the auditory system provides a more accurate estimate of human performance than standard approaches which incorporate internal noise. Our previous work indicates that integrating SDT with a deterministic model of the human auditory system affords theoretical performance predictions that match experimental measurements more closely than traditional approaches. In the work presented here, SDT is integrated with Carney's model of auditory processing [Carney, J. Acoust. Soc. Am. **93** (1993)]. This model incorporates physiologically based models of the basilar membrane and inner hair cell mechanisms, as well as a nonhomogeneous Poisson process to model nerve discharge times. The performance predicted by this model is presented, and its relationship to previous results and to modeling assumptions is discussed. [Work supported by NIH.]

3:30

1pPP6. Signal detection theoretic analysis of auditory processing: Phase uncertainty and multiple channel integration. Lisa C. Gresham and Leslie M. Collins (Dept. of Elec. and Comput. Eng., Box 90291, Duke Univ., Durham, NC 27708-0291, lcollins@ee.duke.edu)

Traditionally, psychophysical data has been predicted either by constructing models of the peripheral auditory system or by using signal detection theory. The resulting theoretical detection performance is often substantially greater than experimental results and has been explained by adding ''internal noise'' to the system in a somewhat *ad hoc* manner. An integrated approach combining both a physiologically based auditory model and classical signal detection theory has been shown to produce more accurate predictions of experimental data. The discrepancy that still exists may be due, in part, to the two unrealistic assumptions that the system is deterministic and that a single channel with the highest signal-to-noise ratio determines detection performance. The work presented here investigates the impact of phase uncertainty and multiple channel integration on predicted detection performance. More accurate predictions of experimental data are obtained, particularly after the first two stages of the auditory model, when phase uncertainty is added to the auditory input signal and the resulting output signals are analyzed using the integrated approach. Integrating information across multiple channels increases the predicted performance, albeit not substantially, suggesting that the single-channel approximation is not a significant source of discrepancy between theoretical and experimental results. [Work supported by NIH.]

3:45–4:00 Break

4:00

1pPP7. A descriptive model of under- and over-shoot patterns in the temporal masking function of a narrow-band noise. C. Formby, J. C. Rutledge (Div. of Otolaryngol.—HNS, Univ. of Maryland School of Medicine, Baltimore, MD 21201), M. G. Heinz (Harvard–MIT, Cambridge, MA 02139), L. P. Sherlock, and I. V. Alexsandrovsky (Univ. of Maryland School of Medicine, Balitmore, MD 21201)

Complex temporal patterns of under- and over-shoot (i.e., enhanced and diminished detection) at onset and offset of a 200-ms, 400-Hz-wide noise masker (2500 Hz center frequency) have been reported for a temporal masking paradigm [Formby *et al.*., J. Acoust. Soc. Am. **102**, 3161(A) (1997)]. The paradigm measured detection of a 5-ms sinusoidal signal (1900 or 2100 Hz) at 43 temporal positions preceding, during, and following the masker. Efforts to model under- and over-shoot patterns in the temporal masking functions will be considered. The modeling strategy is similar to that used in the study of vestibular adaptation [Formby *et al.*, 21st Meeting of Association for Research in Otolaryngology, 1988], and allows for quantification of the relative amplitudes, time constants, and delays of excitatory and adaptive processes. Linear combinations of exponential functions are used to represent the effects of excitatory masking and adaptation. Suppression and adaptation of suppression may also play roles in the masking process. These effects can be evaluated in our scheme by including corresponding exponential functions to represent the processes. The least-squares fitting model will be described and parameter estimates for the model will be presented for a range of masker levels. [Work supported by ONR.]

4:15

1pPP8. Detection of an increment in a single-frequency component of a noise background as a function of increment frequency and duration. C. Formby (Div. of Otolaryngol.—HNS, Univ. of Maryland School of Medicine, Baltimore, MD 21201), M. G. Heinz (Harvard–MIT, Cambridge, MA 02139), and I. V. Aleksandrovsky (Univ. of Maryland School of Medicine, Baltimore, MD 21201)

Increment detection thresholds for a single-frequency component were measured over a range of increment frequencies (250–6000 Hz) and durations ($T=10$–500 ms) in a 500-ms, 6000-Hz-wide noise. The noise spectrum level was either constant ($N_0=40$ dB) or randomly varied (between $N_0=20$ and 40 dB) from interval to interval. Preliminary results reveal performance for each increment frequency to be similar for the fixed- and random-level conditions. Detection thresholds for $T=10$ ms are largely independent of frequency and correspond to S-N_0 values of 17–18 dB, whereas corresponding S-N_0 values measured for long duration increments increase at a rate of 3 dB/oct for increment frequencies between 500 and 4000 Hz. The latter finding is consistent with changes in the critical ratio. This ratio is assumed to be based on an absolute energy cue, which is inconsistent in the random-level condition. Temporal integration (TI) functions fitted to the normalized detection data reveal less integration with increasing increment frequency from 500 to 4000 Hz. When the detection results are corrected for the apparent ''critical ratio'' effect, the TI functions are largely independent of increment frequency. [Work supported by ONR.]

4:30

1pPP9. A unified theory of two-tone suppression and upward-spread of masking. Jont Allen and Deep Sen (AT&T Labs-Res., 180 Park Ave., Florham Park, NJ 07932)

Low-frequency two-tone suppressors (TTS) [Sachs and Abbas], and the upward spread of masking, have traditionally been treated as separate phenomena. We postulate that both these well-known response measures are due to nonlinear (NL) cochlear compression. Furthermore, we attribute this NL compression to a changing basilar membrane (BM) stiffness, which varies inversely with hair cell stimulus. A 65 dB SL low-frequency tonal masker/suppressor excites a broad basal region of the cochlea uniformly, raising the threshold of high-frequency probes throughout this portion of the cochlea. Sufficiently intense low-frequency maskers/suppressors cause a reduction in the stiffness in the base of the BM, resulting in the probe characteristic frequency shifting towards the base. This basal NL shift of the BM response, combined with high-pass filtering by the tectorial membrane, results in an attenuated basal hair-cell response. This change in gain accounts for the expansive slope of the upward spread of masking, having a maximum slope of 2.4 dB/dB. This slope implies a compression proportional to pressure to the $0.8\approx 2/2.4$ power. The same model explains TTS if we treat the CF tone as the probe and the suppressing tone as the masker. The rate of growth of suppression, in this case, is also known to reach a maximum of 2.4 dB/dB [Delgutte].

4:45

1pPP10. Equal-loudness relations at high frequencies: Implications for loudness growth. Rhona Hellman (Auditory Percept. Lab., Dept. of Psych., Northeastern Univ., Boston, MA 02115, hellman@neu.edu), Hisashi Takeshima (Sendai Natl. College of Technol., Sendai, 989-31 Japan), Yoiti Suzuki, Kenji Ozawa, Takanori Yamaguchi (Tohoku Univ., Sendai, 980-77 Japan), Yusaku Sasaki (Sendai Natl. College of Technol., Sendai, 989-31 Japan), and Toshio Sone (Tohoku Univ., Sendai, 980-77 Japan)

Equal-loudness relations spanning a wide range of sensation levels (SL) were measured by the method of adjustment for 12 pairs of frequencies from 1 to 16 kHz. Sensation levels were based on individual thresholds measured prior to the equal-loudness determinations by an adaptive 2IFC procedure. To test for transitivity, three reference frequencies at 1, 3.15, and 5 kHz were used. Below 8 kHz, the matching results were well described by a linear function with a slope of 1.0; at higher frequencies, the functions became increasingly curvilinear. Near threshold and at moderate SLs the loudness of a tone at 12.5 kHz and higher increased more rapidly than the loudness of the lower-frequency reference tone; at high SLs it increased more slowly. The dependence of the slope of the matching function on level was most pronounced at 16 kHz. Moreover, nearly perfect transitivity was observed. The reduction in slope at high SLs is consistent with recent evidence that severely restricting the spread of excitation to high frequencies decreases the tone's rate of loudness growth [R. P. Hellman, J. Acoust. Soc. Am. **96**, 2655–2663 (1994)]. [Work supported by the Japanese Ministry of Education, Science, Sports, and Culture.]

5:00

1pPP11. Global and continuous judgments of sounds with time-varying intensity: Cross-modal matching with a proprioceptive input device. Patrick Susini (Inst. de Recherche et de Coordination Acoustique/Musique (IRCAM), 1 Place Igor Stravinsky, F-75004 Paris, France) and Stephen McAdams (IRCAM, F-75004 Paris, France)

Conventional psychophysical methods such as paired comparisons and psychophysical scaling present problems in the case of real-time perceptual evaluation of long-duration nonstationary sounds. In order to obtain loudness profiles of such sounds and to determine their relation to global loudness judgments, a continuous judgment method was employed using cross-modal matching of loudness with muscular force in a proprioceptive input device. This device was perceptually calibrated with 1-kHz stationary pure tones using S. S. Stevens' cross-modal matching paradigm. The loudness scale is equivalent to that obtained with ratio and magnitude production methods. The device was then tested on pure tones varying in intensity from 10 to 53 s. In order to observe the different memory processes involved in the global evaluation, several types of contours were used with from one to three peaks. The global judgment was influenced both by the duration and the level of the last contour peaks. The continuous judgment task also appeared to entail an auditory and gestual learning of the signal contour.

5:15

1pPP12. On the effects of a subharmonic masker on the loudness of a pure tone. Kenji Ozawa (Res. Inst. of Elec. Commun., Tohoku Univ., 2-1-1 Katahira, Aoba-ku, Sendai, 980-8577 Japan, ken@sone.riec.tohoku.ac.jp), William S. Hellman (Boston Univ., Boston, MA 02215), Yoshinori Inoue, Yôiti Suzuki, and Toshio Sone (Tohoku Univ., 2-1-1 Katahira, Aoba-ku, Sendai, 980-8577 Japan)

Loudness matching functions between a 1-kHz tone in quiet and in the presence of a 70 dB(SPL) 0.5-kHz subharmonic masker were studied to determine the alterations in loudness due to intensity shifts at the stimulus frequency arising from the generation of the second harmonic of the masker. Such effects are understood to be due to a quadratic nonlinearity occurring in the inner ear transduction process. In the experiment, masked thresholds were measured at 30° phase angle intervals of the stimulus. Matching functions were obtained for phase conditions corresponding to minimum and maximum thresholds. Both the shape of the matching functions and their relative displacement from each other were consistent with the predictions of the vector summation model [Clack *et al.*, J. Acoust. Soc. Am. **52**, 536–541 (1972); Schubert, J. Acoust. Soc. Am. **45**, 790–791 (1969)]. Assuming that the threshold shifts were due to constructive and destructive interference between the 1-kHz tone and the second harmonic of the masker, a single effective matching function was constructed by means of the vector summation model. This calculation demonstrated the ability of loudness matching to probe nonlinear cochlear mechanisms. Matching functions generated from several proposed loudness functions were compared with the data.

Session 1pSA

Structural Acoustics and Vibration: Radiation and Scattering From Elastic Structures

Phillip L. Marston, Cochair
Department of Physics, Washington State University, Pullman, Washington 99164-2814

Alain C. Berry, Cochair
GAUS, Department of Mechanical Engineering, Universite de Sherbrooke, Sherbrooke, QC J1K 2R1, Canada

Contributed Papers

1:00

1pSA1. Sound scattering by a fluid-loaded finite cylindrical shell. Aleksander Klauson, Jaan Metsaveer (Dept. of Mech., Tallinn Tech. Univ., Ehitajate tee 5, Tallinn EE-0026, Estonia), Nicolas Touraine, Dominique Decultot, and Gerard Maze (LAUE URA CNRS 1373, Univ. of Le Havre, Pl. R. Schuman, 76610 Le Havre, France)

Let a thin-walled finite cylindrical shell with the hemispherical end caps be immersed in a fluid medium. The shell contains structural inhomogeneities like radial solders and internal ribs and is excited by an axisymmetric pressure load. The paper deals with the numerical and experimental investigation of the sound-pressure field scattered by the shell and its internals. The numerical solution of the problem is based on the finite element method (FEM) with the use of the thin shell theory elements. The Green's function of the structure is obtained by the FEM and is further used in the boundary element solution scheme using both physical and modal coordinates. Flexural and membrane type eigenmodes of the fluid-loaded structure can be determined and identified. A time-domain solution to the problem obtained by the inverse Fourier transform gives echo-signals of the sound scattered by the structure. Individual pulses in the echo-signal help to localize structural inhomogeneities and to appreciate their importance in the scattered sound field. The membrane type vibrations are shown to be substantially modified by the internal solders and substructures. Experimental investigation is performed for a steel target immersed in a water tank and the numerical results correlate well with the experiment.

1:15

1pSA2. Meridional and helical ray contributions to backscattering by tilted cylindrical shells: High-frequency tone burst and wide bandwidth measurements and interpretation. Scot F. Morse and Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814, morse@mail.wsu.edu)

When viewed in the frequency-angle domain, the backscattering by truncated thick cylindrical shells in water exhibit various ridges of high backscatter. These enhancements are associated with leaky waves excited on the shell and can be described in terms of meridional and helical rays. The enhancements are not present for a rigid scatterer of the same shape and may be important for interpreting high-frequency sonar echoes. Previous work [Morse *et al.*, J. Acoust. Soc. Am. (to be published)] has shown that these enhancements are not limited to near broadside incidence but can extend to end-on incidence in the coincidence frequency region. Results of broadband and tone burst measurements of these enhancements are presented. High-frequency tone bursts may be used to separate the meridional ray enhancement from the generally weaker helical ray enhancement. The general angular width of the meridional ray enhancement may be explained using an approximate ray theory [Marston, J. Acoust. Soc. Am. **102**, 358–369 (1997)]. The amplitude of the backscattered meridional ray depends on the leaky wave reflection coefficient at the end of the cylinder which appears to be affected by mode conversion near mode thresholds. [Work sponsored by the Office of Naval Research.]

1:30

1pSA3. A method of target characterization based on the analysis of Scholte–Stoneley and Lamb-type waves scattered by submerged fluid-filled thin-walled shells. Alessandra Tesei, Warren L. J. Fox, Alain Maguer, and Arne Løvik (SACLANT Undersea Res. Ctr., La Spezia, Italy)

The scattering inverse problem is addressed as a powerful aid to target classification. Broadband sonar signals are used to insonify submerged elastic objects, and signal processing techniques are applied to the backscattered response. Cylindrical fluid-filled thin-walled shells are considered. The ka range is (2,50). One characteristic of target echoes playing a major role at low-intermediate frequencies is the resonance contribution, which depends on the target's elastic properties. Here resonance information is extracted by algorithms based on autoregressive (AR) spectral estimation techniques. Assuming a target as a linear acoustic system, AR modeling is used for estimating its poles and, among them, resonance modes which are identified on the basis of resonance scattering theory. Previous investigations on scattering by thin plates and shells demonstrated the existence of fluid-born Scholte–Stoneley waves and shell-born Lamb-type waves. Here, from the study of these waves and internal reflections, a set of equations are formulated which correlate resonance characteristics with geometrical and geophysical target parameters. The algorithms are shown to succeed in estimating parameters such as the shell outer radius, material, and wall thickness. Simulated and real data backscattered by cylinders at broadside are considered. Analyzing simple-shaped objects is seen as preliminary for characterizing more complicated targets.

1:45

1pSA4. Scattering from fluid-loaded elastic targets near penetrable sediments. Garner Bishop and Judy Smith (Naval Undersea Warfare Ctr., Newport, RI 02841)

A T-matrix formalism developed for plane-wave scattering from a fluid-loaded elastic spherical shell near a rough fluid elastic interface [J. Acoust. Soc. Am. **101**, 767–788 (1997)] is generalized to include scattering from targets for which a T-matrix is known and constructed on a spherical wave basis, and from the interface of penetrable sediments for which a T-matrix is known and constructed on a Weyl plane-wave basis. A variety of numerical results are obtained for plane-wave scattering from elastic spheres and elastic spherical shells, as well as a finite elastic cylinder with hemispherical end caps near rigid, soft, fluid, elastic, poroelastic, and layered fluid sediments with planar boundaries that demonstrate some of the effects' target-boundary scattering on resonant and nonresonant scattering.

2:00

1pSA5. Convolution formulation for high-frequency leaky wave scattering enhancements for solids and shells with truncations: Evaluation of the surface integral and experimental and computational tests. Philip L. Marston, Karen Gipson, and Scot F. Morse (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814)

Leaky waves reflected at truncations of elastic objects in water can cause high-frequency backscattering enhancements for ranges of orientation angles [Morse *et al.*, J. Acoust. Soc. Am. (accepted for publication)]. The backscattering can be much greater than for rigid objects having the same shape, size, and orientation. Such enhancements are relevant to the interpretation of sonar images and are usually associated with relatively flat portions of the outgoing wavefront. The typical leaky wave attenuation length can be much less than the largest dimension of the scatterer, rendering global resonances to be unimportant. One approach to approximating these enhancements is to evaluate a two-dimensional convolution integral for the outgoing leaky wave amplitude on the surface [P. L. Marston, J. Acoust. Soc. Am. **102**, 358–369 and 1628–1638 (1997)]. In some cases this integral may be approximated as a one-dimensional integral. It is also necessary to approximate leaky wave reflection coefficients at truncations and to evaluate a propagation integral for the far-field amplitude. Some mechanisms to be discussed give large enhancements from reflected Rayleigh waves on tilted cylinders and cubes and from Lamb waves on cylinders. [Work supported by the Office of Naval Research.]

2:15

1pSA6. Vibroacoustical modeling and modal analysis of a double wall panel. Stan Josef Pietrzko (EMPA Swiss Federal Labs, Ueberlandstrasse 129, CH 8600 Duebendorf, Switzerland)

In this paper results are reported of analytical and experimental vibroacoustical modal analysis of a double wall panel. Both mechanical vibrations of panel walls as well as the modal behavior of acoustical field in cavity were identified. Special interest was focused on identification and modeling of interaction and coupling between vibratory behavior of the walls and acoustical field in the cavity and their influence on drop of the transmission loss of the panel, as measured in the laboratory. The identification was carried out using a set of combined vibratory and acoustical frequency response functions measurements as response due to random noise force and controlled volume velocity source. The identified global complex modes of the panel and an estimated model of the panel employing general parameters were of high quality. It was confirmed by very good orthogonality of the identified vibratory and acoustical mode shapes and high correlation values between the measured and synthesized frequency response functions. This result provides a robust basis for the numerical study of the transmission behavior of the panel in the low-frequency range. Application of this result on optimum placement of force actuators to increase of transmission loss over a lower-frequency range by means of active noise and vibration control will be presented.

2:30

1pSA7. Angular distribution of incident sound energy for estimating the sound transmission loss of multilayered panels. Hyun-Ju Kang, Hyun-Sil Kim, Jae-Seung Kim, Sang-Ryul Kim (Korea Inst. of Machinery and Mater., P.O. Box 101, Yusung, Taejon 305-600, Korea), and Jeong-Guon Ih (Korea Adv. Inst. of Sci. and Technol., Science Town, Taejon 305-701, Korea)

The sound transmission coefficient in the diffuse field depends on the angular distribution of the incident sound energy. Although the field incidence formula has suggested that instead of the random incidence formula, the determination of the limit angle near the grazing incidence is very empirical and various limit angles have been defined by many researchers until now. This angle is especially important when dealing with the sound transmission loss of multilayered panel systems. In order to find the angular energy distribution of the incident sound in the actual test situation, a numerical simulation by using ray tracing is performed for the bounding surfaces of a sphere and a reverberation chamber. The results show that the sound energy distribution of the incidence angle on a bounding surface may be approximated to the Gaussian distribution. Sound transmission losses are predicted by utilizing the concept of the Gaussian distribution of the incidence energy for various single and multilayered panels and they agree very well with experimental results.

2:45–3:00 Break

3:00

1pSA8. An analytical model for bandlimited response of vibroacoustic systems. K. S. Sum and J. Pan (Dept. of Mech. and Mater. Eng., Univ. of Western Australia, Nedlands, Western Australia 6907)

In this paper, an analytical method for predicting bandlimited time-averaged energies in vibroacoustic systems is presented. A vibroacoustic system is considered where a flexible boundary structure is directly excited by some external forces and an enclosed sound field is driven through an acoustic-structural coupling. This method allows for the reduction in computational effort compared to the classical modal coupling method especially for analysis in the medium frequency range where there is a large number of dominant acoustic and/or structural modes. The vibroacoustic problem is also extended to consider cases of structure-borne sound where the sound field is subject to coupling with various vibrating boundary structures. The prediction accuracy of the present method will be compared to the classical modal coupling method and its computational efficiency is demonstrated by means of numerical examples.

3:15

1pSA9. Optimization procedures for strain sensing in active structural acoustic control. Patrice Masson and Alain Berry (G.A.U.S., Mech. Eng. Dept., Univ. de Sherbrooke, Sherbrooke, QC J1K 2R1, Canada, Patrice.Masson@gme.usherb.ca)

Many active structural acoustic control strategies which are based on the minimization of the radiated acoustic power are defined in the wave number domain to exploit the fact that only supersonic structural wave number components are radiating sound. Discrete sensors are more appropriate for this purpose as they allow the ability to distinguish between the structural components on two-dimensional structures. However, the use of such sensors may require extensive real-time computation, as the displacement field has to be reconstructed from the strain field in order to compute the radiated acoustic power. Two approaches are herein proposed to optimize the calculation of the radiated acoustic power. The first one allows the calculation of the radiated power *directly* from the strain field, without requiring any reconstruction of the displacement field. The second approach is directed toward the optimization of the radiated power calculation by using a wavelet transform of the displacement field, thus avoiding the Fourier transform and the integration of the radiating components. Experimental control results are presented using different types of strain sensors mounted on a simply supported beam. It is shown that in that case, a significant gain in computation time can be obtained using the proposed approaches without sacrificing the control performances.

3:30

1pSA10. Measurements and modeling of the transient acoustic field at impacted plates. Antoine Chaigne, Christophe Lambourg (ENST, Signal Dept., 46 rue Barrault 75634 Paris Cedex 13, France), and Staffan Schedin (Luleå Univ. of Technol., Div. of Exp. Mech., S-97187 Luleå, Sweden)

A comparison between measured and simulated transient acoustic fields generated by thin impacted plates is reported. The plate is modeled by using finite difference methods and the radiated pressure is obtained by solving the Rayleigh integral [Ch. Lambourg and A. Chaigne, J. Acoust. Soc. Am. **101**(A) (1997)]. The results are evaluated in the form of the spatial pressure distribution at instants short after the excitation. Experimentally, a lead bullet (0.5 g, 100 m/s) is fired towards a cantilever steel plate ($300\times30\times1$ mm). The duration of the impact is estimated by using a high-speed camera that can register several million frames per second.

An interferometric comparison of the change in optical path length of two holographic recordings, captured before and shortly after the impact, gives a 2-D projection of the acoustic field. A double-pulsed ruby laser, emitting light pulses of short duration (30 ns), is used as light source [S. Schedin, A. O. Waahlin, and P. Gren, J. Acoust. Soc. Am **99**, 700–705 (1996)]. A high degree of similarity is observed between measured and simulated sound field, except in the close vicinity of the impact point. It is presumed that the discrepancies are due to nonlinear effects which are not included in the model.

3:45

1pSA11. Near field acoustic source identification based on numerical models and operational pressure data. Paul Sas and Peter Mas (Dept. of Mech. Eng., Div. PMA, Katholieke Universiteit Leuven, Celestijnenlaan 300B, B-3001 Heverlee, Leuven, Belgium, paul.sas@mech.kuleuven.ac.be)

When pressure measurements are taken in the near field of possible sources, a global inversion of an input–output model of the acoustic medium can be accomplished to calculate the source velocities, and thus to identify the main noise radiating parts. The acoustic medium can be represented by a frequency response function (FRF) matrix which can either be derived analytically based on a point source assumption, or experimentally under laboratory conditions, as well as numerically using a finite or a boundary element model. The operational source velocities can be calculated by multiplying the inverted FRF matrix with the vector of operational pressure measurements. The inversion process, however, is usually an ill-conditioned problem. Regularization techniques based on a singular value decomposition can be used to solve for the ill-conditioned nature of the inverse acoustic problem. In this paper, the applicability of numerically derived FRF models using the boundary element modeling technique, is shown for the calculation of the velocities at the boundary of a source model, based on operational pressure measurements. The results are compared with near field acoustic holography calculations, and the usefulness of an L-curve criterion for the selection of the optimal regularization parameter is indicated based on simulations as well as experiments.

4:00

1pSA12. Visualization of the sound field generated by a plate-cavity coupled system using acoustic holography. Sea-Moon Kim and Yang-Hann Kim (Ctr. for Noise and Vib. Control, Dept. of Mech. Eng., Korea Adv. Inst. of Sci. and Technol., Taejon 305-701, Korea)

In most cases of structural-acoustic problems, it is reasonable to assume that structural vibration is not influenced by the surrounding fluid. In these cases, the vibration of the structure is solved first, then the radiation sound field is calculated by simply applying the Kirchhoff–Helmholtz integral equation. However, this assumption is no longer satisfied when structural stiffness is small or fluid impedance is comparable to it. In this situation the vibration and acoustic fields are to be solved simultaneously. Although many researchers have studied this structural-acoustic coupling problem, there are still difficulties in solving the problem analytically or even numerically. In this study, visualization of sound field by a geometrically simple system (plate-cavity coupled system) is performed experimentally in order to figure out the coupling mechanism between fluid and structure. The system is excited by a speaker, and both internal and external sound pressures are measured. The acoustic holographic method is used to estimate the sound field. The results exhibit that there are frequencies where both plate and cavity are strongly coupled as well as ones where the plate can be considered rigid. Visualization that shows acoustic power flow between the internal cavity and external field enables us to understand the fluid-structure coupling mechanism.

MONDAY AFTERNOON, 22 JUNE 1998 GRAND BALLROOM III (W), 2:15 TO 6:15 P.M.

Session 1pSC

Speech Communication: The State of the Art in Speech Technology

Sadaoki Furui, Chair
Tokyo Institute of Technology, Department of Computer Science, 2-12-1 O-Okayama, Meguro-ku, Tokyo, 152 Japan

Chair's Introduction—2:15

Invited Papers

2:20

1pSC1. Building acoustic models for speech recognition. Francesco Scattone (Dragon Systems, 320 Nevada St., Newton, MA 02160, francesco@dragonsys.com)

This paper reviews the methodology currently used at Dragon Systems to build a speech recognition system for a new language, with particular focus on the training of acoustic models. Issues concerning the selection and use of suitable speech corpora are discussed and experimental results from recent experiences in developing large vocabulary continuous speech recognizers in several European languages are reported.

2:40

1pSC2. Spoken language technology research at Microsoft. X. D. Huang (Microsoft Res., Redmond, WA 98052)

The goal of ongoing work is to develop spoken language technologies for use by a broad audience of computer users and in a wide range of applications. The current work in support of this broad objective falls into four main areas: speech recognition technology code-named Whisper (Windows Highly Intelligent Speech Recognizer); speech synthesis technology code-named Whistler (Windows Highly Intelligent Stochastic Talker); spoken language understanding technology code-named Leap (Language enabled applications); and the Microsoft Speech Software Developer's Kit that facilitates full integration of speech into applications. This paper describes ongoing R&D effort and technical challenges in these areas. Not only are technology components being created that can be shared for the whole spoken language system, but also solutions are being created for application developers to author compelling speech-

enabled applications. To realize this vision of speech enabling the whole computer, not only should critical technology components be improved (i.e., Whisper, Whistler and Leap), but also a tighter integration of these spoken language technology components should be evangelized with broader applications and interfaces to enhance and extend the entire user experience. A number of groups have started developing SAPI-compliant technology components and applications such as telephony, wireless PDA, accessibility, and desktop agent. More information can be found in http://research.microsoft.com/stg.

3:00

1pSC3. Recent activities in spoken language processing at LIMSI. Jean-Luc S. Gauvain (LIMSI-CNRS, BP 133, 91403 Orsay Cedex, France)

This paper summarizes recent activities at LIMSI in speech recognition and its applications. The LIMSI recognizer, which has state-of-the-art performance, uses phone-based continuous density HMM for acoustic modeling, and word and class backoff n-grams for language modeling. Acoustic model adaptation techniques (PMC, MAP, MLLR) are used to reduce the mismatch between test and training conditions, in particular through unsupervised adaptation. While the main goal of speech recognition is to provide a transcription of the speech signal as a sequence of words, the same basic technology can be applied to other areas. This recent work is oriented towards automatic systems for information access and for automatic indexation of audiovisual data. In the former case, the speaker-independent, continuous speech recognizer is embedded in a spoken language dialog system for travel information on an information kiosk (ESPRIT MASK) or by telephone (LE ARISE). User trials highlight the importance of appropriate dialog strategies for satisfactory performance. The latter case is a challenging application for LVCSR as the data contain segments of various acoustic and linguistic natures (prepared/spontaneous speech, studio/telephone quality, background noise/music). Word error rates on this data (about 30%) appear to be sufficient for indexation and retrieval purposes.

3:20

1pSC4. Recent advances in speech recognition at IBM research. Ponani S. Gopalakrishnan (IBM T. J. Watson Res. Ctr., P.O. Box 218, Yorktown Heights, NY 10598)

This paper will review the work going on at the IBM T. J. Watson Research Center on speech recognition systems dealing with real world speech data. Such data includes voice mail messages, broadcast news data, and live interaction with natural language systems. Many problems need to be addressed in order to improve speech recognition accuracy, including speaking style, accents, spontaneous speech phenomena, and different acoustic environments. Some of the methods used for detecting the signal and background condition, fast adaptation to new environments, and for handling pronunciation variations in spontaneous speech will be outlined. Speech recognition results on a variety of materials will be presented and analyzed, with a discussion of the significant problems that remain to be solved in order to obtain acceptable recognition performance on such material.

3:40

1pSC5. Research and development of robust speech recognition. Kiyoaki Aikawa (NTT Human Interface Labs., 1-1 Hikarino-oka, Yokosuka-shi, Kanagawa, 239 Japan)

This paper describes recent research and development activities on robust ASR (automatic speech recognition) in NTT Human Interface Laboratories. ASR system design has been changing from the experimental to the commercial level. A relevant issue in achieving practical ASR is the robustness against environmental noise, and speaker and circuit differences. Adaptation technique has been widely investigated for improving the robustness. It should be noted that customers or users generally prefer simple procedures and quicker response. Thus real time speaker/noise adaptation is getting more and more popular. Confidence measure is also effective in enhancing speech recognition performance by rejecting out-of-vocabulary words, lip, and breath noises. Another relevant issue is how to quickly reach the goal of the human-machine dialog via the telephone without visual information. The voice quality through a telephone is not good enough to provide perfect syllable intelligibility even for a human. This suggests that an appropriate human-machine dialog is needed to support an ASR system in acquiring the information or request intended by the user. Recent research topics including spectral estimation, dialog control, and other new approaches related to above discussion will be shown in the presentation. ASR application examples and related problems will also be shown.

4:00–4:10 Discussion of Papers 1 to 5

4:10–4:25 Break

4:25

1pSC6. 21st century user interfaces within the telecommunication industry: Speech processing is the key. Jay Wilpon (AT&T Labs Res., 180 Park Ave., Rm. D-123, Florham Park, NJ 07932)

As we move into the 21st century, several market factors will play a key role in the success of telecommunication-based services. Such factors include: The merging of the computer and telephony industries is moving ahead at an ever-increasing pace. Barriers-to-entry for new companies wishing to enter traditional telephony businesses are dropping. The union of the computer and telephony industries will lead to effective integration of voice, traditional data, and video, enabling the creation of new services that are more robust, intelligent, and accessible than services that are currently available. People will not be asking for technology, they will be asking for a simpler and easier lifestyle. Given these factors, what will differentiate 21st century telecommunications services from each other are innovative, compelling 21st century user interfaces that create a media center unifying emerging desktop services with

traditional telephony services. Such user interfaces will lead to smarter services. As networks automatically learn their customers' needs and wants, and customers invest their time to learn these intelligent user interfaces, a bond will be formed. Key to this vision is the availability of robust large vocabulary speech recognition, spoken language understanding, and text-to-speech synthesis technologies that allow customers to access information and each other anywhere, anytime.

4:45

1pSC7. Natural speech dialogue systems. Volker Steinbiss (Philips GmbH Forschungslaboratorien, Weisshausstr. 2, 52066 Aachen, Germany)

The application of speech recognition and understanding to real-life situations generates new scientific problems. To illustrate the latter, the example of an automatic switchboard system is given that can be used in a very natural communication style. It is based on a high-end speech recognizer that outputs a word lattice which is subject to further evaluation by a speech understanding and a dialogue component. Although word error rates are above 20% for this task, the integration of additional knowledge sources like the telephone directory and the dialogue history can add to both the acoustic and the language model in order to achieve highly improved performance. As recognition is not perfect, the user has to confirm what the system understood. In order to avoid lengthy confirmations, it was requested to base the level of confirmation (repeated question, explicit or implicit confirmation) on the probability that the request was understood correctly. This was formulated within the statistical framework of speech understanding and successfully implemented into our prototype system. The degree of reliability controls the further dialogue flow. Other useful applications of this reliability measure will be described.

5:05

1pSC8. Recent advances in speech recognition for spontaneous speech translation. Y. Sagisaka (ATR ITL, 2-2 Hikaridai Seikacho Soraku-gun Kyoto, 619-02 Japan)

Speech recognition research for speech translation at ATR will be reviewed. For spontaneous speech recognition, new methods for parameter extraction will be shown that yield acoustic measures to allow discriminative acoustic model training. A new HMM state-sharing algorithm (ML-SSS), fine-acoustic modeling using stochastic segment models, and recurrent neural networks will also be described. The MAP-VFS speaker adaptation techniques will be presented in the context of speaker clustered acoustic models. As for language modeling, variable order N-gram models and task adaptation of language modeling will be described. In addition, further research efforts on pronunciation modeling and training using error characteristics will also be reviewed. Finally, an outline of the continuous speech recognition system (ATR-SPREC) for spontaneous speech will be presented, along with a review of recent recognition experiment results.

5:25

1pSC9. Speech recognition trends and predictions, or, do we need text? Patti J. Price (Speech Technol. and Res. Lab., SRI Intl., EJ147, 333 Ravenswood Ave., Menlo Park, CA 94025, pprice@speech.sri.com)

Speech is the means of communication used first and foremost by humans. Automatic speech recognition capabilities are still far from those of humans, particularly with respect to degradation in the face of noise, nonnative speakers, and casual speech styles. However, significant advances have been achieved in recent years; automatic speech recognition capabilities now permit us to use speech as an interface for dictation and for information access. In such applications, the interactivity of the system allows the person to adapt to the system (and the system to adapt to the user). The real challenge of the information age is using speech not just as a medium for information access, but as itself a source of information, wherever speech occurs. Examples include: voice annotation for later retrieval, sorting and prioritizing voice messages, access to voice recordings of meetings without the need to take minutes, and access to broadcast media. Speech technologies in support of such applications include not just speech recognition, but also speaker recognition, topic spotting, gisting, and summarization. This talk will survey the state of the art and recent trends in speech as a source of information.

5:45

1pSC10. From speech recognition to understanding: Shifting the paradigm to achieve natural human–machine communication. B. H. Juang (Bell Labs, Lucent Technologies, 600 Mountain Ave., Murray Hill, NJ 07974)

Automatic speech recognition has made substantial progress in the last decade due to applications of statistical methods. The task of speech recognition is often described as converting a sequence of sounds into a sequence of linguistic events (words). This is in essence transcription: given an acoustic signal, the system determines which word it represents. The implicit assumption in this traditional formulation of speech recognition is that the speech is well-formed, free from extraneous components such as disfluency, partial words, repairs, hesitation, etc. The statistical framework that supports the formulation is the Bayes decision theory. For natural human–machine communications, however, it is expected that the spoken input to the system will inevitably contain extraneous components, treatments of which have to be developed in a broadened framework utilizing the Neymann–Pearson Lemma. The departure also inspired a new formulation based on the signal detection theory for robust speech understanding. In this talk, the rationale for the new formulation will be discussed and new results towards speech understanding using the detection-based (as opposed to the transcription-based) approach will be presented. It will be shown that the new approach achieves substantial improvements in accuracy for natural but ill-formed utterances over the conventional approach.

6:05–6:15 Discussion of Papers 6 to 10

Session 1pSP

Signal Processing in Acoustics and Speech Communication: Non-Stationary and Wavelet Processing Methods

Gary R. Wilson, Chair
Applied Research Laboratories, University of Texas, P.O. Box 8029, Austin, Texas 78713-8029

Contributed Papers

2:15

1pSP1. Improvement of pitch detection using a signal specific analyzing wavelet. Yoshifumi Chisaki (Dept. of Comput. Sci., Kumamoto Univ., 2-39-1 Kurokami, Kumamoto City, Japan, chisaki@cs.kumamoto-u.ac.jp), Tsuyoshi Usagawa, and Masanao Ebata

Many methods of a speech enhancement system in a noisy environment have been studied. A wavelet transform performs better in analyzing speech than a fast Fourier transform at the point of frequency resolution when the purpose is focused on the pitch detection. Analyzing wavelets, such as Gabor, are derived from a mathematical approach, which have characteristics of low, high, and bandpass filters. In this paper, a modified wavelet transform is proposed. A signal which has characteristics of the target signal is applied to an analyzing wavelet. This signal-specific analyzing wavelet includes not only amplitude and phase for each frequency but also the relationship of harmonic components and a periodic fluctuation of pitches. In addition, the wavelet transform is arranged to use a signal-specific analyzing wavelet for tracking pitches. This method is useful not only for the pitch detection but also for the representation of the fluctuation. Experiments are performed by using both a Gabor function and a signal-specific analyzing wavelet changing both vowels and a band-limited noise. As a result, the ratio of pitch detection was improved by comparing a signal-specific analyzing wavelet with a Gabor wavelet transform. The ratio was recovered over 40% in case of SNR=0 dB.

2:35

1pSP2. Comparison between DyWT- and EGG-based estimation of glottal closure instant for speech signal. Jongwon Seok, Youngho Son, and Keunsung Bae (School of Electron. and Elec. Eng., Kyungpook Natl. Univ., Taegu 702-701, Korea)

Recently, wavelet transform has found many applications. Mallat has shown that, with a proper choice of wavelet function, the local maxima of wavelet transformed signal indicate the position of the signal sharp variation points. Since the glottal closure causes sharp changes in the derivative of air flow in the glottis and transient in the speech signal, wavelet analysis can be used for detection of glottal closure instant (GCI) from the speech signal. It is known that the point of maximum negative value in a differentiated EGG signal agrees well with the closing time of the vocal folds. Due to the limited practical utility of the EGG signal, however, a speech-only-based GCI detection method is still needed. In this paper, GCI detection algorithm based on the dyadic wavelet transform (DyWT) is proposed and experimental results are shown. Using the EGG signal as a reference, experimental results are analyzed and discussed for evaluation of its performance. Preliminary experimental results show that DyWT-based GCI detection method has some speaker dependency and wavelet transformed speech signal at scale 4 gives more reliable GCI information than that of others. [This work has been partially supported by a grant from Korea Science and Engineering Foundation (KOSEF 971-0917-103-2).]

2:55

1pSP3. A novel end-point detection method for noisy speech signal. Jongwon Seok and Keunsung Bae (School of Electron. and Elec. Eng., Kyungpook Natl. Univ., Taegu 702-701, Korea)

This paper presents a novel and robust end-point detection algorithm in the presence of noise. Zero crossing rate and energy of the speech signal have been the most widely used features for locating the end point of an utterance. Most detection algorithms with these parameters work quite well in the high signal-to-noise ratio environment, but they show very poor performance with background noise, i.e., in the low signal-to-noise ratio environment. Therefore robust end-point detection under a noise condition still remains an unsolved problem. A new detection measure based on a discrete wavelet transformed speech signal is proposed. As a detection parameter, the sum of standard deviations of wavelet coefficients in the first and weighted fourth scale are used. Since the new feature parameter enhances high-frequency components of the weak fricative sounds, better discriminating between speech segments and background noise is possible. Experimental results and examples demonstrate that the proposed method is more effective than traditional methods in capturing word boundary even in noisy environments. [This work was supported partially by Korea Telecommunication Inc. and Korea Science and Engineering Foundation (KOSEF: 971-0917-103-2).]

3:15

1pSP4. Speech enhancement by a Kalman filter based on wavelet transformation coefficients. Ryouichi Nishimura, Yôiti Suzuki, Toshio Sone (Res. Inst. Elec. Commun., Tohoku Univ., Sendai-shi, 980-8577 Japan), and Futoshi Asano (Electrotechnical Lab., Ibaraki-ken, 305-0045 Japan)

A new speech enhancement technique based on Kalman filtering and wavelet transformation is proposed. A problem in using a Kalman filter for speech enhancement lies in how to construct and to estimate the state transition matrix. Considering that wavelet transformation can well localize the speech energy in the time-frequency domain in the presence of noise, wavelet coefficients are used here for this purpose. That is, the state and observation vectors are expressed by matrices which have column vectors consisting of wavelet coefficients calculated from densely overlapped frames. Then the state transition matrix is calculated from a single state matrix postmultiplied by the inverse matrix of the preceding state matrix. These matrices are obtained from a speech signal enhanced by existing algorithms such as MMSE. This enhancement process does not need to correctly estimate a speech signal itself but needs just to supply the correct differentials of state matrices. A computer simulation was conducted to evaluate the performance of the proposed method. The result shows that the signal processed by the proposed method generally gives a better subjective impression while physical evaluation by use of the ''signal to deviation ratio'' becomes worse.

3:35

1pSP5. Time variation characterization of a nonstationary time series. Nai-chyuan Yen (CLY Assoc., 2109 Rampart Dr., Alexandria, VA 22308-1537)

The Fourier transform of a nonstationary time series has the difficulty of providing a proper physical interpretation of the signal as the physical attributes and dynamics associated with the time series may consist of several transients that are different in their nature. The time series analysis method developed by NASA [N. Huang *et al.*, ''The Empirical Mode Decomposition and The Hilbert Spectrum for Nonlinear and Nonstationary Time Series Analysis'' (to be published)] has a unique approach to adaptively break down the nonstationary time series into the sum of several simple mode functions. Each single mode function can then be expressed in terms of an analytical signal whose amplitude and phase can both be varied with time. Those time variations can be related to the abrupt changes associated with dynamic systems. Such a time variation characteristic can be used to observe the occurrence of a transient event even in a nonstationary process. Applications of this technique to observe various types of acoustic signals in noisy environments are discussed.

3:55

1pSP6. Comparison of methods for analysis of cyclostationary noise. Karel Vokurka (Tech. Univ. of Liberec, Dept. of Phys., Halkova 6, CZ-461 17 Liberec, Czech Republic, karel.vokurka@vslib.cz)

A great number of machines work cyclically. Hence the noise they generate is cyclostationary. A traditional approach to this noise analysis is based on an assumption of its stationarity. This assumption implies continuous time averaging. Unfortunately in this way a certain amount of information concerning, e.g., time-frequency distribution, is lost. Methods have been looked for to preserve this information. They all are based on periodic time averaging and the resulting statistical characteristics are known as gated spectrum, cyclic correlation, cyclic autospectrum, double autocorrelation, double autospectrum, etc. Each of the methods has its advantages and disadvantages. These may concern processing time, hardware requirements, frequency resolution, occurrence of artifacts (cross terms), and others. In this paper the methods are analyzed and compared and possible improvements in their performance are suggested. [Work supported by the Grant Agency of the Czech Republic and by the European Commission.]

MONDAY AFTERNOON, 22 JUNE 1998 GRAND BALLROOM B (S), 1:00 TO 4:20 P.M.

Session 1pUW

Underwater Acoustics: Sources, Arrays, Tracking, and Localization

Stanley E. Dosso, Chair
School of Earth and Ocean Sciences, University of Victoria, P.O. Box 1700, Victoria, BC V8W 3P6, Canada

Chair's Introduction—1:00

Contributed Papers

1:05

1pUW1. Localized wave pulses in the Keyport experiment. David H. Chambers and D. Kent Lewis (Lawrence Livermore Natl. Lab., P.O. Box 808 L-372, Livermore, CA 94551, chambers2@llnl.gov)

Localized wave (LW) pulses were produced using a standard Navy array in the anechoic tank at NUWC Keyport. The LW pulses used were the MPS pulse first derived by Ziolkowski, and a new type of pulse based on a superposition of Gaussian beam modes. This new type is motivated by a desire to make a comparison of the MPS pulse with another broadband pulse built from solutions to the wave equation. The superposed Gaussian pulse can be described by parameters which are analogous to those describing the MPS pulse. The directivity patterns and the axial energy decay between the pulses are compared. The behavior of the pulses are found to be similar so that the superposed Gaussian could be another candidate in the class of low diffractive pulses known as localized waves.

1:20

1pUW2. Localized wave generation with a standard underwater array. D. Kent Lewis, David H. Chambers (Lawrence Livermore Natl. Lab., 7000 East Ave., L-372, Livermore, CA 94550), and Richard W. Ziolkowski (Univ. of Arizona)

Recent work at the NUWC Keyport has generated localized waves using Navy field equipment. The array, a rectangular arrangement of tonpiltz elements, was driven with precompensated signals to generate sound covering a decade of usable bandwidth. The elements were grouped by radius from the center and excited by signals designed specifically for each radius, for a total of 7 radii. Results of the angular scans show a narrowed beam pattern and lowered side lobes relative to design signals. Results of axial range scans show an increase in directivity of more than an order of magnitude over array specific tone bursts. Frequency analysis shows that the main beam is an octave wide while the surviving side lobes are narrow bandwidth.

1:35

1pUW3. An inexpensive light-weight ocean acoustic research array. Garry J. Heard (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada, heard@drea.dnd.ca)

Current and past research arrays built at the Defence Research Establishment Atlantic (DREA) have been intended for indefinite and repeated use. Building, maintaining, and using these arrays has been expensive in terms of both money and labor due to the rugged construction practices, complicated electronics, and materials. With the reduced resources available in most laboratories today, there is considerable need for cheaper research array systems that are easily maintained and deployable from a variety of platforms: from Zodiacs to full-sized AGOR vessels. This paper describes the initial development of a light-weight research array that is expected to meet the requirements of the Ocean Acoustics Group at DREA. The array uses epoxy resins for most of the structural components, has a serial-digital bus, has low-power requirements, is easily reconfigured, and is essentially disposable, although the intention is to reuse the array in whole or in part as long as it remains serviceable. The prototype array, which is scheduled for sea testing in July 1998, will be a bottom-deployed 500-m aperture horizontal line array with 16 channels sampled at 2048 Hz.

1:50

1pUW4. Optimal array element localization. Stan E. Dosso (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada, sdosso@uvic.ca) and Barbara J. Sotirin (Naval Command, Control and Ocean Surveillance Ctr., San Diego, CA 92152-5001)

Measuring ocean acoustic fields at an array of sensors allows the application of advanced signal processing methods such as beamforming and matched-field processing. However, these methods require accurate knowledge of the positions of individual sensors. Typically, an acoustic survey is required to localize the sensors, a procedure referred to as array element localization (AEL). AEL generally consists of measuring the traveltimes of acoustic signals transmitted from a series of known positions to the sensors to be localized. These traveltime data can then be inverted for estimates of the sensor locations using a linearized inversion algorithm. An important issue in AEL is designing the configuration of acoustic source positions: a well-designed source configuration can produce substantially better AEL results than a poor configuration. In this paper, a procedure is developed to determine optimal AEL source locations for arrays of any configuration. This procedure minimizes the condition number of the Jacobian matrix of partial derivatives on which the linearized inversion is based. Minimizing the condition number in effect minimizes the magnification of data (traveltime) errors into model (sensor position) errors in the inversion. Monte Carlo simulations verify that optimal AEL source configurations can lead to significant improvements in sensor localization.

2:05

1pUW5. Detection and tracking of a moving source with application to real data (SWELLEX). Stacy L. Tantum and Loren W. Nolte (Dept. of Elec. Eng., Duke Univ., Box 90291, Durham, NC 27708-0291)

The optimum uncertain field tracking algorithm (OUFTA) incorporates knowledge of source motion as well as uncertainty in the propagation environment [S. 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)]. It is shown how detection, tracking, and adaptivity to an uncertain environment can be formulated within a unifying and optimal framework. Performance results and trade-offs are presented for simulated data. Results are also presented using this approach with low signal-to-noise ratios and real data (SWELLEX). [Research supported by ONR.]

2:20

1pUW6. Characterization of fault processes in central Arctic ice. Catherine Stamoulis and Ira Dyer (Dept. of Ocean. Eng., MIT, Cambridge, MA 02139)

Fault processes in Arctic ice are characterized by particle motion, propagation, and radiation. Estimates of these can be achieved through analysis of individual acoustic events induced by the faults. Thus data from a central Arctic experiment were analyzed. Particle velocity, displacement, and rise time, which characterize particle motion, were estimated through integration of the acoustic signals in the time domain. Particle displacement was in the range $O(10^{-4})$–$O(10^{-2})$ m, significantly lower than that of earthquake-induced faults; particle velocity was 1–67 cm/s, on average 50% lower than that of a displacement discontinuity in rock. Fracture speed, strike angle, and source dimensions, which characterize fault propagation, were estimated in the frequency domain. Fracture speed estimates were in the range 200–1100 m/s. This wide range suggests either multiple fault mechanisms or a single mechanism at different stages in its propagation. Source dimensions were determined from the displacement spectra of the second integral of the signals, with source length 0.7–100 m, and depth 0.4–4 m. Finally, the radiation characteristics were estimated and tensile and shear fracture were identified, both in the low (10–100 Hz) and mid (100–350 Hz) frequency ranges.

2:35–2:50 Break

2:50

1pUW7. Time-reversal mirror with variable depth focusing. Hee Chun Song, William A. Kuperman, William S. Hodgkiss (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0238), T. Akal, C. Ferla (SACLANT Undersea Res. Ctr., 19138 La Spezia, Italy), and Darrell R. Jackson (Univ. of Washington, Seattle, WA 98105)

Recent time-reversal mirror (TRM) experiments conducted in the Mediterranean Sea demonstrated variable range focusing using the frequency-range invariant property in a shallow-water waveguide [Song *et al.*, J. Acoust. Soc. Am. **102**, 3171(A) (1997)]. The technique involves retransmitting the data at a shifted frequency according to the desired change in focal range in almost real-time fashion. As opposed to the variable range focusing, there does not appear to be an environmentally independent elegant solution for variable depth focusing. However, a method based on knowledge of the environment, which translates to being able to compute the acoustic modal structure, permits a focal depth shift. The method involves projecting out the important modes received at the vertical source–receive array (SRA) and appropriately altering their excitation to focus at another depth. In the May 1997 experiment, some limited results demonstrated that this method is feasible. Comparison of data with simulations and theoretical analysis will be presented and discussed. [Work supported by ONR.]

3:05

1pUW8. A piecewise matched-field tracking algorithm. Michael J. Wilmut (Dept. of Mathematics and Comput. Sci., Royal Military College, P.O. Box 17000 STN Forces, Kingston, ON K7K 7B4, Canada) and John M. Ozard (Defence Res. Establishment Atlantic, Victoria, BC V9A 7N2, Canada)

Efficient matched-field tracking (MFT) algorithms have been successfully applied to both simulated and measured data. They have been used to track sources moving linearly or on a circular path at constant speed and heading. The input to the tracker is a set of ambiguity surfaces, contiguous in time, generated by matched-field processing. Previously described MFT algorithms assume that the track start and end times are known *a priori*; this restriction is removed in the piecewise MFT algorithm described here. In a series of simulations it is shown that the statistically significant tracks can be found provided the track signal-to-noise ratio is greater than 10. Their unknown start and end times are in error by one surface or less most of the time. The algorithm is then applied to find track segments for simulated sources moving linearly, or on a circular path. The path segments were of unknown duration and heading.

3:20

1pUW9. Effects of environmental mismatch on Bartlett sidelobe broadband interference structures. Aaron M. Thode, Gerald L. D'Spain, Hee Chun Song, and William A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., Univ. of California at San Diego, La Jolla, CA 92093-0205)

When a series of perfectly matched range-independent Bartlett surfaces (computed over small increments of frequency) are stacked and then sliced along a constant depth, the resulting sidelobe contours in the range-frequency plane display interference patterns that converge to the correct range [Thode *et al.*, J. Acoust. Soc. Am. **102**, Pt. 2 (1997)]. The existence of these patterns can be derived from waveguide invariant principles [e.g., G. A. Grachev, Acoust. Phys. **39**, 33–35] and their physical meaning interpreted using a time-reversal argument [Song *et al.*, J. Acoust. Soc. Am. **102**, Pt. 2 (1997)]. Several qualitative effects of environmental model mismatch on these patterns have been observed: pattern translation along the range axis, frequency-dependent warping, and ''ghost'' contours that converge to the range origin instead of the true source range. Translational shifts are produced by certain mismatch parameters (e.g., waveguide depth) associated with invariant theory. Warping at low frequencies is generated by a mismatched bottom sound speed, and can be explained

using ''effective depth'' concepts [G. L. D'Spain, 15th Int. Cong. Ac. Proc., 231-4]. ''Ghost'' patterns arise when the modeled environment contains more modes at a given frequency than the actual environment. A time-reversal hypothesis for their convergence to the range origin is suggested. [For Underwater Acoustics Best Student Paper Award.] [Work sponsored by ONR.]

3:35

1pUW10. A parabolic equation-based backpropagation algorithm for matched-field processing. David J. Thomson, Gordon R. Ebbeson, and Brian H. Maranda (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada)

In applications of matched-field processing (MFP) to passive sonar, signals measured on a hydrophone array due to a source at some unknown position $\mathbf{x}_s=(r_s,z_s)$ are correlated with signals predicted by a propagation model (replicas) due to a source at a given position $\mathbf{x}=(r,z)$. This matching is carried out for many assumed $\mathbf{x}$'s within a search region (range, r and depth, z) to form a (normalized) ambiguity surface whose peak value provides an estimate of $\mathbf{x}_s$. For an N-element array, reciprocity is invoked to reduce the computational effort to computing N replica fields at each point on the search grid. In this paper, a matched-field processor proposed by Tappert *et al.* [J. Acoust. Soc. Am. **78**, S85 (1985)] is revisited that combines measured data with PE starting fields to effectively backpropagate an (unnormalized) ambiguity surface outwards from the receiving array. This unnormalized processor generates the ambiguity surface N times faster than the normalized one. The effectiveness of the backpropagation algorithm is examined by applying it to some multi-tonal data obtained during the Hudson Canyon experiment. Issues relating to normalization and incoherent/coherent frequency averaging are discussed.

3:50

1pUW11. Regularized matched-mode processing. Nicole E. Collison and Stan E. Dosso (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada, collison@uvic.ca)

Far-field acoustic propagation can be modeled as a discrete set of propagating normal modes. In this case, the measured fields can be decomposed into their modal components, providing the basis for matched-mode processing (MMP). Modal decomposition is a discrete linear inverse problem that is inherently nonunique and can be unstable. Thus standard MMP generally requires the entire water column to be densely sampled (i.e., more sensors than modes) with relatively low noise. However, the inversion can be stabilized using the method of regularization to include *a priori* information based on the replica excitations at each trial source location of the search grid. The regularized approach gives an optimal match for each grid point for both dense and sparse arrays and noisy data. Regularized matched-mode processing (RMMP), MMP, and matched-field processing (MFP) are compared for a number of different array configurations, including dense and sparse arrays that span all or part of the water column. The comparison is carried out for a realistic shallow-water environment which includes elastic geoacoustic properties and an environmental noise model. [For Underwater Acoustics Best Student Paper Award.]

4:05

1pUW12. A modified grid method for matched-field localization. Ling Xiao, Renhe Zhang, and Lianghao Guo (Natl. Lab. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC, zrh@public.east.cn.net)

A modified grid method (MGM) for matched-field localization is proposed in the paper. It is a practical version of the seed division algorithm (SDA). It adopts the ideal that searches in a coarse grid at first and then regionally fines the grid. The fining process is implemented by SDA. This so-called MGM is tested in the paper by seven two-dimensional global optimization problems. The tests show that MGM is more efficient than the normal grid method. Besides, it is more efficient than some other randomized methods which are currently used today. A numerical test is carried out for matched-field localization in deep ocean. It is based on MGM and WKBZ [Zhang *et al.*, Chin. J. Acoust. **13**, 1–12 (1994)]. The result shows that MGM can greatly upgrade the computational velocity and improve the localization's precision. [Work supported by NSF of China.]

MONDAY AFTERNOON, 22 JUNE 1998 WEST BALLROOM A (S), 5:00 TO 6:00 P.M.

Session 1pPLb

Plenary Lecture

J. Stuart Bolton, Chair
School of Mechanical Engineering, Purdue University, 1077 Ray W. Herrick Laboratories, West Lafayette, Indiana 47907-1077

Chair's Introduction—5:00

Invited Paper

5:05

1pPLb1. Adapted waveform analysis in sound processing. Ronald R. Coifman (Dept. of Mathematics, Yale Univ., New Haven, CT)

Adapted Waveform Analysis [R. Coifman *et al.*, *Wavelet Analysis and Signal Processing*] is a processing toolkit extending the traditional spectrogram and Fourier methods by adapting to local structures in audiosignals. The various libraries of waveforms such as localized cosines and wavelet packets permit a flexible time/frequency analysis and give rise to a ''mathematical musical transcription'' of sound. The waveforms considered could be viewed as musical notes with characteristic duration, pitch, amplitude, and

time localization. They are tied to fast computational algorithms which enable a fast search for an optimally efficient transcription of a clean simple signal as a mathematical ''score.'' More complex signals are then orchestrated as superposition of scores. Various applications to sound, compression, automatic speech segmentation for recognition, structural sound denoising and music restoration [with J. Berger and M. Goldberg, ''Removing Noise from Music Using Local Trigonometric Bases and Wavelet Packets,'' J. Audio Eng. Soc. **42**(10) (1994)] will be provided. Moreover the relations between efficient waveform transcription and fast numerical analysis for acoustic design and simulations [B. Bradie *et al.*, ''Fast Numerical Computations of Oscillatory Integrals Related to Acoustic Scattering,'' Appl. Comput. Harmonic Anal. **1**, 94–99 (1993)] will be indicated.

Session 2aPLa

Plenary Lecture

Lawrence A. Crum, Chair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105-6698

Chair's Introduction—8:00

Invited Paper

8:05

2aPLa1. Noninvasive ultrasonic surgery. Gail R. ter Haar (Phys. Dept., Royal Marsden Hospital, Sutton, Surrey SM2 5PT, UK, gail@icr.ac.uk)

The millimeter and submillimeter wavelengths of megahertz ultrasound in tissue mean that this form of energy can be brought to a tight focus in the body at a distance from the acoustic source. If the energy carried by the ultrasonic beam is sufficiently high, then it is possible to destroy tissue lying within the focal region, while sparing adjacent and overlying tissues. Typical peak intensities at the focus for ablation of tissues using 1–2 s exposure times are 1 kW/cm^2. For a focused bowl transducer with focal length 15 cm, 8 cm in diameter, operating at 1.7 MHz, the focal region is ellipsoidal in form, with axial length 15 cm and diameter 1.5 mm. A number of different exposure methods may be used to ablate clinically useful tissue volumes. Focused ultrasound is capable of selectively destroying tissue at depth without the need to expose the site of interest surgically. The mode of tissue destruction is primarily thermal in origin, temperatures in excess of 60 °C being induced very rapidly. This type of cellular damage is capable of destroying both normal and malignant cells, and of occluding selected blood vessels. Potential applications for this technique lie, among others, in urology, oncology, and fetal medicine.

Session 2aAAa

Opera House Acoustics Workshop—CIARM, MCHA, Architectural Acoustics and Musical Acoustics: Sound Field for Singers and Orchestra, Part 2

Christopher N. Blair, Cochair
Artec Consultants, Inc., 114 West 26th Street, New York, New York 10001

Anders C. S. Gade, Cochair
Department of Acoustic Technology, Technical University of Denmark, Building 352, DK-2800 Lyngby, Denmark

Chair's Introduction—7:45

Contributed Papers

7:50

2aAAa1. On the preferred delay time of a single reflection for cellists. Shin-ichi Sato, Yoichi Ando (Grad. School of Sci. and Technol., Kobe Univ., Rokkodai, Nada, Kobe, 657 Japan), and Saho Ota (Kobe Univ., Kobe, 657 Japan)

In order to attain a knowledge of designing the stage enclosure in a hall, the subjective preference of sound fields with a single reflection for alto-recorder players has been reported [I. Nakayama, Acustica **54**, 217–221 (1984)]. In this study, subjective preference judgments of sound fields with a single reflection were performed by cello players. It is examined whether or not preferred delay time of reflection is dependent on the effective duration (τ_e) of the autocorrelation function (ACF) of the music signals performed by each player and the amplitude of the reflection. [Work was supported by the Grant-in-Aid for Scientific Research JSPS.]

8:05

2aAAa2. A study of improvements to acoustical conditions for singers. Dennis Noson (107 North 77th St., Seattle, WA 98103, noson@worldnet.att.net)

While numerous recent studies have been reported on improving stage acoustics for orchestral performers, the preferred acoustical conditions on performing stages for singers has received limited attention in the past 30

years, with pioneering studies by Nakamura and Marshall. A series of acoustical modifications has been proposed for a Seattle church to improve the acoustics for both the listeners and the performing choir. An investigation was made, within the constraints of the existing acoustical environment of the church, to determine what direction the acoustical modifications should take to improve conditions of ensemble and balance among the singers while simultaneously improving the sound field for the audience. Acoustical conditions were modified both by adding temporary reflecting panels and introducing simulated reflections with strategically located loudspeakers. The results indicate a potential for noticeable improvement in the existing chancel acoustics, with most of the benefit applying to the singers rather than the listeners. Comparison will be made between the measured changes to the sound field and the predicted response of the singers using Ando's theory of subjective preference.

8:20

2aAAa3. Some recent experiences with the acoustical design of orchestra pits. Dennis Paoletti, Kurt Graffy, and Larry Tedford (Paoletti Assoc., Inc., 40 Gold St., San Francisco, CA 94133)

A number of recent projects including work for the San Francisco and Houston Operas have included acoustical designs for new and/or renovated orchestra pits. This paper will review some of those projects, their criteria, design, measurements, and evaluations.

8:35

2aAAa4. The virtual orchestra; pedagogical and creative issues. Frederick W. Bianchi (Humanities Dept., Music Div., Worcester Polytechnic Inst., 100 Institute Rd., Worcester, MA 01609) and Richard H. Campbell (Worcester Polytechnic Inst., Worcester, MA 01609)

The interactive multichannel computer music system known as the ''Virtual Orchestra'' has been used several times in professional opera and theater as an alternative to a live pit orchestra. The technical issues associated with this emerging technology, and the logistical problems of implementing it are presented in a companion paper. This paper describes the equally important issues regarding the creative and pedagogical impact this technology will have on the industry. In particular, the paper suggests that many of the fundamental premises of opera production will change as the industry begins to retool. This would include the way opera is rehearsed and performed, how it is created and disseminated, and how it will adjust to the changing demographics of the country. In addition, the paper addresses the notion of the technologist/musician and examines the role of education and professional training in computer music and musical acoustics.

8:50

2aAAa5. The virtual orchestra; acoustical and audio engineering issues. Richard H. Campbell (Elec. and Comput. Eng. Dept., Worcester Polytechnic Inst., 100 Institute Rd., Worcester, MA 01609) and Frederick W. Bianchi (Worcester Polytechnic Inst., Worcester, MA 01609)

The multichannel MIDI system known as the ''Virtual Orchestra'' has been used several times in the pit for the performance of opera, and recently demonstrated at an ASA regional chapter meeting in Boston. This paper discusses technical issues such as the selection of the types of loudspeakers and their placement in the pit, the amplifier power requirements, and the sound level calibration for each instrumental section based on a mezzo-forte norm. ''Wolf'' tones caused by understage extension cavities and other pit deficiencies can be easily modified within the MIDI computer program or with external equalization. Ideas are presented for the use of this new tool with respect to an acoustic freedom heretofore unavailable to orchestrators of musical theater works. The computer operator's ability to follow in real time the conductor's direction in nuance and tempo guarantees a cohesive performance synchronized with the singers and the dramatic action.

9:05

2aAAa6. The effect of early reflection on players in a concert hall. Kanako Ueno (Dept. of Architecture, Tokyo Univ., 7-3-1 Hongou, Bunkyo-ku, Tokyo, Japan, ueno@iis.u-tokyo.ac.jp), Fumiaki Satoh (Chiba Inst. of Technol., Narashino-shi, Chiba, Japan), Hideki Tachibana (I.I.S. of Tokyo Univ., Minato-ku, Tokyo, Japan), Takahiko Ono (Ono-Sokki Co., Midori-ku, Yokohama-shi, Kanagawa, Japan), and Mariko Senju

In order to investigate the effect of hall response on players, field measurements on the stage of a concert hall and laboratory experiment using digital simulation technique were performed. In the field experiment, the subject, a professional violinist, was asked to play and to make comments on her acoustical impression of five points on the stage. As a physical measurement, impulse responses were obtained at the same points by using omni-directional loudspeakers as a sound source and an omni-directional microphone and directional microphones as receivers. As a result, it has been found that not only the strength of the early reflections but also their direction influences the subject's impression. In the laboratory experiment (anechoic chamber), the sound field was modeled and synthesized by using a 13 channel reproduction system; ambient reverberation judged as being natural was provided by simple digital reverberators and different strength and direction of early reflections were obtained by real-time convolvers. For a constant value of reverberation, several conditions with a different level and direction of the early reflections were created. For each condition, the violin player was asked to make similar judgments as in the field experiment. The results of two experiments were examined.

Session 2aAAb

Opera House Acoustics Workshop—CIARM, MCHA, and Architectural Acoustics: Sound Field for the Audience, Part 2

Yoichi Ando, Cochair
Graduate School of Science and Technology, Kobe University, Rokkodai, Nada, Kobe, 657 Japan

Michael Barron, Cochair
Department of Architecture and Civil Engineering, University of Bath, Bath BA2 7AY, England

Chair's Introduction—9:20

Invited Papers

9:25

2aAAb1. Boxes and sound quality in an Italian opera house. Alessandro Cocchi, Massimo Garai, and Carla Tavernelli (DIENCA, viale Risorgimento 2, 40136 Bologna, Italy, massimo.garai@mail.ing.unibo.it)

The sound field in the boxes of an 18th century Italian opera house was studied by means of binaural measurements and computer simulations. The differences in listening between the boxes and the stall is best described by the values of ITDG, clarity, and spatial criteria resulting from the measurements. The periodic variation of the ITDG along the boxes depends also on the reflections coming from the ceiling. The best boxes for clarity and spatial impression are found in very different positions. Computer simulation supports these conclusions and helps in understanding the two different mechanisms by which the boxes influence the sound field in the hall: the equivalent sound absorption of the boxes and the periodic structure of the box-covered walls. A validated computer model allows one to forecast how the sound field in the historical theater could be if some interior lining were added to the boxes, trying to restore the velvet lining which was supposed to exist in ancient times.

9:45

2aAAb2. Measurements of the polarization of sound energy at the mouth of boxes of an Italian opera house. Domenico Stanzial, Davide Bonsi, and Nicola Prodi (CIARM c/o CNR-Cemoter Acoust. Dept., via Canal Bianco 28, I-44044 Ferrara, Italy)

The measurement of a newly discovered property of sound energy, the polarization [Stanzial *et al.*, J. Acoust. Soc. Am. **99**, 1868–1876 (1996)], has been carried out for the first time inside an opera house. Since energy polarization is due to the distributions of wavefronts passing through the measurement point, the analysis of this quantity allows one to obtain a precise picture of the effect of sound reflections at the receiving location. The highly symmetric architectural structure of an Italian opera house has been thus investigated at the opening of some boxes and the energy polarization has been compared with the sightline from each box to the stage. Furthermore, a third octave band frequency analysis in the range 100–5000 Hz has been accomplished that allows the visualization of the frequency-dependent behavior of the reflections' distribution.

10:05

2aAAb3. Measurements in an opera house: Comparison between techniques and instrumentation. Patrizio Fausti, Roberto Pompoli (Dipartimento di Ingegneria, Univ. of Ferrara, Via Saragat 1, I-44100 Ferrara, Italy), and Angelo Farina (Univ. of Parma, Parma, Italy)

In room acoustics, many objective parameters to quantify subjective impression have been introduced. These quantities can be measured using a wide variety of powerful tools and instrumentation. The results can be influenced by the measurement techniques and instrumentation used. Furthermore, the results also depend on the position and condition of the hall in which the measurements are made. With the aim to obtain a procedure to qualify a room acoustically, and in particular an opera house, which gives comparable and reproducible results all over the world, an extensive measurement campaign was made. The procedure takes into account the measurement techniques, the instrumentation, the measurement positions, and the condition of the audience and the stage. Comparisons were made both between different tools (real-time analyzer, DAT, PC-board) and different kinds of acoustic excitation (stationary, impulsive, pseudorandom). Furthermore, comparisons were made in different conditions of the stage and pit (e.g., empty, with the presence of orchestra tools and/or musicians). In this paper the results of the measurement campaign are presented and a proposal is submitted for the discussion.

Contributed Paper

10:25

2aAAb4. Individual subjective preference of listeners for vocal music sources in relation to the subsequent reverberation time of sound fields. Hiroyuki Sakai (Grad. School of Sci. and Technol., Kobe Univ., Rokkodai, Nada, Kobe, 657 Japan), Hiroshi Setoguchi (Miyama Conceru, Kagoshima, 899-66 Japan), and Yoichi Ando (Kobe Univ., Kobe, 657 Japan)

The purpose of this study is to evaluate subjective preference of a simulated sound field by listeners in changing subsequent reverberation time T_{sub} using vocal music sources. Previously, a great deal of effort has been made on the study using music or speech [for example, Ando *et al.*, Acustica **50**, 134–141 (1982); J. Acoust. Soc. Jpn. **39**, 89–95 (1983)]. In addition to subjective preference judgments, the α-waves' range of the continuous brain waves (CBW) in relation to the specialization of human cerebral hemispheres is also evaluated [C. Chen and Y. Ando, J. Archit. Plann. Environ. Eng., AIJ **489**, 73–80 (1996)]. In this paper, subjective evaluation using vocal music sources, which is often played in concert halls and opera houses, has been examined. Subjective preference tests were conducted changing T_{sub}, which is one of the four objective parameters in relation to subjective preference of sound fields. Individual differences of subjective preference as well as global preference are discussed. [Work was partially supported by the Ministry of Education, Grant-in-Aid for Scientific Research (C), 9838022, 1997.]

TUESDAY MORNING, 23 JUNE 1998 VASHON ROOM (W), 8:00 TO 10:45 A.M.

Session 2aAB

Animal Bioacoustics: Insect, Insectivore and Avian Acoustics

Timothy G. Forrest, Chair
Department of Biology, University of North Carolina, One University Heights, Asheville, North Carolina 28804

Invited Papers

8:00

2aAB1. Scale effects as constraints in insect sound communication. Henry Bennet-Clark (Dept. of Zoology, Oxford Univ., S. Parks Rd., Oxford OX1 3PS, UK)

Calling insects have problems of optimizing range and maintaining the specificity of the call. Higher frequencies are more easily refracted and reflected by objects in the environment. Smaller insects have less muscle power and also, because of the small sound source size, higher frequencies are radiated more efficiently than lower frequencies. In open air or water, the sound spreads spherically and decays by the inverse square law. If the sound can be confined to a sheet it decays as the inverse of range while within a rod it decays due to viscous losses; such calls are usually rather simple pulses and rely on initial time of arrival because of multiple pathlengths in the environment. With airborne sounds, those in the range from 1 to 10 kHz tend to have sustained pure-tone components and a specific pattern of pulses which propagates well; but with higher frequencies pulses tend to become briefer and, once again, to rely on time of arrival of the onset.

8:25

2aAB2. Acoustic detection and identification of insects in soil. Richard Mankin (USDA-Agric. Res. Service, Ctr. for Medical, Agricultural, and Veterinary Entomology, P.O. Box 14565, Gainesville, FL 32604)

There is considerable practical need for user-friendly, inexpensive devices that detect and quantify insect populations in environments hidden from visual observation. One approach that has been used with varying success has been to detect the insects through the sounds or vibrations they generate for communication or through noises that are produced incidentally during feeding and general movement. The ability to precisely monitor soil insect populations is limited by interference from background noise and the high rate of attenuation of sound in soil. This paper describes experiments using different sensors and analysis techniques for detection of insects in soil in an agricultural environment. Results from different sensors are compared, and the spectral and temporal patterns that can be used to distinguish the target insects from noise and nontarget soil organisms are discussed. The use of accelerometers attached to 20–30-cm nails appears to be a low-cost, user-friendly method to locate and monitor the size of infestations of soil insects that tend to group together in clumps.

Contributed Papers

8:50

2aAB3. Two-tone suppression of the ultrasound induced startle response in a cricket. Hamilton E. Farris and Ronald R. Hoy (Sec. Neurobiology and Behavior, Mudd Hall, Cornell Univ., Ithaca, NY 14850, hef1@cornell.edu)

For the ground cricket, *Eunemobius carolinus*, frequencies >20 kHz elicit an acoustic startle response during flight. No startle responses are elicited by single pulses of low-frequency stimuli. However, simultaneous presentation of a low-frequency signal with ultrasound can suppress the startle response. The effect of frequency on suppression was measured for seven frequencies (3–9 kHz) at four intensities of the startle stimulus (40 kHz): 5, 8, 10, and 12 dB above startle threshold. Suppression was tuned to frequencies near 5 kHz. This effect of frequency on suppression diminished slightly at the higher ultrasound intensities, however. Because suppressor growth (change in suppression dB versus change in 40-kHz dB) was >1 for 5–7 kHz and >1 for 3 to 4, 8 to 9 kHz, equal suppression curves flattened at the highest ultrasound levels. Startle threshold de-

creases with increases in ultrasound duration, and temporal integration time constants were estimated to be 32 ms. In nonsimultaneous presentations, however, no suppression occurs when the suppressing tone is >2 ms after and >5 ms before the ultrasound stimulus. These times suggest that across-channel integration time is shorter than within-channel integration. [Supported by NIMH.]

9:10

2aAB4. Synchrony in cricket calling songs: Models of coupled biological oscillators. Timothy Forrest (Dept. of Biol., Univ. of North Carolina at Asheville, Asheville, NC 28804), Joel Ariaratnam, and Steve Strogatz (Cornell Univ., Ithaca, NY 14853)

Snowy tree crickets produce a rhythmic acoustic signal composed of short chirps. The chirp rhythms are very regular and dependent on temperature. When signaling in groups, males synchronize their chirps. The rhythmic signals of these tree crickets were used as a model of a ''biological oscillator.'' The phase responses of males to single probe stimuli similar to the male calling song were measured. The phase response curves (PRC) of males showed both strong (type 0) and weak (type 1) resetting. PRCs differed depending on amplitude, duration, and frequency of the resetting stimuli. In other experiments, males could adjust their natural periods by as much as 15% when phase-locking (1:1) to an entrainment stimulus. Signaling behavior of males was compared with a simple model of coupled oscillators. The phase resetting was not consistent with the simplest resetting model, in which the male's phase shift is completed after one cycle of the chirp period. In addition, a simple iteration of the PRC could not simulate the phase-locking observed in the entrainment experiments indicating that the acoustic coupling between males differs from a simple resetting model.

9:30–9:45 Break

9:45

2aAB5. Functional anatomy of the middle ear of insectivores. Matthew J. Mason (Dept. of Zoology, Univ. of Cambridge, Downing St., Cambridge CB2 3EJ, UK)

An electrical model has been constructed of the middle ear of land mammals [Hemilä *et al.*, Hearing Res. **85**, 31–44 (1995)]. This model can roughly predict the nature of air hearing in an animal, given the values of certain anatomical parameters for that animal. In the present study, light microscopy was used to investigate the anatomy of the middle ears of some representative members of the Insectivora. The nature of air hearing in these mammals was predicted, using an augmented version of the model. The model predicts that within the family Chrysochloridae (golden moles) there are genera (e.g., *Chrysochloris* and *Chrysospalax*) that have ears which appear to be very poorly adapted to the impedance-matching function required for efficient air hearing. The structure of the relatively enormous mallei of golden moles in these genera suggests that their ossicles are primarily adapted towards transmission of gound vibrations. A mechanism of ossicular inertia has been proposed. The electrical model predictions imply that subterranean mammals may be preadapted to evolve a more efficient mechanism for bone conduction of ground vibrations. Certain unusual features of the middle ear apparatus of other subterranean mammals may be interpreted in this light. [Work supported by B.B.S.R.C.]

10:05

2aAB6. Acoustics of the avian vocal tract. Neville Fletcher (Res. School of Physical Sci., Australian Natl. Univ., Canberra 0200, Australia, neville.fletcher@anu.edu.au) and Alex Tarnopolsky (Australian Defence Force Acad., Canberra 2600, Australia)

The avian vocal tract can appropriately be modeled on a source-filter basis, with the pressure-controlled syrinx valve providing the wideband harmonic source and the vocal tract the linear passive filter. This filter consists of four parts: the bronchial tubes above the syrinx valves, the trachea, the mouth with included tongue, and the beak. Each of these can be modeled and its behavior calculated quantitatively using standard methods [Fletcher, *Acoustic Systems in Biology* (Oxford U. P., New York, 1992)]. The long trachea defines the main resonance frequencies, upon which the bronchial resonances impose formant bands. The mouth, the cross section of which can be varied by changing the tongue position, influences particularly the frequencies of the first two tracheal resonances, but also the harmonicity of the upper system resonances. The beak behaves as a slotted tapered transmission duct, and contributes an end-correction to the system, the magnitude of which is a strong function of both beak gape and frequency. This paper reports the numerical results of such a theoretical analysis, compares these with experimental measurements upon appropriate model systems, and discusses the implications of the results for the understanding of avian vocalization.

10:25

2aAB7. Bioacoustic correlates of female choice in European starlings. Timothy Q. Gentner and S. H. Hulse (Johns Hopkins Univ., Dept. of Psych., Baltimore, MD 21218)

The elaborate structure of male European starling song suggests that it is a trait under heavy sexual selection pressures, and the results of several studies in both the laboratory and the field have identified song as the only male characteristic correlated with female choice in this species. However, previous studies have failed to isolate the bioacoustic feature(s) of song that are important for mating decisions and preferences in female starlings. An operant apparatus that reinforces a natural approach response to potential nest sites with different male starling songs is used to demonstrate that female preference is positively correlated with song bout length but not repertoire size (two features that normally covary). Further, this preference for longer song bouts is made in such a way that the absolute length of a bout cannot be the feature upon which the preference is based. A method for quantifying the temporal structure and stereotypy in male starling songs is presented using information theory, and it is shown that these measures correlate very closely with different song bout lengths. The importance of song bout stereotypy for female preference is then tested using synthetic stimuli or operant apparatus.

Session 2aAO

Acoustical Oceanography and Underwater Acoustics: Surf Zone Oceanography and Acoustics I

James F. Lynch, Chair
Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institute, 203 Bigelow Building, Woods Hole, Massachusetts 02543

Chair's Introduction—9:15

Invited Papers

9:20

2aAO1. The Scripps Pier Bubble Experiment of 1997 and overview of rip events and their effects on acoustics measurements. Jerald W. Caruthers, Paul A. Elmore, Stephen J. Stanic (Naval Res. Lab., Stennis Space Center, MS 39529), and Ralph R. Goodman (Appl. Res. Lab., Penn State Univ., State College, PA 16804)

An experiment to determine the dynamics, distributions, and acoustic effects of bubbles in shallow water just offshore from active surf was performed in the Spring of 1997. The region of interest was just North of the pier at the Scripps Institution of Oceanography. An area from a few tens of meters to a thousand meters from the beach was instrumented by several researchers. An effort was made to acquire a comprehensive dataset of coastal dynamics, bubble distributions, and acoustic propagation. A major emphasis was placed on a smaller region approximately 400 m^2 in area and about 300 m offshore, where most of the instruments were clustered. The purpose of this paper is to provide an overview of the experiment, a summary of the runs made over the several days of experiments, and a context for the observed rip events based on an acoustic device known as the ''Delta-Frame.'' Two rip events (Run 5, event starting about 14:48 PST 7Mar97 and Run 7, event starting about 15:38 PST 8Mar97) selected for coordinated analyses and presentations to follow will be highlighted. [Work supported by ONR.]

9:40

2aAO2. Vertical profiles and horizontal scales of bubble clouds in the surf zone as measured with a distributed array of upward-looking sonars. Peter H. Dahl (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98195), Jerald W. Caruthers, and Stephen J. Stanic (Naval Res. Lab., Ocean Acoust. Branch, Stennis Space Center, MS 39529)

A collaborative, multi-institute experiment was recently performed in the vicinity of the Scripps Pier, whose goal was the study of bubble-field generation, transport, and distribution, as influenced by surf zone conditions. The collaboration produced an ensemble of instrumentation at the site for measuring bubbles, ambient noise, temperature and salinity, currents, and surface waves. This paper emphasizes data gathered by APL-UW using a set of four upward-looking transducers (frequency 240 kHz), which simultaneously measured vertical profiles of acoustic volume scattering from bubbles from four locations. The primary locus of measurement activity was defined by NRL-SSC's triangular frame (Delta Frame) and the four transducer locations for the APL-UW system bracketed this region. The transport of bubbles via rip currents emerged as a key feature surf zone bubble environment. Images of volumetric backscattering strength versus time and depth reveal episodic events (by way of increased scattering level) which last O(100) s and can occur in intervals of O(10) min. Time lags for increased scattering at the four locations were consistent with a seaward flow of rip currents equal to $\sim$0.25 m/s. Other features of the data that complement the single-depth measurements made from the Delta Frame will be discussed. [Work supported by ONR.]

Contributed Papers

10:00

2aAO3. Bubble generation and dispersion in the surf zone. Svein Vagle, David Farmer (Inst. of Ocean Sci., 9860 West Saanich Rd., Sidney, BC V8L 4B2, Canada), and Grant Deane (Scripps Inst. of Oceanogr., La Jolla, CA 92093)

A combination of horizontal Doppler imaging, underwater microphotography, and *in situ* acoustical measurements are used to examine wave propagation, wave breaking, bubble cloud formation, and dispersion in the surf zone near Scripps Pier. Wave breaking is interpreted in terms of the evolution in wave properties between the end of the pier and a sensor tripod mounted in the surf zone. The position of the breaking zone is tracked from sharp discontinuities in the horizontal imaging sonar. Bubble size distribution is measured both photographically and with an acoustical resonator. From time to time the imaging Doppler sonar shows bubble clouds moving offshore from the surf. The observations provide a framework for model analysis of bubble generation, advection, and dispersion in the surf zone.

10:15

2aAO4. Bubbles and surf zone oceanography. Eric J. Terrill and W. Kendall Melville (Marine Physical Lab, Scripps Inst. of Oceanogr., 8605 La Jolla Shores Dr., M. S. 0230, La Jolla, CA 92093-0230)

Breaking waves in the surf zone entrain high-density clouds of bubbles which are subsequently transported offshore via rip currents and turbulent diffusion. The presence of the bubbles presents a complex problem for acousticians interested in the propagation of natural and man-made sound in this shallow water regime. To investigate the acoustical properties of this area, a multi-investigator experiment was performed near the Scripps Pier in March 1997. A broadband acoustic technique, developed at the Marine Physical Laboratory, was deployed in a four-channel configuration in water of 6-m depth. Each channel of the system allows the determination of the bubble size distributions at a discrete depth via inversion of the measured sound speed and attenuation. Analysis of the data shows high concentrations of bubbles advecting offshore via rip currents, resulting in acoustic attenuations ranging from O(1)–O(100) dB/m. Vertical gradients

were observed in the bubble concentrations. Simultaneous measurements of the wind, waves, and currents provide a synoptic view of the overlying oceanography during the course of the experiment. Results of the experiment will be discussed in the context of the transport of surf-generated bubbles through the measurement site.

10:30–10:45 Break

10:45

2aAO5. Void fraction and bubble distributions near surf zones. Ming-Yang Su (Oceanogr. Div., Naval Res. Lab., Stennis Space Center, MS 39529)

Extensive measurements of void fraction and bubble size density due to wave breaking both inside and outside surf zones were made by the Naval Research Laboratory during two joint experiments sponsored by ONR: the Scripps Bubbles/Acoustics Experiment at Scripps Institution of Oceanography, March 1997, and the Sandy Duck '97 Experiment at the Army CERC Field Research Facility, Duck, NC, Sept.–Oct. 1997. Using a new technique that employs swing bar void fraction staffs anchored on the sand bottom, the new data show very strong tidal effects on the vertical distributions of void fraction. Data will also be presented on the temporal evolution of bubble size density (radius $=15$–1200 μ) during local storms. Finally, some new results on direct correlation of acoustic noise intensity using a hydrophone with co-located void fraction measurement will be presented.

11:00

2aAO6. Tomographic reconstruction of evolving bubble fields in the Scripps Pier bubble experiment. Daniel Rouseff and Frank S. Henyey (Appl. Phys. Lab., College of Ocean and Fishery Sci., Univ. of Washington, Seattle, WA 98105, rouseff@apl.washington.edu)

The Delta Frame, designed and operated by NRL-SSC, was a major component of the 1997 Scripps Pier shallow water bubbles experiment. The frame supported two sources and eight receivers. Acoustic travel time and attenuation were measured at eight frequencies between 39 and 244 kHz over the resulting 16 intersecting ray paths. Paths ranged in length from 2.5 to 8.6 m. For a given transmission sequence, the measurements at a particular frequency can be combined using tomography to yield a quantitative cross-sectional image of the ensonified region. By generating images for successive sequences, the temporal evolution of the medium can be visualized. A tomography algorithm [D. Rouseff and F. S. Henyey, J. Acoust. Soc. Am. **101**, 3032 (1997)] is applied to the Delta Frame data. The images are produced for time periods during the experiment where the bubble field was probed by other independent measurements. Tomography results are compared to the other measurements. [Work supported by ONR.]

11:15

2aAO7. Effects of bubbles on high-frequency sound propagation in very shallow water. Paul A. Elmore, Jerald W. Caruthers, Stephen J. Stanic (Naval Res. Lab., Stennis Space Center, MS 39529), and Ralph R. Goodman (Appl. Res. Lab., Penn State Univ., State College, PA 16804)

The results of an experiment performed off the research pier at the Scripps Institute of Oceanography that measured the acoustic effects of small bubbles in very shallow water (6 m) are discussed. The results presented are in coordination with researchers who conducted other measurements simultaneously. During the experiment, rip currents passed through a field of measurement instruments 300 m offshore. These rip currents were laden with bubbles created in the surf between the instruments and the shore. Pulse signals, between 39 and 244 kHz, were propagated to 10 m through the bubble clouds. The effects of these rip currents on the spatial distributions of the resulting acoustic attenuation are discussed. From the attenuation data, bubble populations are calculated by an iterative procedure based on the well-known resonant bubble approximation. [Work supported by the Office of Naval Research.]

11:30

2aAO8. A novel approach for obtaining bubble dissolution measurements at sea. Russell Costa (Naval Undersea Warfare Ctr., Newport, RI 02841), Ronald A. Roy (Boston Univ., Boston, MA 02215), Kerry W. Commander, Victor B. Johnson, and Andrew C. Colbert (Naval Surface Warfare Ctr., Panama City, FL 32407-7001)

A hybrid system for measuring the dissolution of gas bubbles in an oceanic environment was developed utilizing the Coastal Systems Station's Light Scattering Bubble Counter (LSBC). The LSBC independently measures flow velocity and bubble size, and is capable of distinguishing bubbles from particulate matter. In the hybrid system, the dynamic range of the LSBC measurements is 50 μm to about 300 μm in radius. The system is filled at depth (30 feet maximum without internal pressurization) and initially purged of all free gas. Pneumatically controlled valves are used to establish a closed recirculating system in which a flow can be generated in either direction using a large-volume peristaltic pump driven at low-volume velocity. Clear air is forced through a ceramic fritted disk to create the initial bubble population. This bubbly mixture is recirculated through the system at a nominal flow rate of 22 l/min, which corresponds to a velocity of 30 cm/s through the LSBC aperture. The bubble radii are continuously measured using the LSBC, resulting in time-dependent bubble size distributions. These distributions are used to estimate bubble dissolution rates for the size range indicated. [Work supported by ONR.]

11:45

2aAO9. Analysis of free bubble cloud scattering data from the Lake Seneca experiment. J. Gregory McDaniel, Ronald A. Roy, Andrew D. Gephart (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215), and William M. Carey (MIT, Cambridge, MA 02139)

In the fall of 1991, Carey, Roy *et al.* conducted a series of experiments at the Seneca Lake test facility which measured the frequency-dependent scattering from submerged bubble clouds. Preliminary results reported by Roy *et al.* [J. Acoust. Soc. Am. **92**, 2993–2995 (1992)] showed that certain features of the high-frequency backscatter data were consistent with single-bubble scattering while the low-frequency results were consistent with collective oscillations of the bubble cloud. A monopole resonance scattering peak was evident that scaled with the bubble cloud volume and the mixture sound speed. Additional analysis showed that the evolving target strength of the cloud as it rose through the insonating acoustic beam showed signs of an interference effect [Carey and Roy, Proc. 2nd Euro. Conf. on Underwater Acoust., European Commission, Luxembourg, 207 (1994)] which could explain the differences between calculated and measured scattering strengths. Advanced signal processing and modeling techniques have been developed to enable the removal of experimental artifacts caused by: (1) uncertainties in the range geometry, (2) coherent electrical noise in the processing electronics, and (3) acoustic reverberation within the test volume. The resulting corrected backscattering target strengths will be reported and discussed. [Work supported by ONR.]

12:00

2aAO10. Sound radiation by a bubble encountering a vortex ring. Ali R. Kolaini (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677, ali@sparc.ncpa.olemiss.edu)

The dynamics of a bubble and subsequent acoustic radiation interacting with a vortex ring created by an underwater nozzle are studied both theoretically and experimentally. The velocity field of the flow containing the bubble may be represented as a ''basic flow'' related to the vortex ring and a ''secondary flow'' related to the bubble motion [G. Chahine, 20th Symposium on Naval Hydrodynamics (1994)]. The ''bubble flow'' is assumed to be irrotational satisfying the Laplace equation. The vortex related velocities are given by the stream function and are taken into account by imposing ''correct'' kinematics and dynamic boundary conditions at the bubble surface. The interaction of the bubble with a vortex ring has been shown to cause severe deformation. Under special circumstances the deformed bubble may be at resonance with these vortex filaments and

force it to radiate sound to the far field. The boundary integral technique (BIT) is employed to solve governing equations. The BIT is utilized to simulate the ''bubble flow,'' its surface distortions and subsequent acoustic radiation. The critical parameters responsible for the acoustic re-excitation of the bubble encountering the vortex filaments are discussed in some details. Some conclusions are arrived at on the acoustical characteristics of the bubble encountering a fully developed turbulent flow field. [Work supported by ONR.]

TUESDAY MORNING, 23 JUNE 1998 GRAND BALLROOM A (S), 9:15 A.M. TO 12:05 P.M.

Session 2aBV

Bioresponse to Vibration/Biomedical Ultrasound and Physical Acoustics: High Intensity Therapeutic Ultrasound I

Pierre D. Mourad, Cochair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

Gail R. ter Haar, Cochair
Physics Department, Royal Marsden Hospital, Sutton, Surrey SM2 5PT, England

Chair's Introduction—9:15

Invited Papers

9:20

2aBV1. Biological effects of ultrasound. Wesley L. Nyborg (Phys. Dept., Univ. of Vermont, Burlington, VT 05405)

Ultrasound has considerable potential as an agent for modifying cells and tissues. This comes partly from its capability to produce temperature elevation in desired spatial and temporal patterns. It also comes partly from its capability to produce cavitation in one or more of its diverse forms and, in addition, partly from the action of radiation forces, radiation torques, and acoustic streaming. In some situations, several mechanisms are involved and may interact with each other; for example, cavitation can produce heat in a medium, and the resulting temperature rise will affect the cavitation activity. Understanding these and other mechanisms has been, and promises to be, helpful in developing applications of ultrasound to therapy, surgery, and biotechnology, as well as in formulation of criteria for safety in medical procedures involving ultrasound. In this review, examples will be given to which the mechanisms are relevant, and recent developments will be discussed.

9:40

2aBV2. Therapeutic ultrasound. Lawrence A. Crum, Michael Bailey, Kirk Beach, Michael Caps, Stephen Carter, Wayne Chandler, Robin Cleveland, Scott Helton, Peter Kaczkowski, George Keilman, Roy Martin, Pierre Mourad, Ronald Roy, Udo Schmiedl, Brent Stewart, and Shahram Vaezy (Univ. of Washington, Seattle, WA 98105)

Although diagnostic ultrasound is now a multibillion dollar a year industry, the therapeutic applications of ultrasound were initially thought to have more promise. Indeed, the Fry brothers were performing brain surgery on human patients in the early 1950s, and with considerable success. However, therapeutic ultrasound has still not yet gained its promising potential. Recently, there has been an enormous resurgence of interest in this topic, and several major projects are underway at various facilities around the world. Our own substantial efforts in this area at the University of Washington are reviewed, with brief discussions of topics in lithotripsy, ultrasound-enhanced drug delivery, HIFU surgery, and acoustic hemostasis. [Work supported in part by DARPA and the NIH under Grants DK-43881 and R43HL57739.]

10:00

2aBV3. Acoustic liver cauterization: A potential tool for bloodless surgery. Roy Martin (Anesthesiology and Bioengineering, Box 356540, Univ. of Washington, Seattle, WA 98195, rmartin@u.washington.edu), Shahram Vaezy, Scott Hilton, Michael Caps, Peter Kaczkowski (Univ. of Washington, Seattle, WA 98195), George Keilman (Sonic Concepts, Woodinville, WA 98072), Steve Carter, Wayne Chandler, Pierre Mourad, Kirk Beach, and Lawrence Crum (Univ. of Washington, Seattle, WA 98195)

Liver surgery is hampered by the difficulty in controlling hemorrhage, particularly the slow oozing from injured hepatic surfaces. Conventional cauterizing fails in this situation. Bleeding from incisions made in animal livers using high-intensity focused ultrasound (HIFU) by volume cauterizing along the surface of the incision has been successfully stopped. The question arose: Could bleeding be avoided during surgery by precauterizing a thick surface around a region to be resectioned prior to the actual resectioning? The livers of anesthetized rabbits (4) and pigs (3) were treated in this manner under a protocol approved for animal studies at the University of Washington. Two HIFU transducers (3.3 and 3.5 MHz focused at 3.5 and 5.5 cm, respectively) were used for the treatment.

Cauterized strips around a region appeared white. The tissue to be resectioned was removed by cutting along the distal border of the strips. No oozing from the liver parenchyma was observed. Bleeding occurred only in the larger vessels (>2 mm inside diameter) present in the cauterized region. Since these size vessels are easily ligated the method appears to be clinically feasible. Survival studies are needed to learn about the possibilities of rebleeding after abdominal closure or unknown healing difficulties.

Contributed Papers

10:20

2aBV4. High-intensity focused ultrasound transducer design for surgical and hemostatic applications. George W. Keilman (Sonic Concepts, 20018 163rd Ave. NE, Woodinville, WA 98072) and Peter J. Kaczkowski (Univ. of Washington, Seattle, WA 98105-6698)

A resurgence in interest in the application of high-intensity focused ultrasound (HIFU) to surgical problems is due in part to significant improvements in imaging modalities for targeting the delivery of ultrasound. The wide variety of potential applications under consideration necessitates a refinement of transducer design to accommodate often conflicting constraints. The need to deliver substantial power to a small target zone requires high focal gain to minimize collateral damage, and yet must be balanced with the need for small transducers in many procedures. In general, the choice of operating frequency, transducer aperture and focal depth, convenient acoustic coupling methods, waveform selection, and amplifier characteristics must all be adjusted to best fit the application. This presentation will discuss such tradeoffs in any HIFU transducer design, and will also address particular techniques for coupling and cooling transducers which deliver hundreds of watts of acoustic power. Transducer efficiencies of over 85% have been routinely achieved by judicious selection of matching layers and tuning networks, while maintaining over 50% relative bandwidth, for center frequencies spanning the range from 1 to 10 MHz. Transducer requirements for array-based systems, motivated by clinical needs, will also be discussed. [Work funded by DARPA/ONR.]

10:35

2aBV5. Acoustic field of a device for noninvasive pressure wave therapy in orthopedy. Eckard Steiger, Thomas Dreyer (Inst. für Hoechstfrequenztechnik und Elektronik/Akustik, Univ. Karlsruhe, D-76128 Karlsruhe, Germany, eckard.steiger@etec.uni-karlsruhe.de), and Juergen Mayer (Storz Medical AG, CH-8280 Kreuzlingen, Switzerland)

The Storz Minilith SL1 is a device specially designed for the treatment of orthopedical indications as tendinosis calcarea, tennis or golf elbow, heel spurs, or supra spinatus syndrome. A cylindrical electromagnetical source generates an intense ultrasonic pulse, which is focused by means of a parabolical reflector. Measurements in the focal region are performed in degassed water using a fiber-optic probe hydrophone. The whole process of pulse radiation from the source over nonlinear propagation including reflection to the focusing is simulated too. The underlying full wave model is based on a specially developed finite difference in time domain method. Excellent agreements of measured and computed pressure wave profiles in the focal region validate the model. The generation of shock waves in dependence on the energy setting of the therapy device is studied. Further investigations are made to consider the interaction of the SL1-pulse with human tissue structures as they occur in medical applications.

10:50

2aBV6. *Ex vivo* studies of the effect of high-intensity focused ultrasound on whole human blood. Sandra Poliachik, Susanah Bloch, Michael Bailey, Peter Kaczkowski, Robin Cleveland, Pierre D. Mourad, Tyrone Porter, Lawrence A. Crum (Univ. of Washington, Appl. Phys. Lab, 1013 NE 40th St., Seattle, WA 98105, pierre@apl.washington.edu), Wayne L. Chandler (Univ. of Washington, Lab. Medicine, UW Medical Ctr.), and George W. Keilman (Sonic Concepts, Woodinville, WA 98072)

A surgical application of high-intensity focused ultrasound (HIFU) is hemostatis. In this application, both tissue and blood will be subjected to HIFU. What will happen to blood? The intent of this study is to explore via experiment some of the bioeffects of HIFU on whole human blood. A 1.1-MHz piezoceramic source with a 60-mm radius of curvature and 100-mm aperture was used. Special acoustically transparent containers were made of handmade, low-density polyethylene tubing. Phenomena studied included hemolysis, morphologic changes in erythrocytes, white blood cell differential, and platelet activation, all as a function of acoustic intensity and dosage. Intensities up to 6000 W/cm^2 were used in cw pulses of one second. The bioeffects noted above were correlated with temperature and a passive acoustic measure of bubble activity within the blood created by the ultrasound. Under the acoustic conditions considered for this work, nonthermal acoustic effects were most strongly associated with bioeffects. [Work sponsored by DARPA.]

Invited Papers

11:05

2aBV7. Shaping of focused ultrasonic beams to expedite thermal necrosis in tumor therapy. Frederic L. Lizzi (Riverside Res. Inst., 330 West 42nd St., New York, NY 10036, lizzi@rrinyc.org)

Ultrasonic tumor therapy employing focused beams shows great promise for treating tumors by producing spatial matrices of thermal-necrosis lesions. A practical problem involves the long time required to treat extended tumor volumes. To treat ocular tumors, our laboratories are investigating beams with asymmetric shapes to expedite the production of lesion matrices. The beams are generated using rectangular strip electrodes plated on spherical-cap ceramic transducers. The phase and amplitude of the strip excitation voltages are controlled to launch beams whose focal-zone patterns consist of a sequence of lobes confined within elliptical, rather than circular, cross sections. A simple theoretical model has shown that thermal diffusion between these lobes can produce lesions with elliptical cross sections; lesion matrices would be readily produced by translating such transducers along the short axis of the lesion cross section, thereby causing necrosis in broad regions. The concept has been verified in *in-vitro* experiments (in liver and the ocular lens) and in *in-vivo* exposures of human tumor explants in nude mice. [Work supported by NIH Grant EY10369.]

11:25

2aBV8. An image-guided ultrasound phased array system for noninvasive surgery. Emad S. Ebbini, Philip D. VanBaren, and Claudio Simon (Dept. of EECS, Univ. of Michigan, Ann Arbor, MI 48109-2122)

An integrated image-guided therapeutic phased array system for noninvasive surgical applications is being developed. The therapeutic array utilizes piezocomposite transducer technology and operates (therapeutically) at 1 and 2 MHz. It has 64 elements on a spherical shell with a geometric center at 100 mm from its apex. The array was shown to be capable of producing well-defined thermal lesions in tissue media at depths from 40 to 60 mm and to scan therapeutic foci up to ±15 mm from its geometric center.

Image guidance is provided by a modified diagnostic ultrasound scanner which, in addition to providing standard B-scan images of the target region, provides real-time images of the temperature rise due to the therapeutic beam. The temperature information is obtained using a correlation based algorithm for echo displacement estimation, which can be directly related to local variation in tissue temperature due to the therapeutic beam. A complete description of the system will be given along with illustrative examples of image-guided tissue ablation and temperature estimation and control. In addition, experimental data demonstrating the use of the therapeutic array in an imaging mode will be presented. [Work funded by NIH Grant CA66602.]

11:45

2aBV9. MRI monitoring and control of focused ultrasound surgery. Kullervo Hynynen, Nathan McDannold, Todd Fjield, and Douglas Daum (Div. of MRI, Dept. of Radiology, Brigham and Women's Hospital, Harvard Med. School, 221 Longwood Ave., Boston, MA 02115)

The clinical and animal tests of focused ultrasound surgery have shown significant variation between the effectiveness of the same power sonications. Therefore, it has become important to seek methods to monitor the *in vivo* ultrasound exposures. One of the most promising methods is the use of magnetic resonance imaging (MRI) to detect the temperature elevation. In this paper the calibration and accuracy of MRI thermometry will be presented. The *in vivo* mapped temperature distributions allow the thermal exposure of the tissue to be estimated. These estimates have been compared with the tissue coagulation produced in muscle and brain tissues *in vivo*. The results show that the lesion boundary is predicted with the resolution of the MRI voxels. This has significant clinical value since the MRI thermometry can be used to control the treatment on-line. These experiments allowed comparison of the accuracy of the theoretical models for predicting the temperature elevation and tissue necrosis *in vivo*. Similarly, the MRI thermometry was used to investigate the temperature buildup during multiple sonications and evaluate the performance of new applicators such as phased arrays *in vivo*.

TUESDAY MORNING, 23 JUNE 1998 EAST BALLROOM A (S), 9:15 TO 11:45 A.M.

Session 2aEA

Engineering Acoustics: Ultrasonic Transducers for Applications in Air

Dehua Huang, Chair
Arichell Technologies, Inc., Border Street, Newton, Massachusetts 02165

Chair's Introduction—9:15

Invited Papers

9:20

2aEA1. Influence of air loading and dissipation on plate transducer amplitude. Arthur Ballato (U.S. Army Commun.-Electron. Command, Fort Monmouth, NJ 07703-5201, a.ballato@ieee.org)

Air-loaded plate transducers driven piezoelectrically are used in a variety of modern applications such as optical modulators, microactuators, and certain types of resonant acoustic sensors and frequency control elements. For these uses it is necessary or desirable to know the maximum mechanical amplitudes obtainable, and how these are limited by the various internal dissipation mechanisms and the loading of the ambient fluid. Internal conduction, dielectric, and acoustic viscosity losses are incorporated, and, by introducing of a complex piezoelectric coupling coefficient, changes in critical frequencies and amplitudes are traced with variations in type and size of dissipation mechanism, and the ambient loading. The analysis also shows that the unavoidable interfacial stresses existing between the crystal plate surfaces and the exciting electrodes can become exceedingly large at the frequencies commonly used for frequency control of cellular communications. Because the various devices covered usually form components of an overall circuit design, a number of alternative equivalent electrical circuits are also provided whereby the physics of these structures may be rendered and readily interpreted in network terms.

9:45

2aEA2. Absolute amplitude measurements using a capacitive receiver. M. A. Breazeale (Univ. of Mississippi, Natl. Ctr. for Physical Acoust., University, MS 38677)

Absolute acoustical amplitudes can be measured in solids with a capacitive receiver. Amplitudes as small as 10^{-4} Å$=10^{-14}$ M have been measured. A system is described with dimensions used for frequencies between 5 and 60 MHz. Two special innovations are described: low-temperature measurements which require adjustment of the spacing of the capacitor and high-sensitivity measurements (amplitudes smaller than 10^{-14} M) which are accomplished by using a dielectric to prevent electrical breakdown.

10:10

2aEA3. Silicon micromachined ultrasonic transducers. B. T. Khuri-Yakub (Edward L. Ginzton Lab., Stanford Univ., Stanford, CA 94304-4085, khuri-ya@ee.stanford.edu)

This paper presents a novel type of ultrasonic transducer that is made by silicon micromachining. Transducers for both air-borne and immersion applications are made from parallel plate capacitors whose dimensions are controlled through traditional integrated circuit manufacturing methods. Typical dimensions of the capacitors are: gap =1 micron (vacuum or air), membrane thickness =1 micron (silicon nitride or polysilicon), and diameter =50 microns. A large number of small elements are connected in parallel to make a transducer. Transducers for air-borne ultrasound applications have been operated in the frequency range of 1–11 MHz, while immersion transducers have been operated in the frequency range of 1–20 MHz. Theoretical modeling calculations of the transducers and their electrical input impedances will be presented. The theory predicts a dynamic range of over 140 dB for both air and immersion applications. Different processing methods for fabricating the transducers and the merits of the different approaches will be presented. Finally, the results of transducer operation both in air and water with over a 100-dB dynamic range will be presented.

10:35

2aEA4. Air transducers with high acoustic impedance. Lawrence C. Lynnworth (Panametrics, Inc., 221 Crescent St., Waltham, MA 02154, lynnworthl@panametrics.com)

High-Z air transducers evolved from several industrial transducers. These predecessors include (a) acoustic emission and angle-beam NDT contact transducers; (b) flowmeter transducers for high-pressure methane, hot refinery gases, and corrosive flare gases; and (c) transducer arrays for monitoring hot corrosive gases flowing at Mach 0.1 in smokestacks of ϕ3–13 m. This peculiar ancestry accounts for their unusual high-acoustic-impedance construction. By not employing low-Z air-backed radiating membranes, transducer bandwidth, response time, and sensitivity are sacrificed. The solid, robust construction, however, offers some compensating features: tolerance to wide ranges in pressure and temperature, including rapid rates of change (thermal shock, depressurization); corrosion resistance; operable with standard lab equipment, without bias voltage; mounting options where the transducer forms part of the pressure boundary, or where it is outside the pressure boundary. Clamp-on air flow applications at one bar include small wind tunnels and plastic pipes. Acoustically isolated pairs measure secondary flow components (crossflow, circulation) in a plane perpendicular to the pipe axis. Flow applications include gases such as air, mild steam (pending) or other hot, pressurized, or corrosive fluids. Air-coupled measurements of transmission characteristics of wood or other low-density solids, and air-ranging, are readily demonstrated with the low-Z transducers.

Contributed Papers

11:00

2aEA5. Ultrasonic capacitance transducers for flow and temperature tomography in air. David A. Hutchins (Dept. of Eng., Univ. of Warwick, Coventry CV4 7AL, UK, dah@eng. warwick.ac.uk), William M. D. Wright (Univ. College Cork, Cork, Ireland), David W. Schindel (N.R.C., Ottawa, ON K1A 0R6, Canada), and Peter W. Carpenter (Univ. of Warwick, Coventry CV4 7AL, UK)

Capacitance transducers have been constructed that can operate at frequencies of up to 2 MHz in air, over a wide bandwidth, and with good efficiency. These have been used to propagate signals in air across regions containing variations in air flow and temperature. A source and receiver pair were scanned along a set of projections under computer control, allowing various ray paths across a chosen cross section to be selected. For each location, the received waveform was digitized and recorded, and the propagation time determined from cross-correlation techniques. The data were then used to form tomographic images, which corresponded to variations in either temperature or flow speed, both of which affect the local acoustic velocity. In one experiment, an air-jet was imaged at various distances from the nozzle. In another, the temperature variations in air above a hot metal tip were determined. Images will be shown for both sets of experiments, and the usefulness of the technique discussed.

11:15

2aEA6. Ultrasonic air-coupled capacitance transducers using thin mica films. Paul Ingleby (Dept. of Eng., Univ. of Warwick, Coventry CV4 7AL, UK, es2041@eng.warwick.ac.uk), David A. Hutchins, Andrew G. Bashford (Univ. of Warwick, Coventry CV4 7AL, UK), and David W. Schindel (N.R.C., Ottawa, ON K1A 0R6, Canada)

Micromachined air-coupled capacitance transducers have been manufactured using anisotropically etched silicon backplates and commercially available mica films. Devices developed using 20-μm-thick insulating layers have yielded frequencies of up to 400 kHz. However, new devices using mica with a thickness of approximately 4 μm will deliver characteristics similar to existing transducers employing polymeric dielectric films. This will result in wideband behavior and bandwidths extending from approximately 100 kHz to 2 MHz. Mica-based devices have several advantages over capacitance transducers with polymer films due to the robust physical characteristics exhibited by this type of material. Improved thermal characteristics and lower porosity allow this type of device to be employed in a variety of conditions in which the performance of the polymer films would be impaired. Results of experiments performed in air at a range of temperatures from ambient to 400 °C will be presented. This will demonstrate the versatility of such devices in terms of both sensitivity and bandwidth.

11:30

2aEA7. Recent developments in air-coupled ultrasonic capacitive transducers for NDE. Craig S. McIntyre, David A. Hutchins (Dept. of Eng., Univ. of Warwick, Coventry CV4 7AL, UK, esrna@eng.warwick.ac.uk), and David W. Schindel (Inst. for Aerosp. Res., Natl. Res. Council, Ottawa, ON, Canada)

A new design of an air-coupled capacitive ultrasonic transducer has been developed that could have potential for use in NDE experiments, particularly when the pulse-echo method is employed. The design is based on a standard capacitive transducer, except the newly developed device has two concentric, backplate elements with a common membrane. Each of the elements is electrically isolated from the other, allowing one to be used as a source of ultrasound, while the other is used as a detector. This yields a device containing both a source and receiver of ultrasound within the same structure. Experiments have shown that the device performance depends on which of the elements is selected to be source and receiver, size of the bias voltage employed, surface of the backplates, and membrane material utilized. Initial studies have shown that the device is capable of detecting the front face echo from various material samples, and these results are contained within this paper. There has also been some investigation of the signal quality when capacitive devices have been placed within an enviroment containing a range of pressures above standard atmospheric pressure. Applications of this work include measurements in high-pressure gas pipelines.

Session 2aMU

Musical Acoustics: Physics and Materials of Musical Instruments I

Shigeru Yoshikawa, Chair
Department of Acoustical Design, Kyushu Institute of Design, 4-9-1 Shiobara, Minami-Ku, Fukuoka, 815 Japan

Invited Papers

7:45

2aMU1. Fluid dynamic aspects of human voice and brass instruments: Implications for sound synthesis. Pelorson Xavier (Institut de la Commun. Parlee, ESA CNRS 5009, 46 av. F. Viallet F-38031 Grenoble, France), Msallam Regis (Institut de Recherche et de Coordination Acoustique Musique, F-75004 Paris, France), Gilbert Joel (Institut d'Acoustique et de Mecanique, F-72017 Le Mans Cedex, France), and Hirschberg Avraham (Tech. Univ. of Eindhoven, NL-5600 MB Eindhoven, The Netherlands)

The challenge of physical models of musical instruments is to provide a global description of a highly complex instrument which has to be accurate but simple enough to be numerically implemented. For this purpose, a logical approach is to extract, from a theoretical or experimental study, what are the most important phenomenon involved which are thus to be taken into account. Two examples of such an approach are given by the physical models of human voice and brass instruments presented here. Although both instruments share in common the same basic principle, a self-oscillating source coupled with a resonator, crucial physical aspects are quite different. It is shown that, for brass instruments, the lips' self-sustained oscillations are mainly driven by the acoustical coupling, therefore the simpliest one degree of freedom model for the lips is sufficient. In the case of voicing, the acoustical feedback, although effective, is not sufficient to maintain self-sustained oscillations. A more elaborated vocal-folds model is needed. Regardless of the sound source, the description of the resonator is also discussed. While, for voice simulations, significant improvement can be gained using a two-dimensional linear theory, nonlinear propagation is sometimes to be taken into account for brass instruments.

8:10

2aMU2. The effect of wall materials on the timbre of brass instruments. Robert W. Pyle, Jr. (11 Holworthy Pl., Cambridge, MA 02138-4509)

It has long been thought by players and instrument makers that the material from which a brass instrument is made influences the tone quality and playing characteristics of the instrument. For instance, the brochure of a maker of custom trombones describes bells made of three alloys: yellow brass (70% Cu, 30% Zn), having the ''clearest sound, sharpest articulation,'' gold brass (85% Cu, 15% Zn), with ''warm sound, rounded articulation,'' and red brass (90% Cu, 10% Zn), with ''warmer sound, covered articulation.'' The bells are made in three metal thicknesses: light weight, with ''flexible tone, rapid response,'' standard weight, with ''balanced tone, even response,'' and heavy weight, with ''firm tone, stable response.'' Most players agree with these statements. This model of trombone is made so that bell sections can be interchanged while leaving the remainder of the instrument intact. The present paper reports the results of measurements under playing conditions on several different bells. Comparisons of steady-state behavior are made via the transfer function between sound pressure in the mouthpiece cup and an on-axis point outside the bell. An attempt will also be made at the rather more difficult task of characterizing the attack transient.

8:35

2aMU3. The effect of wall properties on the timbre of organ pipes. Judit Angster (Fraunhofer-Institut fur Bauphysik, Nobelstr. 12, D-70563 Stuttgart, Germany, angster@ibp.fhg.de), Gyvrgy Paal, Walter Garen (Fachhochschule Ostfriesland, Constantiaplatz 4, D-26723 Emden, Germany), and Andras Miklss (Fraunhofer-Institut fur Bauphysik, D-70563 Stuttgart, Germany)

The influence of the wall vibrations on the timbre of flue organ pipes is one of the most disputed questions of organ building. Wall vibrations of metal and wooden labial pipes have been investigated by using miniature piezoelectric accelerometers and optical noncontact vibrometers. Both free and forced vibrations have been studied. Two sets of diapason pipes that were built with different wall properties but identical dimensions have been voiced to very similar speech and timbre, then vibration and sound spectra and attack transients were recorded. While large differences have been found in vibration spectra, hardly any change in the sound signal was observed. Similar investigations have been performed on wooden pipes. In this case the voicer ''tuned'' the side wall resonances in order to make the sound richer in overtones. The effect of this ''tuning'' on the vibration and sound spectra has been investigated. The ''tuning'' resulted in a significant increase of the vibration amplitude, while the radiated sound-pressure level decreased and the overtones became relatively stronger.

9:00

2aMU4. Acoustically driven sinuous instability of a planar air jet. A. W. Nolle (Dept. of Phys., Univ. of Texas, Austin, TX 78712-1081)

The sinuous instability wave of a planar air jet is excited by localized acoustic flow across the nozzle. Phase velocity and the growth exponent are found from synchronous hot-wire measurements made beyond the excited region, where the profile is approximately sech squared. In the observed range of scaled radian frequency, 0.02–1.33 (the stability limit), results agree with real-

frequency (spatially growing) analysis but not with complex-frequency (temporally growing) analysis, which predicts smaller phase velocity at low frequencies. In further tests, the acoustic driving signal is independent of downstream distance, as in an organ pipe. The jet deflection is then the sum of acoustic convection and of the instability wave, summing to zero at the nozzle, as proposed by Fletcher, Elder, and others. The instability-wave theory applies to linear behavior in the inviscid limit and therefore to a hypothetical nonspreading jet. The local velocity profile width must be considered in relating to a physical jet. In a flue organ pipe oscillating at equilibrium amplitude the jet deflection at the lip exceeds the range of validity. Direct sound generation by the jet is investigated briefly.

9:25–9:35 Break

9:35

2aMU5. Jet wave amplification in organ pipes. Shigeru Yoshikawa (Dept. of Acoust. Des., Kyushu Inst. of Design, 4-9-1 Shiobara, Minami-Ku, Fukuoka, 815 Japan)

The amplification factor of an air jet waving over the mouth is estimated from flow visualization of the smoked jet, where the velocity ranges from 7–33 m/s, the Reynolds number from 1000–5000, and the sounding frequency from 130–580 Hz. The jet length is 15.8 and 10.2 mm in two models, while the jet thickness is 2.2 mm in both. The estimation is carried out by comparing the visualized envelope of jet waves with a simplified theoretical one. The envelope is visualized by the OR summation of the digitally memorized frames showing the instantaneous deflection taken with a speed of 1440 frames per second. Furthermore, the strength of mouth field defined as the acoustic displacement amplitude is derived from the envelope-based estimation. A hot-wire anemometer is then applied to measure this mouth-field strength, whose value shows a good agreement with the estimated result. The amplification factor of organ jets was estimated as 0.26 to 0.18 (1/mm); the mouth-field strength as 0.5 to 1.5 mm. A nondimensionalized representation of the relation between the amplification factor and the blowing velocity in our organ-pipe jets roughly agrees with the theoretical curve based on the linearized spatial analysis of jet instability.

Contributed Papers

10:00

2aMU6. Interpretation of shear material properties of vocal fold mucosal tissues with Fung's quasilinear viscoelastic theory. Roger W. Chan and Ingo R. Titze (Natl. Ctr. for Voice and Speech, Dept. of Speech Pathol. and Audiol., The Univ. of Iowa, Iowa City, IA 52242, roger-chan@uiowa.edu)

Fung's quasilinear viscoelastic theory (QLV) [Y. C. Fung, *Biomechanics* (Springer-Verlag, New York, 1993), pp. 277–292] has been successfully applied to model the time- and history-dependent properties of a variety of biological soft tissues in the passive state. In particular, the hysteresis or damping curves of many soft tissues remain relatively insensitive to the rate of deformation (strain rate or frequency) across several decades of variation. This insensitivity to frequency cannot readily be accounted for by traditional discrete models of linear viscoelasticity. Rather, Fung's QLV provides a mathematically powerful framework in which a continuous stress relaxation spectrum can be used to model the relatively flat damping characteristics. The shear properties of human vocal fold mucosal tissues were measured with a rotational rheometer. Data on complex dynamic shear modulus, complex dynamic viscosity, and damping ratio as functions of frequency were obtained. It was found that vocal fold mucosal tissues show the shear-thinning behavior and have a relatively flat damping curve typical of biological soft tissues. Theoretical predictions based on the QLV matched well with the empirical data. The results suggest that the vocal fold mucosa can be modeled by the QLV under small-amplitude oscillation conditions. [Work supported by NIH Grant No. P60 DC00976.]

10:15

2aMU7. Theoretical and experimental investigation of the air-driven free reed. James P. Cottingham (Phys. Dept., Coe College, Cedar Rapids, IA 52402)

The free reed is the excitation mechanism for several families of instruments, including the reed organ and harmonium, the harmonica, the accordion and concertina, and the Asian free reed mouth organs. While most studies of reed behavior have concentrated on reeds which are either inward-striking (closing) and outward-striking (opening), free reeds do not fit neatly into either category and can exhibit behavior characteristic of either opening or closing reeds, depending on certain details of the reed configuration. Some characteristics of free reed behavior, in particular the relation between frequency and blowing pressure, have been modeled by adapting Fletcher's analysis of musical instruments driven by a reed mechanism [N. H. Fletcher, Acustica **43**, 63–72 (1979)]. In addition, a variety of measurements of reed motion have been made, including experimental determination of measurements of reed position and phase using proximity sensors. These are compared with observations of the near-field sound as well as with predictions based on theoretical considerations.

10:30

2aMU8. Material and obsolescence on flute tone quality. Alessandro Cocchi and Lamberto Tronchin (DIENCA— CIARM Univ. of Bologna, Viale Risorgimento, 2 40136 Bologna, Italy, tronchin@ciarm.ing.unibo.it)

The rule of material and obsolescence on flute tone quality has been discussed for a long time among musicians and flute makers. Mozart himself made use of crystal flutes in some of his works (Rondo', K617, Adagio, K356). From an acoustical point of view, since Boehm's work of 1832, the effect of material on flute tone quality has not been investigated, until researches of recent years [Coltman (1971); Chapuis (1991)]. In this work the research has been carried out measuring impulse responses inside the tube of the flute. Two different flutes have been utilized, i.e., a light-alloy flute and a silver flute. The MLS sequences, with homemade software, have been utilized, gathering impulse responses in two different positions in the tube, in order to verify the variation of sound field. From the impulse responses a frequency analysis has been obtained, which pointed out that the silver flute contains more high frequency and shorter transient than the light-alloy flute. Starting from the impulse responses, sound example of difference in sound quality has finally been obtained.

Session 2aNSa

Noise and Architectural Acoustics: Acoustics of Buildings

Mei W. Wu, Cochair
Colin G. Gordon and Associates, 411 Borel Avenue, Suite 425, San Mateo, California 94402

Colin G. Gordon, Cochair
Colin G. Gordon and Associates, 411 Borel Avenue, Suite 425, San Mateo, California 94402

Chair's Introduction—8:25

Contributed Papers

8:30

2aNSa1. Active attenuation of fan noise in an air-conditioning duct. Ondrej Jiricek and Petr Konicek (CTU-FEE, Dept. of Phys., 166 27 Prague, Czech Republic)

In this paper, the results of an investigation on using ANC systems in a duct will be presented. This system was used to reduce the noise of the axial-flow fan used as the noise source. The system was based on the adaptive digital signal processing controller. The attenuation of noise at low frequencies was achieved. This research represents a part of the project focused on the improvement of working conditions in operating theaters.

8:50

2aNSa2. Studies on a new type of silencer for air ducts. Fathy B. Shenoda (Dept. of Acoust., Natl. Inst. for Standards, Giza-Egypt, P.O. Box 136 Giza, Code No. 12211) and Reda N. Haroun (Dept. of Acoust., Natl. Inst. for Standards, Giza, Egypt)

The research work deals with a study of a specific type of silencer previously designed and tested by Shenoda. It consists of a rectangular duct, which is treated at its side walls with a gradually variable flow resistance. It was realized by increasing the perforated area along the duct wall combined with an homogeneous sound absorbing material. Based on the assumptions that only plane waves propagate in the duct, ignoring the effect of the mean flow velocity (very low) and applying linear acoustic equations, the sound propagation in such a silencer was studied. Therefore the wave impedance, the sound reflection coefficient at the input of the silencer, and the transmission loss introduced by it were calculated, measured, and compared with a silencer of comparable dimensions which is treated with a constant flow resistance along its length. To optimize the design different configurations were theoretically studied, computed, and compared. An important result: The silencer performance depends on a unique design parameter, namely, the silencer characteristic area. The methods of realizing a variable wall flow resistance were introduced. The effect of using different kinds of sound absorbing materials and changing the graduality of the variable flow resistance were measured.

9:10

2aNSa3. Noise of cleanroom recirculation systems. Mei Q. Wu and Colin G. Gordon (Colin Gordon & Assoc., 411 Borel Ave., Ste. 425, San Mateo, CA 94402)

Recirculation air systems are used by the microelectronics industry as the primary source of unidirectional air flow in cleanrooms. The commonly used recirculation systems include: (1) packaged air handlers, which include centrifugal or plug (plenum) fans, generally located at fan deck level above the cleanroom; (2) fan-tower recirculation systems, consisting, most often, of vertically mounted vaneaxial fans arranged along the cleanroom perimeter; (3) individual fan-filter units, in which the HEPA or ULPA filters are combined with small direct drive blowers, supported by the ceiling suspension system; and (4) minienvironments, which create superclean conditions around individual tools within a less clean ''ballroom.'' In this paper, the commonly used recirculation systems will be described, typical sound power levels and spectrum characteristics of the systems will be presented, the natural noise attenuation along the air flow paths will be listed, the theoretical models for predicting noise levels in a cleanroom will be introduced, and noise mitigation means which can be integrated into different systems will be discussed. Also presented will be the typical noise levels measured in operating cleanrooms.

9:30–9:45 Break

9:45

2aNSa4. Noise and vibration characteristics of cleanroom fan filter units. Mei Q. Wu and Colin G. Gordon (Colin Gordon & Assoc., 411 Borel Ave., Ste. 425, San Mateo, CA 94402)

Fan filter units (FFU's) are used increasingly by the microelectronics industry to provide recirculation air in cleanrooms. Such units usually consist of one or more direct-drive blowers located in a small discharge plenum and a HEPA or ULPA filter. There may be several hundred FFU's in a large cleanroom. The noise and vibration generated by FFU's differ greatly between manufacturers. Since it is difficult to apply external noise and vibration mitigation, it is important to select an FFU model with appropriate noise and vibration levels, so that a cleanroom served by these units will meet the prespecified noise and vibration criteria. Unfortunately, FFU manufacturers usually do not provide enough information for predicting the FFU generated sound-pressure spectrum or vibration in a cleanroom. In this paper, simple noise and vibration measurement methods will be described, which can be used as a basis for specifying FFU performance without special testing facilities. Measured data will be presented for a number of units. Procedures will be demonstrated to predict the performance of a cleanroom which uses FFU's as the primary source of recirculation air.

10:05

2aNSa5. Design and acoustic characteristics of an anechoic chamber with high-power heating or cooling and sound field free transform system. Weicheng Yang and Jiahua Yang (Res. Inst. of Household Elec. Appliance of China, No. 6 Yuetan Beixiaojie, Beijing 100037, PROC)

In this paper, the design and the acoustic characteristics of an anechoic chamber with high-power heating or cooling condition and sound field free transform system are summarized. Moreover, during the construction of the anechoic chamber, the following technical problems have been solved successfully. (1) Under 15 kW heating or cooling energy load being released by the sound source in the anechoic chamber, the adjusting to the temperature and the humidity can be completed. (2) Through the

moving of the mobile sound absorption plane, two anechoic chambers can be combined into one anechoic chamber, one semi-anechoic chamber, or one octant semi-anechoic chamber. Its sound field characteristics conform to the standard ISO 3745. (3) During the normal state, the background noise in the anechoic chamber is less than 23 dBA.

10:25

2aNSa6. Hand-held analyzer systems for building acoustics and sound intensity/sound power. Peter Larsen (Bruel & Kjaer, Skodsborgvej 307, DK-2850 Naerum, Denmark, plarsen@bk.dk)

Two new applications have been developed for a hand-held analyzer: Building acoustics and sound intensity/sound power. The building acoustics application is designed for measurements of airborne and impact sound insulation and reverberation time. The sound intensity application is used for sound power measurements, noise source location, noise mapping, and building acoustics. Both applications we used with complete systems consisting of all the necessary parts to simplify the measurement task and provide the user with all results immediately at the measurement site. There is no need for postprocessing. The measurements can be done according to a wide selection of international and national standards. The paper will describe the use of both applications in solving a problem in building acoustics. The building acoustic application will be used to detect a problem with poor sound insulation of a wall. Then, the intensity technique is used to determine the cause of the problem. Both applications use a built-in noise generator in the hand-held analyser with wireless transmission to an amplifier/loudspeaker. The paper will focus on the various design features that have been incorporated to ensure easy, secure, and flexible use.

TUESDAY MORNING, 23 JUNE 1998 CASCADE BALLROOM I, SECTION C (W), 8:00 TO 9:55 A.M.

Session 2aNSb

Noise: Community Noise Annoyance from Entertainment and Other Sources

Jim Griffiths, Chair
Symonds Travers Morgan Limited, Mead House, Canteloupe Road, East Grinstead, West Sussex RH19 3DG, England

Chair's Introduction—8:00

Invited Paper

8:05

2aNSb1. A Hong Kong approach to control noise from outdoor entertainment activities. P. S. Ng, Sam W. H. Wong, Y. K. Kam, and K. S. Chan (Environ. Protection Dept., The Government of the Hong Kong Special Administrative Region of the People's Republic of China, 46th Fl., Revenue Tower, Wan Chai, Hong Kong)

With 6.5 million people living on some 1000 square kilometers of land, about 70% of which are hilly terrains, one could imagine how difficult it would be in planning compatible land use. High-rise residential blocks surrounding sports ground are not uncommon in Hong Kong. They should co-exist harmoniously if the latter is intended solely for recreational sports. Hong Kong, being a cosmopolitan city, hosts many kinds of outdoor entertainment activities of cultures and many are held frequently in these sports grounds owing to the absence of purpose-built outdoor venues for these activities. With increasing environmental expectations, complaints against noise from these activities are on the rise. Special considerations will have to be struck between efficient use of land and adequate noise protection of the nearby noise-sensitive receivers. This paper describes the various outdoor entertainment activities being held in Hong Kong, their noise generation characteristics, and the approaches taken to control the noise.

Contributed Papers

8:25

2aNSb2. Annoyance and health effects of entertainment noise. C. G. Rice (Inst. of Sound and Vib. Res., The Univ., Southampton SO17 1BJ, UK)

There is no doubt that over the past 20 years awareness in the possible adverse annoyance and health effects which can result from entertainment and leisure activities has increased, and the topic is becoming more worthy of research effort in its own right. A review of the literature has established certain knowledge base gaps in the causal relationships between such noise and its adverse effects. Indeed, apart from noise-induced hearing loss there appears to be a lack of in-depth research from which unequivocal quantitative conclusions can be drawn. While damage to hearing from certain activities has been established, in general nonauditory health effects are not apparent although annoyance effects and provoked violence are becoming increasing causes of major concern. However, bearing in mind the relatively low integrated noise exposures which personally accrue from participation in and exposure to entertainment and leisure noise activities, care should be taken not to overemphasize the adverse effects to the prejudice of the other positive benefits created by participating in such activities.

8:40

2aNSb3. Proactive concert sound management: The essential ingredient for community acceptance of outdoor concert venues. Richard G. Cann (Grozier Tech. Systems, Inc., 157 Salisbury Rd., Brookline, MA 02146-2032) and William J. Cavanaugh (Cavanaugh Tocci Assoc., Inc., Sudbury, MA 01776)

Residents living adjacent to outdoor music venues have often objected to sound emissions. With increased sound power now available from loudspeakers, the situation is exacerbated, threatening the existence of music venues. Venues have demonstrated that by using a fixed installation sys-

tem to manage and control concert sound emissions effectively, the friction with the community can be reduced substantially. An essential ingredient of the amelioration process is making automatic sound emission reports available to amphitheater sound engineers and management as each concert proceeds. Sharing these same reports with citizen activists immediately after each concert helps develop trust. Case histories will be presented.

8:55

2aNSb4. Entertainment park noise control devices. Alessandro Cocchi and Giovanni Semprini (DIENCA, V.le Risorgimento 2, 40136 Bologna, Italy, alessandro.cocchi@mail.ing.unibo.it)

A big entertainment park is operating with noisy attractions near the Garda lake, a pleasant land where people go not only for amusement but also to relax, especially during holiday time. Noise measurements have been made during different periods of daytime and a noise map has been developed taking into account the difference of level between the top of the noisiest attraction and the windows of the most exposed house. Starting from this computerized model, and knowing the emission spectrum of the most relevant sources, it has been possible to design some noise control device useful for reducing the noise exposure level restricted to the exposed houses alone: This was possible combining earth barriers and a hot windstream.

9:10

2aNSb5. Comparison of predicted and measured noise levels for refinery units. Frank H. Brittain (Bechtel Corp., 50 Beale St., San Francisco, CA 94105) and Mark M. Gmerek (The Boeing Co., P.O. Box 3707, Seattle, WA 98124)

Predicting noise levels from new refinery units is a vital part of environmental assessment and designing units to meet noise limits. The accuracy of those noise predictions is a very important concern. The simplest way to assess the accuracy of predictions is to compare predicted and measured noise levels. This is usually difficult, because measured levels are strongly affected by noise from adjacent units and by atmospheric effects on sound propagation. Further, actual noise levels of significant sources often deviate from expected levels used in the noise prediction model. Thus to meaningfully compare predicted and measured levels, the actual noise source levels for the major sources, atmospheric conditions, and noise levels from adjacent units must all be accounted for. Predicted and measured levels are compared for two large refinery units. Measurements were made at locations where noise from adjacent units has little effect and close enough so that atmospheric conditions have little impact on the measured levels. Measured operational noise levels of major sources were used to update the noise prediction model. Accuracy of the predictions is evaluated.

9:25

2aNSb6. Acoustic modeling and simulation of an urban substation. Eduardo Bauzer Medeiros and Gia Kroeff (Dept. de Eng. Mecânica da UFMG, Av. Antônio Carlos 6627, 31270-901, Belo Horizonte-MG, Brazil, flugzbau@oraculo.lcc.ufmg.br)

Demographic growth has often resulted in a disorganized occupation of large urban areas. Regions which were previously empty started to be populated, frequently resulting in an increasingly more common proximity between residential construction and industrial installations. As a consequence, very serious environmental problems have occurred, including those caused by acoustic emission from industrial installations. A common problem these days is the disturbance caused by electric substations closely positioned to residential buildings. The disregard for acoustic effects during the design stage can still be observed in some of the older substation designs, particularly in areas previously far from city centers. In some countries, such as Brazil, a systematic acoustic evaluation of substations only started in the 1980s, with the main bulk of the work still under development. This paper describes some of the joint work being carried out on noise control by the researchers of Universidade Federal de Minas Gerais and CEMIG (the state power administration). The initial definition of a series of acoustic elements, followed by the discussion of the governing equations, is considered. The computational implementation of the resulting model is presented, aimed at the definition of the most convenient noise control strategy. Finally, a few results obtained at a real substation are considered, including some experimental results obtained during actual operation.

9:40

2aNSb7. Environmental pollution and noise control of two power generator sets placed into a recreation area. Victor Rastelli, Nila Montbrun, and Alexis Buoza (Departamento de Mecánica, Universidad Simón Bolivar, Caracas 1080-A, Venezuela)

This paper deals with the noise control of two power generator sets of 545 kW each, located in a Venezuelan resort. The generator sets made a noise level of 73 dB(A) at a distance of 47 m from the source, where are located the guest rooms. This level was far from the legal regulations for allowable environmental conditions. A noise enclosure with air silencers was proposed to reach a final level of 53 dB(A) at the guest accommodations. Two aspects must be taken into account, the noise reduction and the ventilation specifications of the generator sets, in order to keep the operational temperature. The final noise level was measured at 54 dB(A) and the pressure loss of the silencers was well placed into the ventilation specifications.

Session 2aPAa

Physical Acoustics: Ducts and Tubes

James P. Chambers, Chair
National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677

Chair's Introduction—9:25

Contributed Papers

9:30

2aPAa1. The effect of mass transfer on sound propagation in cylindrical tubes using the low reduced frequency approximation. Craig J. Hickey and Richard Raspet (Dept. of Phys. and Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677, chickey@olemiss.edu)

The theory for sound propagation in a gas/vapor mixture contained in a cylindrical tube is investigated. The tube is assumed to have a rigid wall covered by a thin film of water. This thin film of water acts as a source/sink for the mass and the heat associated with the vapor. This problem has been previously solved in the high- and low-frequency limits using the Rayleigh eigenmode method [Y. Mao and J. M. Sabatier, J. Acoust. Soc. Am. **96**, 3254(A) (1994)]. However, the interpretation of these results is limited by their complexity. This formulation of the problem parallels the low reduced frequency approximation work of Tijdeman [J. Sound Vib. **39**, 1–33 (1975)]. In contrast to the earlier results an easily interpreted analytical solution for the propagation constant is obtained. The parameters governing the propagation of the sound waves are the shear wave number, Prandtl number, Schmitt number, and the reduced frequency. The limits of small and large shear wave number provide useful insight into the behavior of the gas/vapor mixture. Examples of air/water and helium/water mixtures will be discussed. [Work supported by ONR.]

9:45

2aPAa2. Measurement of the radiation impedance of a pipe with a circular flange. Jean-Pierre Dalmont (Laboratoire d'Acoustique, Université du Maine, 72085 Le Mans Cedex 9, France) and Cornelis J. Nederveen (Acacialaan 20, The Netherlands)

The radiation of an open pipe with an infinite flange, as well as that of an unflanged pipe has been explicitly calculated by, respectively, Norris and Sheng [J. Sound. Vib. **135**(1), 85–93 (1989)] and Levine and Schwinger [Phys. Rev. **73**, 383–406 (1948)]. The radiation of real tubes, i.e., with a finite wall thickness or a finite circular flange, is not accurately known. The aim of the present measurements is to determine the radiation of tubes with various flanges and their frequency dependence. The radiation impedance has been deduced from the measurements of the input impedance of a short tube. As the differences between the different situations can be very small, the impedance sensor must be accurately calibrated. Results are reproducible and are in good agreement with a calculation using a boundary elements method (software, Rayon). It is observed that the curves obtained are not limited by the case of the infinite flange and the case of the unflanged pipe. This is probably caused by a resonance effect due to the diffraction at the edges of the flange.

10:00

2aPAa3. Boundary-integral-equation methods for accurate calculation of acoustic fields. James B. Mehl (Dept. of Phys., Univ. of Delaware, Newark, DE 19716-2570, jmehl@udel.edu)

Acoustic fields near solid boundaries can be approximated by solutions of the Helmholtz equation subject to Neumann boundary conditions on the boundaries. For closed, axisymmetric geometries, the boundary-value problem can be expressed as a one-dimensional integral equation. Boundary shapes can be represented exactly in this formalism. With current code, numerical solutions can be obtained for shapes formed by rotating plane figures made up of mixtures of straight and circular line segments. Numerical solutions to the integral equation were determined by using cubic-spline approximations to the velocity potential Ψ. This procedure enabled both Ψ and the tangential acoustic velocity Ψ' to be determined with high precision. In calculations of the fields within ducts near the orifices, with modest ($\sim$200) numbers of boundary elements, approximate values of Ψ' were found which were smooth over five orders of magnitude. The technique was used to calculate fields near the duct ends and to evaluate inertial and resistive end effects, for both baffled and unbaffled duct ends. The effects of rounding sharp edges has been evaluated through the use of circular-arc boundary elements. [Work supported in part by the Office of Naval Research.]

10:15

2aPAa4. Numerical study of sound field radiated from a circular duct with an open end. Zhichi Zhu, Anqi Zhou, Dongtao Huang, and Qing Guo (Dept. of Eng. Mech., Tsinghua Univ., Beijing, 100084, PROC, zzc-dem@mail.tsinghua.edu.cn)

A dispersion-relation-preserving (DRP) scheme for computational acoustics has been developed in recent years. This scheme is of high accuracy and can preserve the original dispersion relations of governing equations so that a good global acoustic solution can be ensured. However, the scheme is usually not only very complex but also problem dependent. In this paper, a unified DRP scheme for both internal and boundary points is developed for solving the sound field radiated from a circular duct with an open end. This unified DRP scheme is further applied to obtain the sound fields numerically in three different cases, namely, without flow and flange, without flow but with flange, and with flow and flange. In addition, relevant reflection efficiencies are also predicted at various Helmholtz numbers. Satisfactory comparisons of the predicted reflection efficiencies with data from other studies are illustrated. It is also shown that the numerical method with the DRP scheme presented in the paper is successful.

10:30

2aPAa5. Experimental investigation on nonlinear standing waves by two tube structure. Ke Liu and Dah-You Maa (Inst. of Acoust., Academia Sinica, P.O. Box 2712, Beijing 100080, PROC)

Nonlinear standing waves have been studied in a structure consisting of two lengths of tubing with different diameters. The measurement system of the experiment is presented. A new high-power loudspeaker driver (300 W) is used as a sound source to excite the large tube of 45 mm diameter through a small hole near one end which is closed. Preliminary nonlinear theory is proposed, and a resonance diagram of the tube structure is given. The main properties of nonlinear standing waves were obtained. At the end of smaller tube, a sound-pressure level (SPL) of nearly 177 dB has been obtained with a low distortion. The phenomena of saturation and properties of the output at the end of the smaller tube as well as their relation with the input were investigated. Increasing the diameter of

the structure directly with a modern high-power loudspeaker produces high intensity in a more extended space. Experiments are in progress. [Work supported by the National Natural Science Foundation of China.]

10:45

2aPAa6. Influence of airflow on sound propagation in a curved tunnel. Hiroyuki Imaizumi and Takehiro Isei (16-3 Onogawa, Tsukuba, Ibaraki, 305 Japan)

Measurements were made to examine influences of airflow on propagation attenuation of sound in a curved tunnel. The speed of airflow varied in width direction of the cross section of tunnel, and the values were from around 1 to 3 m/s. Sound attenuation upwind was less than that downwind at high frequencies up to the receivers that are in sight of the source. The tendency agrees with sound attenuation measured in a straight tunnel with airflow [H. Imaizumi *et al.*, J. JSSE **31**, 228–239 (1992)]. Difference of the attenuation gradually became small in the same frequency region at receivers that are out of sight from the source. Numerical calculations by means of a ray tracing method were applied to propagation attenuation of sound in a curved two-dimensional plane. Transmission losses of sound energy calculated only from reflection times of the sound rays on the inner wall were compared under each propagation condition. The transmission loss downwind was larger than upwind up to the receivers that are in sight, while the difference of both became small at the further receivers due to the rapid increase of the reflection times of rays. The calculation results qualitatively agree with the experimental measurements.

11:00

2aPAa7. Acoustic propagation in the presence of an arbitrary flow. C. J. Peyret (ONERA/MFE/DSNA, 29 Ave. de la division Leclerc, 92322 Chatillon, Cedex, France)

The study of acoustic propagation in a lined duct in the presence of an arbitrary flow remains open. Duct acoustic propagation in the presence of flow is generally computed assuming some particular properties of the flow and simplifying the duct geometry to find an appropriate wave equation. For instance, the Lilley or Phillip equation can be deduced this way. However, a unique partial differential equation based on the acoustic displacement vector allows a nonhomogeneous flow in an arbitrary duct geometry to be taken into account. This equation was first established by Galbrun in 1931. After recalling the Galbrun equation, a method to solve it is presented. The method uses the Lagrangian properties in order to apply the finite elements method. Different boundary conditions may be introduced. The solution of the Galbrun equation is compared to the solution of the Lilley equation in the case of modal propagation in a lined duct in the presence of a shear flow. Both results are in good agreement, and validate the method used to solve the Galbrun equation.

11:15

2aPAa8. Propagation of vibration waves in a pipe. Mihail-Dan F. P. Simion (Dept. of Machines — Tools, Univ. Polytechnica Bucharest, 313 Splaiul Independentei, 77206, Bucharest), Gabriela-Cristina D. Simion, and Florian-Paul Simion (Univ. Polytechnica Bucharest, 313 Splaiul Independentei, 77206, Bucharest, simion@freddy.cib.pub.ro)

Our studies concerning propagation of vibrations inside a pipe presume a behavior analysis for a continuous medium, but also a dynamic study of elastic medium. A diagram for describing the behavior of a viscoelastic medium is represented with emphasizing properties of material. In our case energy is transmitted through propagation of longitudinal pressure waves. Every particle of medium is vibrating around its equilibrium position. The relation for deformations dependence on tensions are determined, taking into account a medium with negligible viscosity. Propagation velocity for longitudinal and transversal plane waves are computed, by aiding Navier equations. A ratio between longitudinal waves propagation velocity and transversal waves propagation velocity equal to 1.87 for $v = 0.3$ is obtained.

TUESDAY MORNING, 23 JUNE 1998 ASPEN ROOM (S), 9:25 A.M. TO 12:00 NOON

Session 2aPAb

Physical Acoustics: Topics in Thermoacoustics

Anthony A. Atchley, Chair

Graduate Program in Acoustics, Pennsylvania State University, Applied Science Building, P.O. Box 30, State College, Pennsylvania 16804

Chair's Introduction—9:25

Contributed Papers

9:30

2aPAb1. Simple model for temperature gradient formation in a short stack. Ralph T. Muehleisen (Dept. of Civil, Environ., and Architect. Eng., Univ. of Colorado, Boulder, CO 80309) and Anthony A. Atchley (Penn State Univ., P.O. Box 30, State College, PA 16804)

There has been much analytical, numerical, and experimental work investigating steady-state thermoacoustic phenomenon. However, much less work has been done on time-dependent problems such as the evolution of the thermoacoustic temperature gradient in a short stack of plates. If the time scale of the acoustics is much shorter than that of the thermodynamics of the stack itself, the acoustics can be assumed to be in steady state while the temperature gradient in the stack is formed. With this assumption, the thermoacoustic heat flux between the stack plates will be of the form $Q = (C_1 + C_2 \partial T/\partial x) p^2$. This gives rise to a temperature gradient that rises exponentially in time with a time constant proportional to the thermal conductivity of the stack and the acoustic pressure squared. There is an additional exponential temperature rise from viscous heating. The model is shown to match measured temperature gradients very well. [This work was supported by the Office of Naval Research and the American Association for Engineering Education.]

9:45

2aPAb2. Performance measurements of a thermoacoustic refrigerator driven at high amplitudes. Matthew E. Poese and Steven L. Garrett (Grad. Prog. in Acoust., P.O. Box 30, State College, PA 16804, matt@sabine.acs.psu.edu)

Since the power density in a thermoacoustic device is proportional to the square of the acoustic Mach number, there is strong motivation to design thermoacoustic refrigerators to operate at larger pressure amplitudes. Measurements will be reported of a modified version of the space thermoacoustic refrigerator (STAR) [Garrett *et al.,* ''Thermoacoustic refrigerator for space applications,'' J. Thermophysics Heat Transfer **7**, 595–599 (1993)] driven at pressure amplitudes two to three times as large as reported in 1993. The results of these measurements will be compared to both a DELTAE computer model of the low-amplitude (linear) performance and one that includes turbulent flow in the duct regions of the device. An overview of the instrumentation, including a measurement of exhaust heat with an absolute accuracy of 65 mW, will also be presented. [Work supported by Office of Naval Research.]

10:00

2aPAb3. Extended performance measurements for SETAC. Anat Grant and Steven L. Garrett (Penn State Univ., P.O. Box 30, State College, PA 16804, anat@sabine.acs.psu.edu)

The Shipboard Electronics ThermoAcoustic Cooler (SETAC) was built at the Naval Postgraduate School in Monterey, CA. In a demonstration at sea aboard the U.S.S. DEYO in April 1995, SETAC was used to provide 419 watts of useful cooling power across a stack temperature difference $T_{\text{hot}}-T_{\text{cold}}=30\ ^{\circ}\text{C}$ with an acoustic input power of 227 watts. New data spanning a wider range of temperature difference and heat load will be presented and compared to a DELTAE model. [Work supported by the Office of Naval Research.]

10:15

2aPAb4. Solar/heat-driven thermoacoustic engine. Reh-lin Chen and Steven L. Garrett (Grad. Prog. in Acoust. and Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804)

With its abundance and cleanliness, solar energy has been harnessed to generate power using various techniques. A thermoacoustic engine is described which was built, instrumented, and tested to demonstrate use of solar power to generate acoustic power. Sunlight is collected using a 3-ft-diam acrylic Fresnel lens and focused on one end of a ceramic ''stack.'' This engine has a total length of 40 cm and resonates at around 420 Hz. To permit testing on the abundant cloudy days in Pennsylvania, an electrical heater was also built and installed to provide a reliable thermal energy source for indoor measurements. Experimental results are in reasonable agreement with a DELTAE model. When running in solar mode, as a $\frac{1}{4}$ wavelength open end resonator, the measured acoustic SPL can reach 120 dB easily at 1 m away from the open end on a clear day. The onset of sound generation can be achieved even with a 1-ft-diam Fresnel lens. While an impressive demonstration in terms of sound output, despite modest efficiency, real solar energy conversion will be improved with inert gas mixtures at higher pressures. [Work supported by ONR, lens provided by G. W. Swift.]

10:30

2aPAb5. Thermoacoustic refrigeration demonstration. Ray Scott Wakeland and Steven L. Garrett (Grad. Prog. in Acoust., Penn State Univ., State College, PA 16801, garrett@sabine.acs.psu.edu)

A portable device for demonstrating thermoacoustic cooling is developed using off-the-shelf components: the Morel MW-142 5-in. woofer and the Selenium HL14-25 horn. Acoustic peak pressure amplitudes of 19 kPa (19% of mean pressure) have been measured (without stack) using atmospheric air as the working fluid. Measurements of spectra, quality factors, stack temperature differences, and heat pumping capacity are compared to DeltaE models. Use of off-the-shelf components leads to design considerations not encountered in full-scale devices. For example, the high mechanical resistance of the driver, which was intended for audio reproduction, dictates against making the driver coresonant with the resonator, as is common when the driver is designed for operation at a specific frequency. [Work supported by NSF and ONR.]

10:45–11:00 Break

11:00

2aPAb6. Precision measurements of thermoacoustics in a single pore. Gabriela Petculescu and Larry A. Wilen (Dept. of Phys. and Astron., Ohio Univ., Athens, OH 45701, wilen@helios.phy.ohiou.edu)

Continuing experiments designed to look at thermoacoustic effects in a single pore will be described. A volume-modulation technique is employed to measure the complex compressibility of a gas in the pore directly. When a temperature gradient is imposed on the pore, the complex compressibility is sensitive to thermoacoustic effects which can be measured to high precision. These effects depend on both thermal and viscous properties of the gas. Measurements on different gases and mixtures of gases will be presented and compared to predictions of the theory. [Work supported by Ohio University Research and Sponsored Programs.]

11:15

2aPAb7. Time-average temperature distribution in a circular thermoacoustic pore. George Mozurkewich (Ford Motor Co., Maildrop 3028, P.O. Box 2053, Dearborn, MI 48121-2053, gmozurke@ford.com)

The time-average temperature distribution in a thermoacoustic pore, such as one pore of a thermoacoustic stack, is calculated. The analysis applies to straight pores of arbitrary shape but is limited to the acoustic and short-stack approximations. A partial differential equation for temperature is found by applying conservation of energy to the standard expression for time-average energy flux. Previous solutions for parallel-plate pores [G. Mozurkewich, J. Acoust. Soc. Am. **103**, 380–388 (1998)] are extended to pores of circular section. A heat-transfer coefficient is defined by $h_1=q(z)/\Delta T(z)$, where q is the rate of heat transfer into the pore interior per unit area of pore wall, ΔT is the temperature difference between pore wall and pore centerline, and z is the axial coordinate. This coefficient is independent of z provided z is not too close to an end of the pore. Results will be given for the shapes of the radial eigenfunctions and for the variations of eigenvalues and h_1 with pore size and Prandtl number.

11:30

2aPAb8. Linear resonant duct thermoacoustic refrigerator having two regenerator stacks. Yong Tae Kim, Sang Joon Suh (Acoust. and Vib. Group, Korea Res. Inst. of Standard Sci., P.O. Box 102, Yusong, Taejon 305-600, Korea), and Min Gon Kim (Dept. of Phys., Hankuk Univ. of Foreign Studies, Wangsan, Mohyun, Yongin, Kyungki 449-791, Korea)

The measurements of acoustically generated temperature difference between the two ends of a TAC (thermoacoustic couple) in a resonant tube are made using the similar method suggested by Atchley *et al.* [J. Acoust. Soc. Am. **88**, 251–263 (1990)] with a ceramic TAC cylinder. The linear duct was made of PVC pipe with an inner diameter of 30 mm and 680 mm long. Measurements are made for a frequency range of 150–350 Hz in atmospheric air along the TAC length from 2.5 to 10 cm with a 1-cm step. The unknown information of ceramic can be obtained from nonlinear fitting of experimental data with the equation reported by Atchley *et al.* From the result of this experiment and prediction by theory it is found that the direction of acoustic heat transport is altered at the point of a quarter of a wavelength. From the hint of the above experiment, a newly designed thermoacoustic refrigerator having two regererator stacks is tested.

11:45

2aPAb9. Theory of nonlinear acoustic waves in a thermoacoustic prime-mover. Pierrick Lotton, Vitalyi Gusev, Hélène Bailliet, and Michel Bruneau (Lab. d'Acoustique, IAM—UMR 6613, Univ. du Maine, av. O. Messiaen, 72085 Le Mans Cedex, France, lotton@laum.univ-lemans.fr)

An equation describing stabilization of oscillations in a thermoacoustic prime-mover was derived: $\partial V/\partial t - V(\partial V/\partial \tau) + (A-D)(\partial V/\partial \tau) = \sqrt{2/\pi} A \int_{-\infty}^{\tau} (\partial V/\partial \tau')[d\tau'/(\tau-\tau')] + V + D(\partial^3 V/\partial \tau^3)$. In this equation, particles velocity V, ''slow'' time t, and ''fast'' time τ are normalized; parameters A and D depend on temperature distribution inside thermoacoustic stack. The third term on the left-hand side (lhs) describes a shift in resonance frequency of a half-wavelength resonator caused by a stack installation. The third term on the right-hand side (rhs) describes sound velocity dispersion when the stack is acoustically thin. The first term (if $A>0$) and the second term on the rhs describe sound amplification due to quasiadiabatic interaction of acoustic and thermal waves. The first term on the rhs contributes also to dispersion. Particular analytical solutions were found and computer modeling was applied for the analysis. Conditions when a shock front formation caused by acoustic nonlinearity (the second term on the lhs) is the dominant mechanism for the saturation of wave amplitude were determined. [Work supported by DGA.]

TUESDAY MORNING, 23 JUNE 1998 GRAND BALLROOM II (W), 7:45 TO 10:45 A.M.

Session 2aPP

Psychological and Physiological Acoustics: From Ear Canal to Cortex—Clear Cuts and Old Growth (Poster Session)

Robert S. Schlauch, Chair
Department of Communication Disorders, Univerity of Minnesota, 115 Shevlin Hall, Minneapolis, Minnesota 55455

Contributed Papers

All posters will be on display from 7:45 a.m. to 10:45 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 7:45 a.m. to 9:15 a.m. and contributors of even-numbered papers will be at their posters from 9:15 a.m. to 10:45 a.m. To allow for extended viewing time, posters will remain on display until 12:00 noon on Wednesday, 24 June. A cash bar will be set up near these posters from 6:15 p.m. to 7:30 p.m. on Tuesday, 24 June.

2aPP1. Temporal versus spectral cues in AM detection. Martin E. Rickert and Neal F. Viemeister (Dept. of Psych., Univ. of Minnesota, 75 East River Rd., Minneapolis, MN 55455)

Several recent studies have shown that modulation thresholds for detecting sinusoidal AM are approximately constant out to much higher modulation frequencies for tonal carriers than for broadband noise carriers. It has been suggested that with tonal carriers there is a range of modulation frequencies for which both temporal (envelope) cues and spectral cues are used for detection. The present study attempts to assess the relative importance of these cues. Modulation thresholds for AM, quasifrequency modulation (QFM), and beats were measured for modulation frequencies of 8–512 Hz and carrier frequencies of 2 and 5 kHz. The results indicate that at intermediate modulation frequencies, AM thresholds are approximately constant while those for QFM and for beats decrease with increasing modulation frequency. This indicates that the detection of AM is not based solely on temporal cues. The implications of these findings for envelope detection and modulation filter-bank models will be discussed. [Work supported by NIDCD Grant No. DC00683.]

2aPP2. Estimates of temporal resolution at low and high spectral frequencies. David A. Eddins (Psychoacoust. Lab., Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405, deddins@indiana.edu)

The auditory system must encode dynamic acoustic stimuli throughout the audible spectrum. This coding is limited by the temporal resolution of the ear. Unfortunately, the results of several investigations disagree on the way in which temporal resolution varies with stimulus frequency. While gap-detection thresholds measured with sinusoidal signals change little over the frequency range of 200–2000 Hz, gap and amplitude modulation detection with narrow-band noise is more variable. Eddins [J. Acoust. Soc. Am. **93**, 470–479 (1993)] showed that modulation detection is not frequency dependent from 500 to 4400 Hz, however, there are few measures of modulation detection above this frequency range. In the present experiment, thresholds for modulation depth were obtained as a function of modulation frequency. Narrow-band noise carriers had upper- cutoff frequencies ranging from 3200 to 12 800 Hz and bandwidths from 400 to 3200 Hz. Temporal modulation transfer functions (TMTF) demonstrated the typical low-pass characteristic. The form of the TMTF did not vary with noise bandwidth or frequency region from 3200 to 9600 Hz. Sensitivity to modulation increased with increasing bandwidth. For lower modulation frequencies, thresholds were slightly reduced at 12 800 Hz. These results indicate that modulation detection varies little over a very broad range of frequencies.

2aPP3. Detecting irregularity of envelope modulation. Stanley Sheft (Parmly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626)

The ability to detect periodic irregularity in the temporal envelope of a modulated stimulus was measured with a 2IFC procedure. Irregular amplitude modulation (AM) was generated with a periodic phase modulation (PM) of the sinusoidal modulator of wideband noise. The modulation index of the PM was adaptively varied in terms of delta envelope rate to determine the threshold to just detect the PM of the sinusoidal amplitude modulation (SAM). Base SAM rate ranged from 4 to 256 Hz. For each SAM rate, the PM rate ranged from 2 to half the SAM rate. The PM waveshape was either sinusoidal or had linear ramps with slope varied across conditions. At low SAM rates, PM waveshape had little effect with

thresholds across conditions rising with PM rate. At higher SAM rates, an effect of waveshape was obtained; thresholds were highest at lower PM rates with a steeply sloping linear-ramped PM. For the other PM waveshapes at the higher SAM rates, thresholds were roughly constant across change in PM rate. Results suggest two rate regions in processing AM with the regions separated at roughly the corner frequency of the temporal modulation transfer function. [Work supported by NIH and AFOSR.]

2aPP4. Effects of carrier level on the processing of highly detectable envelope fluctuations. Magdalena Wojtczak and Neal F. Viemeister (Dept. of Psych., Univ. of Minnesota, 75 East River Rd., Minneapolis, MN 55455)

A modified 2AFC matching procedure was used to measure the modulation depth that yielded similar perceived envelope fluctuations for different carrier levels. In one case, the standard was a 60-dB SPL 1-kHz tone that was amplitude-modulated at 16 Hz, with a modulation depth that produced highly detectable envelope fluctuations. The comparison had the same carrier and modulation frequencies as the standard, and its level ranged from 15 to 85 dB SPL. The data indicate that the higher the carrier level, the less modulation was needed to produce the same perceived fluctuation depth. For broadband noise carriers, however, the data indicate that approximately the same modulation depth was needed to obtain subjectively similar envelope fluctuations for all carrier levels tested. These results are similar to those observed for detection of AM, and they suggest that the level effects seen for detection extend to suprathreshold AM processing, i.e., Weber's law for broadband noise and a ''near-miss'' for sinusoidal carriers. [Work supported by NIDCD Grant No. DC00683.]

2aPP5. Variations of the leaky-integrator model and validation with detection of sinusoidal and percentage duty-cycle modulated noise. John A. Nelson (Commun. Sci. and Disord., Jesse Jones Commun. Ctr., The Univ. of Texas, Austin, TX 78712-1089, john-a-nelson@uiowa.edu) and Ruth A. Bentler (The Univ. of Iowa, Iowa City, IA 52242)

The leaky-integrator model is commonly used to describe auditory temporal processing. In the current study, the low-pass filter and decision device of the model were investigated. Maximum-to-minimum envelope magnitude, rms envelope power, envelope fourth moment, and envelope crest factor were implemented as potential decision devices [E. A. Strickland and N. F. Viemeister, J. Acoust. Soc. Am. **99**, 3638–3646 (1996)]. Model predictions demonstrated interaction between the decision device and low-pass filter cutoff frequency. To validate model variations, detection thresholds of sinusoidal and percentage duty-cycle modulated broadband noise were measured at octave frequencies 2–512 Hz for listeners with normal hearing. To ensure subjects responded exclusively to temporal cues, the modulated noise was compensated for overall average power. Further, an intensity roving technique was used between observation intervals. The data show low-pass transfer function characteristics across all modulation functions. A two-line regression fit of the data revealed that the cutoff frequency, slope, and overall gain varied as a function of the duty-cycle modulation function. The nonlinear growth of the various decision devices as a function of the low-pass filter is used to explain individual variability among subjects. [Work supported, in part, by the American Academy of Audiology Student Investigator Award.]

2aPP6. Comodulation masking release with tone pips added out of phase to the valleys of a sinusoidally amplitude modulated pure tone. Deborah A. Fantini, Roel Delahaye, and Ray Meddis (Dept. of Psych., Essex Univ., Wivenhoe Park, Colchester CO4 3SQ, England, deb@essex.ac.uk)

Two studies of comodulation masking release (CMR) investigated the influence of temporal position of brief signals added to a modulated on-frequency band (OFB) [Grose and Hall III, J. Acoust. Soc. Am. **85**, 1276–1284 (1989); Moore *et al., ibid.* **87**, 1683–1694 (1990)]. The former reports mixed results as to whether flanking bands (FBs) modulated to the same depth as the OFB elicit more CMR than 100%-modulated FBs. The latter reports that when antiphase signals are added to valleys of 50%-modulated OFBs, thresholds with correlated FBs are *poorer* than with quadrature-phase FBs. In the present paper, signals were three antiphase tone pips added to valleys of a 50%-modulated 700-Hz pure tone. Thresholds were obtained with the OFB and 6 FBs, modulated at 10 Hz. FBs were comodulated or in quadrature phase with the OFB, at either 50% or 100% depth. With 50% FBs, subjects performed *better* with comodulated than with quadrature-phase FBs. With 100% FBs, performance in the comodulated condition improved, and in the quadrature-phase condition was impaired, relative to 50% FBs. These results indicate that CMR *can* occur when antiphase signals are added to valleys of 50%-modulated OFBs.

2aPP7. Comodulation masking release: The effect of the characteristics of noise bands presented before and after a signal. Lee Mendoza (Dept. of Commun. Sci. and Disord., Louisiana State Univ., 163 M&DA, Baton Rouge, LA 70803, lmendoz@unix1.sncc.lsu.edu), Joseph W. Hall III and John H. Grose (Univ. of North Carolina, Chapel Hill, NC 27599)

CMR was measured as a function of masker duration in three conditions. In the first condition, a 400-ms pure-tone signal was presented in the temporal center of comodulated masking bands ranging in duration from 400 to 1200 ms. No effect of masker duration was observed: CMR was equivalent to that found in continuous maskers. In the second condition, continuous independent bands identical in spectra to the comodulated bands were gated off before, and on following, the comodulated bands. The presence of the independent bands reduced CMR considerably. Increasing the duration of the comodulated bands increased CMR somewhat; however, CMR was still substantially smaller than that observed without the independent bands. In the third condition, band limited noise spanning the same range as the comodulated bands was substituted for the independent bands. The bandlimited noise had little effect on CMR. In a second experiment, independent bands were gated on for various durations preceding spectrally identical 400-ms comodulated bands masking a 400-ms signal. As the duration of the independent bands increased, CMR decreased. Based on these results, it appears that CMR is not entirely the result of processing only information concurrent with the signal. [Work supported by NIH-NIDCD and Louisiana Board of Regents.]

2aPP8. Perception of roughness in harmonic multitones. Andrzej B. Dobrucki and Maurycy J. Kin (Inst. of Telecommunication and Acoust., Wroclaw Univ. of Technol., Wyb. Wyspianskiego 27, 50-370 Wroclaw, Poland)

This paper presents results and problems of perception of mixed modulation (MM) in harmonic multitone for higher modulating frequencies (70 and 200 Hz), and a 5-harmonics complex. Results indicated that perception of frequency modulation at the presence of the threshold values of amplitude modulation is different than the perception of amplitude modulation at the presence of the threshold values of frequency modulation. It means that two mechanisms of perception for these kinds of modulation exist. It turns out that for perception of this kind of modulation, the hearing system uses two sidebands produced by modulation process, but the high sideband seems to be more important than the low one. Another phenomenon is the perceptual grouping mechanism: for some of the modulated partials MM can be perceived as an additional pitch occurring in multitones, which usually is out of tune for fundamental frequency of the complex. The summation of the partials in one critical band takes place, which suggests that a specific pitch for that partial could be changed

and this is a simple way to detect these simultaneous fast changes in the amplitude and frequency of the modulated components from the musical point of view.

2aPP9. A model of auditory image flow. I. Architecture. Neil P. McAngus Todd (Dept. of Psych., Univ. of Manchester, Manchester M13 9PL, UK, todd@hera.psy.man.ac.uk)

A model is presented here based on the theory that the auditory cortex computes the spatio-temporal power spectrum of the auditory image flow [D. Heeger, ''Model for the extraction of image flow,'' J. Opt. Soc. Am. **4**, 1450–1470 (1987)]. A computational implementation of the model has the following components: (1) a gammatone filter-bank spaced on the ERB rate scale; (2) an orthogonal, ERB spaced modulation filter-bank which generates a 2-D modulation map which is then low-pass filtered; and (3) a population of causal 3-D spatio-temporal Gabor filters which compute the spatio-temporal power spectrum of the image sequence. The cortical filters are organized into families such that each member of the family occupies the same spatial-frequency band but has a different spatio-temporal orientation. The spatial frequencies range from 0.05 cycles erb^{-1} to 0.4 cycles erb^{-1}. The temporal frequencies range from ± 0.5 to ± 64 Hz. The spatio-temporal power spectrum may be visualized by replacing each image pixel with the output of all filters centered at the pixel arranged in frequency and orientation bands, somewhat like orientation columns in the visual cortex. This model may be applied to modulation perception and auditory grouping.

2aPP10. A model of auditory image flow. II. Detection of amplitude and frequency modulation. Neil P. McAngus Todd (Dept. of Psych., Univ. of Manchester, Manchester M13 9PL, UK, todd@hera.psy.man.ac.uk)

In order to evaluate the cortical model proposed in Todd [these proceedings] it was applied to the phenomenon of modulation detection. It was assumed that a discrimination task involves the correlation of the spatio-temporal power spectra of two image sequences, such as may occur in a typical 2A-2IFC procedure. The correlation was computed by $k(1-r)$, where r is the product moment correlation, for four different modulation frequencies, 4, 16, 64, and 256 Hz at four modulation depths for both AM and FM of a 1000-Hz carrier of 1000-ms duration against the image power spectrum of an unmodulated 1000-Hz 1000-ms tone. The obtained measures were then linearly regressed (with no intercept) as a function of the square of the modulation index. In all cases $k(1-r)$ gave a good fit ($R^2 > 0.95$) for both AM and FM. For each of the modulation frequencies the slope of the regression could be used to estimate the model thresholds. The model shows the correct behavior for AM and FM detection. This mechanism, it is suggested, may provide an account of modulation detection interference since low spatial-frequency filters span several ERBs.

2aPP11. Role of pinna cavities in median plane localization. Kazuhiro Iida (AVC Res. Lab., Matsushita Comm. Ind., 600 Saedo, Tsuzuki, Yokohama, 224 Japan, kiida@adl.mci.mei.co.jp), Motoki Yairi, and Masayuki Morimoto (Kobe Univ., Nada, Kobe, 657 Japan)

This paper investigates the role of pinna cavities in median plane localization. First, sound image localization tests were carried out under the conditions of progressively occluding three pinna cavities: scapha, fossa, and concha. Second, HRTFs were measured under the same conditions as those of the localization tests. Finally, the results of the localization tests and those of HRTFs measurements were compared. The results of the localization tests and the HRTFs measurements show: (1) Concha plays a great role in median plane localization regardless of the source elevation. (2) The amplitude spectrum of HRTFs shows a significant change in cases of concha occlusion. The comparisons of these results suggest that (3) the spectral cue in median plane localization is mainly formed by concha. (4) However, the degree of localization error which was caused by the occlusion of the pinna cavities depends on the subjects.

2aPP12. The influence of different speech segments on spatial localization. Brian L. Karlsen (Ctr. for Personkommunikation, Aalborg Univ., Frederik Bajersvej 7, DK-9220 Aalborg, Denmark, blyk@get2net.dk)

Numerous researchers have investigated to what degree humans are able to localize different sounds in the horizontal plane. Much of this work has focused on human response to simple sounds, like pure tones, white noise, and click trains, the underlying assumption being that responses to more complex stimuli could be inferred from a combination of responses to the simple stimuli. It was decided to investigate this for speech segments. Specifically, it was hypothesized that subjects would display a difference in localization ability for CVs containing some consonant manner groups rather than others. A number of Danish CVs were used as target stimuli in a horizontal localization task. The CVs were simultaneously masked by a diffuse white noise field. All subjects were normally hearing. The subjects were asked to determine from which clock position they heard a speaker, and to respond as fast as possible by pressing the corresponding button on a clock face. Final results are expected to confirm the hypothesis in full and indicate that there is indeed a difference in localization ability between the different consonant manner groups.

2aPP13. Underwater sound localization: Role of interaural time and intensity differences. Pierre Bovet (Lab. de Psych. Cognit., Univ. de Genéve, 9 route Drize, 1227, Carouge, Switzerland), Carolyn Drake (CNRS URA 316, Univ. Rene Descartes, 75006, Paris, France), François Bernaschina (Univ. de Genéve, Carouge, Switzerland), and Sohpie Savel (CNRS URA 316, Paris, France)

In theory, underwater sound localization is very limited compared with that in air. Most cues are greatly reduced underwater: Interaural time differences are divided by four (due to faster sound propogation in water) and intensity differences are greatly reduced because signal conduction to the inner ear mainly occurs through skull bone condition, making the head transparent. However, preliminary studies indicate some localization abilities in scuba divers. An azimuth identification task was conducted in Lake Leman. At a depth of 10 m, 14 divers indicated which of eight transducers located every 45 was generating a sound (400 Hz, 6 kHz, or white noise). Results indicate: (a) correct estimates were highest for lateral positions (left and right) and poor for median positions (front and back), (b) performance was high for 400 Hz and white noise but low for 6-kHz signals, and (c) head movements did not appear to facilitate localization. Thus the best localization abilities were observed in positions with highest interaural time differences, suggesting that, contrary to theoretical predictions, temporal binaural cues must be at least partially functional. However, the fact that localization was poor for high-frequency signals suggests that interaural intensity differences do not play a significant role.

2aPP14. A new model for binaural signal detection. Jeroen Breebaart, Steven van de Par, and Armin Kohlrausch (IPO-Ctr. for Res. on User–System Interaction, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands)

Most present-day binaural models make use of coincidence detectors following an internal delay, so-called crosscorrelation models. We present a new modeling approach based on a subtractive mechanism as found in excitatory–inhibitory type neurons in the auditory pathway. This approach transforms, for each auditory filter, time-domain waveforms into a time-varying two-dimensional activity pattern, where each unit within that pattern has a characteristic interaural time delay and intensity difference. Limits of resolution are modeled by adding the same amount of internal noise to each unit. From this activity pattern, several features of the stimu-

lus can be extracted, such as the apparent lateralization and the presence of a test signal within a masker. For modeling binaural detection it is assumed that a test signal must induce a significant change within the internal representation in order to be detectable. Depending on the condition, detection is limited by internal noise ($NoS\pi$) or by external fluctuations ($N\rho S\pi$, $\rho \leqslant 0.95$). This model can be used as an artificial observer in the same way as described by Dau *et al.* [J. Acoust. Soc. Am. **99**, 3615–3622 (1996)] for monaural signals.

2aPP15. Further evidence for the influence of peripheral compression on binaural detection. Steven van de Par and Armin Kohlrausch (IPO-Ctr. for Res. on User–System Interaction, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands)

Binaural detection at high frequencies depends on the envelope statistics of the masker. Van de Par and Kohlrausch [J. Acoust. Soc. Am. **101**, 3104 (1997)] proposed that this results from the compressive response of the basilar membrane which can be shown to reduce interaural intensity differences (IIDs) strongest in stimuli with flat masker envelopes. To further test the effect of compression, the same experiment was repeated at 500 Hz. Results show that binaural detection at low frequencies is much less dependent on masker envelope statistics as compared to high frequencies. This was expected because at low frequencies, besides IIDs, interaural time delays can be used for signal detection, which are not affected by compression. In addition, psychometric functions for the discrimination between 4-kHz NoSo and $NoS\pi$ stimuli were measured as a function of the S/N ratio for stimuli of 20 and 65 dB SPL using a 250-Hz-wide masker. The psychometric functions were rather symmetric around a 0 dB S/N ratio at low SPLs, but asymmetric at 65 dB SPL favoring discrimination at a negative S/N ratio. Such an asymmetry at high stimulus levels is predicted, if the detection cue is a change in the interaural correlation of the compressed envelopes.

2aPP16. A model of auditory localization in noise using neural networks. Scott K. Isabelle (Dept. of Psych., Wright State Univ., Dayton, OH 45435), James A. Janko (Ball Aerosp. and Technologies Corp., Fairborn, OH 45324), and Robert H. Gilkey (AL/CFBA, Wright-Patterson AFB, Dayton, OH 45433-7901)

Janko *et al.* [*Binaural and Spatial Hearing in Real and Virtual Environments*, edited by Gilkey and Anderson (Erlbaum, Hillsdale, NJ, 1996), pp. 557–570] used neural networks to describe auditory localization and found that data in quiet were insufficient to discriminate among models. To further constrain the models, localization in the presence of a spatially fixed masker is considered here. The wideband targets and maskers were filtered by head-related transfer functions, then by a gammatone filter bank. A model for binaural interaction similar to that of Lindemann [J. Acoust. Soc. Am. **80**, 1608–1622 (1986)] was used to process the filterbank outputs. The inhibited cross-correlation output of the binaural processor was sampled and used as input to a neural network (e.g., 24 correlation lags at each of 13 frequency channels for 312 input nodes). The network was trained using back propagation to estimate the target location in quiet, then tested on targets in noise at several values of signal-to-noise ratio (SNR). The network was also trained across several values of SNR. Results are compared to localization in noise by humans [Good and Gilkey, J. Acoust. Soc. Am. **99**, 1108–1117 (1996)]. [Work supported by AFOSR F49620-95-1-0106 and the Ohio Board of Regents.]

2aPP17. On the difference between absolute and relative auditory distance perception. Hae-Young Kim, Yôiti Suzuki, Shouichi Takane, Kenji Ozawa, and Toshio Sone (Res. Inst. Elec. Comm., Tohoku Univ., Sendai-shi, 980-8577 Japan)

Absolute distance perception in the absence of information about intensity and reflections from boundaries is quite difficult [B. G. Haustein, Hochfrequenztech. u. Elektroakustik **78**, 46–57 (1969)] but is possible to achieve if a sound source exists within the distance of around 1 to 2 m from a listener. In the presence of multiple sound sources, however, some cues can be derived by comparing information from different sources. From this point of view, it is important to study relative distance perception between sounds from multiple sources to investigate general distance perception in auditory system. Thus using the same experimental setup, two kinds of listening experiments about the absolute and relative distance perception were conducted. In the experiment for relative distance perception, a subject was asked to report the perceived distances of two sound sources presented successively with an interval of 2 s by using a rating scale with seven categories. The results of absolute distance perception showed that the perceived distance of a sound image significantly increases with the increase of actual distance of the sound source up to around 1.2 m. On the other hand, the results of the relative auditory distance perception showed that the perceived distance significantly increases up to 2 to 3 m.

2aPP18. Auditory motion aftereffects with a two-component adapter. Hisashi Uematsu and Makio Kashino (NTT Basic Res. Labs., 3-1, Morinosato-Wakamiya, Atsugi, Kanagawa, 243-01 Japan)

Auditory motion aftereffects have been used to investigate mechanisms used to detect changes in interaural time differences (ITDs). In this paper, auditory motion aftereffects with two-tone adapters were used to examine how the ITD-change information from each frequency band is integrated. In the experiment, a complex tone with 250- and 500-Hz components was used as an adapter. The magnitude of the auditory motion aftereffect for a physically stationary 500-Hz test tone was measured using various combinations of direction and adapter component. Apparent sound movement was produced by varying only ITDs. The results showed no difference in the magnitude of the aftereffect between the 500-Hz tone adapter and the complex adapter when the two components moved toward the same direction. On the other hand, the aftereffects disappeared when the two components of the adapter moved in opposite directions. This suggests that the integration of ITD-change information from each frequency band is not a simple linear summation. Mechanisms of across-frequency integration will be discussed.

2aPP19. Echolocation by moving and stationary listeners. Michael S. Gordon, Lawrence D. Rosenblum, and Luis Jarquin (Dept. of Psych., Univ. of California, Riverside, Riverside, CA 92521, rosenblu@citrus.ucr.edu)

It has long been known that human listeners can echolocate a sound reflecting surface as they walk towards it [e.g., Supa *et al.*, Am. J. Psychol. **57**, 133–183 (1944)]. There is also evidence that stationary listeners can determine the location, shape, and material of nearby surfaces from reflected sound [Rice, Science **155**, 656–664 (1967)]. The current research tested whether there is an advantage of listener movement for echolocating as has been found for localization of emitted sounds [Ashmead *et al.*, J. Exp. Psychol.: Hum. Percept. Perform. **21**, 239–256 (1995)]. Two groups of blindfolded subjects were asked to echolocate a 3×6 ft wall while either moving or stationary. After echolocating, the wall was removed, and subjects were asked to walk to where the wall had been. Subjects in the moving and stationary groups echolocated for the same amount of time and for each of four distances. Results showed that moving group subjects were more accurate than stationary group subjects for some distances. These results might be a function of useful time-to-arrival acoustic information available to moving listeners.

2aPP20. Reflectance measurements of the contralateral acoustic reflex input–output function. M. Patrick Feeney (Ohio Univ., School of Hearing and Speech Sci., Lindley Hall 218, Athens, OH 45701) and Douglas H. Keefe (Boys Town Natl. Res. Hospital, Omaha, NE 68131)

Available data suggest that the acoustic reflex reduces the transmission of sound through the middle ear in humans at frequencies below 1000 Hz. Wideband (0.2 to 10 kHz) reflectance, complex admittance, and absorbed power were used to measure the input–output function of the contralateral acoustic reflex in four adults. Broadband noise was used as a reflex activator. Measurements were taken at reflex threshold and in 5-dB increments above up to 112 dB SPL in one subject. The power transmission loss grew approximately 2.5 dB for every 5-dB increase in the activator above reflex threshold. Admittance measurements generally mirrored these changes below 1000 Hz with decreases in equivalent volume, conductance, and complex admittance. There was an increase in admittance with the acoustic reflex at frequencies from 1000 to 4000 Hz in agreement with Rabinowitz [Ph.D. thesis, MIT, 1977]. An unexpected decrease in energy reflectance was observed for one subject in a notch centered at 5600 Hz, near the zero crossing for equivalent volume, which suggests that the activation of the acoustic reflex increased the power delivered to the middle ear at that frequency. [Work supported by Ohio University Research Committee Grant No. 97-14, Ohio University CHHS SA Award, and NIH Grant No. DC00520.]

2aPP21. Vibration measurement of the tympanic membrane using the time-averaged speckle pattern interferometry. Hiroshi Wada, Masataka Takeuchi, Koji Hozawa (Tohoku Univ., Sendai, Japan), Takashi Gemma (Nikon Corp., Tokyo, Japan), and Makoto Nara (Nikon Technologies, Inc., Tokyo, Japan)

In this study, using the time-averaged electronic speckle pattern interferometry, the sinusoidal phase modulation technique by a direct modulation of laser injection current, and the digital image processing technique, an attempt was made to develop a system for measuring a small vibration in nanometer order. The vibratory responses of the guinea pig's tympanic membrane were measured. At the frequency of 1500 Hz, the whole part of the tympanic membrane vibrated almost in phase. The maximum amplitude was about 30 nm. At the frequency of 2500 Hz, the tympanic membrane had two local maxima in the posterior and inferior portions. At the frequency of 3000 Hz, the four peaks around the manubrium were seen. The number of the peaks increased with an increase in frequency and the vibration mode became more complicated. The results of this study indicate that the vibration amplitude is detectable in nanometer order by this system even when the stimulus level is small, e.g., 60 dB SPL. Consequently, the system developed by our group has a potential for clinical applications. [Work supported by the Ministry of Education of Japan.]

2aPP22. Filter shapes for brief signals as a function of preceding stimulation. Elizabeth A. Strickland (Dept. of Audiol. and Speech Sci., Purdue Univ., West Lafayette, IN 47907)

The threshold for a brief probe tone in a broadband noise was measured as a function of probe frequency (1 or 4 kHz) and masker level (0–40 dB spectrum level). The probe was always presented 2 ms after the onset of the masker, and the masker was preceded either by silence, or by a precursor with the same spectrum and level as the masker. For each subject there were some conditions in which probe threshold was higher with no precursor than with a precursor (overshoot), and some conditions in which probe threshold did not differ with the presence of a precursor. Filter shapes for the same brief probes were then measured using the notched noise method. The probe was always presented with a 2-ms delay from the onset of the notched-noise masker. The masker was preceded by silence or by a precursor. The probe level was set 10 dB above the threshold measured for a given condition in the overshoot experiment. A comparison will be made of filter shape and the presence or absence of overshoot. [Work supported by NINCDS.]

2aPP23. Variable-duration notched-noise experiments in a broadband-noise context. Jim J. Hant, Brian P. Strope, and Abeer A. Alwan (Speech Processing and Auditory Percept. Lab., Dept. of Elec. Eng., UCLA, 66-147E Eng. IV, 405 Hilgard Ave., Los Angeles, CA 90095)

To model the confusion of short-duration speech signals in noisy environments, it is important to characterize how humans resolve the spectral components of short-duration signals in background noise. Toward that end, a variable-duration notched-noise experiment was conducted in a noise context. Broadband noise preceded and followed a tone and notched noise of similar duration. Thresholds were measured at four durations (10, 30, 100, and 300 ms), two center frequencies (0.6, 2.0 kHz), and five relative notch-widths (0.0, 0.1, 0.2, 0.4, 0.8). At 0.6 kHz, 10-ms thresholds decrease 6 dB across notch-widths while 300-ms thresholds decrease over 35 dB. These durational trends are similar but less pronounced at 2 kHz. In a second experiment, the short-duration notched noise was replaced with a flat noise, which provided an equivalent amount of simultaneous masking to that of the first experiment, and thresholds dropped by as much as 20 dB. A simple filtering model with either a linear or nonlinear combination of simultaneous and nonsimultaneous masking is unable to predict these results. Instead, it appears that the elevated thresholds at short durations are dependent on the spectral shape of the simultaneous masker, implying reduced frequency selectivity. [Work supported by NIH-NIDCD and The Whitaker Foundation.]

2aPP24. Comparison of psychophysical and single-unit thresholds in response to electrical stimulation of the cochlea in guinea pig. Bryan E. Pfingst, Lisa M. Montney, and Sanford C. Bledsoe (Univ. of Michigan, Kresge Hearing Res. Inst., 1301 East Ann St., Ann Arbor, MI 48109-0506, bpfingst@umich.edu)

Recent studies have shown that the spatial pattern of currents used on individual channels of cochlear implants can have a large effect on speech recognition and other measures of electrical-stimulus perception. This study is one of a series that seeks to understand the neural patterns of response underlying these effects. Guinea pigs were deafened in one ear by local perfusion of an antibiotic and implanted with an electrode array in the scala tympani. Psychophysical strength-duration functions (detection threshold versus electrical pulse duration) were obtained in guinea pigs trained using positive reinforcement, operant conditioning techniques. Single-unit data were obtained from inferior colliculus in the same animals, or in guinea pigs deafened and implanted in a similar manner. In these preparations, thresholds for neural responses were near the psychophysical thresholds at all phase durations, and the thresholds for the most sensitive neurons were several dB below the psychophysical detection thresholds. Thus, the paradoxical discrepancies between neural responsiveness and behavioral stimulus detection observed in previous comparisons made across animal models do not occur in this preparation. This seems to be a useful preparation for determining the spatiotemporal patterns of neural activity that underlie electrical-stimulus perception. [Work supported by NIH-NIDCD Grant No. DC00274.]

2aPP25. Physiological correlates of the ''mean-term'' auditory adaptation. N. G. Bibikov (N. N. Andreyev Acoust. Inst., Moscow 117036, Russia)

In psychophysics, adaptation is defined usually as a decline of loudness of steady stimulus as a function of time. It occurs over periods ranging from seconds to minutes. Recent psychophysical data show that during a ''mean-term'' adaptation period (the first 30 seconds of the sound) the detection of amplitude changes considerably improves. Physiological correlates of ''mean-term'' adaptation in single units located in different regions (cochlear nucleus, superior olive, *torus semicircularis*) of the auditory system of the immobilized frog *Rana temporaria* are explored. Driven responses to long-duration pure tones and AM tones decreased over time periods ranging from milliseconds to minutes. The adaptation

increased at the upper levels of the auditory system. The response to pure tone stimuli declined more rapidly than the response to AM tones. For AM stimuli the adaptation increased as the modulation depth decreased. The considerable enhancement of the phase locking to the sound envelope could be observed during the first 10–15 s of the AM tones with low modulation indexes. The effect of phase-locking enhancement during the mean-term adaptation matched some psychoacoustical findings. [Work supported by the RFBR.]

2aPP26. Mechanisms underlying recovery from forward masking in electrical stimulation. Monita Chatterjee (Dept. of Auditory Implants and Percept. Res., House Ear Inst., 2100 W. Third St., Los Angeles, CA 90057)

In normal hearing, adaptation is hypothesized to be closely related to recovery from masking. In electrical stimulation, adaptation at the inner hair cell synapse is absent. However, refractory mechanisms can result in an adaptationlike decrement at high stimulation rates. Shannon [J. Acoust. Soc. Am. **88**, 741–744 (1990)] found that normalized psychophysical forward masking functions were similiar in normal and electric stimulation, and suggested that the mechanisms underlying forward masking are primarily retro-cochlear. Recovery from forward masking was studied in five Nucleus-22 cochlear implant users. Results were found to be strongly subject-dependent. In general, a rapid decay in the first 15 ms after masker offset, and a slow decay persisting up to 200 ms after masker offset, were observed. Under some conditions, plateaus and nonmonotonicities were observed in the region of transition between the two processes. Recovery was independent of masker level. Masker duration, pulse rate, and masker-probe spatial separation strongly influenced the relative weight of the rapid component. The amount of masking near masker offset did not determine the time course of recovery. It is hypothesized that the rapid mechanism is due to auditory nerve refractoriness, while the slow mechanism is likely to be more central in origin. [Work supported by NIDCD.]

2aPP27. Neurophysiological and psychophysical measures of duration discrimination in normal-hearing adults and adults with cochlear implants. Patricia G. Trautwein, Curtis W. Ponton, Betty Kwong, and Michael D. Waring (House Ear Inst., 2100 West Third St., Los Angeles, CA 90057)

The ability to discriminate duration contrasts (a first approximation of voice onset time differences) was evaluated in adult cochlear implant users and normal-hearing adults using a neurophysiological response (the mismatch negativity—MMN). The MMN is an evoked response generated by a deviant stimulus when embedded in a series of standard homogeneous stimuli. Magnitude of the MMN correlates with psychophysical discrimination thresholds indicating that the MMN provides a measure of perceived acoustic change [Kraus *et al.* (1996)]. The MMN thus provides a neurophysiological correlate of acoustic discrimination. The present study compared an objective statistical measure of MMN magnitude to psychophysical measures in response to duration differences of brief auditory stimuli. A ten-click train was paired with shorter trains to create an oddball sequence. The selected duration differences were similar to voice onset time differences that distinguish voiced from voiceless speech phonemes. The MMN magnitude was measured at four scalp-electrode locations. The relationship between neurophysiological and psychophysical thresholds was determined and comparisons were made between the normal-hearing adults and adults with cochlear implants. While some similarities were found, the results suggest that the sensitivity to duration differences differs between normal-hearing and cochlear implant subjects.

2aPP28. Spectro-temporal influence on auditory cortical evoked potential thresholds. Ann Clock Eddins (Auditory Physiol. Lab., Dept. of Speech & Hearing Sci., Indiana Univ., Bloomington, IN 47405, aeddins@indiana.edu)

Previous studies from our laboratory have shown that auditory cortical evoked potential thresholds (N1-P2) decrease with increasing stimulus duration. Similar to behavioral studies of auditory temporal integration, N1-P2 threshold improvement was greater for 1000-Hz than 4000-Hz stimuli. It is unclear whether spectral splatter at short durations contributed to the threshold differences across frequency. The present study investigated spectro-temporal influence on N1-P2 thresholds by manipulating the stimulus rise–fall time, total duration, and stimulus type. Thresholds were estimated in three temporal conditions to compare stimuli with equal rise–fall (2 ms) but different total durations (8, 36 ms), and stimuli with different rise–fall (2, 16 ms) but equal total durations (36 ms). Broadband noise bursts and 1000-Hz tones were used to test each temporal condition. Preliminary results show that thresholds for both noise and tonal stimuli decrease with increasing total duration. No significant differences in threshold were observed as a function of rise–fall time nor were significant differences observed as a function of the stimulus type. The results suggest that N1-P2 thresholds are determined primarily by changes in total duration and that thresholds for short durations at 1000 Hz are not adversely affected by spectral splatter. [Work supported by NSF.]

2aPP29. Psychometric functions for gap detection. Mary Florentine (Commun. Res. Lab., Dept. of Speech-Lang. Pathol. and Audiol. (133FR), Northeastern Univ., Boston, MA 02115), Søren Buus, and Wei Geng (Northeastern Univ., Boston, MA 02115)

To determine the form of psychometric functions for gap detection, three normal listeners were tested at octave frequencies from 0.25 to 8 kHz using a constant-stimulus procedure with a cued yes–no paradigm. The gaps were carried by bandpass noises with an overall level of 85 dB SPL and a bandwidth equal to three auditory-filter bandwidths. To avoid audibility of spectral splatter, the bandpass noise was presented in a complementary notched-noise masker with a spectrum level 10 dB below that of the signal and a bandwidth equal to nine auditory-filter bandwidths. Gap durations were chosen randomly from one of six fixed values (seven at 0.25 kHz) ranging from 0.5 Minimal Detectable Gap durations (MDG) to 2 MDGs (2.8 MDGs at 0.25 kHz). An inaudible gap of 0.5 ms was included as a control condition. Although the false-alarm rate varied among the listeners, the slopes of the psychometric functions were quite similar. The slopes increased as the test frequency increased up to 2 kHz, but remained constant at the higher frequencies. These results indicate that stimulus variability is not important in determining gap-detection thresholds at 2 kHz and above.

2aPP30. Suppression and the upward spread of masking. Andrew J. Oxenham (Northeastern Univ., 360 Huntington Ave., Boston, MA 02115, oxenham@ipo.tue.nl) and Christopher J. Plack (Sussex Univ., Brighton BN1 9QG, UK)

For signals well above a masker in frequency, signal level at threshold grows more rapidly than masker level, leading to the upward spread of masking (USM). It is unclear whether USM is due to a relative increase in masker excitation or an increase in suppression with increasing masker level. Thresholds were measured using a 4-kHz signal, gated with a 10-ms Hanning window. The masker was a 500-Hz-wide random Gaussian noise, centered at either 4 kHz (on frequency) or 2.4 kHz (off frequency) and gated with 2-ms ramps. Masker levels ranged from 50–90 dB SPL. In simultaneous-masking conditions, the signal was temporally centered in a 400-ms masker. In nonsimultaneous conditions, where no suppression should occur, the signal was in a brief gap between two 200-ms maskers. Results are qualitatively consistent with an explanation based on the characteristics of the basilar-membrane input-output function. At higher signal levels, the differential compression of masker and signal produces USM

even in nonsimultaneous conditions (no suppression). At low signal levels (below about 40 dB SPL) the basilar-membrane response to single tones is nearly linear and nonlinear USM is only observed in simultaneous-masking conditions. [Work supported by the Wellcome Trust and the Royal Society.]

2aPP31. Intensity discrimination of ramped and damped tones. Robert S. Schlauch, Dennis T. Ries, Jeffrey J. DiGiovanni, Sara Elliot (Univ. of Minnesota, Minneapolis, MN 55455), and Shari L. Campbell (Univ. of Georgia, Athens, GA 30602)

Intensity discrimination of 1.0-kHz tones that had exponentially ramped or damped envelopes was measured as a function of duration and level. Standard levels ranged from near threshold to 90 dB SPL and durations ranged from 10 to 200 ms. Just-noticeable differences for some subjects were smaller for damped tones than for ramped tones for the same duration and level. For these conditions, damped tones showed more temporal integration than ramped tones. This result is somewhat surprising given that the subjective duration of ramped tones increased by a factor of 100 for stimuli ranging from 10 to 200 ms, whereas the subjective duration of damped tones increased by a much smaller factor of 28 over the same range of durations. Possible physiological and cognitive explanations for these differences will be discussed. [Work supported by NIH-NIDCD R29 DC01542.]

2aPP32. Preferential detection of rising versus falling intensity. John G. Neuhoff and Jennifer Wess (Dept. of Psych., Lafayette College, Easton, PA 18042)

An increase in intensity caused by the approach of a sound source can signal an ecological threat or opportunity, and therefore a potentially important event. A receding source produces a pattern of decreasing intensity change and a potentially less important event. Therefore preferential responding by an organism to rising intensity stimuli may provide a selective advantage. To investigate the hypothesis that listeners respond preferentially to rising intensity stimuli, listeners were presented with white-noise samples that either rose or fell dynamically 20 dB (SPL) in intensity. Stimulus duration ranged from 0.5 s to 1.5 s, and onset of intensity change ranged from 0 s to 1 s. In a speeded 2AFC task listeners had to determine whether intensity rose or fell. Results show that listeners detected increases in intensity more quickly and more accurately than equivalent decreases. The findings are consistent with the interpretation that the auditory system processes some acoustic events according to a set of ecologically defined priorities.

2aPP33. Detection of differences in rate of frequency change in gliding tones. John Madden (Dept. of Commun. Sci. and Disord., Univ. of North Dakota, Grand Forks, ND 58202)

The purpose of the study was to measure discrimination for glides in frequency, based on discrimination of rate of frequency change. Changes in frequency transition span and duration were eliminated as confounding cues by using signals of equal frequency span and duration that smoothly changed in rate at the half-way point in their duration. That is, each signal consisted of a first segment at one rate of frequency change, a brief rate transition segment, and a second segment at another rate of frequency change. The listeners compared a standard signal during which there was no change in rate (a linear glide) with a target signal during which the second segment changed at a greater rate than the first. The task was to indicate which signal changed in rate. Discrimination thresholds were obtained for signals with a range of center frequencies, frequency spans, and durations. Both up- and down-glides were used. Preliminary results indicate that discrimination thresholds are relatively constant for center frequencies from 0.5 kHz to 2 kHz, but increase by a factor of 2.5 at 6 kHz. [Work supported by Research Grant Number 1 R15 dc 02662-01 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.]

2aPP34. Contribution to the study of the audiometrical scores sources of 20 variation. Marielle Bruyninckx, Saefda Hamache, and Bernard Harmegnies (Dept. of Speech Commun., Univ. of Mons, 18 place du Parc B7000, Mons, Belgium, marielle.bruyninckx@umh.ac.be)

This paper summarizes the development and the presentation of a discrimination test of the relative height of two pure sounds. Sixty pairs of sounds have been presented to 68 subjects. The elaboration of these pairs was based on a selection of ten initial stimuli that were paired according to three different intervals: octave, 1/3 octave, and 11/24 octave. These stimuli were then presented by pairs to the subjects in a random way, according to an ascending or a descending sequential structure. The aim of this work is to test: whether the paired sounds' presentation order (ascending or descending) has, or has not, an effect on the subjects' results; whether the kind of frequential interval is a determining factor in the subjects' discriminatory abilities; whether the musical education plays a role on performances (in comparing musicians' and nonmusicians' performances). Together with measurement of reliability, the results show that: it is easier to discriminate sounds in an ascending structure, whatever the frequential intervals may be; there is no significant difference between 1/3 octave, octave, or 11/24 octave intervals; the musicians present better discriminatory abilities than the nonmusicians concerning the ascending structure pairs, but worse abilities when the structure of the pairs is descending.

2aPP35. Hawkins and Stevens revisited at low frequencies. Adrianus J. M. Houtsma (IPO, Eindhoven Univ. of Technol., P.O. Box 513, 5600 MB Eindhoven, The Netherlands, houtsma@ipo.tue.nl)

The classical data on masking of pure tones by white noise [J. E. Hawkins, Jr. and S. S. Stevens, J. Acoust. Soc. Am. **22**, 6–13 (1950)] yield a nonmonotonic critical ratio function that increases at low and at high frequencies. Modern bandwidth estimates of auditory filters, however, yield equivalent rectangular bandwidth (ERB) functions that are monotonic with frequency. This apparent discrepancy has sometimes been interpreted as evidence for frequency-dependent detection efficiency in the auditory system. Another possible reason for the discrepancy, however, could be the fact that the classical data were measured with headphones that needed substantial correction at low frequencies to compensate for acoustic leakage. To test this, the Hawkins and Stevens experiment was repeated with Etymotic ER-2 insert phones over a frequency range from 30 to 8000 Hz. An adaptive forced choice, as well as a Bekesy tracking procedure, was used. The resulting critical ratio function shows monotonic dependence on frequency, down to 90 Hz, similar to the ERB function. This appears to make critical ratio data, in principle, consistent with ERB data measured with notched-noise techniques, and largely eliminates the need for a frequency-dependent detection efficiency in the auditory system.

2aPP36. The perception of three tone patterns by Indian and American listeners. Nandini Iyer, Evelyn M. Hoglund, and Mari Reiss Jones (Ohio State Univ., Dept. of Speech and Hearing Sci., 51 Pressey Hall, 1070 Carmack Rd., Columbus, OH 43210)

Serial identification of three tone patterns by Indian and American listeners was investigated in three experiments. Six pitch patterns combined a high-frequency (H) and a low-frequency tone (L) (e.g., H, H, L or L, L, H). The first experiment presented all patterns at each of three different tempi (fast, medium, and slow trial blocks) with blocks presented in either ascending or descending order of tempi. Each pattern occurred six times (randomly) within a block. The second experiment investigated

the differences in pattern perception between musicians and nonmusicians among the previously tested Indian and American listeners. In the third experiment, frequency separation between the high and low tones was investigated at two different tempi. A repeating cycle of three tone patterns was presented for identification and streaming thresholds were determined for Indian and American listeners. Preliminary results indicate that the two groups of listeners differ in their ability to identify three tone patterns. These differences were affected by the listeners' musical background. Frequency separation and tempo also differentially affected pattern perception among the two groups of listeners.

2aPP37. Tempo perception: The multiple-look strategy is limited by a short temporal window. Nathalie Panissal-Vieu and Carolyn Drake (Lab. de Psych. Exp., CNRS URA 316, Univ. Rene Descartes, 28 rue Serpente, 75006, Paris, France)

Previous studies [Drake and Botte (1993)] indicated that sensitivity to tempo changes in isochronous sequences improves as the number of intervals in the sequence increases. Further studies suggested that this multiple-look process only works within a limited temporal window of 0.5 or 2 s depending on the tempo of the sequence [Panissal-Vieu *et al.* (1995)]. This temporal window represents the time required to extract all the necessary temporal information (mean interval duration and dispersion). Two experiments investigated how this mental representation is used to make temporal judgements. In a 2 AFC paradigm, trained listeners identified the slower of two isochronous sequences. The number of intervals was varied in either the first or second sequence, with the temporal window saturated in the other sequence. The number of intervals required to saturate the temporal window of the first sequence varied with tempo, as predicted. However, only three intervals were required in the second sequence. These results suggest a buildup of a stable and accurate memory trace during the first sequence which is subsequently compared with the first two intervals of the second sequence. There is no need to recalculate the memory trace for the second sequence.

2aPP38. Investigations and analysis of assessment word for sound quality attaching importance to high-order sensation. Tomoharu Ishikawa, Shingo Fuyuki, and Makoto Miyahara (School of Information Sci., Japan Adv. Inst. of Sci. and Technol., 1-1 Asahidai, Tatsunokuchi-machi, Nomi-gun, Ishikawa-Ken, 923-12 Japan)

The standpoint of our research is to evaluate subjectively the sound quality of sound-reproducing systems by evaluating deep emotions which are brought by the sounds of music and speech. In evaluating sound qualities, assessment words that express the emotive meanings which have arisen in an evaluator's mind are used. The importance of sound reproduction is the transmission of ''mind'' that a musician wants to send. Therefore the assessment words which can represent the high-order sensations of the sender need to be obtained. Not only conventional assessment words but also assessment words which represent high-order sensation have been investigated and then analyzed. The method is as follows, (1) 35 representative assessment words by the KJ method from 1322 assessment words which were obtained by surveying audio publications over the past three years were obtained. (2) Psychological distances among 35 representative assessment words have been analyzed by the MDS method, and four clusters by hierarchical clustering have been obtained. (3) Independent to (2) above, representative assessment words by the public attention evaluation method have been scored and ranked. (4) The important assessment words which are strongly correlated with the overall sound quality have been selected.

2aPP39. Factor analysis of the perception of the evaluation of environmental sounds using adjectives describing sound quality, emotional state, and information carried by sounds. Toshio Sone, Kôji Abe, Kenji Ozawa, and Yôiti Suzuki (Res. Inst. of Elec. Comm., Tohoku Univ., Sendai-shi, 980-0812 Japan)

In order to investigate into the perception of environmental sounds, a psychoacoustic experiment was carried out. The experiment was performed with the SD (semantic differential) method using 66 environmental sounds for 39 pairs of the descriptive adjectives in Japanese. These pairs of descriptive adjectives are roughly divided into three groups. One of them was adjectives describing sound quality that had been used in similar studies reported so far. The other groups were adjectives describing emotional states of evaluators due to the sounds and information carried by sounds in addition to them. The new adjectives were selected for explaining a more general state of perception of the environmental sounds. As the result of a factor analysis, five factors were obtained. Three of them were factors describing ''aesthetic state,'' ''volume,'' and ''brightness'' of sounds. These factors correspond to the three well-known factors describing sound quality. Other than these, two new factors were derived: a factor concerning information on environment and a factor concerning the significance of sound. Moreover, it was shown that emotional states produced by environmental sounds correlate with the factor concerning the aesthetic state.

2aPP40. Sound quality estimation using hearing sensation. Soon-Hyob Kim (Kwang Woon Univ., Wolgaedong, Nowongu, Seoul, Korea, inphase@chollian.dacom.co.kr)

This paper is on a study of objective sound quality estimation by using human hearing sensation. Until now, the sound quality estimation has been done by subjective methods by people. Unfortunately, it depends on environment and condition of the human ear. So it was difficult to tell exactly how sound quality was. This study was performed to make objective estimation of sounds, regardless of the factors. In this paper, six parameters were defined, which are related to the sound quality: loudness, sharpness, fluctuation strength, roughness, pleasantness, and annoyance, and signal processing methods for them were developed. In this way, a simulator was implemented that measures human hearing sensation. To prove objectiveness of the simulator, the simulator results were compared with subjective test results which were given by people. As a result, the average error rates were 8.1% in loudness, 12% in sharpness, 20% in fluctuation strength, 18.5% in roughness, 5.5% in pleasantness, and 19% in annoyance. The overall average error rate was 13.7%, which showed that the simulator is very useful.

2aPP41. Analysis of annoyance of low-frequency beat inside a vehicle. Klaus Genuit (HEAD acoustics GmbH, Ebertstrasse 30a, 52134 Herzogenrath, Germany)

Because station wagons have a greater volume than sedan vehicles, low-frequency acoustic effects (''humming'') very often occur in the frequency range between 30 and 50 Hz. This is not only unpleasant, but can lead to physiological changes in the driver due to specific properties of the acoustic signals. The mechanisms may be very varied: Excitation frequencies can be produced by the powertrain, exhaust system, or wheel suspension and are transmitted via the mounting points. Resonances in the regarded frequency range may well be present along these transmission paths. The spectral components thus arriving at the structure are radiated into the car interior via various ''loudspeakers,'' such as floor panels, roofs, or doors. Moreover, for particular frequencies some loudspeakers may show special resonances or the volume of the vehicle interior results in amplification. During field trials, excitation mechanisms and transmission paths were investigated using the combination of Artificial Head and multichannel measurement technology that allows subjective evaluation of various models. Playback of the Artificial Head signals, however, is not only via headphones, but via a special ''sound car'' which also creates low-frequency excitation of airborne and structure-borne noise. This en-

ables a connection to be made between the frequencies at the driver's ear giving cause for complaint and the various excitation points. Moreover, the use of a two-dimensional laser vibrometer enables a connection to be made between vibrating structures and the annoying sound components a driver is hearing.

2aPP42. Spectrum contrast and noise annoyance. Anna Preis (I.I.S. Univ. of Tokyo, 7-22-1 Roppongi, Minato-ku, Tokyo, 106 Japan)

According to the computational view, the identification of a sound event by the auditory system consists of processing the information it extracts from the temporal pattern and frequency spectrum. The fine structure of the frequency spectrum facilitates, among other things, the determination of the relative amplitudes of the partials. The auditory system uses the values of these amplitudes to compute the total loudness of the sound event by summing the specific loudness of each partial across a set of critical bands. This correlates well with experimental data of perceived loudness. The present study investigates whether the auditory system makes use of other information contained in this spectrum. The tested hypothesis stated that the auditory system detects change in the relative amplitude of the partials. It was assumed that these changes can be traced in experimental data of perceived annoyance of noise. The experimental results show that the best correlation between the reported annoyance and calculated values was obtained for the average absolute value of the second derivative of the weighted specific loudness calculated as a function of the critical-band rate scale. This specific loudness was weighted by the function introduced in sharpness calculation [W. Aures, Acustica **59**, 130–141 (1985)].

2aPP43. Detection and discrimination of amplitude-modulated signals in noise. King Chung (Northwestern Univ., 2299 N. Campus Dr., Evanston, IL 60208/House Ear Inst.), Fan-Gang Zeng (House Ear Inst., Los Angeles, CA), and Beverly A. Wright (Northwestern Univ., Evanston, IL)

Comodulation masking release demonstrated improved detection of pure-tone signals in amplitude-comodulated noise bands [Hall *et al.*, J. Acoust. Soc. Am. **76**(1), 50–56 (1984)]. However, in most real-world situations, signals such as speech sounds are comodulated and noises such as cocktail party and traffic noises are not. Two experiments were designed to examine whether comodulation of signals can lead to improved signal detection in noise. Experiment I used a 2AFC method to measure detection threshold of a 2/3-oct-spaced, five-tone complex in pink noise. The complex tones were either amplitude comodulated or randomly modulated. Results from three normal-hearing listeners showed no difference in detection threshold between amplitude-comodulated and randomly modulated conditions. Experiment II used a method of constant stimuli to measure the psychometric function of discriminating six vowel-like stimuli in pink noise. These vowels consisted of three formant-frequency sinusoids which were unmodulated or amplitude-comodulated or randomly modulated by a 125-Hz sinusoid simulating the voice pitch. While no difference was observed between comodulated and randomly modulated conditions, the psychometric function for the nomodulation control condition was shifted by 4–6 dB toward lower signal-to-noise ratios. A possible explanation is that auditory fibers phase-lock to formant-frequencies of the unmodulated stimuli and to the fundamental frequency of the modulated stimuli.

TUESDAY MORNING, 23 JUNE 1998 EAST BALLROOM B (S), 9:15 A.M. TO 12:00 NOON

Session 2aSA

Structural Acoustics and Vibration: New Challenges in Structural Acoustics

J. Gregory McDaniel, Chair
Aerospace and Mechanical Engineering Department, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Chair's Introduction—9:15

Invited Papers

9:20

2aSA1. Product design and structural acoustics. Richard H. Lyon (RH Lyon Corp, 691 Concord Ave., Cambridge, MA 02138)

If, by structural acoustics the analysis and understanding of noise and related problems through an understanding of mechanical vibration is meant, then structural acoustics has a very long history in product design through the application of isolators and structural damping to products. But the greatest amount of research and methods development has been focused on fairly large commercial and military structures like submarines, aircraft, and spacecraft. The resources for, attention to, and methods for structural acoustics that has been available in these and the automotive industries has not been present for the industries that make most of the products each of us deals with daily. Structural acoustics as applied to hairdryers, vacuum cleaners, and lawnmowers presents challenges in both understanding and cost that the technical community can respond to. For example, designers would like to predict sound from CAD drawings. Such a need can be an inspiration to some of us. This presentation will provide examples of the need and some possible responses.

9:45

2aSA2. Wind noise challenge in the automobile industry. Jen-Yuan Her (Ford Motor Co., Dearborn, MI 48121) and William B. Coney (BBN Technologies, Cambridge, MA 02138)

Aerodynamically generated noise during vehicle operations is of increased importance for automobiles as operating speeds increase and design improvements reduce noise contributions from the powertrain and road tire interaction. Wind noise reaches the interior through a variety of mechanisms, including the aeroacoustic penetration and aerodynamic excitation of vibration, and its radiation of sound from vehicle greenhouse (roof, windshield, side glasses, and backbite) and underbody. Although vehicle wind noise

performance keeps improving, the fundamentals of wind noise generation and transmission paths still remain uncertain. Instead of relying on measurements on hardware, the recent development of the practical CAE approach, which is different from direct CFD plus FEA simulation, may shed light on solving the puzzle from the statistical energy analysis (SEA) point of view. This paper raises issues to be solved in vehicle-shape-related wind noise root-cause understanding and prevention using SEA. Wind noise mechanisms such as vehicle body wall pressure fluctuations (pseudonoise) generation, its scaling law, and interaction with body structure are examined in detail. Challenging issues such as the identification of low wave number contribution versus convective ridges of pseudonoise to structure excitation and how sound is radiated to driver's and passengers' ears are discussed.

10:10

2aSA3. Brake-noise challenge in the automobile industry. Shih-Emn Chen and Jen-Yuan Her (Ford Motor Co., Dearborn, MI 48121)

Brake noises generated during vehicle operation are caused by nonlinear structural behavior that is difficult to model analytically. Noises, predominantly contact related, falling in this group can be classified into two categories based on the frequencies at which they occur: (1) brake squeal in kHz ranges; (2) brake moan and groan in 100–300 Hz range. These noises may often be accompanied by structural instability caused by some, if not all, of the following phenomena: (1) lining deformation; (2) negative slope in friction-velocity characteristic; and (3) stick-slip. In addition to the aforementioned phenomena, coupled resonance of brake components is also a necessary condition of the brake-noise generation. Currently, vehicle brake-noise design prevention is still based on a trial and error basis. No reliable design prevention tool is yet established. Although computer simulations were sometimes used to predict the structural instabilities, no clear understanding of the link between structural instabilities and radiated noise has been established. The major challenge to automobile noise control engineers will be to establish a noise radiation model which could identify the influence of the key design parameters on the noise level.

10:35

2aSA4. Integration of small transducers in commercial products. Jarmo Hietanen (Nokia Res. Ctr., P.O. Box 100, FIN-33721 Tampere, Finland, jarmo.hietanen@research.nokia.com)

Miniaturization of electronic devices requires smaller components as well as a higher level of integration. Whenever audio acoustics is involved, as in mobile phones, it is exactly the acoustics components that offer the largest challenge in this respect. There are physical limitations to how small those components can be made, and with present methods of integration, even the smallest acceptable size is still too large. Progress in material research is expected to offer new possibilities for solving problems in further integration of components, e.g., microphones. In the development of silicon microphones, new materials and processes are utilized already. Decreasing the size of components and systems makes the problem of vibration coupling through structures more severe. Furthermore, in smaller components there are a number of physical effects to be accounted for, not only in practice but also in numerical simulation models of these systems. An example of this is the effect of viscosity when predicting acoustic performance of the system. Normally viscosity is neglected, but in very narrow ducts, for example, viscosity causes significant effects. Very small components and systems cannot be expected to operate physically, as do their larger predecessors.

Contributed Papers

11:00

2aSA5. Complex resonances of elliptic cylinders. Stéphane Ancey, Antoine Folacci, and Paul Gabrielli (URA 2053 CNRS, SDEM, Université de Corse, Faculté des Sci., BP 52, 20250 Corte, France)

Complex resonances of elliptic cylinders of arbitrary eccentricities (including the limit cases of the circular cylinder and the strip) are exactly determined by solving the Helmholtz equation with the Sommerfeld's radiation condition at infinity and various boundary conditions on the scatterer. Resonances are then fully classified along the two irreducible representations A_1 and A_2 of the symmetry group C_2 of the scatterer. This spectrum of resonances is also asymptotically recovered (i) by using the ordinary geometrical theory of diffraction, and (ii) by applying the method of the Gutzwiller–Voros spectral determinant. This last technique could be developed in acoustic multiple scattering. Indeed, in this context, because of the complexity of the system of rays and its chaotic behavior, ordinary geometrical theory of diffraction fails to provide accurate results.

11:15

2aSA6. Acoustic scattering by a three-cylinder scatterer: Theoretical and experimental study of resonances. Paul Gabrielli, Yves Decanini, Antoine Folacci, and Elisabeth Fournier (URA 2053 CNRS, SDEM, Université de Corse, Faculté des Sci., BP 52, 20250 Corte, France)

Acoustic scattering by a system of three identical parallel cylinders, located at the vertices of an equilateral triangle, is theoretically, numerically, and experimentally studied by emphasizing the role of the symmetries of the scatterer. Incident and scattered fields are expanded over the different irreducible representations of C_{3v}, the symmetry group of the scatterer. From the boundary conditions, an infinite set of four linear complex algebraic equations is obtained, where the unknown coefficients of the scattered fields are uncoupled. As a consequence, the form function is calculated for various incidence and observation angles, and the positions of the scatterer resonances are numerically determined in the complex plane of the reduced frequency. New types of resonances appear. A physical interpretation of these resonances by the phase-matching of geometrical and surface waves along closed paths is given by using the geometrical theory of diffraction. These resonances are experimentally observed for elastic cylinders immersed in water. The work is a step toward a complete study of chaos (associated with the system of rays) in the context of multiple scattering.

11:30

2aSA7. Bandpass filters for the evaluation of short reverberation times. Malte Kob (Inst. of Tech. Acoust., RWTH Aachen, Templergraben 55, D-52056 Aachen, Germany, kob@akustik.rwth-aachen.de)

In building acoustics, evaluating the sound insulation of building elements suffers from rather bad reproducibility. The reason is the energy flow between the building element and the surrounding structure of the test laboratory. To get more reproducible measurements, it is desirable to measure the loss factor of the element, e.g., by evaluation of the reverberation time of the structure. Unfortunately, the evaluation of decreasing reverberation times T_{60} of building elements, for instance in one third octave bands is distorted by a growing influence of the applied bandpass filter. In literature a minimum permissible product of the bandwidth and the reverberation time of the filter $B \cdot T = 16$ can be found but no error in the estimation of T_{60} beyond this limit is given. A maximum admittable error depending on the $B \cdot T$ product is derived for several filters used in

real time analyzers. An optimized filter according to the IEC 1260 standard is developed and results of simulated and measured short reverberation time evaluations are presented. Other possibilities of the investigation of the loss factors and their practical use are discussed.

11:45

2aSA8. Piezoelectric thin shell theoretical model and eigenfunction analysis of radially polarized ceramic cylinders. D. D. Ebenezer and Pushpa Abraham (Transducer Group, Naval Physical and Oceanograph. Lab., Naval Base, Kochi 682 004, India, root@drnpol.ren.nic.in)

A piezoelectric thin shell theoretical model of radially polarized ceramic cylinders is first developed. An eigenfunction approach is then used to determine the displacement, stress, electric field, and input electrical admittance of the cylinder. In the model, the effect of bending strain is included and it is assumed that the electric field varies linearly between the electrodes. The electrostatic condition is then used to express the electric field in terms of the displacement. Next, the equations of motion and the boundary conditions are derived by using Hamilton's variational principle. They have the same form as the corresponding equations for a thin elastic shell and are solved after expressing the frequency-dependent displacement as the sum of the static displacement and a weighted sum of the eigenfunctions of the short-circuited cylinder. The eigenfunctions are determined here by using a general method and, therefore, cylinders with arbitrary boundary conditions can be analyzed. It is shown that experimental resonance frequencies and input electrical admittance of free-free cylinders are in good agreement with those computed here and that the electric field is in good agreement with results obtained by other investigators who used a numerical method to solve the equations.

TUESDAY MORNING, 23 JUNE 1998 FIFTH AVENUE ROOM (W), 8:00 TO 10:35 A.M.

Session 2aSC

Speech Communication, Psychological and Physiological Acoustics and Musical Acoustics: Rhythm in Music and Speech

Robert F. Port, Cochair
Department Linguistics, Indiana University, 330 Memorial Hall, Bloomington, Indiana 47405

J. Devin McAuley, Cochair
Psychology Department, University of Queensland, Brisbane, QLD, 4072 Australia

Chair's Introduction—8:00

Invited Papers

8:05

2aSC1. Limit cycle dynamics in prosody. Fred Cummins (Dept. of Linguist., Northwestern Univ., Evanston, IL 60208)

Compared to music, juggling, dancing, etc., speech may seem to be only weakly rhythmic. New experimental evidence suggests, however, that rhythmic constraints on the timing of certain speech events can be quite strong. This is particularly clear when subjects repeat a phrase in time with a periodic stimulus. Mode locking, mode switching, and hysteresis have all been documented, thus showing hallmarks of the hierarchical entrainment among two periodic cycles. These cycles are coextensive with independently motivated prosodic units, the foot (or prosodic word), and the intonational phrase. It appears that this repetition task induces a second-order limit cycle dynamic in each of these units. [Work supported by the Indiana University Graduate School and the Andrew W. Mellon Foundation.]

8:25

2aSC2. Temporal correlates of foot-level timing in English and Japanese. Keiichi Tajima and Robert F. Port (Indiana Univ., 322 Memorial Hall, Bloomington, IN 47401)

Languages of the world are felt to be spoken with different kinds of rhythm. But what is the basis for this impression? Notions like ''stress-timed'' and ''syllable-timed'' have been difficult to verify experimentally because rhythm in spoken language rarely yields consistent temporal regularities. Experiments were conducted in which native English and Japanese speakers repeated phrases using a specified rhythm in time with an auditory metronome. Results demonstrate that temporal correlates exist for prosodic units approximately the size of the ''foot,'' consisting of one relatively strong syllable followed by zero or more weak syllables. Strong (foot-initial) syllables show greater temporal stability than do weak (foot-internal) syllables when the timing of syllables was perturbed by manipulating inherent segmental durations. This is found for both English and Japanese. However, in English, increasing the foot size by the addition of weak syllables is accompanied by compensatory shortening only within the foot, while in Japanese, compensation occurs both within and between adjacent feet, suggesting a weaker role for this unit. By eliciting speech in a rhythmic context, such as ''speech cycling,'' both language-general and language-specific properties of speech rhythm can be directly measured.

8:45

2aSC3. Attention, rhythmicity, and expectancy. Mari Reiss Jones (Dept. of Psych., 142 Townshend Hall, Ohio State Univ., Columbus, OH 43210)

It seems plausible that body gestures created by ourselves and other living things may derive from some underlying periodic processes. Motor rhythms are ostensibly measurable. The case for *attending having a rhythmic aspect* is more challenging, in part because attention is not directly observable. So the question is asked: ''How does an approach that posits internal (nonobservable) oscillators actually explain attention?'' Three steps are considered. The first considers that attending primarily involves selectivity. Visual attention theories often relate selectivity to space; attention is focused at one place not another. Significantly, few theories of auditory attention exist. But surely in speech and music, attention can operate selectively in time; it can be targeted to one temporal location, not another. The second step requires that nonlinear oscillators describe selective attending in time. That these models deliver is evident in several recent oscillator models, including an elegant one developed by Ed Large. Step three involves connecting oscillator activities to measurable behaviors that reflect attending. Listeners are primed with auditory patterns, only some of which display rhythms predicted to facilitate temporal expectancies accounted for by oscillator entrainment. The talk centers on these experiments, their outcomes, and implications for a theory of dynamic attending.

9:05–9:20 Break

9:20

2aSC4. Context effects in time discrimination. J. Devin McAuley (Dept. of Psych., 142 Townshend Hall, Ohio State Univ., Columbus, OH 43210)

Music and speech processing exemplify skilled perception of temporal patterns, as both require an ability to make fine-grained comparisons of relative duration. Clock theories of time perception have been used to model the discrimination of isolated durations successfully, by assuming that temporal resolution is determined by the statistical variability of an independent central timing process. However, the explanatory power of the clock model has been shown to be limited when applied to more naturalistic tasks such as the detection of a temporal deviation embedded within a temporal pattern. This paper discusses an entrainment theory of time discrimination that is based on the concept of coupled oscillators. The theory accounts for a number of contextual effects in time discrimination which have been problematic for clock models. The theory also applies equally well to the discrimination of isolated durations. An important link is demonstrated between fast and slow tempos, short and long durations, and early and late onset of a comparison sequence which only becomes clear when one assumes that the relevant psychological variable in time discrimination is relative phase.

9:40

2aSC5. Categorization of temporal intervals. Edward W. Large (Dept. of Psych., Univ. of Pennsylvania, Philadelphia, PA 19104)

The rhythms of music and speech comprise sequences of temporal intervals separating the onsets of component events (e.g., notes and syllables). The perception of rhythmic patterning in such sequences involves the identification of distinct interval categories. Notating a musical rhythm, for example, relies upon the perception of nominal relationships between time intervals (e.g., 2:1, 3:1). Such categorization appears robust to perturbations that alter the absolute duration of temporal intervals. Thus it makes sense to ask whether the perception of categorical temporal relationships is ''categorical perception.'' The answer depends not only on the ability to make category judgments, but also on the differential discriminability of temporal intervals. Entrainment models propose that perceived onset time depends upon onset phase relative to an internal oscillation [Large and Kolen, Conn. Sci. **6**, 177–280 (1995); McAuley, Proc. Cog. Sci. Soc., 615–620 (1996)]. Such models predict discrimination performance, however, they do not address categorization. This paper presents an extension of entrainment models that addresses categorization: In addition to relative phase dynamics, it explicitly considers the amplitude dynamics of internal oscillations. It captures existing data on interval categorization and makes new predictions regarding categorization and discrimination performance. Tests of the predictions are discussed.

10:00

2aSC6. Models for entrainment of neural rhythms. Bill Baird (Dept. of Mathematics, Univ. of California, Berkeley, CA 94720)

Psychological results suggest that motor and perceptual behavior are organized by neural rhythms with periods from seconds to tenths of a second. It appears that their entrainment to external speech and music or to internal motor coordination rhythm is essential for effective performance. It is proposed that theta-alpha sensory-motor cycles (5–15 Hz) form the basis of these rhythms. They generate and entrain rhythmic motor output or temporal perceptual expectancies. It is hypothesized that cortical EEG and event-related potentials (ERPs) reflect thalamic biasing currents. The cortical evoked potential is a 5–15 Hz signal whose peaks and troughs correlate with stages of cortical information processing in sensory-motor tasks. Cortical rhythms are found in visual, auditory, and somatosensory areas as well as the hippocampus and reticular system that entrain to periodic inputs. If the arrival of an auditory stimulus is regular, it has been shown that 5–12 Hz auditory activity becomes phase locked *in advance of the target stimulus* to the expected arrival time. Our model of rhythmic expectation accounts for this since the theta-alpha rhythmic cycles constitute a fast clock that adjusts phase and frequency to match stimulus arrival, and longer period cycles are derived from this base clock.

10:20–10:35 Discussion

Session 2aSP

Signal Processing in Acoustics: Implementation Issues of Acoustical Signal Processing in Real Time Systems I

Stergios Stergiopoulos, Chair
Environmental Medicine, DCIEM, Defence Civil Institute, P.O. Box 2000, North York ON M3M 3B9, Canada

Chair's Introduction—9:15

Invited Papers

9:20

2aSP1. Implementation of real-time acoustic communications systems. David Herold, Robert Nation, Geoffrey Edelson, and Eric Will (Sanders, A Lockheed Martin Co., Adv. Systems Directorate, P.O. Box 868, Nashua, NH 03061-0868)

Over the past several years, the field of underwater acoustic (UWA) communications has been focusing on bandwidth efficient phase coherent modulation techniques. In this application, practical communication algorithms, and thus their implementations, are still maturing. In addition to the computer-intensive algorithms used to track and decode these acoustic data, the systems may incorporate compression, coding, and error-detection schemes to provide usable communication links. Several programs are currently underway to utilize these developments in fieldable systems. In this paper, an overview of the signal processing requirements for UWA communications is provided, as well as some constraints placed on these systems by specific applications. The effect of these requirements on overall system design and implementation is outlined.

9:50

2aSP2. An adaptive processing structure for integrated active-passive sonars deploying cylindrical arrays. Amar C. Dhanantwari (Dept. of Elec. and Comput. Eng., Univ. of Western Ontario, London, ON N6A 5B9, Canada) and Stergios Stergiopoulos (Defence and Civil Inst. of Environ. Medicine, North York, ON M3M 3B9, Canada)

The paper details the development and implementation of an adaptive processing structure for integrated active–passive sonars deploying cylindrical arrays. The main concept includes decomposition of the computationally intensive multidimensional beamformer into two simple modules, which are line and circular array beamformers. Thus, the multidimensional beamforming process can now be divided into coherent subprocesses which lead to efficient implementation in real-time sonar systems. Furthermore, the application of spatial shading to reduce the side-lobe structures can now be easily incorporated. Moreover, the new approach makes the implementation of adaptive schemes in multidimensional sensor arrays practically achievable. The proposed adaptive processing concept has been implemented in an integrated active–passive real-time sonar deploying a cylindrical array. Real data results from the adaptive and conventional beamforming outputs of the cylindrical array sonar system demonstrate the superior performance of the adaptive beamformer in suppressing the reverberation and cluttering effects in active sonar applications. Moreover, for passive sonar applications, the adaptive processing provides substantially improved angular resolution performance as compared with that of the conventional beamformer. Both these two performance improvements for a cylindrical array sonar are of particular importance for mine hunting operations.

10:20–10:35 Break

10:35

2aSP3. Sound ranging of impulsive sources in air. Brian G. Ferguson (Defence Sci. and Technol. Organisation, P.O. Box 44, Pyrmont 2009, Australia)

A passive sound ranging method based on wavefront curvature is described where the impulsive sound of a rifle firing is received at three spatially separated collinear sensors. Estimating the differential time-of-arrival (or relative time delay) of the acoustic wavefront at each pair of adjacent sensors enables the range and bearing of the source to be estimated. The temporal variability in the polar coordinates of the source is presented for real data and the system errors affecting the accuracy of the measurements are discussed. The time delay estimation technique is then applied to the passive ranging of an impulsive sound source located in the far field, where the wavefront curvature technique is no longer appropriate as the arriving wavefront is planar. For the plane-wave arrival case, the range of the sound source is estimated from the angle-of-arrival estimates at two widely spaced known locations. The source position estimated by the acoustic method is compared with the source position determined by standard surveying methods. The nonstationary behavior of the sound propagation medium leads to variability in the source localization parameters which is presented for the acoustic data collected during artillery firings.

11:05

2aSP4. Active target detection in the ocean: Optimizing performance and cost. Zoi-Heleni Michalopoulou (Ctr. for Appl. Math. and Statistics, New Jersey Inst. of Technol., Newark, NJ 07102)

The goal of this work is to design a processor for underwater target detection with superior performance to the commonly used matched filter without the efficiency penalties of the optimal filter. The theoretically optimal filter for active detection in the ocean is a correlator between the received signal and the convolution of the transmitted signal and the impulse response of the ocean. This filter is seldom employed because of the considerable computational overhead. The practically preferred method is a standard matched filter that correlates the transmitted signal to the received signal. It is computationally efficient but analytically suboptimal because it ignores the effects of the propagation medium on the transmitted signal. The standard filter is also preferable because of its attractive property of portability across environments, this feature stemming from deliberate disregard of the physics of the propagation medium. Here, the potential for improving detection performance is explored through modified matched filters that retain the portability property but do not resort to the exhaustive multiparameter grid searches of the optimal filter. [Work supported by ONR Ocean Acoustics, through Grant No. N000149710600.]

11:35

2aSP5. Implementation of a broadband, environment-sensitive detector in real-time systems. Georgios Haralabus (SACLANT Undersea Res. Ctr., Viale San Bartolomeo 400, 19138 La Spezia, Italy)

Active systems, such as Low-Frequency Active Sonars (LFAS), have extensively used the matched filter method for detection purposes. This model-free technique is designed to maximize the signal-to-noise ratio of known signals irrespective of the nature of the waveguide. However, shallow water and dense multipath conditions often distort and obscure the matched filter output to a point at which reliable detection is not possible. The objective of this project is to investigate the implementation of a Channel Sensitive Processor (CSP) in real time systems. The performance improvement of the conventional matched-filter processor by exploiting characteristics of the medium is demonstrated both in an ambient noise and a reverberation-limited environment. It is clearly shown that environmental mismatch poses the primary problem in applying the CSP in real-time systems. Environmental uncertainty can be overcome by utilizing global search methods in conjunction with the CSP detector. In implementing such a scheme, the balance between potential advantages and concomitant disadvantage of inefficiently long real-time processing is examined.

TUESDAY MORNING, 23 JUNE 1998 METROPOLITAN BALLROOM (S), 9:15 A.M. TO 12:15 P.M.

Session 2aUW

Underwater Acoustics: Propagation, Ambient Noise, and Imaging (Précis Poster Session)

Evan K. Westwood, Chair
Applied Research Laboratories, University of Texas, EVG, P.O. Box 8029, Austin, Texas 78713-8029

Chair's Introduction—9:15

Contributed Papers

In this session, an oral summary of each paper will be presented followed by a poster session. All posters will be on display, and all contributors will be at their posters, from 11:10 a.m. to 12:10 p.m.

9:20

2aUW1. Comparisons of simultaneous signal reception over long range among hydrophones with unlike acoustic geometries. H. M. Walkinshaw (Box 72, Peapack, NJ 07977)

Propagation measurements were made to six hydrophones installed offshore Bermuda at depths from 430 to 2450 fm. Four were bottomed onslope or at basin depth; two were suspended, one at the sound channel axis and one 30 m above the deep bottom. Both impulse sources (charges detonated at 18 and 240 m) and a continuous wave (cw) transducer towed at 120 m were used. The 800-nm track began 40-nm south of Bermuda at midpoint of the ensemble of hydrophones and ended off St. Thomas beyond the Puerto Rican Trench. The bottom receiver at 430 fm had the lowest transmission loss; the deep basin hydrophone the highest. Arrival order identification from shot processing for all hydrophones showed performance depended on the degree to which the local bottom slope enabled a receiver to intercept both up and downcoming ray paths. Suspension *per se* was not significant. Energy partition among arrival orders indicated one arrival carried about 50% of received signal strength. About five arrivals were needed before 90% of the total signal was collected, suggesting broad ocean coverage is unavoidably linked to coping effectively with multipath propagation. [Work supported by NavElecs.]

9:25

2aUW2. The form of the normal mode that ensures escape with certainty from a surface channel. Edward R. Floyd (10 Jamaica Village Rd., Coronado, CA 92118-3208)

Surface channels trap those rays that are reflected internally before reaching a deep sound channel according to classical geometric ray tracing. In the wave representation, some leakage of trapped normal modes does occur. Recent advances in wave propagation [E. R. Floyd, Anal. Fond. L. Broglie **20**, 263–279 (1995)] have shown that a wave function having subbarrier energy but with a particular compound modulation in amplitude and wave number does tunnel through a barrier with certainty and without reflection. Herein, a normal mode with a particular compound

modulation is shown to escape from a surface channel with certainty and without reflection. These normal modes with compound modulations, which may be unfamiliar to many workers, are solutions to the wave equation. The investigation considers a bottomless sound velocity profile, $C(z)$, whose reciprocal square, $C^{-2}(z)$, forms a surface channel over the deep sound channel by two piecewise-continuous linear segments. The findings are generalized to include various surface and deep sound channels. While the quantum tunneling problem was originally solved in a trajectory representation [analogous to rigorous ray tracing, cf. E. R. Floyd, J. Acoust. Soc. Am. **80**, 877–887 (1986)], herein a wave representation is used.

9:30

2aUW3. Spatial coherence of sound in convergence zones and shallow zones in the South China Sea. Dinghua Guan, Renhe Zhang, Zhenge Sun, and Yan Wang (Natl. Lab. of Acoust., Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC, zrh@public.east.cn.net)

Some results of a horizontal spatial coherence experiment of sound propagation in the South China Sea were presented. Average water depth is over 4000 m and the depth of axis of the underwater sound channel is 1200 m. Explosive sources were discharged at depths of 7, 100, 200, and 1000 m. Signals were received at a horizontal receiving array of seven hydrophones with an overall length 410 m at a depth of 200 m. Distances between hydrophones are 10, 20, 40, 80, 160, and 1000 m. It is found that at the direct arrival zone and the convergence zones the sound field has good horizontal coherence. At the direct arrival zone the radius of coherence is larger than 200 m. At the first convergence zone the radius of coherence is larger than 250 m. At the second, third, and fourth convergence zones for frequencies lower than 1000 Hz the radii of coherence are larger than 250 m. When the sound source is placed at the shadow zones, the radius of horizontal coherence is less than 100 m, much lower than at the convergence zones. This experiment shows that a horizontal array can be used for beam forming with high spatial gain and good directivities at convergence zones.

9:35

2aUW4. Pulse propagation in random media. Dalcio K. Dacol (Naval Res. Lab., Washington, DC 20375-5320)

Wave propagation in randomly fluctuating media is a problem of continuing interest in ocean acoustics. Usually the work done in this field is carried out in the frequency domain since there are powerful techniques to study wave scattering in the frequency domain, while time-domain formulations are notoriously more difficult to use and obtain results. Nevertheless, our intuition about wave phenomena is much better when discussing such phenomena in a space-time formulation than in a space-frequency formulation. In this presentation wave propagation in a weakly and randomly fluctuating medium is discussed entirely within the space-time formulation. McDonald and Kuperman derived from the wave equation an approximate time-domain progressive wave equation which is the equivalent in the space-time formulation of the well-known parabolic wave equation in the space-frequency formulation. A path integral representation for the solution of this equation is used here to study wave propagation in random media. The weakly inhomogeneity assumption yields an approximate evaluation of the path integral for the Green's function thus allowing an analysis of the time-dependent statistics of acoustic fields propagating through random media. [Work supported by ONR.]

9:40

2aUW5. Time series analysis of propagation in range-dependent shallow-water environment. Finn B. Jensen, Francesco Bini-Verona, Peter L. Nielsen (SACLANTCEN, Viale S. Bartolomeo 400, 19138 La Spezia, Italy, jensen@saclantc.nato.int), and Peter Gerstoft (UCSD, La Jolla, CA 92093-0704)

One track of the Formiche_97 shallow-water experiment in the Mediterranean involved broadband propagation over a sloping bottom with water depths varying between 120 m at the 16-element receive array and 600 m at the end of the 35-km track. The 0.82-kg SUS charges were detonated at two depths, 18 and 90 m, and at range intervals of around 1.2 km. Signals were recorded over a bandwidth of 50–5000 Hz on hydrophones covering the full water column. Detailed environmental measurements were also performed. Using state-of-the-art numerical modeling tools it was demonstrated the degree of accuracy with which broadband signals can be modeled in complex ocean environments, after having first fine-tuned the sediment parameters via a broadband acoustic inversion. The forward model used is a modified version of the ORCA code [Westwood *et al.*, J. Acoust. Soc. Am. **100**, 3631–3645 (1996)], which uses adiabatic mode theory combined with frequency-interpolation techniques to efficiently compute the broadband frequency response of the propagation channel. The broadband inversion was done using the SAGA code [Gerstoft, SACLANTCEN Rep. SM-333 (1997)], which does parameter optimization via a directed Monte Carlo search based on genetic algorithms. [Work was supported by the EC under the MAST-III project PROSIM, Contract No. MAS3-CT95-0008.]

9:45

2aUW6. A two-way parabolic equation for a fluid/elastic waveguide. Joseph F. Lingevitch and Michael D. Collins (Naval Res. Lab., Washington, DC 20375, jfl@abyss.nrl.navy.mil)

The parabolic equation (PE) method is applied to acoustic propagation in a fluid above an elastic layer. Two-way PEs have previously been developed for purely fluid or elastic waveguides [M. D. Collins, J. Acoust. Soc. Am. **93**, 1815–1825 (1993)]. These methods divide a range-dependent environment into a sequence of range-independent regions and apply the single scattering method to obtain the transmitted and reflected fields at each interface. Recently, these ideas have been extended to mixed media such as fluid/poro-acoustic waveguides. Improved Padé approximations for the interface conditions are useful in stabilizing the iteration formula for the transmitted and reflected fields [F. A. Milinazzo *et al.*, J. Acoust. Soc. Am. **101**, 760–766 (1997)]. In this paper, a fluid/elastic waveguide is examined and examples of the coupling between the fluid compressional wave and elastic waves at vertical interfaces are presented. [Work supported by ONR.]

9:50

2aUW7. Spatial coherence in range-dependent shallow water environments. Ilya Rozenfeld (Rensselaer Polytechnic Inst., Troy, NY 12180), William M. Carey (MIT, Cambridge, MA 02139), James F. Lynch (Woods Hole Oceanograph. Inst., Woods Hole, MA 02543), Peter Cable (BBN Systems and Technologies, New London, CT 06320), and William L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180)

In this presentation, effects of small random fluctuations on the spatial coherence of the sound field in shallow water with multilayered sediments will be examined. These fluctuations occur within the water volume and the sediment layers as well as at the layer interfaces. Most previous work has dealt with either range-independent or weakly range-dependent environments with random variabilities. However, strong range dependence often occurs in sound-speed profiles and at layer interfaces. For example, the region of the ACT III experimental series in the Straits of Korea shows range variability arising from sloping bathymetry and a coastal front. Consequently, spatial coherence for range-dependent waveguides with small random fluctuations will be treated. Numerical simulations of the coherence function will be presented. Results will be compared with coherence lengths of the field extracted from ACT III data.

9:55

2aUW8. Thin-sediment shear-induced effects on low-frequency broadband acoustic propagation in a shallow continental sea. Dag Tollefsen (Norwegian Defence Res. Establishment, Div. for Underwater Defence, P.O. Box 115, N-3191 Horten, Norway, dt@ffi.no)

Low-frequency acoustic broadband propagation loss data collected in areas of a shallow continental sea where the seabed is composed of hard substrates covered by thin deposits of sediment is presented. Data are

shown to exhibit several characteristics consistent with loss due to coupling to interface waves at the sediment–substrate interface; in instances of a physically thin sediment cover, loss is attributed to guided elastic waves in the sediment [S. J. Hughes *et al.*, J. Acoust. Soc. Am. **88**, 283–297 (1990)]. The interpretations are supported by numerical modeling of propagation loss using a recently developed wave number intergation code for horizontally layered fluid–solid media, and geoacoustic models consistent with geophysical data from the sites of acoustic experiments.

10:00

2aUW9. Elimination of branch cuts from the normal mode solution using gradient half-spaces. Evan K. Westwood and Robert A. Koch (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029)

A method for eliminating the branch cuts and branch line integrals from the normal mode solution to the range-invariant acoustoelastic wave equation has been developed and implemented. In the usual normal mode formulation, evaluation of the vertical wave number in the ocean waveguide's lower half-space involves a square-root operation and gives rise to a branch point at the critical angle. When the Pekeris cut is chosen, energy steeper than the critical angle is represented by leaky modes, and lateral-wave energy near the critical angle is represented by the branch line integral. When a mode lies near the branch cut, the branch line integral can be significant at all ranges. Branch line integrals are cumbersome to compute, and, for range-dependent problems, cannot be included in adiabatic and coupled-mode algorithms. By inserting a (small) positive attenuation gradient in the lower half-space, the plane wave solution is replaced by a complex-valued Airy function solution, the branch point is eliminated, the Pekeris branch cut is replaced by a series of modes, and the leaky modes become bounded at infinite depth. Implementation and applications of this approach using the normal mode model ORCA are given. [Work supported by ONR 321OA.]

10:05

2aUW10. Improving a practical broadband adiabatic normal mode model by including untrapped modes. Robert A. Koch, Evan K. Westwood, Julienne E. LeMond, and David P. Knobles (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029)

An adiabatic normal mode model that neglects untrapped mode contributions will give inaccurate transmission loss predictions for environments with bottom slopes when modes transition between being trapped and untrapped. For example, initially untrapped, but strongly excited, modes that subsequently become trapped away from the source may contribute significantly to water column fields. Results of Tindle and Zhang [Oceans '93 Proceedings, I-81] show that including untrapped modes in an adiabatic description significantly improves transmission loss predictions for the ASA benchmark wedge problem. A practical implementation of an adiabatic broadband propagation model for both trapped and untrapped modes will be described. Difficulties associated with the transition region between trapped and untrapped modes near the critical angle will be discussed, and the results of inserting a gradient in the lower half-space to solve the problem will be presented. [Work supported by ONR.]

10:10

2aUW11. Parabolic equation models for poroelastic transversely isotropic sediments. Andrew J. Fredricks, William L. Siegmann (Dept. of Math Sci., Rensselaer Polytechnic Inst., Troy, NY 12180), and Michael D. Collins (Naval Res. Lab., Washington, DC 20375)

Anisotropy in ocean sediments can have a significant effect on transmission loss. A parabolic equation (PE) model for problems involving transversely isotropic elastic sediments was recently developed [A. J. Fredricks *et al.*, J. Acoust. Soc. Am. **101**, 3182(A) (1997)]. Biot's theory of poroelasticity, which is appropriate for some sediments, is a generalization of elasticity that includes effects due to fluid-filled pores. Propagation models have been developed for poroelastic media, but only for the isotropic case. A poroelastic PE model is extended to handle transversely isotropic sediments. A description of the model will be presented along with examples to illustrate the effects of anisotropy. [Work supported by ONR.]

10:15

2aUW12. Inclusion of continuum effects in coupled-mode theory using leaky modes. David P. Knobles, Robert A. Koch, Evan K. Westwood, and Steven A. Stotts (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029)

A difficulty with coupled-mode algorithms for acoustic propagation in sloping, shallow-water regions is properly including the modal continuum. The disadvantage of the standard false-bottom approach is that a large number of modes are needed to represent the continuum, which makes the coupled-mode solution less practical. A leaky mode decomposition requires significantly fewer modes to represent the continuum. However, because the leaky modes are not bounded in the lower half-space, the overlap integrals involved in computing the coupling coefficients are undefined. In addition, the branch line integral, which is *not* included in the usual coupled-mode formalism, can become significant as modes undergo the transition between being trapped and being leaky. In this work, the approach of inserting a small attenuation gradient in the lower half-space is taken. The half-space gradient makes the mode functions have Airy-function solutions, eliminates the branch cut, replaces it by a series of modes, and makes the leaky modes well behaved at infinite depth. Using the differential form of the coupled-mode equations, overlap integrals can be computed analytically using well-known formulas for Airy functions. The accuracy and efficiency of the approach is evaluated using the ASA benchmark wedge problem as a test case. [Work supported by ONR.]

10:20

2aUW13. A search algorithm to predict the resonance frequency of a shallow-water soliton packet. Stanley A. Chin-Bing, David B. King, Alex Warn Varnas, and Robert A. Zingarelli (Naval Res. Lab., Stennis Space Center, MS 39529-5004)

Recently a number of researchers have performed acoustic simulations of shallow-water regions that contain large amplitude internal waves (solitons). Their results have confirmed earlier findings: (a) mode conversions, due to acoustic interactions with soliton packets, can produce a large loss in acoustic transmission; (b) this loss occurs within a narrow band of acoustic frequencies (about the resonance frequency); and, (c) the transmission loss is most prominent when there is strong coupling between the lower-order (water-borne) propagation modes and the higher-order, very lossy (bottom interacting) modes. Using this knowledge, development of a search algorithm has begun to predict the resonance frequency of a shallow-water soliton packet. In this proof-of-concept work, a shallow-water soliton model was used to generate replicas of the soliton fields in the Strait of Messina. The allowable acoustic mode conversions were calculated and correlated with the dominate spatial wave numbers of the soliton packet. This was used to predict the resonance frequency. The resonance frequency was later confirmed by making hundreds of acoustic runs. The search algorithm will be discussed and examples shown. [Work supported by ONR/NRL and by a Federal High Performance Computing DoD grant.]

10:25

2aUW14. Numerical model for low-frequency sound propagation in inhomogeneous waveguides. Alexander G. Voronovich (NOAA/ERL/ETL and Univ. of Colorado/CIRES, 325 Broadway, Boulder, CO 80303, agv@etl.noaa.gov)

Sound propagation in a shallow-water environment can result in the appearance of steep enough rays. Sufficiently accurate treatment of such situations with the help of the parabolic equation (PE) method can be complicated. On the other hand, propagation of low-frequency sound in inhomogeneous waveguides can be conveniently considered in terms of local modes. An algorithm for sound field calculation in the inhomoge-

neous waveguides was suggested recently [A. Voronovich, J. Comp. Acoust. **4**, 399–423 (1996)]. This algorithm is not based on parabolic approximation, it includes backpropagating field, and for 2-D situations is principally exact. Similar to the PE method it is also based on ''marching,'' however, in this case it proceeds in the backward direction. In contrast to other exact approaches which invoke piecewise constant (''staircase'') approximation of the inhomogeneous waveguide, this algorithm takes into account the mode interaction within each horizontal step. This allows one to increase the value of the horizontal step significantly making the algorithm very efficient. Generalization of this algorithm for the case of the uneven liquid bottom is considered. It is shown that due to sharp discontinuities at the bottom a bigger number of modes should be taken into account for sufficiently accurate calculations of the field in this case.

10:30

2aUW15. Broadband propagation over randomly varying, range-dependent elastic sediments. Michael Jaye (Rensselaer Polytechnic Inst., Troy, NY 12180), Xin Tang, Mohsen Badiey (Univ. of Delaware, Newark, DE 19716), and William L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180)

Acoustic experiments conducted over different tracks at the Atlantic Generating Station (AGS) site have shown significant transmission loss variability. These fluctuations cannot be accounted for entirely by bathymetric and water volume variations. Consequently, as suggested by previous studies, geoacoustic variations must be responsible for significant portions of the variability. We can take advantage of the relatively extensive core data that are available at the AGS site to model both the larger-scale range dependence and the smaller-scale variability. Uncertainties in the geoacoustic fields are modeled here by varying the vertical correlation length of the core profiles which changes the coefficients of representations by empirical orthogonal functions. The stochastic field representations are used for propagation modeling with both elastic and fluid sediments. Statistics of broadband signal transmission loss fluctuations are determined in terms of geoacoustic variabilities. Sensitivities of transmission loss predictions to frequency dependence of geoacoustic parameters are examined. Simulation statistics are compared with experimental data. [Work supported by ONR.]

10:35

2aUW16. The existence and coherence of ray bundles in a refractive ocean. Ivan P. Smirnov, Alexander I. Khil'ko (Inst. of Appl. Phys., Nizhny Novgorod, 46, Ulianov, Str., 603600, Russia), and Jerald W. Caruthers (Naval Res. Lab., Stennis Space Center, MS 39529)

Due to the existence of relatively smooth local extrema in the dependencies of ray-cycle lengths on initial path angles, different kinds of weakly diverging bundles (WDB) of rays can arise in refractive oceanic waveguides. As Brekhovskikh *et al.* have shown [Acoust. J. (in Russian) **36**, 5, 824–831 (1990)], such bundles transporting high-level acoustic intensity can be interesting with respect to these applications in ocean acoustic tomography (OAT). Conditions for the existence of WDB for piecewise linear approximations of sound velocity profiles are formulated in this work. Based on these results, the possibilities for the existence of WDB for different regions of the world's oceans are discussed. Important for tomographical application, the coherence of WDB is investigated under assumptions that the acoustic signals are emitted by noise sources with limited lengths of spatial coherence and the oceanic environment is randomly inhomogeneous in spatial coherence, as described by a Gaussian distribution. Numerically, it is shown that, because within a WDB ray parameters differ only slightly, coherence in WDB is maintained for propagation to long distances. In contrast are the rays outside the WDB. The optimal space/time filtering for OAT on the basis of the developed multiscale coherence approach is discussed.

10:40

2aUW17. Vertical noise coherence measurements in shallow water using Lagrangian drifters. Fancine Desharnais (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada, desharnais@drea.dnd.ca), Blair R. MacDonald, and Kenneth J. Mah (Seimac, Dartmouth, NS B3B 1W6, Canada)

The vertical coherence of an acoustic noise field in shallow water (at frequencies dominated by sea surface noise) can be used to infer certain ocean bottom properties such as the compressional sound speed of the upper sediment layer. As coherence is a normalized measurement, it is independent of background noise level, or wind speed/sea state, and remains a stable measurement over a wide range of environment conditions. A set of three Lagrangian ambient noise drifters (LANDs) were developed for vertical coherence measurements. The drifters (drogued at 15-m depth) follow the near-surface currents, and take hourly measurements of coherence from their two hydrophones located at 50-m depth. The hydrophone separation is 1 m, and the measurements are made over a usable frequency range of 100 Hz to 1 kHz. The drifters were deployed at Western Bank (Scotian Shelf, Canada) and drifted for up to 2.5 days. The sea conditions varied from calm to 3–4-m waves. The stability of the coherence data through the various sea states, and the impact of shipping noise on the measurements will be discussed.

10:45

2aUW18. Prediction of ambient noise in the ocean, the missing components. Douglas H. Cato (Defence Sci. and Technol. Organisation, P.O. Box 44, Pyrmont, NSW 2009, Australia, doug.cato@dsto.defence.gov.au)

The main components of ambient noise in the ocean were established by pioneers in ambient noise [V. O. Knudsen, R. S. Alford, and J. W. Emling, J. Marine Res. **7**, 410–429 (1948); G. M. Wenz, J. Acoust. Soc. Am. **34**, 1936–1956 (1962)] and prediction methods are generally based on the results of these studies. Most measurements were made in regions of high shipping densities where traffic noise usually obscured other components of ambient noise at low frequencies (below 200 Hz), so these do not appear in most prediction methods, even though their presence was noted in the early studies. Biological noise was also recognized as important, and widely studied, but is also not usually included. Studies of ambient noise near Australia, where shipping densities are generally lower, have shown the importance of low-frequency wind-dependent noise. Biological noise is also a major component in this part of the world, often being the dominant noise in tropical waters. Biological choruses are widespread and occur with diurnal regularity. These results are not inconsistent with measurements in other parts of the world. This paper discusses important components of ambient noise that are often missing from prediction methods.

10:50

2aUW19. Application of adaptive filter in influence extraction. S. Veerabhadraiah (Naval Sci. and Technol. Lab., Ministry of Defence, Govt. of India, Viyana Nagar, Visakhapatnam-530 027, India)

The very-low-frequency ocean ambient noise changes from place to place and time to time drastically due to various factors like thermal agitation, sea state, traffic, seismic factors, and biological activity, wind speed, etc. The surface and underwater mobiles generate an influence signal within the same frequency band of ambient noise but with extremely low magnitude. The influence signal gets buried in the high ocean ambients, yielding an extremely low signal-to-noise ratio, and are as such indistinguishable in the original time traces. In certain applications, perfect localization of such mobiles is extremely essential for extracting the influence signal buried in ocean ambients. This can be done using a self-optimizing adaptive filter known as a LMS (least mean square) filter. The filter automatically adjusts its parameters with the arrival of each new

ambient sample with a superior convergence rate at an appropriate step size. This technique, discussed in the paper, has been developed and successfully implemented to extract the desired signal from high ambients. Some sample graphical outputs depicting a combination of signals buried in ambient and extracted signal at the filter output. Necessary theory, a flowchart for implementing the LMS algorithm, and hardware realization of the same are also included in the paper.

10:55

2aUW20. Global autofocusing techniques for acoustic imaging. Peter T. Gough and Richard L. Lane (Dept. of Elec. and Electron. Eng., Univ. of Canterbury, Private Bag 4800, Christchurch, New Zealand, gough@elec.canterbury.ac.nz)

When images of acoustic fields are computed from measured data, the quality of the image is sometimes impaired by the acoustic wavelength limits or by turbulence in the media. However, more often, the image is impaired by random displacements in the position of the detecting array. Techniques exist to optimize the image on the basis of *a priori* knowledge. Since this knowledge is often statistical rather than deterministic, the image computed is the most probable image given the *a priori* knowledge. There can be no guarantees that what is computed is the exact image since there is an infinite combination of acoustic fields and tracks able to produce the same measured field. Using examples from some simulations of an existing synthetic aperture sonar, it is shown (by appealing to Bayesian statistics) that statistical information about the towfish behavior can be used to improve the quality of the image. Inherent in this technique is a cost function to be minimized and it is the precise nature of the cost function chosen that produces the most probable image. It is techniques such as this that hold out the prospect of SAS and passive towed array sonars without the need for on-board navigation units.

11:00

2aUW21. Up-to-date state and outlook for the intensity measurement method in underwater acoustics. Vladimir Shchurov (Lab. of Ocean Acoust. Noise, Pacific Oceanological Inst., Far Eastern Branch, Russian Acad. of Sci., 43 Baltiyskaya St., Vladivostok, 690041, Russia)

The acoustic intensity measurement method consisting of four components of the acoustic field [acoustic pressure $p(t)$ and three orthogonal components of the particle velocity vector $\mathbf{V}(t)\{V_x(t), V_y(t), V_z(t)\}$ simultaneous measurements] did not occupy a suitable place in underwater acoustics until now. Simultaneous measurement of four components of the acoustic field enables one to construct a point receiving system possessing direct characteristics in a wide frequency range (for example, from 10 to 100 Hz). By now unconventional results of investigations of underwater acoustic noises, reverberation, point source capacity, seismic noises, and underwater navigation were obtained. The intensity measurement method has given one a chance to research the acoustic field characteristics that have never been revealed during measurements of only acoustic pressure, i.e., characteristics connected with energy flux density vector. For example, a compensation of opposite energy fluxes of ambient noise and tone or signal similar to noise; horizontal energy flux of dynamic noise connected with propagational direction and the extent of wind near-surface roughness; splitting of ambient noise total field energy density into an isotropic component connected with the isotropic (diffuse) part of the field and an anisotropic one connected with the anisotropic (coherent) part of the field related to the energy transfer in the ocean. The isotropic component (possessing zero averaged flux) existence in the field of ambient noise allows for decreasing the threshold noise level for receiving systems measuring acoustic intensity, thus enables one to research a thin structure of the acoustic field of the noise and signal. Measurements of only acoustic pressure cannot give one such information. The acoustic intensity measurement method development will allow for constructing the most up-to-date acoustic devices for oceanologic processes dynamic investigation on the bases of ambient noise and sounding signal vector characteristic measurements.

TUESDAY MORNING, 23 JUNE 1998 GRAND BALLROOM III (W), 11:00 A.M. TO 12:00 NOON

Session 2aPLb

Plenary Lecture

Henry E. Bass, Chair
National Center for Physical Acoustics, University of Mississippi, University, Mississippi 38677

Chair's Introduction—11:00

Invited Paper

11:05

2aPLb1. Acoustics of outdoor surfaces. Keith Attenborough (Faculty of Technol., The Open Univ., Milton Keynes MK7 6AA, England)

While the development of absorbing materials for barrier surfaces is attracting considerable research interest, relatively little effort is being devoted to modifying ground surfaces for optimum ground effect. Indeed current schemes for predicting outdoor noise levels due to transportation and industry distinguish only between two types of ground; acoustically hard and acoustically soft. Ground effect phenomena are described with reference to relevant laboratory and outdoor data which indicate the influences of surface characteristics including porosity, small-scale roughness, elasticity, vegetation, extended reaction, and topography. Theories for the propagation of curved wavefronts near to porous ground and above rough ground predict the existence of surface waves. The relationship between these waves is discussed. The robustness of ground effects to meteorological influences is considered. Optimum characteristics for porous pavement with regard to propagation effects are described. Other possibilities for passive noise control reviewed include deployment of rough pavements and poroelastic materials, modifications to railway ballast, roadside cultivation, and landscaping. [Work supported by EPSRC and BBSRC (UK).]

Session 2pPLa

Plenary Lecture

James E. West, Chair
Lucent Technologies, Room 2D-338, 600 Mountain Avenue, Murray Hill, New Jersey 07974

Chair's Introduction—1:00

Invited Paper

1:05

2pPLa1. Trends in electromechanical transduction. Ilene Busch-Vishniac (Mech. Eng. Dept., Univ. of Texas, Austin, TX 78712 and Aero. and Mech. Eng. Dept., Boston Univ., Boston, MA 02215)

Electromechanical transducers (sensors and actuators) are devices which link electrical and mechanical energy domains. Ideal sensors are transducers in which the state of a system is monitored without being affected. Ideal actuators are transducers in which a system state is imposed, regardless of the load imposed. In recent years there has been growing attention on electromechanical transducers for a number of reasons: they are becoming more common in consumer items instead of being limited to laboratory equipment and industrial facilities; the demands placed on transducers, particularly with regard to size, speed, cost, and sensitivity, are becoming progressively harder to meet; the applications arenas for transducers are expanding to new domains; there is a push for miniaturization of sensors and actuators; more attention is being focused on actuators, as they are very difficult to design; and the time from discovery of new materials and techniques to the application in transducers is shrinking significantly. These trends, supported with reviews of research in the area, will be examined. Repercussions of these trends, current and in the near future, will also be discussed.

Session 2pAAa

Opera House Acoustics Workshop—CIARM, MCHA, and Architectural Acoustics: Opera House Design Case Histories

Robert D. Essert, Jr., Cochair
Arup Acoustics, Boston House, 36-38 Fitzroy Square, London W1P 5LL, England

Roberto Pompoli, Cochair
Dipartimento di Ingegneria, via G. Saragat, 1, 44100 Ferrara, Italy

Chair's Introduction—1:00

Invited Papers

1:05

2pAAa1. Acoustics in the competition for the construction of ''La Fenice'' opera house. Maria Ida Biggi (Universita' Ca' Bonvicini, Venezia, Italy) and Roberto Pompoli (Dipartimento di Ingegneria-Via G. Saragat, 1, 44100 Ferrara, Italy)

In 1789, a group of nobles wished to build a new large theater in Venice. They announced a competition, in which 29 architects from all over Italy took part, presenting many different projects. Although the prize was given to one project, a different one was chosen to be executed. The author of the latter project was Giannantonio Selva and was the cause of much criticism, but it was appreciated for the results it obtained from the acoustic point of view. In his memoirs, Selva explained how he adapted the shape of the theater so that it became a semicircle with extensions on both sides and added that there were no precise theories regarding the building of a ''harmonic'' theater. He also appeared to be well informed on the debate about different curves to use in the hall. He suggested leaving an empty space under the orchestra pit, because ''it looks good.'' This paper deals with the description of the acoustical requirements in the competition for the construction of ''La Fenice'' and with the analysis of some of the projects presented taking into account the knowledge of room acoustics in the XVIII century.

1:25

2pAAa2. Preliminary project for the rebuilding of ''LA FENICE'' theater of Venice: Acoustic aspects. Mauro Strada (Istituto Architettura Universita' Venezia, 191 Santa Croce Tolentini 30135 Venezia, Italy) and Roberto Pompoli (Facolta' Ingegneria Universita' Ferrara, Via Saragat 1, 44100 Ferrara, Italy)

After a fire destroyed ''La Fenice'' Theater of Venice, the Town Council decided it should be rebuilt, being a place of memories for Venetians and people worldwide, committing to the Prefect, named Delegate Commissary, the management of the operation, compiling of a Regulated Contract until reconstruction is finished. He set up an expert workgroup for the drawing up of the Contract which is of main importance in the competition for judging the works. They decided to start the rebuilding of the theater as it was and where it was, preserving at least the internal and external aspects of the ''SaleApollinee'' halls, and the ''SalaGrande'' hall. They also decided to give freedom to the designers of the companies invited to the competition, concerning the stage aspects and all the facilities areas (halls for rehearsal, dressing-rooms, etc.). The problem has been the acoustic characteristics of the ''Sala Grande,'' which had to appear aesthetically as it was, nevertheless the materials were not going to be the same as before the fire, due to newer safety fireproof requirements. This article deals with all acoustic problems and the several stages of historical-critical deepening of the acoustics of the theater before the fire, brought into effect during the compiling of the Regulated Contract.

1:45

2pAAa3. The restoration of La Fenice in Venice: The consultant's viewpoint. Daniel E. Commins (commins acoust. workshop, 350 Fifth Ave., New York, NY 10118, comminsacoustics@compuserve.com)

Following the destruction by a fire of La Fenice in Venice, an extremely complex process was set in motion to attempt to rebuild the opera house as soon as possible. In this instance, acoustics is of utmost importance but it tends to be overshadowed by many other technical economical, and political requirements. Other major problems have surfaced, among which are the following: Very few opera houses, including La Fenice, are correctly documented from the point of view of acoustics; the leaders of such projects are not always aware of the impact of acoustics; the science of opera house acoustics is incomplete, to say the least. In the case of La Fenice, the decision was made to reconstruct it ''as was'' through a series of discussions with experts and users, and a well organized and efficient competition was set up. The present paper will attempt to analyze the strengths and weaknesses of this process, from the point of view of the ''acoustics'' referee.

Contributed Papers

2:05

2pAAa4. The acoustics of the Italian opera house, ''Teatro di San Carlo'' in Naples, Italy. Luigi Maffei, Gino Iannace, Carmine Ianniello, and Rosario Romano (DETEC–Univ. of Naples Federico II, Piazzale Tecchio 80, 80125 Naples, Italy, maffei@unina.it)

Teatro di San Carlo is reported to be the oldest existing European opera house. This horseshoe-shaped theater was opened on the 4th of November 1737, 41 years before the opening of La Scala in Milan and 51 years before the opening of La Fenice in Venice. Surprisingly, no objective acoustical data were found in the open literature for San Carlo. Therefore, the authors were stimulated in carrying out measurements of modern objective criteria in the above-mentioned opera house. As the theater hosts both symphonic/classical and opera concerts, two sets of measurements were carried out. The first one with the stage fitted for performing symphonic and classical music, and the second one with fitted scenery and the orchestra pit open, for performing opera concerts. Measurements in the unoccupied hall, yielded now classical, objective criteria (RT, EDT, C80, G, LE, and LF) at specific locations both on the main floor and in the boxes. The sound source was located along the longitudinal axis of the stage under the proscenium arch.

2:20

2pAAa5. Acoustics of the historical opera house of Bordeaux: Comparison between the results of objective and subjective surveys. Catherine Semidor and Aline Barlet (ERIAC, Ecole d'Architecture et de paysage de Bordeaux, Domaine de Raba, F 33405 Talence Cedex, France, catherine.semidor@bordeaux.archi.fr)

This paper deals with the relationships between objective and subjective results of an acoustic survey in the historical opera house of Bordeaux. The investigations were conducted with ordinary opera-goers during four performances of ''Traviata'' in order to have available statistical population answers. The results of an earlier experiment [C. Semidor, Proc. IOA **19**, Part 3, 29–38 (1997)] carried out in a dual-purpose hall with a test questionnaire built with simple words which explain the sensations of the listeners, have shown that the questions are correctly understood. The comparison of the ordinary listeners' answers with the objective results shows that the auditive sensations are fitting with the measured criteria obtained in the opera configuration with the opened pit and sceneries on the stage. For example, for the majority of people asked, the sound in this hall seems clear and coming directly from the stage (no surrounding by the sound sensation), and the sound attack impression is very clear, which is correlated by the weak values of reverberation time and EDT. In spite of comfort defects due to the age of the opera house, such as the lack of space between the seats, the architecture of the Grand Theatre of Bordeaux is very well appreciated.

2:35–2:45 Break

2:45

2pAAa6. Acoustic conditions of the San Francisco War Memorial Opera House and its renovation. Dennis Paoletti, Kurt Graffy, Larry Tedford, and Red Wetherill (Paoletti Assoc., Inc., 40 Gold St., San Francisco, CA 94133)

The War Memorial Opera House is one of the most culturally significant performing arts venues in San Francisco. The San Francisco Opera organization (which includes the performers, designers, systems, and facilities) is known worldwide for the quality of its performances. The Loma Prieta earthquake which occurred just south of San Francisco in October 1989 was the impetus for the major renovation to the War Memorial Opera House currently in progress. This paper will report on the history of the San Francisco Opera House and the current renovation. It will also touch on some of the peripheral facilities and projects resulting from the renovation. Paoletti Associates has enjoyed a long-standing professional relationship with the San Francisco Opera organization, and the San Francisco War Memorial Board of Trustees, having been involved in a number of various studies for them over the past 20 years. They recognize the importance of acoustics, and they are quite concerned about the historical and cultural reputation of their performing arts facilities.

3:00

2pAAa7. A unique setting for an opera orchestra. Ewart A. Wetherill (Paoletti Assoc., Inc., 40 Gold St., San Francisco, CA 94133)

In 1995 the San Francisco Opera vacated the Opera House for one winter season and two summer seasons to allow for substantial renovation necessitated by damage caused during the 1989 earthquake. Rejecting the proposal of canceled opera seasons, the company tackled the more difficult option of utilizing the 4500 seat San Francisco Civic Auditorium augmented by two of San Francisco's older theaters. The cavernous and noisy Civic Auditorium, which is basically an arena with a large U-shaped balcony, presented a challenge in several dimensions. A complete stage, main floor seating, lighting, and backstage facilities had to be created. Lacking an orchestra pit, the opera designers placed the opera orchestra above dressing rooms, 16 ft above stage level and 60 ft behind the stage apron in a temporary orchestra enclosure. The entire undertaking is seen as a triumph of resourcefulness and willingness to face the challenge of severe physical limitations and budget constraints. Although complicated by an awkward plan configuration and limited overhead access, as well as conflicting rigging for lighting and scenery, the orchestra enclosure nevertheless evolved into a performance element that was both acoustically and visually appropriate. The development of the orchestra enclosure design is discussed in this paper.

3:15

2pAAa8. Virtual reflecting walls for improving the acoustics of defective halls. Jean-Paul Vian and Xavier Meynial (Acoust. Group, CSTB, 24 rue Joseph Fourier, F 38400 Saint-Martin d'Hères, France, vian@cstb.fr)

Many performing art centers and multipurpose halls with poor acoustics cannot be improved for economical reasons by architectural correction works or by mechanical moving devices. The principle of virtual acoustic wall [Guicking *et al.*, J. Acoust. Soc. Am. **78** (1985)] has been successfully applied to control the reflectivity and diffusivity of concert halls existing walls. Several independent active cells composed of a microphone, an electronic unit, and a loudspeaker form a virtual wall. By appropriately controlling the open loop gain in the cells of each virtual wall it is possible to create different sound fields fully coherent in space and time with the preexisting one. The first reflections can thus be reinforced and a natural reverberation enhancement effect is obtained by the acoustic exchange between the virtual walls. Computer sound simulation of three virtual walls in a concert hall was performed and experimental results, both objective and subjective, are shown for two concert halls in which a virtual wall system was installed.

3:30

2pAAa9. The next steps: Four examples of Artec's evolving room acoustics philosophy. Russell Johnson, Robert W. Wolff, Damian J. Doria, and Ashley Goodall (Artec Consultants, Inc., 114 W. 26 St., New York, NY 10001-6812, info@artec-usa.com)

This paper examines four Artec projects that have opened in 1997, and relates them to a room acoustics philosophy that has been evolving in Artec's designs since 1970. An illustrated discussion of the Winspear Centre (Edmonton, Alberta), the Living Arts Centre in Mississauga (Ontario), the Chan Centre (Vancouver, British Columbia), and the renovations to the stage acoustics at Music Hall, Cincinnati, highlights the importance of seat-count, complete freedom from exterior noise, room shaping, and a considerable amount of adjustability in order to achieve acoustical success in performing arts facilities.

3:45

2pAAa10. La Cité de la Musique in Paris. Daniel E. Commins (commins acoust. workshop, 350 Fifth Ave., New York, NY 10118, comminsacoustics@compuserve.com)

Fifteen years have been necessary to program, design, and build the Cité de la Musique in Paris. The author has been personally involved in all phases of the project and has assumed the principal acoustical responsibilities, with assistance from many acousticians. This music complex has two parts: the National Conservatory of Music and Dance, with some 250 music rooms, and the concert hall wing. The main concert hall, which has been in service for a few years, has a capacity of 900 to 1200 only, but it is designed in such a way that it can be used for various orchestral configurations: classical, in the round, central, etc. This means, of course, that acoustical reciprocity is a must: hearing conditions and musicality must be adequate for any pairs of points, in spite of the elliptical shape of the concert hall. The present paper will describe the means used to achieve a correct acoustical result in the concert hall and will give some insight into the characteristics of some of the other rooms.

TUESDAY AFTERNOON, 23 JUNE 1998 GRAND FOYER (W), 1:00 TO 4:45 P.M.

Session 2pAAb

Opera House Acoustics Workshop—CIARM, MCHA and Architectural Acoustics: Opera House Acoustics Student Design Competition

Robert C. Coffeen, Chair
Architectural Engineering, University of Kansas, Kansas City, Missouri 66045

The 1998 Student Design Competition involves the schematic design of a multipurpose opera hall with the design emphasizing room acoustics and noise control. Entries from college and university students interested in building acoustics will be on display from 1:00 p.m. to 4:45 p.m. on Tuesday, 23 June and from 9:15 to 5:00 p.m. on Wednesday.

Session 2pAB

Animal Bioacoustics and Acoustical Oceanography: Fish Acoustics

Joseph A. Clark, Chair
Carderock Detachment, Naval Surface Weapons Center, Code 734, Bethesda, Maryland 20085

Invited Paper

1:30

2pAB1. An experimental study of the peripheral auditory mechanics in the goldfish (*Carassius auratus*) and oscar (*Astronotus ocellatus*). Mardi C. Hastings, James J. Finneran, and Corrie Derenburger (Ohio State Univ., Dept. of Mech. Eng., Columbus, OH 43210, hastings.6@osu.edu)

Dynamic mechanical responses of peripheral auditory organs in the goldfish (*Carassius auratus*), considered a hearing specialist because of its coupling via the Weberian ossicles between the swimbladder and inner ear, and the oscar (*Astronotus ocellatus*), an auditory nonspecialist, were measured *in vivo* using a noncontact, noninvasive ultrasonic system. Motion of the auditory organs was induced under water using pure tones at frequencies ranging from 12.5 to 3000 Hz. Measurements were made in both species with swimbladders intact and swimbladders deflated. Results from 23 goldfish and 38 oscars show that the swimbladder controls auditory bandwidth not only in hearing specialists, but also in nonspecialists which have no special coupling between the swimbladder and inner ear. Dual resonances routinely observed in the response of the anterior swimbladder, and motion of the Weberian ossicles transmitted to the saccular otolith at higher frequencies, contribute to the relatively broad auditory bandwidth of the goldfish. Although data were recorded only in the medial-lateral direction, they clearly indicate that displacement of the saccular otolith relative to its sensory epithelium in both the goldfish and oscar correlates with auditory thresholds for these animals reported in the literature. [Work supported by ONR Grant No. N00014-94-1-0337.]

Contributed Papers

1:55

2pAB2. A cohesive lumped parameter analysis of the mechanics of the goldfish peripheral auditory system. James J. Finneran (Biosciences Div. Naval Command, Control and Ocean Surveillance Ctr., RDT&E Div., 49620 Beluga Rd., San Diego, CA 92152-6266) and Mardi C. Hastings (Ohio State Univ., Columbus, OH 43210)

A mathematical analysis has been formulated to model the dynamic behavior of the goldfish peripheral auditory system, including the swimbladders, Weberian apparatus, and saccule. The two-chambered goldfish swimbladder is treated as a two degree-of-freedom mechanical system consisting of two coupled mass-spring-damper arrangements. The swimbladder is coupled to the Weberian ossicles using a phenomenological analysis of the anterior swimbladder tunica externa which permits both stretching and sliding. Analysis of the saccule features only a single degree-of-freedom, corresponding to the direction of orientation of the ciliary bundles. Inputs to the saccule consist of the transverse canal fluid motion and the motion of the animal's head (assumed to match the local acoustic particle motion). Mechanical properties required for the system equations were estimated from published literature and direct measurements. Results of the analysis compare favorably to experimental data for the motions of the swimbladders and Weberian ossicles. The results indivate that the Weberian apparatus has a significant impact on the hearing ability over a broad frequency range and that the saccule functions as a displacement sensor above approximately 300 Hz. [Work supported by ONR Grant No. N00014-94-1-0337 and the ASA Hunt Fellowship.]

2:15

2pAB3. Directional acoustic startle response in the goldfish. Thomas N. Lewis and Peter H. Rogers (The George W. Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405)

Fish exhibit a characteristic Mauthner-initiated startle response to sudden acoustic stimuli. This escape response is thought to be triggered by predatory attacks since an acoustic wave is launched ahead of a striking predator. The nature of acoustic stimuli which elicit the startle response was examined in free-swimming goldfish (*Carassius auratus*). The results showed that a threshold level of acoustic acceleration (-30 dB *re*: 1 m/s^2), not pressure, was required to elicit a startle and that the threshold was invariant of bearing to the source. Near threshold, the fish turned in a random direction, but as the stimulus level increased, the fish was more likely to turn away from the source. The data indicated that the lateral component of the acoustic acceleration was the relevant parameter determining directionality. A proposed model uses the utricle to detect acoustic acceleration directly, the saccule to detect acoustic pressure indirectly, and logical wiring to perform the neural computation which explains the fish's behavioral performance.

2:35–2:50 Break

2:50

2pAB4. Detection of ultrasound and marine mammal echolocation clicks by the American shad (*Clupeidae*). David A. Mann (Tucker-Davis Technologies, Inc., 4550 N.W. 6th St., Gainesville, FL 32609), Zhongmin Lu, and Arthur N. Popper (Univ. of Maryland, College Park, MD 20742)

Clupeid fishes (herrings and shads) are among the main prey of echolocating cetaceans. Recent studies have shown that some clupeids are repelled by intense sounds from ultrasonic transducers. The hearing thresholds of a clupeid, the American shad (*Alosa sapidissima*) was determined using classical conditioning. Shad could detect sounds from 0.2 to 180 kHz, with two regions of best sensitivity, one from 0.2 to 0.8 kHz and another from 25 to 150 kHz. American shad were also able to detect simulated bottlenose dolphin (*Tursiops truncatus*) echolocation clicks with a threshold that suggests they could detect an echolocating dolphin at distances up to 187 m. These results have practical implications for the design of ultrasonic devices that could be used for repelling clupeids from power plant intakes and dams, as well as for the design of ultrasonic pingers that may reduce dolphin bycatch without affecting the catch of clupeids. [Work supported by NSF.]

3:10

2pAB5. Morphological responses of fish to low-frequency sound. Joseph A. Clark (CDNSWC, Code 734, Bethesda, MD 20084), Jane A. Young (CDNSWC, Bethesda, MD 20084), Amrit N. Bart, and Yonathan Zohar (COMB, Baltimore, MD 21202)

Results of an earlier study [Clark *et al.*, J. Acoust. Soc. Am. **100**, 2709(A) (1996)] suggested that low-frequency vibrations could have a positive effect on the growth rate of fingerlings grown in tanks. This talk describes a more extensive study with a larger sample of fish (640 Tlapia Nilotica grown in four control tanks and four treated tanks) for a longer period of time than the previous six week period. The control tanks were mildly acoustically quieted with foam rubber tank coatings that are normally applied to reduce heat conduction into the tanks. The treated tanks were excited continuously with shakers in the frequency range of 30–60 Hz for the duration of the test. Average weight and length versus time for fish in treated and control tanks were measured at 3 week intervals. Differences in the two groups and their possible causes are discussed.

3:30

2pAB6. Evening fish choruses near coral reef systems in the Great Barrier Reef, Australia. Robert D. McCauley (Dept. of Marine Biol., James Cook Univ., Townsville, Queensland 4811, Australia) and Douglas H. Cato (Defence Sci. Technol. Organisation, Pyrmont NSW 2009, Australia)

The nighttime ambient sea noise on and around coral reef systems in northern Australia is characterized by choruses produced by nocturnally active fishes. These choruses are produced year round, from dusk to dawn. Choruses are composed of many ''pops'' of between 5 and 28 ms length, with a decay envelope suggestive of a lightly damped swimbladder and with source levels calculated at 154–157 dB *re*: 1 μPa^2 at 1 m peak to peak. Sources are found mostly concentrated over 6–20 m water depth. Choruses have a mean spectral peak at 550 Hz (normal range 500–750 Hz) but may influence sea noise from 100 Hz to 3 kHz. The highest chorus levels occur just after dusk or in the early evening, reaching up to 35 dB above background levels. The recorded levels of choruses displayed distinct daily patterns, lunar patterns with higher levels around new moon periods, and seasonal patterns of highest level during the austral summer. Choruses occur over large spatial scales, with active sources heard at up to 15 km from reef systems and chorus influence on ambient sea noise extending considerably further. It is believed that nocturnal planktivorous fishes, such as those from the families Holocentridae and Priacanthidae, are responsible for choruses. [Work supported by the Australian Defence Science and Technology Organization.]

3:50

2pAB7. Mathematical model of echo-location of fish-eating flying bats. Anatoli Stulov (Dept. of Mech. and Appl. Math., Inst. of Cybernetics, Akadeemia tee 21, Tallinn, Estonia, EE0026)

The ability of flying bats *Vespertillionidae* to detect fish under the water surface using echo-location is surprised, because only one-thousandth of the acoustical pulse energy penetrates into water. Supposing the fish (air bubble) reflects the whole signal backwards, the bat must detect the echo-signal which is a million times less than the transmitted one. This fantastic level of the perception sensitivity of the flying bat lies on the bound of its ability. It seems incredible that the main method of getting food is provided for by the maximum effort of bat senses. The method of the bat echo-location is explained here in terms of the interaction of the frequency-modulated acoustical pulse with the air bubble under the water surface. It was shown that in the presence of the bubble the amplitude of the echo reflected by the water surface depends on the signal frequency. In the usual range of the echo-location, 50–100 kHz (the resonance frequency of the bubble lies near 1 kHz), there are at least four regions of frequency for which the amplitude of the signal reflected back to the bat varies up to 25%. Thus, the bat can detect fish in water easily.

TUESDAY AFTERNOON, 23 JUNE 1998 WEST BALLROOM B (S), 2:15 TO 5:35 P.M.

Session 2pAO

Acoustical Oceanography and Underwater Acoustics: Surf Zone Oceanography and Acoustics II

Grant B. Deane, Chair

Scripps Institute of Oceanography, Marine Physical Laboratory 0238, La Jolla, California 92093-0238

Invited Paper

2:15

2pAO1. The generation of ambient noise due to breaking surf. James H. Wilson (Dept. of Oceanogr., Naval Postgrad. School, Monterey, CA 93943, wilson@vsinc.com), Robert H. Bourke (Naval Postgrad. School, Monterey, CA 93943), and Josette P. Fabre (Neptune Sci., Inc., Slidell, LA 70458)

The past predominance for deep-water acoustic studies has dictated that little consideration be placed on the role of the surf zone as a source of ambient noise. Recent emphasis on propagation in coastal waters has increased the awareness of the importance of breaking surf to the noise field in littoral waters. Two recent studies, in Monterey Bay, CA, and at Duck, NC, have provided for the first time information on the source level density of surf-generated ambient noise. Inverse techniques were developed to infer the source level density from the total ambient noise field as measured at several offshore hydrophones. Source level spectra from 50–1000 Hz demonstrate a -5-dB/oct slope with absolute levels varying by 10–15 dB dependent on wave height and distance from the surf zone. Surf noise was found to contribute 80% (95%–99%) of the total offshore ambient noise during light (heavy) surf conditions. The inverse technique is highly dependent on accurate modeling of the transmission loss seaward from the surf zone, being predominantly a function of the geoacoustic properties of the sea bed. Additional comments will address breaker types and factors affecting breaker height and the noise due to waves breaking on the ice edge of the Greenland Sea.

2:35

2pAO2. Measurements of the space–time–frequency distribution of the noise in the surf zone. Richard M. Heitmeyer, Steven L. Means, Stephen C. Wales (Naval Res. Lab., Code 7121, Washington, DC 20375), Ellen S. Livingston (Office of Naval Res., Arlington, VA 22217), Jeffrey A. Schindall, and Marshall H. Orr (Naval Res. Lab., Washington, DC 20375)

The space–time–frequency distribution of the noise generated in the surf zone was measured in the SandyDuck'97 experiment during September and October 1997. The noise measurements were obtained on a 17-phone, 138-m, bottom-mounted, linear array located along a line perpendicular to the shore at a depth of one to three meters. Concurrent measurements of the location, size, and time evolution of the individual breaking waves directly above the array were made using a video camera mounted on a nearby tower. The measurements of the time history of the noise at individual phones along the array indicate that in regions of wave breaking, the noise contributions are broadband bursts, isolated in time with significant variations in level. A significant number of these broadband bursts exhibit a downward frequency sweep. Similarly, measurements of the time history of the noise observed along the array at selected frequencies show significant variations in both the range extent and the propagation speeds of the noise sources. The paper will present an interpretation of this variability in terms of the space–time distribution of the breaking waves obtained from the video measurements. [Work supported by ONR.]

2:50

2pAO3. SANDYDUCK97 angle-time-frequency measurements of breaking wave noise. Ellen S. Livingston (Code 321OA, Ocean Atmosphere and Space S&T Div., Office of Naval Res., 800 N. Quincy St., Arlington, VA 22217-5660, livinge@onr.navy.mil), Joal J. Newcomb (Naval Res. Lab., Stennis Space Center, MS 39529-5000), Richard Heitmeyer, Stephen C. Wales, and Steven L. Means (Naval Res. Lab., Washington, DC 20375-5350)

Low-frequency acoustic measurements (0 to 2400 Hz) were taken in October 1997 as part of the SANDYDUCK97 experiment. These measurements were made on a horizontal bottom-mounted line array deployed in the surf ''transition'' zone at about 500 m offshore. The array consisted of 64 elements oriented end fire to shore in 6 m of water. Video cameras, mounted on a shoreside tower, monitored the breaking wave events over the transition zone array and in the nearby surf zone during data collection periods. The beamformed array data provide an angle-time-frequency distribution of the breaking wave noise as detected in the transition zone. Comparisons will be made of the angle-time-frequency characteristics from noise fields generated during low-, medium-, and high-surf conditions, with particular attention to the differences between inshore and offshore characteristics. Results from a similar data set obtained during DUCK94 will be presented for comparison as well.

3:05

2pAO4. Characteristics of the low- and high-frequency ambient sound from breaking waves in the surf zone. Ali R. Kolaini (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, Coliseum Dr., University, MS 38677) and Jeffrey A. Nystuen (Univ. of Washington, Seattle, WA 98105)

An experiment was conducted along the Scripps Institution of Oceanography Pier to study the characteristics of the ambient sound in the surf zone. The role of the single large bubble's (order of a few centimeters) oscillations versus the collective oscillations of bubble clouds as possible mechanisms of generation of the low-frequency sound, especially the relatively broadband peak that appears around 20 Hz in the sound spectrum, is discussed. The frequency band at which ambient sound from breakers in the surf leaks out is identified. The acoustic attenuation which makes the surf zone extremely quiet above 2 kHz, and a possible method to use ambient sound to infer bubble size distribution in shallow water near the surf zone, are discussed. [Work supported by ONR.]

3:20

2pAO5. A theoretical investigation of the low-frequency sound generated by breaking waves. Mark R. Loewen and Cesar E. Farell (Dept. of Mech. and Industrial Eng., Univ. of Toronto, Toronto, ON M5S 3G8, Canada, loewen@mie.utoronto.ca)

The important influence of breaking waves on the dynamics of the upper ocean has motivated research into developing techniques for quantitatively measuring the effects of wave breaking in the field. Breaking waves are a major source of sea surface sound and breaking events produce sound at frequencies as low as 10 Hz. Collective bubble cloud oscillations have been shown to be the dominant source of the low-frequency sound. Laboratory experiments have shown that the low-frequency sound generated by breaking is correlated with the amount of wave energy dissipated and the volume of air entrained. Therefore it may be possible to use passive acoustic measurements in the field to obtain estimates of the volume of entrained air and energy dissipated by individual breaking events. Development of a model of the low-frequency sound generated by individual breaking waves is an important step in achieving this goal. A simplified model of the low-frequency sound generated by breaking has been developed based on the approach of Oguz [J. Acoust. Soc. Am. **95**(4), 1895–1912 (1994)]. The importance of the simplifying assumptions used in the model and the use of the model as an aid to interpreting field and laboratory measurements will be discussed.

3:35

2pAO6. Acoustic propagation in the ocean surf zone. Steven L. Means, Richard M. Heitmeyer, Stephen C. Wales, Thomas J. Hayward (Naval Res. Lab., Code 7121, Washington, DC 20375), Ellen S. Livingston (Office of Naval Res., Arlington, VA 22217), and Jeffrey A. Schindall (Naval Res. Lab., Washington, DC 20375)

Broadband acoustic propagation measurements were conducted during the SandyDuck'97 experiment to provide a means of interpreting noise data taken both within and outside of the surf zone. The source was located outside of the nominal surf zone region at a distance of 470 m from the shore in 6 m of water. The receivers were spaced at 6-m intervals 45 to 183 m offshore in 1 to 3 m of water. Both the source and the receivers were located along a line orthogonal to the shore. The cw line sets were transmitted to obtain a description of the temporal fluctuations in the narrowband propagation induced by the temporal variability of the medium, the surface, and the entrained bubble distribution, as a function of range and frequency. The cw pulses and biphase-modulated PR sequences were transmitted to obtain a description of the spatial and temporal variability of the multipath structure. Measurements show significant time variations even under low surf conditions. The measured results are interpreted through simulations obtained using a FEPE propagation model with near-concurrent estimates of the bathymetry and geoacoustic parameters derived by Fabre and Wilson [IEEE J. Ocean. Eng. **22**, 434–444 (1997)]. [Work supported by ONR.]

3:50–4:05 Break

4:05

2pAO7. Wind-generated acoustic spectral effects in the surf zone in the presence and absence of rainfall at Duck, North Carolina. John R. Proni (NOAA/Atlantic Oceanograph. and Meteorological Lab., Ocean Acoust. Div., 4301 Rickenbacker Cswy., Miami, FL 33149, proni@aoml.noaa.gov) and John C. Wilkerson (NOAA/Natl. Environ. Satellite Data and Information Service, Washington, DC 20746)

Observations of wind-induced effects upon concurrently measured rainfall-generated spectra in the surf zone off Duck, North Carolina were made in November, 1992. Rapid changes in wind speed from about 5 m/s to 14 m/s were accompanied by a frequency-dependent reduction in rain-

fall spectral level. The reduction initiated at about 50 kHz (the highest frequency recorded) and progressed to lower frequencies with time. When the wind burst diminished, spectral levels increased. Rainfall rate was monitored using conventional rain gauges so that rainfall rate change effects could be distinguished from wind change effects in the measured acoustical spectra. The general surf-noise spectrum level changes were also observed to be affected by wind speed changes.

4:20

2pAO8. Acoustic observations of suspended sediments in the Changjiang Estuary. Shuying Zhang (Shanghai Acoust. Lab., Acad. Sinica, 456 Xiao-Mu-Qiao Rd., Shanghai 200032, PROC)

An acoustic backscatter instrument of dual frequencies (0.5 and 1.5 MHz) was developed for suspended sediment measurements and for sediment transport studies. The instrument measures the vertical concentration profiles of suspended sediments with a temporal resolution of 1 s and a spatial (depth) resolution of 10 cm for 0.5 MHz or 2 cm for 1.5 MHz. The data of backscattered acoustic energy are sampled at a rate of 75 kHz for a 450-s burst each time, so each burst comprises 450 vertical profiles. Coincidentally, six calibration points for the water column are taken through water sampling. Further data processing for each burst, such as compensation for sound transmission losses, *in situ* calibration of acoustic scattering intensity to real concentration of suspended sediments, and ensemble averaging over adjustable time–depth sizes permits three-dimensional concentration profiles to be derived. To increase the accuracy of concentration measurement, a multisegment procedure to form an appropriate compensation and calibration curve is applied. Through observations at the Changjiang Estuary four patterns of vertical suspension profiles are identified at flood and ebb tides. In addition, lutoclines and interfacial waves within the near-bed high concentration suspensions are revealed.

4:35

2pAO9. Motivations for using a pulsed full spectrum Doppler to measure bedload and near-bottom suspended sediment transport. Peter Traykovski, James D. Irish, and James F. Lynch (AOP&E Dept., WHOI, Woods Hole, MA 02543, ptraykovski@whoi.edu)

Acoustic observations of bedform migration and suspended sediment transport at the New Jersey LEO-15 site revealed that bedload and near-bottom suspended load could be the dominant sediment transport mode over wave-formed ripples on a sandy bottom. Bedforms observed during storms using a rotary sidescan sonar were found to be wave orbital scale ripples which migrated in the onshore direction, forced by asymmetrical wave orbital velocities. Estimates of suspended sand transport were calculated from observations of acoustic backscattering and water velocity profiles. Suspended sand transport was also forced by asymmetrical wave velocities, and was found to occur primarily during the weaker offshore phase of wave motion. This net offshore suspended sediment transport was an order of magnitude less than the flux associated with onshore ripple migration. Thus it is hypothesized that ripple migration is most likely forced by unobserved bedload and near-bottom suspended transport. Spatially resolving bedload transport from the stationary bed is difficult since the motion occurs within a few grain diameters of the bed. Acoustic Doppler-based techniques, which are ideal for this type of measurement since the sediment velocity can be resolved from the stationary bed in the frequency domain, are being developed.

4:50

2pAO10. Changes in acoustic impedance of marine sediment covered with liquid pollutants. Henning Harms, Rainer Matuschek, and Volker Mellert (Dept. of Phys., Univ. of Oldenburg, D-26111 Oldenburg, Germany)

Acoustic properties of porous material are slightly changed by a coating layer, which is thin compared to the acoustic wavelength. These changes affect mainly the spherical wave reflection coefficient at grazing incidence. The phase change of a surface wave component is observed and enables the easy detection of the thin layer by measuring the interference of direct and guided surface waves. This method is investigated for screening observations in order to monitor liquid pollutants on marine sediment. Since sediment is usually not plain but more or less irregularly grooved, the coating by liquid pollutants will occur in patches. Laboratory experiments are carried out with rough surfaces of sediment to investigate the sensitivity of detecting an average change in impedance. The results are compared with a numerical model incorporating Biot theory [Tooms *et al.*, J. Acoust. Soc. Am. **93**, 171–181 (1998)], which was successfully applied to model experimental data of plain sediment (Harms *et al.*, 134th meeting of ASA). Since the rough surface produces a significant backscattering, the liquid patches are detectable in the backscattered signal as well. Experimental results of this contribution are presented and the possibility of using this method for monitoring purposes discussed. [Work supported by German Ministry of Technology BMFT.]

5:05

2pAO11. Scattering by two spheres: Theory and experiment. Irina Bjørnø and Leif Bjørnø (Dept. of Industrial Acoust., Tech. Univ. of Denmark, Bldg. 425, DK2800 Lyngby, Denmark)

Extensive studies of scattering of acoustical signals by targets of different regular shapes have formed a useful background for attempts to develop procedures for remote monitoring of suspended materials in marine environments as, for instance, measurements of characteristic parameters of suspended sediments. The scattering properties of single regular-shaped particles have been studied in depth by several authors in the past. However, single particle scattering cannot explain all features of scattering by suspended sediment. When the concentration of particles exceeds a certain limit, multiple particle scattering becomes important. As a first step in the investigation of mutual interactions between several particles, the acoustical scattering by two spheres has been studied theoretically and experimentally and the results are reported in this paper. The study has mainly been focused on three issues: (1) to develop a simplified theory for scattering by two elastical spheres; (2) to measure the scattering by two spheres in a water tank, and (3) to compare the theoretical/numerical results with the measured data. A number of factors influencing multiple scattering, including the geometrical distribution of particles, their shape, size, and material properties are taken into consideration in the studies to be reported. Broadband signals were used for the experimental studies.

5:20

2pAO12. Acoustic propagation affected by internal solitons in a coastal area. E. C. Shang, Y. Y. Wang, and L. Ostrovsky (Cooperative Inst. for Res. in Environ. Sci., Univ. of Colorado/NOAA/ETL, Boulder, CO 80303)

The NOAA/Environmental Technology Laboratory (ETL) conducted the ''Coastal Ocean Probe Experiment'' (COPE) in the Oregon coastal area (with water depth of 150 m) in September, 1995. Strong internal soliton activities have been observed. The isotherm records measured by FLIP demonstrate that the maximum amplitude of isotherm depression can reach ~30 m. Based on these internal soliton data, numerical simulation on acoustic propagation through solitons has been conducted. Some interesting features of the acoustic field are found: (1) The internal solitons can cause significant mode coupling, but there is no abnormal transmission loss (TL) appearing in the frequency range 50–1000 Hz. (2) Mode coupling strongly depends on frequency, mode number, and the direction of propagation. (3) The modal horizontal refraction (MHR) caused by the internal solitons is significant and offers additional useful information for acoustic monitoring of the internal soliton activities. [Work supported by NOAA.]

Session 2pBV

Bioresponse to Vibration/Biomedical Ultrasound and Physical Acoustics: High Intensity Therapeutic Ultrasound II

Gail R. ter Haar, Cochair
Physics Department, Royal Marsden Hospital, Sutton, Surrey SM2 5PT, England

Pierre D. Mourad, Cochair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

Contributed Papers

2:10

2pBV1. Optimal acoustic parameters for induced hyperthermia from focused MHz ultrasound: Phantom measurements with fluid flow and bubble activity. R. Glynn Holt and Ronald A. Roy (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, rgholt@bu.edu)

The effective coagulation of flowing blood by ultrasound, or acoustic hemostasis, is a complex process that encompasses several physical, biophysical, and biochemical facets. Motivated by this ultimate goal of efficient and predictable hemostasis, we have undertaken a series of laboratory experiments on a variety of phantoms insonified by 1-MHz focused ultrasound at MPa amplitudes. The goal of the present study is to quantify and characterize the physical response of a viscous, particulate-laden Newtonian flow to high-intensity focused ultrasound (HIFU). Experimental hyperthermia in a tissue/flow phantom as a function of acoustic parameters (pulse width, duty cycle, amplitude), flow parameters (volume flow rate, vessel size, flow geometry), and bubble content in the flow phantom is investigated. Spatiotemporally resolved temperature measurements in the flow and surrounding tissue phantom are reported. Time permitting, visualization of the modified flow due to acoustic streaming and radiation pressure will be undertaken to correlate flow events with the hyperthermia results. [Work supported by DARPA.]

2:25

2pBV2. Numerical simulations of tissue heating created by high-intensity focused ultrasound. Francesco P. Curra, Peter Kaczkowski, Pierre D. Mourad, Lawrence A. Crum (Univ. of Washington, Appl. Phys. Lab., 1013 NE 40th St., Seattle, WA 90105, pierre@apl.washington.edu), and Vera A. Khokhlova (Moscow State Univ., Moscow, Russia 119889)

Modeling of high-intensity focused ultrasound (HIFU) propagation and heating effects in tissue has usually been studied under the assumption of linear acoustics. Nevertheless, nonlinear propagation can lead to important differences in heat deposition depending on the degree of nonlinearity achieved in a HIFU field. To be presented are calculations of the spatial patterns of acoustic intensity (I) and heat generation [$H=-(\partial/\partial z+1/r\partial/\partial r)I$ for strongly nonlinear acoustic waves and $H=2\alpha I$ for weakly nonlinear acoustic waves, where α is acoustic absorption] associated with focused weakly and strongly nonlinear acoustic wave propagation in biological tissue. The propagation is modeled using a KZK equation method with special attention to correct modeling of shock generation and biologically relevant attenuation, as well as using initializing data based on existing designs of transducers for acoustic hemostasis. The heat generation term forces the ''bio-heat'' equation, which predicts the generation, movement, and dissipation of heat within biological tissue. The tissue model contains blood flow in an artery or vein and is perfused by a capillary bed. For this talk the differences in spatial patterns of acoustic intensity created by weakly and strongly nonlinear CW 1-Mhz acoustic wave propagation are shown. How those differences translate into differences in heat generation and therefore bioeffect within biological tissue are then demonstrated. [Work sponsored by DARPA.]

2:40

2pBV3. Occlusion of blood vessels using high-intensity focused ultrasound. Shahram Vaezy (Anesthesiology, Box 356540, Univ. of Washington, Seattle, WA 98195, adasi@u.washington.edu), Roy Martin, Peter Kaczkowski (Univ. of Washington, Seattle, WA 98195), George Keilman (Sonic Concepts, Woodinville, MA 98072), Steve Carter, Michael Caps, Pierre Mourad, and Lawrence Crum (Univ. of Washington, Seattle, WA 98195)

Vascular injury may result from both trauma and invasive medical procedures. High-intensity focused ultrasound (HIFU) was used to occlude surgically exposed blood vessels of anesthetized rabbits, as a mechanism of hemostasis applicable to vascular injury. A HIFU transducer, with a spherically curved aperture area of 6.34 cm^2, a focal length of 4 cm, and a frequency of 3.3 MHz, was used. The transducer was equipped with a conical housing filled with water, providing a coupling medium to the vessels. Femoral arteries and veins were occluded after a HIFU treatment of approximately 30–60 s, at a focal intensity of 3000 W/cm^2. The vessels were severed after the treatment to confirm complete occlusion. No bleeding was observed. Gross and microscopic examination of the vessels revealed two major mechanisms of occlusion, dissection of the vessel wall and shrinkage of the vessel. In the dissection mechanism, a flap of the inner wall of the vessel peeled as a result of the HIFU treatment and caused a blockage of the blood flow, while in the shrinkage mechanism a mechanical constriction of the vessel lumen caused the blockage. HIFU may provide a valuable method for hemostasis of injured vessels by treating a region of the blood vessel, proximal to the hemorrhaging site.

Invited Paper

2:55

2pBV4. Occlusion of blood flow by high-intensity focused ultrasound. Ian Rivens, Ian Rowland, Gail ter Haar (Phys. Dept., Royal Marsden Hospital, Sutton, Surrey, UK, ian@icr.ac.uk), Mark Denbow, and Nicholas Fisk (Royal Postgraduate Med. School, Inst. of Obstetrics and Gynaecology, Queen Charlotte's and Chelsea Hospital, London, UK)

This work investigates the possiblity of occluding blood flow (*in vivo*) using high-intensity focused ultrasound surgery (FUS). Such an effect could be used, for example, in the noninvasive treatment of foetal dysfunctions. A high-power 1.7-MHz, 150-mm focal length, piezoelectric transducer was used to expose rat femoral vessels to a free field $I_-^{SP}=4660$ W cm^{-2} for 2 s under terminal

anaesthesia. Initially, eight separate exposures were used. Later experiments were aimed at determining an optimal single exposure to produce flow occlusion. Before and after exposure, magnetic resonance FISP 3-D angiograms (MRA) with submillimeter spatial resolution were acquired using a Siemens Vision MR System (1.5T) and extremity coil. In all animals, MRA showed reduced blood flow both in the treated region and distal to it. The results of single-dose optimization will be discussed. The ability to curtail blood flow using FUS allows the possibility of treating Foeto-Foetal Transfusion Syndrome by occluding the placental shunt vessels responsible for vascular inbalance in identical twins. Currently available interventional treatments (surgery or intrauterine laser therapy) have significant mortality and morbidity. The FUS has the advantage of being noninvasive. [Work was funded by the UK Medical Research Council.]

Contributed Paper

3:15

2pBV5. Measurements of sound speed in excised tissue over temperatures expected under high-intensity focused ultrasound conditions. Susannah H. Bloch, Michael R. Bailey, Lawrence A. Crum, Peter J. Kaczkowski (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105, peter@apl.washington.edu), George W. Keilman (Sonic Concepts, Inc., Woodinville, WA 98072), and Pierre D. Mourad (Univ. of Washington, Seattle, WA 98105)

Noninvasive monitoring of high-intensity focused ultrasound (HIFU) treatment by diagnostic ultrasound is currently limited by a poor understanding of the bioacoustic effects of tissue heating. Sound-speed measurement has been proposed as a means of noninvasive thermal mapping. The sound speed in most tissues changes monotonically but nonlinearly with temperature to around 60 °C. Above that temperature, at which the tissue proteins begin to denature, irreversible changes occur in the tissue and its corresponding acoustic properties. Sound-speed data around and above the range in which the tissue ''cooks'' is a challenge to obtain because of changes in tissue shape, compressibility, and overall dimensions associated with the cooking process. In this research, an apparatus to measure sound speed during a controlled temperature change was used to study the effects of temperature, thermal dose, and thermal dose rate on the speed of sound in various mammalian tissues. The apparatus, potential sources of error, and results will be presented, and implications for forward and inverse HIFU propagation modeling and acoustic remote sensing of HIFU-induced temperature changes will be discussed. [Work supported by DARPA.]

3:30–3:45 Break

Invited Papers

3:45

2pBV6. Role of cavitation during high intensity focused ultrasound treatment of prostate tissue. Narendra T. Sanghvi (Focus Surgery, Inc., 3940 Pendelton Way, Indianapolis, IN 46226)

The use of high intensity focused ultrasound (HIFU) has been well established to produce deep seated tissue necrosis using a short time ultrasound exposure. The scavenging system of the body replaces the necrosed tissue over time from the treated organ. However, in the case of benign prostatic hyperplasia (BPH) treatment, it is essential to remove treated tissue immediately after the treatment to offer improved urine flow rate and relief from the symptoms. Therefore, it was necessary to develop an ultrasound treatment system that not only provides thermal lesions that are deep in the tissue but also mechanical disintegration of urethra and surrounding tissue. Once the urethra is mechanically disintegrated, the necrosed tissue is sloughed with urine and provides an improved flow rate. The mechanical defect in tissue is produced using the cavitation process. The acoustic cavitation is produced by inserting a silastic catheter in the urethra. The catheter blocks the HIFU energy and produces standing waves which produce cavitation, and converts HIFU in a thermal dose larger than that produced by only absorption of acoustic energy by tissue. This method of treatment yields far greater results. The system is now being used in human BPH treatment.

4:05

2pBV7. Transrectal high-intensity ultrasound therapy of localized prostate cancer. Jean Y. Chapelon (INSERM U281, 151 Cours A. Thomas, Lyon, F69424, France, chapelon@lyon151.inserm.fr), Albert Gelet (E. Herriot Hospital, Lyon, F69434), and Emmanuel Blanc (Edap Technomed, Vaulx en Velin, F69120)

Prostate cancer is the most common cancer in men. The Ablatherm project has been developed in order to achieve a minimally invasive treatment of localized prostate cancer using transrectal high-intensity focused ultrasound (HIFU). The method uses a sharply focused wideband transducer operating in the frequency range 2.25–3.0 MHz to induce a localized thermal necrosis with a shot of 4.5-s exposure. This therapy transducer focused at 40 mm is coupled to a retractable 7.5-MHz biplane imaging transducer. A computer-controlled positioning device moves the probe in the rectum to ablate the volume defined with the image by repeating the shots every 5 s. For patient safety, the system includes a power monitoring circuitry combined with a real time A-mode ultrasound detection of the rectal wall. Since 1993, 44 patients have been treated in Lyon: A complete response was obtained in 26 and only 18 numerous advantages: It is minimally invasive requiring only a short hospital stay; it is repeatable unlike radiotherapy and additional treatment centered on insufficiently treated area can be performed. [Work supported by Technomed and Rhone Alpes Futur.]

4:25

2pBV8. Concept of biological focal field and its importance in tissue resection with high-intensity focused ultrasound. Zhi B. Wang, Feng Wu, Zhi L. Wang, and Chuan Liu (Inst. of Ultrasonic Eng. in Medicine, Chongqing Medical Univ., Box 153, 1 Medical College Rd., Chongqing 400046, PROC, chhifu@public.cq.sc.cn)

Many studies showed that *in vitro* the size and shape of the necrosed tissue volume generated by high-intensity focused ultrasound (HIFU) were dependent upon the transducer size and radius of curvature, operating frequency, input power, pulse duration, and delay time between each pulse. In this research the shape and size of necrosed tissue volume produced by HIFU with 1.6-MHz focused transducer were carefully checked and compared in rabbits and pigs. This volume was generated by a single exposure or focal planes. It was shown that with the same therapeutic parameters, HIFU can selectively necrose one region of liver *in vivo*. However, its shape and size differed between *in vitro* and *in vivo*. These changes implied that the necrosed tissue volume *in vivo* depended upon not only HIFU conditions, but also the structure, function, and movement of target tissue. In contrast with the physical focal region, necrosed tissue volume *in vivo* could also be termed the biological focal field of HIFU. Theoretically, it resembled more closely the living tissue damaged by HIFU. In order to avoid the side effects, it is important to set up the data about biological focal fields of different organs in tissue resection with HIFU.

4:40

2pBV9. Changes in ultrasonic images of tissue damaged by high-intensity focused ultrasound *in vivo*. Feng Wu, Zhi B. Wang, and Zhi L. Wang (Inst. of Ultrasonic Eng. in Medicine, Chongqing Medical Univ., Box 153, 1 Medical College Rd., Chongqing 400046, PROC, chhifu@public.cq.sc.cn)

In order to clarify the effects of high-intensity focused ultrasound (HIFU) on the liver *in vivo*, the morphological changes in a large animal model were detected by quantitative analysis of ultrasonic image. One region of the liver was selected as a target tissue by ultrasonic imaging. Then, using the method of focal planes, the target part of the liver was resected by ultrasound-monitored HIFU with a 1.6-MHz extracorporeal spherical-bowl transducer. Ultrasonic images of the liver at the selective region were acquired before, immediately, and consecutively after HIFU treatment. The bit difference of ultrasonic image at the selective damage of the liver *in vivo* before and after HIFU treatment was detected by a quantitative analysis system. It was observed immediately after focused extracorporeal pyrotherapy. Sharp changes in ultrasonic image were also detected by consecutive monitoring after treatment. It corresponded to the histological characteristics of the changes of the damaged tissue. These results suggest that as a method to detect necrosed tissue, the quantitative analysis of ultrasonic image may also be used as an indicator of morphological changes after HIFU treatment.

4:55

2pBV10. Ultrasound-positioned high-intensity focused ultrasound in the ablation of superficial bladder wall in miniswines. Zhi B. Wang, Feng Wu, Zhi L. Wang, and Chuan Liu (Inst. of Ultrasonic Eng. in Medicine, Chongqing Medical Univ., Box 153, 1 Medical College Rd., Chongqing 400046, PROC, chhifu@public.cq.sc.cn)

In order to simulate the ablation of a superficial bladder tumor, the effectiveness of ultrasound-positioned high-intensity focused ultrasound (HIFU) on the bladder wall was determined. One part of the bladder wall in 10 miniswines was selected as a damage region using ultrasonic imaging. Then, it was treated with a 1.6-MHz extracorporeal spherical-bowl transducer. Animals were killed immediately after treatment and the bladder wall and rectal tissue were examined macroscopically and histologically. As a result, brighter echoic regions at the selective damage of the bladder wall were detected by ultrasonic imaging immediately after treatment. These brighter echoes corresponded to the acute bladder wall damage which caused obvious necrosis on the superficial bladder wall. Also, histological analysis of the treated areas revealed epithelial denudation and irreversible changes of submucosal and superficial muscular layer under light and electron microscope. Two of the ten animals had microscopic hematuria. No macroscopic hematuria and rectal perforation occurred except for one case of a small burn on the rectal wall. It was concluded that as an extracorporeal approach HIFU can be used successfully to sonicate the selective region of the bladder wall in a large animal model *in vivo*.

5:10

2pBV11. Effects of high-intensity focused ultrasound on H_{22} liver tumor: An evaluation with different therapeutic programs. Feng Wu, Zhi B. Wang, and Zhi L. Wang (Inst. of Ultrasonic Eng. in Medicine, Chongqing Medical Univ., Box 153, 1 Medical College Rd., Chongqing 400046, PROC, chhifu@public.cq.sc.cn)

Most research shows that high-intensity focused ultrasound (HIFU) offers an extracorporeal approach to local removal of tumors. In order to choose the optimal conditions for clinical trials, this study was to evaluate the effects of different therapeutic programs of HIFU with 1200 W/cm^2 on H_{22} liver tumors. Seventy-three ICR tumor-bearing mice were divided into a control group and HIFU therapeutic groups. The therapy programs were defined by the numbers of pulses and pulse durations of HIFU with a 1.6-MHz focused transducer. The various parameters of tumor growth, x-ray, histology, electron microscopy, and mice nutritional state were used to assess the results of HIFU treatment. The data showed that HIFU had an inhibitory effect on H_{22} tumor growth in all therapeutic groups. With the same input power, the local effects of tumor removal were dependent upon the pulse durations and numbers of pulse. When the total number of pulses increased within a short time, the inhibitory effect was stronger. However, the nutritional state of animals in this group became worse. It implies that severe stress response and catabolism induced by an increase of HIFU pulse in a short time would occur.

Session 2pEA

Engineering Acoustics: Topics in Engineering Acoustics

P. K. Raju, Chair
Mechanical Engineering Department, Auburn University, Ross 201, Auburn, Alabama 36849-5341

Contributed Papers

2:15

2pEA1. Ultrasonic echo signals obtained in the outdoor environment. Michael J. Anderson, Jeffrey Luke, John Canning, and Dean B. Edwards (Dept. of Mech. Eng., Eng. Phys. Bldg. 324O, Moscow, ID 83844-0902)

Many, if not most, applications of ultrasonic transducers in air are for ranging systems used in a relatively structured environment. We describe the application of ultrasonic transducers to range sensing for a robotic autonomous vehicle that is intended to operate outdoors in a forested environment. Unlike the indoor environment, application of ultrasonic ranging outdoors involves a much more cluttered environment. The outdoor environment contains a large number of categories of objects and features, as well as a wide variation within each category. In this talk, a statistical summary of echoes obtained from features including ground contour, roughness elements, and obstacles common in the forest environment is presented. Information that can be extracted from the echo data that is useful to an autonomous controller is then discussed.

2:30

2pEA2. Modeling of electrostatic ultrasonic transducers with micro air-gap structure. Li-Feng Ge (Dept. of Phys., Anhui Univ., 3 feixi Rd., Helfei 230039, PROC)

A theoretical model developed by previous papers, i.e., the plate-on-air-spring model or short-tube model (or DK model) [L.-F. Ge, J. Acoust. Soc. Am. **100**, 2809 (1996), and Chin. Sci. Bull., Academia Sinica **42**, 22 (1997)], has been successfully extended to a generalized case, in which diaphragms of transducers are loaded also by in-plane tensile forces. Based on the dynamic analysis of such an electrostatic transducer, its diaphragm can be considered as a tensile thin plate supported by air-spring, or as a short tube enclosing air, terminated by a flexible tensile diaphragm and grooved backplate as a rigid wall. The physical model and mathematical formulation describing the transducer behavior has been further developed. A comparison between the tensile-plate-on-air-spring model (or TDK model) and the plate-under-tension model (may be termed TD model) [M. J. Anderson *et al.*, J. Acoust. Soc. Am. **97**, 262–272 (1995)] is discussed in the paper by an example of predicting the resonant frequencies of a U-grooved backplate transducer with $T=100$ N/m and $Vp=250$ V. The fundamental frequency is estimated by the new model at 349 kHz, corresponding pistonlike motion of the diaphragm, and second at 515 kHz, corresponding to the first bending vibration mode. The predictions are more consistent with the measurements reported. [Work supported by the Anhui Education Council Foundation.]

2:45

2pEA3. Control of particles in a standing wave field using ultrasonic vibration. Shinfuku Nomura, Kouichi Murakami, Yukinori Yoshikawa, and Junji Ochi (Ehime Univ., Dept. of Mech. Eng., 3, bunkyo-cho Matuyama, 790 Japan, nomu@en1.ehime-u.ac.jp)

In order to control the motion of micromachines by acoustic power, the behavior of particles falling in several liquids with a standing wave field was studied experimentally. A PZT transducer with a frequency of 28 kHz was fixed to the bottom of a cylinder. Several small tubes were inserted in this cylinder, and glass particles with a diameter of about 1.5–3.5 mm were dropped in these tubes, respectively. Since cavitation bubbles are observed by applying ultrasonic vibration to a liquid, cavitation intensity as estimated from the erosion loss of an aluminum sheet was measured to investigate the effect of cavitation generation on the radiation pressure. Cavitation intensity in degassed water was weaker than that in ion-exchanged water; however, the decrease in velocity of the glass particles by applying ultrasonic vibration exhibited the greatest deceleration in the degassed water. Ultrasonic power effectively acts on a particle as a radiation force when cavitation does not occur and there is almost no damping of the ultrasonic wave. When varying the tube diameter, as the ratio of the diameter of tube to the particle diameter increased, so did the deceleration. All particles were levitated by the ultrasonic radiation force in a ratio of more than 0.6.

3:00

2pEA4. Miniaturized ultrasonic sensor system for liquid analysis. Bernd Henning, Peter C. Daur, Alexander Wolf (Institut für Automation und Kommunikation e.V. Magdeburg, Steinfeldstrasse 3, D-39179 Magdeburg-Barleben, Germany, hen@ifak.fhg.de), Peter Hauptmann, and Alf Puettmer (Otto-von-Guericke Univ. IPE, Pf 4120, D-39106 Magdeburg, Germany)

At the moment, known commercial devices to measure the sound velocity in liquids are limited in their application because the relation between velocity and concentration of substances is often not clear, for instance, if it is a parabolic curve. Additionally, the temperature dependence shows similar behavior. This contribution will present an ultrasonic sensor system consisting of a miniaturized multisensor arrangement for the comprehensive acoustic characterization of liquid mixtures. The sensor is based on segmented cylindrical piezoceramics and measures velocity, impedance coefficient, attenuation coefficient, and temperature. If it is a known liquid mixture it is possible to use the measuring quantities to calculate the material properties acoustical impedance and absorption or the concentration of selected substances in complex liquids. The sensor system is able to adapt itself to different and changeable absorption conditions which is typical of industrial processes. The measuring chamber is optimized to investigate small volumes of liquid specimens. In this contribution the sensor design and the developed approach for the continuous and simultaneous measurement of acoustic properties of liquids will be described. Application examples of the sensor in the chemical and food industry will show the advantages and limits of our new sensor system.

3:15

2pEA5. The multicomponent composition of the low acoustic impedance matching layer for the ultrasonic transducer operating in the air. Tadeusz Gudra (Inst. of Telecommunication and Acoust., Wroclaw Univ. of Technol., Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland)

Piezoceramic transducers with matching layers are usually used to generate an ultrasonic wave in the air within the frequency range above 100 kHz. The paper presents a new material composition of low acoustic impedance named ITAKOM, whose acoustic characteristics makes it applicable as a matching medium between a piezoelectric transducer and a gas medium in a wide range of frequencies. Porosity is the common feature of all the lowest acoustic impedance materials. The results of the studies of different material compositions show that pore diameters and

their percentage in the composite volume have greater importance than the composite make-up and render possible the elaboration of a composite material whose parameters can be defined by the percentage interdependence of components. Due to the comparatively wide range of density and wave velocity materials of acoustic impedance ranging from $0.2-1\times10^6$ Pa·s/m can be produced. An example of an application of this material to construct the focusing ultrasonic transducer operating in the air at the frequency $f=1.3$ MHz in the setup of phase sensitive scanning acoustic microscopy as well as the results of object measurements are presented in the ultrasonic journal [Gudra *et al.*, Ultrasonics **34**, 711–719 (1996)].

3:30

2pEA6. High-intensity ultrasound for dewatering of slurries. Luis Elvira-Segura, German Rodriguez-Corral, Enrique Riera F. de Sarabia, and Juan A. Gallego-Juarez (Instituto de Acustica, CSIC, Serrano, 144, 28006 Madrid, Spain)

Dewatering is a process in which the water is removed from a product without changing its phase. Fine particle slurries are difficult to dewater and several processes such as filtration and centrifugation are conventionally applied. These processes, which are efficient to remove free water, are ineffective in removing water from interstitial spaces. High-intensity ultrasound can be used to assist conventional filtration processes in order to increase the rate of dewatering. Particularly, the application of ultrasonic energy has shown to be useful as a postfiltration process, once the free water has been removed. In this paper a new system is presented in which the combination of vacuum filtration and ultrasonic energy results in very efficient dewatering of very fine slurries constituted by particles sized between 0.1 and 10 μ. The system is based on the application of direct coupled ultrasonic energy to the cake formed on the surface of a ceramic vacuum filter by means of specially designed plate transducers. The obtained results demonstrate that the application of ultrasound significantly increases the interstitial water removal and, consequently, it diminishes the energy consumption of the whole process (dewatering and drying).

3:45–4:00 Break

4:00

2pEA7. Parametric array in air: Distortion reduction by preprocessing. Thomas D. Kite (Dept. of Elec. and Comput. Eng., Univ. of Texas, Austin, TX 78712-1084, tom@vision.ece.utexas.edu), John T. Post (Univ. of Texas, Austin, TX 78713-8029), and Mark F. Hamilton (Univ. of Texas, Austin, TX 78712-1063)

In a parametric array, highly directional low-frequency sound is generated by the self-demodulation of an intense, amplitude-modulated high-frequency sound beam as a result of nonlinear propagation effects. The term ''audio spotlight'' was introduced by Yoneyama *et al.* [J. Acoust. Soc. Am. **73**, 1532 (1983)] for a parametric array in air used to generate directional audio frequency sound with an ultrasonic primary beam. Berktay's far-field solution [Berktay, J. Sound Vib. **2**, 435 (1965)] predicts a demodulated secondary waveform along the axis of the beam that is proportional to the second time derivative of the square of the modulation envelope. The secondary wave is therefore generated with high levels of harmonic distortion, even at moderate modulation indexes [Blackstock, J. Acoust. Soc. Am. **102**, 3106(A) (1997)]. Integrating the modulation signal twice and taking the square root removes this distortion; however, the resulting reduction in distortion due to taking the square root is severely limited by the bandwidth of the primary beam transducer. Here, the characteristics of the transducer are taken into account in developing an appropriate predistortion scheme to achieve high quality reproduction. Relationships between distortion, modulation index, secondary bandwidth, and secondary sound-pressure level are also discussed, and results of simulations are presented. [Work supported in part by ARL:UT IR&D.]

4:15

2pEA8. Acoustical imaging in air by a low-cost system working in the audio band. Andrea Trucco (DIBE–Univ. of Genoa, Via Opera Pia 11A, 16145 Genoa, Italy)

In order to generate a two-dimensional acoustic image of an aerial scene, a system working in the audio band and composed of low-cost conventional electroacoustic and computation devices has been developed and tested. Such a system uses a loudspeaker for the emission of acoustic pulses, eight microphones for the reception of backscattered waves, a generation/acquisition digital board, and a personal computer. As in medical echography, the adopted algorithm for the image formation is a wide-band coherent beamforming followed by a scan conversion in order to display the image in a Cartesian format (as a top view of the imaged scene). Robotic vision and remote measurement are potential applications of the system developed. Two major problems that were faced are the calibration of the microphones and the study of the best configuration in terms of pulse waveform and microphone placement. A simple calibration algorithm was achieved by comparing the obtained point spread function with the theoretical expectations, whereas the best configuration was obtained by taking into account the constraints imposed by the instrumentation and the application addressed. The obtained images are of good quality, also considering the low-cost instrumentation and the highly reverberant environment in which the experiments were performed.

4:30

2pEA9. Numerical analysis of high-intensity ultrasonic processing systems. Cleofe Campos-Pozuelo (Instituto de Acustica (CSIC) Serrano, 144 28006 Madrid, Spain), Bertrand Dubus (Institut d'Electronique et de Microelectronique du Nord, 59046 Lille Cedex, France), and Juan A. Gallego-Juarez (Instituto de Acustica (CSIC) Serrano, 144 28006 Madrid, Spain)

High-intensity applications of ultrasonic energy in industrial processing are based on the performance of the power transducers and their coupling with the processing chamber. In fact, the right distribution of the acoustic field inside the treatment chamber and the efficiency of the system transducer-chamber are determinant factors for the feasibility of the applications. Therefore, the knowledge of the influence of geometry and dimensions of the processing chamber as well as the effect of the excitation transducers on the pressure field distribution is essential for the development of practical systems. This paper deals with the finite element modeling of a high-intensity ultrasonic processing system which is composed of a flexural-plate transducer and a cylindrical tube with characteristic dimensions much longer than the wavelength. A three-dimensional numerical procedure has been developed for modeling the finite amplitude acoustic field. Fundamental and second harmonic pressure distribution are calculated for several chamber geometries and excitation conditions. It has been found that the efficiency of the system is dependent on the geometry and dimensions of the chamber as well as on the efficiency, dimensions, and vibration distribution of the exciting transducer. The results give a guideline for the design of practical processing systems.

4:45

2pEA10. A method of interpolating the diffractive information of the sphere-baffled microphone in the sound field of the spherical wave. Kimitoshi Fukudome (Dept. of Acoust. Design, Kyushu Inst. of Design, Fukuoka, 815 Japan)

The sphere-baffled microphone (SBM) has been used for estimating DOA of progressive plane waves and their extractions, but it is now planned to construct the diffractive information directory (DID) of the SBM with DOA entry and source_distance entry in the sound field of spherical wave. This paper deals with the interpolation methods of getting information associated with entries not appearing in DID. Numerical computation of the diffractive information is done for the 2-element SBM (right and left microphones at both ends of a diameter of 17.68 cm) at frequency points corresponding to the sampling frequency 22.05 kHz and 1024 point DFT. Two data sets of DOA interval of 6° [0,6,12, . . .] and [3,9,15, . . .] are used for interpolation and evaluation of errors. Interpo-

lations and their error evaluations are done in the regions of both the log magnitude versus frequency and the group delay versus frequency. Both the bilinear interpolation and cubic interpolation method result in significant interpolation errors up to 10 dB and 0.14 ms for the shadow region. Finally, a new method is presented with high accuracy of interpolation based on the landmarks-loci of the diffractive information.

5:00

2pEA11. A conceptual experiment of direct converting digital microphone. Yoshinobu Yasuno (Matsushita Commun. Industrial Co., Ltd., 600 Saedo, Tsuzukiku, Yokohama, 224 Japan) and Yasuhiro Riko (Riko Assoc. 1988-34, Kitahassaku, Midoriku, Yokohama, 226 Japan)

A transducer which directly converts signals from analog acoustic to digital electric is described. This consists of a subtracter, a sampling and holding circuit, a sigma-delta modulation circuit as a comparator, an accumulator, and a local digital-to-analog converter similar to a typical electric analog-to-digital converter. The subtracter is an electrostatic device which has a diaphragm, driving electrode, and detecting electrode. The driving electrode is divided into corresponding digital signal bits by the ratio of each area which enables one to drive the diaphragm digitally and is regarded as an electroacoustic digital-to-analog converter. The detecting electrode makes electric signals proportional to the displacement of the diaphragm driven by subtracting the incoming acoustic signal and electrically driven force. This is regarded as the subtracter. This detected signal is amplified and sampled-held and modulated by the sigma-delta procedure and generates subtracted signal to $+$ or -1 bit which is added to the accumulator memory by a very high-speed clock rate. The output of the accumulator is the digital signal output and is also given to the driving electrodes. A 4-bit conceptual device was developed for an experiment. Some results will be reported.

5:15

2pEA12. Acoustic simulation for a loudspeaker using FEM/BEM. Veijo Ikonen, Antti Suutala, and Seppo Pohjolainen (Tampere Univ. of Technol., Dept. of Math., P.O. Box 692, FIN-33101 Tampere, Finland, ikonen@alpha.cc.tut.fi)

Although the theory of acoustics is well known, computers and numerical methods have not been effectively utilized in the field. Formulating a sufficiently good physical model is often difficult. The wave equation for pressure can easily be derived, but the boundary conditions are not so obvious. The structure of the walls of an acoustic cavity and the behavior of surface materials can be quite complex. In addition, under the influence of an acoustic field the structures can start vibrating, and this vibration is then coupled to the acoustic field. Mathematically this leads to a system of two partial differential equations. This paper studies the problem of computing an internal acoustic field of a loudspeaker box. The acoustic properties of walls, made of MDF, are taken into account by using the impedance boundary condition. The coupling between the structure and the fluid is also considered. Both finite-element and boundary-element methods (FEM/BEM with SYSNOISE) are used for numerical computations and the results are compared with each other and with analytical solutions. When the wavelength is decreased, the element grid must be refined and the number of equations becomes very large. This is why partial differential equation models are more suitable in low frequencies.

5:30

2pEA13. Vibration characteristics of high frequency and complex vibration systems of 160 kHz to 1 MHz for ultrasonic wire bonding. Jiromaru Tsujino and Hiroyuki Yoshihara (Kanagawa Univ., Faculty of Eng., 3-27-1 Rokkakubashi, Yokohama, 221 Japan)

Vibration characteristics of high frequency and complex vibration systems of 160 kHz to 1 MHz for ultrasonic wire ball bonding are studied. High-frequency and complex vibrations have been shown to be very effective for ultrasonic wire bonding of IC, LSI, or various electronic devices. Required vibration displacement amplitude for successful bonding of a 600-kHz system is about 1/30 that of a conventional 60-kHz system. For installing a conventional ceramic capillary 1.8575 mm in diameter required for ball bonding, longitudinal to complex transverse vibration systems 5.5–7.0 mm in diameter are proposed. Vibration velocity distributions along longitudinal, transverse vibration systems and ceramic capillaries and vibration loci are measured by laser Doppler vibrometers. The ceramic capillaries are measured vibrating in transverse vibration modes with one, two, or three vibration nodes along them at 160 kHz to 1 MHz, and the conventional capillaries are revealed applicable in such high-frequency systems. The free edge of the capillary vibrates in elliptical or circular locus with the two longitudinal vibration systems driven simultaneously. Maximum vibration velocity at a capillary tip is over 0.5 m/s (peak-to-zero value) which is enough for bonding of thin wires successfully at high frequencies.

TUESDAY AFTERNOON, 23 JUNE 1998 GRAND CRESCENT (W), 1:00 TO 4:15 P.M.

Session 2pMU

Musical Acoustics: Physics and Materials of Musical Instruments II

Shigeru Yoshikawa, Chair

Department of Acoustical Design, Kyushu Institute of Design, 4-9-1 Shiobara, Minami-Ku, Fukuoka, 815 Japan

Invited Papers

1:00

2pMU1. Composite materials for musical instruments: The maturity. Charles Besnainou (Lab. d'Acoustique Musicale, Univ. Pierre et Marie Curie, 4, place Jussieu-75252 Paris Cedex 05, France)

Nowadays, musical instruments made of composite materials allow comparison with good instruments made by craftsmen. These instruments result from 20 years of accumulated experiments carried out in laboratories. The first substitute of wood was a composite of wood: plywood. Over 70 years with this new material, honorable results were obtained for an industrial line of products: piano, harp, guitar, double bass. However, composite materials made of fibers and resin are designed to imitate the structure of wood. So a specific knowledge is needed. The first attempts were on bows because it seemed easy; good results were obtained with fiberglass. The next step was to replace the neck of an electric guitar. However, it was a long time before acoustic elements such as a soundboard

were designed. One needs first of all to develop a methodology to measure the mechanical parameters of wood: density, Young's modulus, and damping; and to cross over the mechanical results with the subjective choices of expert craftsmen. Of course, assimilated composite techniques are required. And then to learn, as in traditional apprenticeship, how to feel the quality of an instrument. Paradoxically, in terms of sound qualities, new materials need the same knowhow used by traditional instrument makers. Maturity means that the updated composite techniques are now very simple to be assimilated by craftsmen and manufacturers too; the prices of instruments are sufficiently low and the sound qualities have reached a good level. Some instruments, namely, the violin, lute, and guitar are demonstrated.

1:25

2pMU2. Nonlinear effects in a vibrating string (wire). Roger J. Hanson, H. Kent Macomber, and Michael P. Kassakatis (Dept. of Phys., Univ. of Northern Iowa, Cedar Falls, IA 50614, roger.hanson@uni.edu)

Many unusual effects are observed in vibrating strings or wires as a consequence of tensional changes and longitudinal motion which yield coupled nonlinear equations for transverse motion in orthogonal directions. For a wire driven sinusoidally in a transverse direction, notable effects observed at driving frequencies near a natural resonance include the generation of motion perpendicular to the plane of the driving force, sudden jumps in amplitude and phase, whirling motion, and hysteresis [R. J. Hanson, J. M. Anderson, and H. K. Macomber, J. Acoust. Soc. Am. **96**, 1549–1556 (1994)]. Further work demonstrates the generation of second, third, and higher harmonics with amplitudes sometimes comparable to the first. These harmonics, present in both the driving and perpendicular directions, contribute to a very complex total motion. All of the effects are very sensitive to asymmetries which produce slight differences in free vibration frequencies in two orthogonal directions. For the large amplitude free vibration of a plucked string, there are also significant consequences of the nonlinear effect of the tension changes. These effects for a plucked string will be related to several musical instruments. Phenomena for both the driven and plucked cases will be qualitatively demonstrated with video displays of the signals from optical motion detectors.

1:50

2pMU3. On the acoustics and psychology of piano touch and tone. Anders Askenfelt (Dept. of Speech, Music and Hearing, KTH, Box 70014, S-10044, Stockholm, Sweden), Alexander Galembo, and Lola L. Cuddy (Queen's Univ., Kingston, ON K7L 3N6, Canada)

By way of ''the touch,'' a pianist is able to achieve two degrees of freedom in the control of the key motion. Besides the final hammer velocity, the pianist can control, to some extent independently, the force history between the key and finger. By controlling the acceleration history of the key, while keeping the final hammer velocity constant, the pianist can vary the duration of the key and hammer motion and the forces between the interacting parts of the action. This will influence the intensity of vibrations in the hammer shank and of a noise precursor in the radiated sound. This report addresses the pianist's ability to change the touch parameters and to control them (at will) for musical purposes. The traditional controversy between physicists and some pianists raises a psychological question: ''Why do some pianists think they are able to control a timbre of the individual tone by touch?'' One hypothesis is that variations in touch change the perception of a tone for the pianist, but not for a listener, because for the pianist the tone perceived starts at finger–key contact, while for a listener it starts when the hammer hits the string.

Contributed Papers

2:15

2pMU4. Dynamic mechanical measurements on violin wood and audience evaluations of violin tone quality. Edwin R. Fitzgerald (Johns Hopkins Univ., 3400 N. Charles St., Baltimore, MD 21218), Carleen M. Hutchins, and Morton A. Hutchins (Catgut Acoust. Soc., Montclair, NJ 07042)

At the 125th Acoustical Society of America meeting in Ottawa, Canada, values of complex shear compliance ($J^* = J' - iJ''$), modulus ($G^* = 1/J^*$), and shear wave velocity and attenuation, measured over a continuous frequency range from 10 to 10 000 Hz at 20 °C and 50% RH, were described for samples of spruce and maple wood used in the construction of six violins. Four wood samples were measured for each violin, viz., samples of spruce and maple with and without cross grain. Relation of the violin wood dynamic mechanical properties to plate tuning of the violins also was discussed. Following these presentations, two members of the Ottawa Symphony orchestra each played two short pieces on each of the six violins, and the audience was asked for written comments on the tone quality of each violin. The evaluations display some variations as to the violin(s) with the best tone, but one of the six was rated as good or best tone by none, and the worst tone by many. The dynamic mechanical shear properties of this violin's wood differ in several respects from those of the other violins. Similar property differences for this wood relative to the others also exist in extensional moduli determined from flexural vibrations.

2:30

2pMU5. Vibrational modes of a tenor steel pan. Uwe J. Hansen (Phys. Dept., Indiana State Univ., Terre Haute, IN 47809, phhanse@scifac.indstate.edu) and Thomas D. Rossing (Northern Illinois Univ., DeKalb, IL 60115)

The tenor or lead pan has 28 or 29 notes covering the range of C_4 to E_6. In spite of its being the lead voice in a typical steel band, little research has been reported previously on its modes of vibration or its acoustical behavior. Using electronic TV holography, we have studied the modes of vibration of the various note areas in tenor pans by two different makers. Note areas generally have one or two overtone modes tuned harmonically to the fundamental [Rossing *et al.*, Phys. Today **49**(3), 24–29 (1996)]. Special attention has been paid to the strong acoustical and mechanical coupling between the note areas as they vibrate.

2:45–3:00 Break

3:00

2pMU6. Observation of the brass player's lips in motion. R. Dean Ayers (Dept. of Phys. and Astron., California State Univ., Long Beach, 1250 Bellflower Blvd., Long Beach, CA 90840, rdayers@csulb.edu)

Mouthpieces with flat windows are used to examine the lip reed's motion during the steady parts of notes played on brass instruments. Magnified stroboscopic images show different parts of the upper lip moving

out of phase with each other. Sighting directly along the channel between the lips, one can observe a Rayleigh wave propagating downstream in the flesh of the upper lip. Downward crests of this disturbance reach maximum amplitude at the downstream end of the channel, where they provide most of the valving action on the air flow. The lower lip moves more as a lumped element, and mostly in the longitudinal direction. A computer model for this behavior will be complicated and may not show much improvement over the two-dimensional model of Adachi and Sato [J. Acoust. Soc. Am. **99**, 1200–1209 (1996)]. Of potentially greater value is a more accurate mental image for performers. Muscles controlling the embouchure just establish the boundary conditions for the motion of a thin, passive layer of flesh. This knowledge may help beginners to avoid such practices as jamming the mouthpiece against the lips and tensing muscles against each other unnecessarily.

3:15

2pMU7. Number of degrees of freedom of a lip-reed in brass-instrument playing. Xavier Boutillon (Dept. of Phys., Univ. of Michigan, Ann Arbor, MI 48109-1120)

A lip-reed is a double continuous system that potentially exhibits several degrees of freedom. Numerical simulations for sound synthesis as well as fitting a model with measurements naturally improve while increasing the number of degrees of freedom that are put into the model. However, this parameter is worth investigating as such. In a first approach, lips are coupled to a Helmholtz resonator whose resonance factor is actively reinforced, as described in Chen and Weinreich [''Nature of lip reed,'' J. Acoust. Soc. Am. **99**, 1227–1233 (1996)]. The dimensionality of the attractor or permanent regime of this nonlinear dynamical system is computed with embedding techniques of the corresponding time-series and information-based measurements [Gershenfeld *et al.*, *Time Series Prediction* (Addison-Wesley, Reading, MA, 1994)]. Preliminary results indicate a dimension of 1, 2 or 3 for different couplings between the lips and the resonator. The link with the number of freedom of the lip-reed will be discussed. [Work supported in part by CNRS and NSF.]

3:30

2pMU8. Annular time-domain acoustic source for horn acoustics. Joaquim B. Agullo and Ana Barjau (Dept. of Mech. Eng., Diagonal, 647, 08028 Barcelona, Spain, agullo@em.upc.es)

The time-domain description of horns is needed in time-domain simulations of acoustical systems such as wind instruments and is very helpful to understand their transient behavior. Unidimensional horn wave equations have been used as a first approximation. The Webster horn equation, however, is too unrealistic for horns with intense flare. The improved horn equation developed by Keefe *et al.* [*Proceedings of the SMAC-93* (1993), pp. 496–500] may be promising. A discretized approach that assumes the horn is made of cylindrical and conical elements with plane and spherical waves, respectively, has been widely used. More general alternatives based on the Kirchhoff integral are attractive because they can handle 2 and 3-D acoustic fields and lead to a surface discretization instead of volume discretization (as required in FEM). For axisymmetric horns, the use of an annular time-domain unit source, not implemented in commercial codes, dramatically reduces the number of elements. This paper presents the development of such a source and its application to time-domain horn acoustics. An example is presented for a cathenoidal horn and its results are compared to time-domain measurements.

3:45

2pMU9. Determining resonance frequency changes due to bore irregularities in woodwinds. Cornelis J. Nederveen (Acacialaan 20, 2641 AC Pijnacker, The Netherlands) and Jean-Pierre Dalmont (Université du Maine, 72017 Le Mans Cedex, France)

To determine tuning changes due to deviations from an ideal cylindrical or conical bore constituted by holes and by diameter variations, these deviations are modeled by local changes in compressibility and inertance. The magnitude of a compressibility change, being proportional to the volume change of a perturbation, is easily determined, but inertance changes pose more problems, since they may be less than proportional to the volume change. Determining these inertance changes is possible in various ways. An approximation (with a systematic error) is obtained by depicting the 3-D flow as stratified and integrating 2-D results obtained from Schwarz–Christoffel transformations, for which various accurate analytical and numerical solutions are available. Exact results can be obtained with finite difference methods: they are simple, have no systematic error, and only demand a large computing power. Results will be presented for correction impedances at hole junctions, hole ends with various flanges, and toroidal bends. A method of experimental verification and some results will be discussed.

4:00

2pMU10. Modeling in 3D of directional radiation of curved woodwind instruments. Rene E. Causse and Carole Lheureux (IRCAM, 1 place Igor Stravinsky, 75004 Paris, France)

In a previous presentation [Causse *et al.*, J. Acoust. Soc. Am. **100**(4), Pt. 2, 2812 (1996)], the directional pattern of some instruments of the woodwind family was modeled and compared. These instruments, oboe, clarinet, and flute, were chosen for their straight tube and for the fact that the simplification that the holes of the lattice are aligned and oriented in the same direction did not introduce big errors. A second class of instruments of this family are instruments with curved tubes, for example the bassoon, the bass clarinet, or the alto saxophone. The original model was modified to take into account the curves but also the real orientation of the holes. The far-field radiated pressure was calculated at a constant radius on a sphere enclosing the instrument. The main difficulty with this model is expressing the directivity of each opening (holes and bell) in the same coordinate system. In the cases where this leads to excessive calculation, some simplifications are proposed. Model results were compared with experimental measurements of actual instruments. Comparisons of the directional radiation between these instruments and also with straight ones will be presented. Implications for the perception of these instruments will be discussed.

Session 2pNSa

Noise and Architectural Acoustics: Environmental Noise from Combined Sources

Brigitte Schulte-Fortkamp, Cochair
Carl Von Ossietzky University, Physics-Acoustics, Carl Von Ossietzky Street, D-26111 Oldenburg, Germany

Truls Gjestland, Cochair
SINTEF Telecom and Informatics, N-7034 Trondheim, Norway

Chair's Introduction—12:50

12:55–1:05 Discussion

Invited Papers

1:05

2pNSa1. Reactions to aircraft noise. Truls Gjestland (SINTEF Telecom and Informatics, N-7034 Trondheim, Norway, truls.gjestland@informatics.sintef.no)

The data from three Norwegian surveys on reactions to aircraft noise have been reanalyzed. Data from more than 9000 respondents have been included. The response to a general annoyance question is compared to the response to questions on special effects concerning speech communication, sleep and rest disturbance, ambient noise, behavioral patterns, etc. The established dose–response relationships will form the basis for new Norwegian guidelines for assessing aircraft noise.

1:30

2pNSa2. The assessment of environmental noise—ISO 1996. Ian H. Flindell (Inst. of Sound and Vib. Res., Univ. of Southampton, UK) and Nicole D. Porter (Natl. Phys. Lab., Teddington, UK)

The concept of a grand ''general noise annoyance model'' which takes different noise sources and different situations into account by appropriate adjustments or correction factors is very attractive. On the other hand, it is well known that different people respond to different components of the overall situation (*acoustic features*) at different times, and that this selective attention process can be essentially unpredictable. In order to make progress in this field, it is essential that any new methods of assessment of environmental noise must take this inherent flexibility into account. This paper describes an input-process-outcome model which the authors have been working on as the underlying framework for the current revision of ISO 1996 ''Description and measurement of environmental noise.'' In essence, there are three possible outcomes of assessment (no impact, intermediate impact, and unacceptable impact) and three comparison processes (change in noise, new noise compared against existing background noise, and absolute criteria) which can be applied to the various inputs to determine the assessment outcome. The model allows sufficient flexibility to accommodate most existing national noise assessment procedures, reflects current scientific thinking, and provides a sound basis for choosing the best balance between the costs and benefits of noise control action.

1:55

2pNSa3. Loudness of combined noises derived from singular and concurrent community noises. Birgitta Berglund and Mats E. Nilsson (Inst. of Environ. Medicine, Karolinska Inst. and Dept. of Psych., Stockholm Univ., Stockholm, Sweden)

It seems logical to assume that the total loudness of combined noises should correspond to the arithmetic sum of the loudnesses of the singular noises as perceived within the combined noise. Although postulated by Powell in the mid 1970s, this hypothesis has not yet been empirically tested. In real life situations, a particular community noise is always perceived within a complex of environmental noises, and therefore, both its loudness and the annoyance it evokes will be associated with the noise immission rather than the noise emission from the singular source. Binary combinations of community noises were tested at up to seven levels of loudness and subjects estimated the loudness of combinations as well as the singular sounds with the aid of magnitude estimation. The loudness summation did not follow Powells arithmetic summation hypothesis. The loudness of singular noises differs when heard alone compared to when heard within the noise combination. As a consequence the type of interaction between loudness of community noises in real life situations (heard within) clearly has implications for the development of better models for total loudness as well as for total annoyance. Earlier models hvae relied on data produced from singular noise sources when heard alone.

2:20

2pNSa4. Psychological evaluation of sound environment with mixed sources. Sonoko Kuwano, Jiro Kaku, Tohru Kato, and Seiichiro Namba (Dept. of Environ. Psych., Osaka Univ., 1-16 Machikaneyama, Toyonaka, Osaka, 560 Japan)

In order to make the experimental situations closer to actual situations, experiments have been conducted in field conditions, where various sounds were audible. The results suggested that generally speaking, L_{Aeq} was a good measure for the evaluation of environmental noises. However, there were some cases where L_{Aeq} did not show good correspondence with subjective impressions. On the basis of the results of the field experiments, an experiment was conducted in a laboratory situation with precisely controlled stimuli in order to examine the applicability of L_{Aeq} to sound with mixed sources. Sounds which consisted of aircraft noise, road traffic noise, and train noise were used. Three events of each noise source were included in 15-min stimuli and the duration of road traffic noise was varied in three stimulus conditions. Subjects were asked to judge instantaneous loudness using the method of continuous judgment by category. After instantaneous judgment was over, overall impression of loudness of each sound source as well as the overall loudness of 15-min stimuli were judged. The following analyses was conducted; (1) comparison among three sound sources, (2) comparison among three experimental conditions, and (3) overall evaluation of sound environment with mixed sound sources.

2:45

2pNSa5. Reaction to combined sources of noise may depend on the respondents' interpretation of the questions. R. F. Soames Job (Dept. of Psych., Univ. of Sydney, NSW 2006, Australia, soamesj@psychvax.psych.su.oz.au)

Reaction to multiple sources of noise has proven difficult to explain. The unexpected results of assessment of total reaction to noise may not arise from inadequate integration within the subject nor from complexities of the perception of the noise. This abstract suggests that responses to questions regarding reaction to combined sources may reflect the subjects' interpretation of the relevant questions. Requests for reaction to the total noise load may be interpreted as requesting total reaction to the two sources, or average reaction to the two sources, or total or average reaction to the total noise environment. The consequences of such interpretations by the subjects are not easily predicted. Predictions would be influenced by: whether reaction is a unipolar negative scale versus a bipolar scale from negative to positive; whether the combined reaction is an average or a total; and whether the average or total should include consideration of the times for which each noise is not present. Finally, it should be noted that different subjects may interpret the same question in different ways in relation to these alternatives. Examples of the role of these variables are identified, and a set of models of reaction based on the various alternatives proposed.

3:10

2pNSa6. Combined qualitative and quantitative measurements to evaluate noises from combined sources. Brigitte Schulte-Fortkamp (Dept. of Phys., Oldenburg Univ., D-26111 Oldenburg, Germany, brigitte@aku.physik.uni-oldenburg.de)

Previous studies on assessing combined noises have shown the rating of environmental noises from combined sources as a very complex and difficult task. Disturbance and annoyance caused by the influence of noises require in their complexity an adequate procedure of evaluation. There are different theories to explain those rating results of an overall noise situation which is almost rated lower than the most annoying single source. If sounds are synergistic, the criteria and kinds of sources have to be found which are dominant for the person who assesses the noises. In traditional laboratory studies the context is faded out by the evaluation process, while the context effects in field studies play a decisive role and the evaluation of annoyance caused by noise allows the consideration of different factors in a situation where noises are combined. In a current socio- and psychoacoustical field study the process of assessing the noises and the noises as well are precisely described in interviews and acoustical diaries. In another combined procedure of qualitative-explorative and objective-standardized measurements in a laboratory setting specific noise components causing annoyance and disturbance have been identified. The special methodical procedure and results will be presented.

3:35–3:50 Break

3:50

2pNSa7. Annoyance accumulation modeling in a community noise annoyance expert system. Dick Botteldooren (Dept. of Information Technol., Acoust. Group, Univ. of Gent, St. Pietersnieuwstraat 41, B9000 Gent, Belgium)

When the state of the sound environment in a region is monitored by tracing the percentage of people who are annoyed by noise, simulation models for community noise annoyance are needed. They allow, for example, the prediction of the impact of various noise control measures. Building such simulation tools is not an easy task because information is scarce and relations between quantities involved are uncertain. Therefore, an expert system was built that makes optimal use of the scarce information and combines as many expert opinions as possible. Consensus mechanisms, representation of uncertain data, and model uncertainty are key factors in this

model. Combination of annoyance caused by different sources is an important part of the system. Models proposed by various scientists are combined and compared to obtain the most likely global annoyance percentage. Both psychophysical and perceptual models can be applied at the same time. The probability approximation that is used for this task will be introduced. The results of this approach, which combines different opinions, will be shown for a number of example situations.

Contributed Papers

4:15

2pNSa8. Characterization of an urban area's acoustical comfort. Jacques Beaumont (ENTPE-DCGB/URA CNRS 1652, rue M. Audin, F 69518 Vaulx-en-Velin Cedex, France, jacques.beaumont@entpe.fr), Aline Barlet, Christophe Louwrse, and Catherine Semidor (ERIAC, Domaine de Raba, F 33405 Talence Cedex, France)

In this paper a global method to characterize acoustic comfort in urban areas is presented. This method takes into account the relationships between urban forms (shapes, sizes, and material surfaces), the opinions of town-users, and noise descriptors. These descriptors are deduced from a method based on time automatic detection threshold to classify four types of noise. This system allows the study of sound situation in a self-contained way over a long period (several hours). The influence of building design is estimated according to the visualization of the evolution of these descriptors in 2D and 3D maps all over the studied area. The first results show that noise levels fit well with the architectural characteristics of the buildings (facade irregularities, materials,...). To validate this method, experiments were conducted on a Bordeaux central district. The effects of an urban area's morphology on listeners' noise perception are analyzed from investigations carried out with people living or working in the tested site. The aim of this work is to develop a decision-assisting tool in the acoustic design of urban areas. Environmental considerations, such as noise, must be included in urban planning not only to avoid disturbance, but to provide better satisfaction in living in the city.

4:30

2pNSa9. Sound perception of the urban environment through an audio-visual approach. Stephanie Viollon, Catherine Lavandier (Universite de Cergy, IUT Dept. Genie civil, rue d'Eragny, Neuville sur Oise, 95 031 Cergy Pontoise Cedex, France, viollon@u-cergy.fr), and Carolyn Drake (Laboratoire de Psychologie Experimentale, 75 006 Paris, France)

Sound and visual environments of cities are usually studied separately without taking into consideration possible interactions between the two sensory modalities. The present research aims at developing an audio-visual approach for assessing the sound perception of the urban environment. It means an innovative approach which takes into account the influence of visual information on the sound judgment. In a preliminary *in situ* questionnaire, participants rated real-life urban environments along pairs of descriptive adjectives all related to the auditory modality. A Principle Components Factor Analysis revealed four perceptive factors characterizing the perception of the urban environment and also revealed two auditory variables which were the most influenced by the visual setting. Then, with a simulation experimental setting involving the creation of audio and visual artificial fields, a series of semantic differential tests was carried out. Listeners were required to rate the two auditory variables: appeasing–stressing and pleasant–unpleasant for eight sound urban environments and for four different visual settings. The results point out that judgments about the two sound variables are significantly influenced by the visual setting. Moreover, improvements in the experimental procedure suitable for testing audio–visual interactions in an urban context are discussed.

TUESDAY AFTERNOON, 23 JUNE 1998

CASCADE BALLROOM I, SECTION C (W), 1:00 TO 3:15 P.M.

Session 2pNSb

Noise: Damage Risk Criteria for Impulse Noise

James H. Patterson, Jr., Cochair
U.S. Army Aeromedical Research Laboratory, P.O. Box 620577, Fort Rucker, Alabama 36362-0577

Guido F. Smoorenburg, Cochair
Vasseveldlaan 13A, Soest 3768GK, Holland

Chair's Introduction—1:00

Invited Papers

1:05

2pNSb1. The effects of exposure of intense free-field impulse noise on humans wearing hearing protection: Implications for new criteria. James H. Patterson, Jr. (U.S. Army Aeromedical Res. Lab., P.O. Box 620577, Fort Rucker, AL 36362-0577, jhpinal@snowhill.com) and Daniel L. Johnson (EG&G, MSI, P.O. Box 9100, Albuquerque, NM 87119-9100)

In a series of studies, human volunteers wearing earmuffs were exposed to high-intensity free-field impulse noise which simulated various military weapons noise. By detonating increasing amounts of explosive material, the exposure level was varied from one at which no effects were expected to a maximum level just below the threshold of injury to the lung and upper airway. The number of blasts was varied from 6 to 100. Temporary threshold shifts (TTS) measured 2 to 6 min after each exposure were used as the measure of effects on hearing. The highest level for each number of blasts at which 95% of the exposed population would not show a

significant TTS was used to establish the maximum safe exposure levels as measured in the open and under the earmuffs. These maximum safe exposure levels exceed the exposure limits used in the United States and other countries. Also, these results suggest that the number-intensity trading rule is more shallow than that used in most criteria. The exposures measured under the earmuff provide a good indication of the hazard. Implications of these results for impulse noise exposure criteria will be discussed.

1:30

2pNSb2. Hearing protectors and hazard from impulse noise: Melding method and models. G. Richard Price and Joel T. Kalb (US Army Res. Lab., APG, MD 21005-5425, dprice@arl.mil)

Studies agree that the protected ear tolerates surprising amounts of energy at very high levels; but with no theory to explain these results, practical application of such findings is like traversing an acoustic minefield. Our modeling of auditory hazard [G. R. Price and J. T. Kalb, J. Acoust. Soc. Am. **100**, 2674 (1996)] suggests the physical basis for the ear's resilience is a function of interaction between the hearing protector's effect on the waveform arriving at the ear and the ear's own complex responses to intense stimulation. We believe that the primary physical basis for hazard is mechanical stress at the level of the hair cell; hence assessment needs to take place in the time domain. Calculation begins either with the pressure history measured under a hearing protector or one derived from a free-field pressure and attenuation measures on the protector. The ear model includes an active middle ear muscle system, an amplitude-limited stapes displacement, and an algorithm for calculating hazard in the cochlea which integrates peak upward stresses at the level of the basilar membrane. In the end, the protected ear's apparent resilience is predictable and can be understood as part of a general theory of auditory hazard.

1:55

2pNSb3. Results of human studies with linear and nonlinear earplugs: Implications for exposure limits. Armand L. Dancer and Pascal J-F. Hamery (French-German Res. Inst. of Saint-Louis, 5 rue du General Cassagnou, BP 34, 68301 Saint-Louis Cedex, France, dancer@nucleus.fr)

Soldiers equipped with linear and nonlinear earplugs have been exposed to large impulses (weapon noises) and audiometric tests have been performed just before and after the exposures. In particular, new nonlinear perforated earplugs (allowing speech communication, detection and localization of acoustic sources) which have been developed at the French–German Research Institute of Saint-Louis are efficient for repeated exposures up to 187-dB peak when they are well fitted [Johnson and Patterson (1996) personal communication]. From various TTS studies it is concluded that: (i) the peak pressure attenuation (measured with a specially designed ATF) grossly underestimates the protection afforded by the earplugs when used in conjunction with the classical criteria for weapon noises (CHABA, MIL-STD, criterion of Pfander, criterion of Smmorenburg). The risk corresponding to an exposure to a slow rise-time impulse (i.e., under an earplug) is much lower than the risk corresponding to an exposure to an impulse of the same peak pressure with an instantaneous rise time, (ii) the L_{Aeq8} attenuation values based on objective insertion loss measurements (ATF) give in most exposure conditions a good evaluation of the auditory hazard, but in some other cases they still underestimate the efficiency of the hearing protectors.

2:20–2:25 Chair's Summary

2:25–2:35 Break

Contributed Papers

2:35

2pNSb4. Modeling the effect of a hearing protector on the waveform of intense impulses. G. Richard Price and Joel T. Kalb (US Army Res. Lab., APG, MD 21005-5425, jkalb@arl.mil)

Implicit in microphone-in-real-ear (MIRE) or real-ear-attenuation-at-threshold (REAT) tests with a hearing protector is the potentiality that such tests would establish its ability to ameliorate hazard. However, because the sound field under the protector is very different from the free field for which traditional DRCs customarily predict hazard, it is not clear how such measures should be used. The theoretically based mathematical model of auditory hazard we proposed by us as the basis for such assessments calculates motions of structures in the ear [Price and Kalb, J. Acoust. Soc. Am. **100**, 2674 (1996)]. Thus phase relationships in waveforms matter and the model must accurately couple them to the ear. Where only REAT data and free-field pressure measures are available, single path conduction through the protector is assumed and minimum-phase processing to calculate waveforms under protectors is used. Empirical tests with free-field impulses and an artificial head test fixture indicate that such an approximation is surprisingly good. With such calculated waveforms or MIRE measures, a Thevinin-equivalent source to drive the ear which allows the full calculation of auditory hazard with the model is created.

2:55

2pNSb5. A new nonlinear earplug for use in high-level impulse noise environment. Armand L. Dancer and Pascal J-F. Hamery (French-German Res. Inst. of Saint-Louis, 5 rue du Gener. Cassagnou, BP 34, 68301 Saint-Louis Cedex, France, dancer@nucleus.fr)

A nonlinear earplug allows speech communication, detection and localization of acoustic sources, and prevents hearing hazard from high-peak pressure impulses. Improved physical characteristics of a nonlinear earplug are explained: nonlinear attenuation starting with relatively small level (110 dB peak) and improvement of nonlinear attenuation for high levels. Such an earplug is especially designed against weapon noises (up to 190 dB peak). The study is based on experimental approach. Moreover, a theoretical model has been developed to explain the experimental results and to optimize the acoustic nonlinear filter dimensions and configuration. The nonlinearity of such a filter is due to one (or several) little hole(s), the acoustic impedance of which is essentially a viscous resistive one and depends on the particulate velocity in its center.

Session 2pPAa

Physical Acoustics: Acoustic Characterization and Manipulation of Material Properties

Sameer Madanshetty, Cochair
Mechanical Engineering Department, Kansas State University, Rathbone Hall, Room 339, Manhattan, Kansas 66506-5106

D. Keith Wilson, Cochair
U. S. Army Research Laboratory, AMSRL-IS-EE, 2800 Powder Mill Road, Adelphi, Maryland 20783-1197

Chair's Introduction—2:15

Contributed Papers

2:20

2pPAa1. Directional microphone array for measurement in an anechoic wind tunnel. Robin J. Alfredson and Alison K. Wilson (Monash Univ., Clayton, Vic, Australia)

An anechoic wind tunnel in which full-scale acoustic measurements may be conducted on automobiles at speeds up to 200 km/h has been constructed. The wind tunnel uses an open jet which is collected and returned in the closed-loop wind tunnel. This design allows for microphones to be positioned outside of the air flow typically some 4 to 5 m from the vehicle. Thus problems associated with extraneous wind noise on the microphones are largely eliminated but this has resulted in the need to develop microphone arrays which have directional characteristics in order to identify noise sources on the vehicle. A wide range of microphone configurations was examined recognizing the various constraints imposed by cost, data collection facilities, frequency range, and the geometry of the wind tunnel and vehicle. The design that produced an optimum solution was based on an arced array of 10 microphones. The output from each was summed having first been weighted. The array could be moved along the vehicle. Data will be presented indicating the characteristics of this and other arrays as well as the cost of construction.

2:35

2pPAa2. Phenomenological theory of the translational relaxation times in gases. Allan J. Zuckerwar (NASA Langley Res. Ctr., MS 236, Hampton, VA 23681-0001)

The classical equations governing the translational dispersion and absorption of sound in a gas yield a solution of such complexity as to obscure the relaxational character of the associated viscous and thermal transport processes. The approach taken here is to solve the secular equation by the method of Pade' approximation, which even to the relatively low order R_{11} yields a remarkably close approximation to the exact solution over a wide range of frequency/pressure ratios. As a result, translational relaxation can be ascribed the attributes of a conventional relaxation process such as relaxation times, relaxation strength, collision numbers, additivity relations, etc. To extend the theory to extremely high values of frequency/pressure ratio, a model is proposed to account for the noncontinuum behavior of the transport coefficients (viscosity and thermal conductivity) as the molecular mean free path approaches the acoustical half-wavelength. The theory shows good agreement with measurements in argon over the classical, Burnett, and transition regions as defined by Meyer and Sessler [E. Meyer and G. Sessler, Z. Phys. **149**, 15–39 (1957)].

2:50

2pPAa3. An automated apparatus for measuring sound speeds in hazardous gases. John J. Hurly (Natl. Inst. of Standards and Technol., Bldg. 221, Rm. A111, Gaithersburg, MD 20899)

Our laboratory has developed an automated apparatus for measuring the sound speed in hazardous gases such as those used in processing semiconductors. The sound speed is deduced from measurements of the acoustic resonance frequencies of a gas-filled, cylindrical cavity. After the cavity is calibrated with argon, the uncertainty in the sound speed is less than 0.01%. The $u(P,T)$ data span the temperature and pressure ranges 200–475 K and 100–1500 kPa. From the measured sound speeds we derive the ideal gas heat capacities from the zero-pressure intercepts with a typical uncertainty of only 0.1%. The virial equation of state is also deduced from $u(P,T)$ providing the temperature-dependent density virial coefficients. In comparisons with Burnett measurements, we find that the densities calculated with the derived virial equation of state have uncertainties of less than 0.1% in density. Some recent measurements in SF6, CF6, and C2F6 will be used to illustrate the systems capabilities. Work in progress on BCl3 will also be discussed.

3:05

2pPAa4. Ultrasonic flowmeter for the accurate metering of gases for semiconductor processing. Adam M. Calabrese (Natl. Inst. of Standards and Technol., Bldg. 220, Rm. A55, Gaithersburg, MD 20899, adam.calabrese@nist.gov) and Stuart A. Tison (Natl. Inst. of Standards and Technol., Gaithersburg, MD 20899)

Semiconductor production processes such as chemical vapor deposition rely upon accurate control of the flow rate (in the range of 1–1000 standard cm^3/min) of several ultrapure, yet highly toxic and/or reactive gases. Currently, thermal mass flow meters (TMFMs) are widely used in industry for these processes. Complications with the use of TMFMs arise because the flow must be divided (due to limited dynamic ranges) and calibration is normally performed with inert gases. In the present research, an ultrasonic Doppler shift flowmeter has been developed to measure gas flows. Measurements of the flow-induced phase shift of an acoustic plane wave propagating in an unobstructed tube, combined with the speed of sound of the gas, are used to determine the mean gas velocity. The combination of the mean gas velocity with measurements of the pressure and temperature allow calculation of the mass flow rate. The experimental system uses stainless steel tubes of diameter 0.05–0.1 cm, and frequencies near 100 kHz. No division of the flow is necessary. Prototype meters indicate that sensitivities on the order of 0.75°/standard cm^3/min are possible. Experimental data of two prototype systems using different geometries and receivers will be presented. [Work supported by NIST.]

3:20

2pPAa5. Ultrasonic methods for the on-line real-time characterization of mixing in multiphase systems. Leonard J. Bond and Hans Matthiesen (Univ. of Denver Res. Inst. and Dept. of Eng., 2050 E. Iliff Ave., Denver, CO 80208)

There is currently an increasing interest in developing ultrasonics to provide data that can be used for both process monitoring and control. These developments are being driven by the need to better understand the fluid dynamics and multiphase media interaction phenomena encountered. This is needed to give insights into the process and quantitative data for use in models. This paper will report results obtained through the analysis of dynamic B-scan and digital ultrasonic A-scan data, which are being used for the characterization of mixing in multiphase systems. Results obtained with a model system of sugar mixing into water, under a variety of conditions, will be presented. The insights that are given into the mixing process through the use of a B-scan system will be shown on video. The processes being used to extract data from (I) B-scan images and (II) rf (time-amplitude traces) measured with A-scan systems, to enable the implementation of a mixing index will be reported. Particular attention will focus on the issues of measurement scale and characterization of inhomogeneity. Data will be compared with theoretical model results for ultrasonic properties of homogeneous systems, in particular for suspensions, colloidal systems, and reactions including phase changes.

3:35

2pPAa6. Water jets and a self-focusing shock pulse source. Bruce J. P. Mortimer (School of Elec. Eng., Cape Technikon, P.O. Box 652, Cape Town 8000, South Africa) and Beric W. Skews (School of Mech. Eng., WITS Univ., South Africa)

The interaction of a liquid pressure pulse with a liquid–air boundary can produce surface deformation and water jet eruption if the incident pressures are high enough. This is important when analyzing damage mechanisms in lithotripsy and in our efforts to produce pulsed water jets [Mortimer and Skews, J. Acoust. Soc. Am. **100**, 3548–3553 (1996)]. High-pressure pulses can be generated by either using focusing lens refraction and convergent reflectors, or using self-focusing sources. In this work experimental results using an electromagnetic shock wave source (EMAS) and a new self-focusing source (FEMAS) as discussed. A high-speed digital camera was used to study the water jet production. It was found to consist of two regions; first a high-speed initial jet and second a slower cumulative jet due to the hydrodynamic flow initiated by the source. Results for the FEMAS show a contained cumulative jet as opposed to an unstable jet observed with the EMAS. These results indicate that the device concentrates its on-axis pressure and velocity, a result confirmed by our pressure measurements. However, the operation of the FEMAS was also found to be rather complex in behavior. [Work supported by the Foundation for Research Development.]

3:50–4:05 Break

4:05

2pPAa7. Measuring longer tubular objects using acoustic pulse reflectometry. David B. Sharp and D. Murray Campbell (Dept. of Phys., JCMB, King's Bldg., Univ. of Edinburgh, Edinburgh, EH9 3JZ, UK, d.b.sharp@ed.ac.uk)

Acoustic pulse reflectometry is a noninvasive technique for measuring the internal dimensions of tubular objects. A sound pulse is produced using a loudspeaker and injected into the object under investigation via a source tube. The resultant object reflections are recorded by a microphone embedded in the wall of the source tube. Analysis of the reflections yields information about the bore profile of the object. At present, there is a restriction on the length of object that can be measured using the standard reflectometer. This restriction arises because after the object reflections pass the microphone they undergo further reflection at the loudspeaker. These source reflections return to the microphone as unwanted signal. Hence, there is only a finite time over which the object reflections can be accurately recorded which, in turn, limits the length of object that can be measured. In this paper, a method for removing the length restriction is described and preliminary results are presented. The method involves the cancellation of the incoming object reflections at the loudspeaker, thus preventing the formation of source reflections. [Work supported by EPSRC.]

4:20

2pPAa8. Temperature-scanning ultrasonic measurements in aqueous systems: What is new? Dmitri P. Kharakoz (Inst. of Theoretical and Exp. Biophys., RAS, 142292, Pushchino Moscow reg., Russia)

Ultrasonic measurements are the only high-precision approach to experimental studies of the mechanical part of free energy in liquids, the part directly related to the volume, pressure, and compressibility. During the last decade, a significant progress has been made in the studies of these parameters in aqueous solutions and dispersions of biomolecular materials. The following aspects are considered in this paper: (1) temperature dependence of volume and compressibility as a means to characterize hydration of solutes in water. Important information on the structure and dynamics of bound water can be obtained. (2) Single-frequency simultaneous measurements of sound velocity and absorption and the investigations of fast kinetics. A first-order phase-transition in lipid vesicles (aqueous dispersion) is considered as a process for which unique and very important kinetic information can be obtained by means of such measurements. (3) Sound velocity, compressibility, and volume as a means to characterize conformational states of macromolecules in aqueous solution. Temperature-induced phase transitions in macromolecules (proteins and other polymers). (4) A puzzle of the intrinsic compressibility and nonlinearity of native protein molecules. (5) The acoustic and calorimetric temperature-scanning measurements: A comparative aspect. [Work was supported in part by Russian Foundation for Basic Research, Grant No. 95-04-12226a.]

4:35

2pPAa9. Acoustic monitoring of the curing process in cement and concrete. Bernhard R. Tittmann, Hasson M. Tavossi (Penn State Univ., Dept. of Eng. Sci. and Mech.), and Frederic Cohen-Tenoudji (Univ. of Paris-7, Denis Diderot, Paris, France)

Acoustic wave velocity and attenuation measurements are used to monitor cement and concrete during hardening. The measured acoustic parameters can then be used to investigate the curing process. The goal of this work is to determine the time required for a given mixture of concrete to be hard enough for the removal of the molds, in order to speed up the construction process. Currently this removal time is unpredictable. In addition concrete, mortar, and cement, before hardening, can be considered as examples of granular materials. The shear, rigidity moduli, and acoustic wave velocity of these materials can differ from the values obtained by applying effective-medium theories and Biots theory. The complete description of the mechanical properties of these materials requires the consideration of discrete nature of the solid constituents and the effects of contacts between the grains. These effects are included in the theoretical description of the process. The methods of acoustic microscopy are also applied to investigate the surface and subsurface physical changes of cement during hardening. Theoretical and experimental results of this investigation will be presented.

4:50

2pPAa10. Model experiments on the evaluation of the cement bond with ultrasonic pulse-echo techniques. Wenxiao Qiao, Guangsheng Du, and Shihong Chi (Dept. of Petroleum Resource, Univ. of Petroleum, East China, Dongying, Shandong 257062, PROC)

Previous research shows that when the transducer is immersed in slurry, it is nearly impossible to determine the cement bond quality of the interface between the cement annulus and formation (Interface II) using the ultrasonic pulse-echo technique. Because strong reflection occurs at the interface between the slurry and casing, most of the portion of the pulse energy is reflected back. In this paper, methods of cement bond evaluation are studied in four different model experiments, in which the transducer is contacted closely to the casing. In these experiments, no matter whether the ultrasonic pulses are normally or obliquely incident to the casing, obvious signals corresponding to the Interface II are observed. Considering the practical implementation of the technique, a thin steel plate is inserted between the transducer and casing. The reflection signal can still be achieved. The model experiments indicate that the transducer-contact-casing system can eliminate the influence of the slurry, improve the ratio of signal-to-noise, evaluate the interface between the casing and cement (Interface I), and determine the cement bond quality of the Interface II if the Interface I is well bonded.

5:05

2pPAa11. Acoustic measurements of condensation/evaporation and crystal growth. Andi G. Petculescu and Larry A. Wilen (Dept. of Phys. and Astron., Ohio Univ., Athens, OH 45701, wilen@helios.phy.ohiou.edu)

Using a volume-modulation technique developed to study thermoacoustics, the complex compressibility of a liquid in coexistence with its vapor is measured. Such a system exhibits a distinctive behavior. At low frequencies, the inverse complex compressibility vanishes because the pressure remains at the equilibrium vapor pressure as the total volume is changed slowly. The volume change is accompanied by an exchange of molecules between the two phases. At higher frequencies, however, this exchange is too slow, and the pressure response approaches that which would be expected in the absence of the liquid phase. The presence of the liquid acts effectively as a high-pass filter whose time constant depends on the rate of condensation/evaporation. The application of this technique to the measurement of condensation coefficients of liquids will be discussed. The same technique can also be applied to solid/vapor systems where the results provide a measure of the crystal growth rate. A unique feature of this technique is that the crystal growth rate can be monitored continuously as a function of temperature. [Work supported by Ohio University Research and Sponsored Programs.]

5:20

2pPAa12. Anomalous excess noise in inhomogeneous elastic and piezoelectrical solids. Alexander M. Dykhne (Federal Sci. Ctr. ''Trinity,'' Troitsk, Moscow oblast', 142092, Russia)

The description of the macroscopic behavior of a microinhomogeneous material and the determination of its internal structure may be obtained from its response to external action. The parameters related to the second moments of the field are the simplest to determine. This is the work of deformation. Additional information about the material can be deduced from the higher moments of the respective field. These higher moments are most important in the case of a field with a highly nonuniform spatial dependence. For certain types of the heterogeneity geometry, some points of the elastic field concentration can exist where the field strength increases faster than a certain power function of the distance. The fourth moments of the field can diverge, which will lead to abnormally high excessive noise. The measurement of the excessive noise gives information about the structure of the material. A simple method for determining the high distribution moments of a field is based on measurements of the excessive noise arising from the fluctuations in material constants such as elasticity moduli. This is related to, for example, the deformation of a solid under a fixed external load or the eigenfrequencies of elastic vibrations. In the case of piezoelectrical materials, the excess noise may originate from both elastic and electrical external fields. Their measurement gives more information about the internal structure of the material.

5:35

2pPAa13. Acoustic pulse diagnostics of relaxation media. German A. Maksimov and Vladimir A. Larichev (Moscow Eng. Phys. Inst., Kashirskoe sh. 31, Moscow, 115409, Russia, maximov@dpt39.mephi.ru)

The investigation of propagation of acoustic fields in different media is an important source of information about nonideal properties of these media. For study of small-amplitude acoustic waves, the most essential of the nonideal properties are the dissipative–dispersive properties, which are characterized by the frequency dependence of the absorption factor and the phase velocity. These properties can be described in terms of relaxation processes, which arise in the field of acoustic waves. Conventional ways to determine the relaxation times spectrum (RTS) parameters are acoustic spectroscopic measurements of temperature–frequency dependence of absorption and phase velocity of sound in the medium. Currently, such problems can be solved by applying the pulse methods of measurements. In the report the possibilities of acoustic pulse diagnostics of relaxation media are discussed. Parameters of media (such as five moments of RTS) are defined which can be determined experimentally by methods of pulse acoustic diagnostics. The criteria for a pulse excitation source and its registration distances are obtained which provide the most efficient pulse acoustic diagnostics. The comparison of developed approach with the first experiments is fulfilled.

Session 2pPAb

Physical Acoustics: Half-Spaces and Plates

Paul E. Barbone, Chair
Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Chair's Introduction—2:15

Contributed Papers

2:20

2pPAb1. Optical detection of leaky-Rayleigh waves at air–solid interfaces. Laszlo Adler, Christophe Mattei (Adler Consultants, Inc., 1275 Kinnear Rd., Columbus, OH 43212), Michel de Billy, and Gerard Quentin (Univ. Paris 6 and 7, CNRS U.R.A. 17, Paris Cedex 05, France)

Due to the difficulties associated with ultrasonic generation and detection in air in the megahertz region, air-coupled nondestructive evaluation has only recently been explored. In this work, a shear transducer is mounted directly over the edge of a solid specimen so that it can generate a bulk mode in the interior of the solid as well as a Rayleigh wave along the surface. A recently developed optical heterodyne interferometer proved to be sensitive to the leaky-Rayleigh wave in air more than a centimeter away from the solid surface. Results of velocity and attenuation measurements of leaky-Rayleigh waves for several solids will be presented. The technique can also be applied to plates to measure leaky-Lamb velocity and attenuation in air. This method can be used for low optical reflectivity solid surfaces: Thus it can be extended for noncontact nondestructive evaluation applications.

2:35

2pPAb2. Theoretical and experimental study of surface waves on solid–fluid interfaces when the fluid sound velocity is larger than the shear wave velocity in the solid. F. Padilla, M. de Billy, and G. Quentin (Groupe de Physique des Solides, Tour 23, Univ. Paris 7, Denis Diderot, 2 place Jussieu, 75251 Paris Cedex 05, France, Unit assoc. au CNRS n 17, padilla@gps.jussieu.fr)

Two surface waves can propagate on a solid–fluid interface: the Rayleigh and the Stoneley waves. The problem is well known when the fluid sound velocity cF is lower than the velocities in the solid (cS and cL). The Scholte–Stoneley dispersion equation has been studied when cF is greater than the shear wave velocity cS. The numerical results obtained for PVC–water and Plexiglas®–water show that there might exist two surface waves whose velocities are lower than cS. The two corresponding roots of the dispersion equation tend to the Rayleigh root of the solid–vacuum interface as the density of the fluid tends to zero. The structure of the two waves has been studied using the formalism of the evanescent plane waves. They are theoretically propagating without attenuation along the interface and exponentially decaying along the direction normal to the interface inside both media. These two waves have been experimentally observed. They exponentially decay along the direction normal to the interface inside both media. The experimental arrival times are in very good agreement with the theoretical ones.

2:50

2pPAb3. A modal reduction technique for the finite element formulation of Biot's poroelasticity equations in acoustics applied to multilayered structures. Franck C. Sgard (LASH, DGCB URA CNRS 1652, ENTPE-Rue Maurice Audin, 69518 Vaulx-en-Velin Cedex, France, Franck.Sgard@entpe.fr), Noureddine Atalla, and Raymond Panneton (GAUS, Univ. de Sherbrooke, QC J1K 2R1, Canada)

Lately, the authors have presented a general technique, referred to as *selective modal analysis* allowing for a reduction of setup time, computer storage, and solution time of large finite element models involving poroelastic structures. This method uses a dual basis associated with the skeleton *in vacuo* and the fluid phase in the rigid skeleton limit, respectively. A compact system in terms of modal unknowns is obtained by projecting the coupled system over the dual basis. This technique has been previously tested in the case of the determination of surface impedances and quadratic velocities of single porous materials excited both mechanically and acoustically. It is proposed in this talk to investigate the efficiency of the *selective modal analysis* technique in the case of multilayered structures made up from a combination of elastic, poroelastic, and acoustic media under various boundary and loading conditions. The theory behind the technique will be recalled together with numerical examples illustrating its performance.

3:05

2pPAb4. Sound propagation in gas-filled rigid-framed porous media: General theory and new experimental and numerical data. Denis Lafarge, Michel Henry (LAUM, UMR CNRS 6613, 72085 Le Mans, France), Mouaouia Firdaouss, and Jean-Luc Guermond (LIMSI, UPR CNRS 3251, BP 133, 91403 Orsay, France)

The acoustic properties of gas-filled porous media have been the subject of extensive theoretical and experimental work. However, even in the simplest case of rigid-framed materials, most of the theoretical developments found in the acoustical literature remain largely empiric or semi-phenomenologic. This is due to the complicated microstructure of porous materials in general. Starting with first principles and using a simple two-scale analysis, a general formulation of the problem of (linear) long-wavelength sound propagation in gas-filled media is provided in terms of two independent permeabilities. One of these—the viscous dynamic permeability—is well known. Its thermal analog, the thermal dynamic permeability, is new in the acoustical context. It can be related to a notion of ''mean survival time'' in diffusion-controlled reactions. The exact connection between the microgeometry and the frequency-dependent dynamic permeabilities is obtained in terms of independent geometrical distribution functions. Useful analogies with the dispersion of electric and magnetic permeabilities are noted. New experimental and numerical data demonstrate the usefulness of the scaling functions proposed by Wilson *et al.* and Pride *et al.* Different limitations of the general modeling and the mentioned scaling functions are also noted.

3:20

2pPAb5. Nonlinear propagation of laser-generated acoustical pulses in granulated media. Konstantin Naugolnykh (CIRES, Univ. of Colorado/Environ. Technol. Lab., Boulder, CO 80303), Igor Esipov, Konstantin Matveev (N. Andreev Acoust. Inst., Moscow 117 036, Russia), and Viktor Zosimov (Appl. Acoust. Inst., Dubna 141 980, Russia)

Laser generation of high-intensity broad frequency band acoustical pulses in different fluids and their propagation in water-saturated cobalt-manganese crust (CMC) samples are investigated. Acoustic pulses of amplitude up to 3 MPa and duration about 1 s are generated in water, oil, and benzene by the action of CO_2 on a laser pulse. The propagation of the short acoustic pulses through the water-saturated CMC samples is considered. The complexity of the CMC structure leads to specific nonlinear acoustical features of the process—nonlinear distortion of the acoustic pulse and amplitude-dependent attenuation. Abnormal sound attenuation in the high-frequency band is observed with low-energy probing sound pulses. This effect could be connected with the sound energy capture and its localization as oscillating modes of the structure. This effect vanishes with increasing intensity of the probing pulses, apparently due to the changes in crust structure. At the same time, specific nonlinear distortion of the sound pulse takes place—increasing of the acoustic pulse amplitude leads to decreasing of the negative phase of the passed acoustic pulse. [Work supported in part by Russian Fund for Fundamental Research.]

3:35

2pPAb6. High-resolution spectroscopy of the Lamb modes of a plate near normal incidence. G. Durinck, W. Thys (Katholieke Universiteit Leuven Campus Kortrijk, Interdisciplinair Res. Center, 8500 Kortrijk, Belgium), J. L. Izbicky, and P. Rembert (Universite du Havre, 76610 Le Havre, France)

Using a high-resolution digital oscilloscope and advanced datahandling plate modes near normal incidence, other than those identified by Lamb, were studied. This search was led by the complex roots of the dispersion relation for a plate in water. Solutions with a complex wave number component in the propagation direction indeed reveal near normal incidence the existence of plate modes no one has ever observed to our knowledge. In this paper experimental results are presented at the cutoff frequencies of the S1 and S2 modes of a 1.5-mm-thick aluminum plate, and at the cutoff frequencies of the A1 and A2 modes of a 1.5-mm-thick glass plate. The extra resonance predicted by the theory is resolved by looking at the time history of the radiating plate. This is done either by displacing the time window far into the free emission regime or by displacing the receiving transducer in the propagation direction.

3:50

2pPAb7. Electric-field influence on Lamb and SH wave properties in $LiNBO_3$ plates. B. D. Zaitsev, S. G. Joshi, and I. E. Kuznetsova (Inst. of Radio Eng. and Electron. of RAS, Saratov Branch, Zelyonaya str., 38 410019 Saratov, Russia)

Presently, acoustic devices using an electrically controlled time delay are highly promising. They include acoustic phase shifters, amplitude and phase modulators, and transversal filters with electrically controlled amplitude-frequency characteristics. Lithium niobate is one of the most promising materials with the anisotropic nonlinear electroacoustic interaction studied for bulk (BAW) and surface (SAW) acoustic waves. This paper is devoted to the theoretical investigation of electrically controlled characteristics of acoustic waves in plates of lithium niobate. The main crystallographic cuts and various thicknesses of plate under the action of different orientations of an external electric field have been investigated. It was confirmed that the dependence of velocity on an electric-field intensity has a complex nature as well as for SAW and BAW. Therefore, even and odd portions of analyzed dependencies have been introduced and studied separately. It was found that mentioned portions are functions not only of wave propagation direction and electric-field orientation but of plate thickness. It allows the changing of the nature of velocity dependence on electric-field intensity from a quadratic law to a linear one. Another advantage is the possibility of providing a significant electric-field intensity by using low voltage due to thin thickness of plate.

4:05

2pPAb8. Effects of sound and vibrations on tensions and deformations of a stratificated plate. Florian-Paul I. Simion (Dept. of Theoret. Mech., Univ. Politehnica of Bucharest, 313 Splaiul Independentei, 77206, Bucharest) and Gabriela-Cristina D. Simion (Univ. Politehnica of Bucharest, 313 Splaiul Independentei, 77206, Bucharest, simion@freddy.cib.pub.ro)

Effects due to sound and vibrations phenomena in bodies are multiple and more complex than those in the case of sound propagation in air. In the present paper deformations and tensions are analyzed in the case of a plate with eight layers and a solicitation in flexion. Our plate is subjected to a uniform pressure. Let us consider a Cartesian frame $(O, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ for our multilayer plate from composite material. Axis $\mathbf{x}_3$ is normal to the medium surface of the plate. This plate is composed of eight unidirectional layers and the angles between fibers of a layer and axis $(O|\mathbf{x}_1)$ have the following values: 0, $\pi/2$, 0, $\pi/2$, $\pi/2$, 0, $\pi/2$, 0. Matrix of rigidities for a layer is determined in a rectangular base $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ in case $\mathbf{x}_1$ parallel to the direction of fibers. A formula of medium deformated surface, arrow in the middle of plate and flexion deformation is determined. Particular application efforts in each layer are computed and their graphs are plotted.

Session 2pPP

Psychological and Physiological Acoustics: Aspects of Hearing Development

Lynne A. Werner, Chair
Speech and Hearing, University of Washington, 1417 NE 42nd Street, Seattle, Washington 98105-6246

Chair's Introduction—1:00

Invited Papers

1:05

2pPP1. Development of active and passive processes in gerbil cochlea. David M. Mills and Edwin W. Rubel (Dept. of Otolaryngol., Head & Neck Surgery, Univ. of Washington, Seattle, WA 98195)

Distortion product otoacoustic emissions were monitored in neonatal gerbils during furosemide intoxication, to examine the ontogeny of different mechanisms underlying cochlear responses to acoustic stimuli. Results indicate: (1) Above 2 kHz, active processes were functional at adult levels at all ages. (2) In the youngest animals (14 days) the maximum frequency seen for the active response was 16 kHz, while in adults it was 40 kHz. This limitation on the active frequency response is believed to be due to the underlying passive cochlear mechanical response, and is consistent with the known place code shift. (3) The length of the active amplification zone at the extreme base decreased from 1.2 mm at 15–19 days to 0.6 mm at maturity. (4) At midfrequencies, there was an increase in net active gain from 15 to 19 days, then a subsequent decrease to reach a uniform 25-dB gain across frequency by maturity. Overall, these results are consistent with an increase in the rate of gain (dB/mm) from 15 to 19 days as the endocochlear potential increases to mature levels, followed by a decrease in the length of the active amplification zone throughout the cochlea with maturity. [Work supported by NIH DC00395.]

1:35

2pPP2. Development of human cochlear function. Carolina Abdala (House Ear Inst., 2100 W. 3rd St., Los Angeles, CA 90057)

The time course for functional development of the human cochlea is not clear. The advent of otoacoustic emissions (OAE) methodology has made it possible to noninvasively probe cochlear function in humans. Distortion product OAE isosuppression tuning curves indirectly reflect cochlear frequency selectivity. They are generated by presenting a third suppressor tone simultaneously with the stimulating tones f1 and f2 and increasing the level of the suppressor until a criterion amount of suppression is achieved. When the suppressor level required to achieve criterion reduction in DPOAE amplitude is plotted as a function of frequency, a DPOAE suppression tuning curve (STC) can be generated. DPOAE (2f1−f2) STCs have been generated in adults and neonates to investigate development of cochlear tuning in humans [Abdala *et al.* (1996); Abdala and Sininger (1996)]. Results of these studies suggest that the human cochlea is resolving frequency in an adultlike fashion by term birth. Other findings also support this contention [Bargones and Burns (1988); Eggermont *et al.* (1996)]. However, results of a recent study conducted with very premature neonates [Abdala (1997)] suggest that STC width and growth of suppression is not adultlike in premature neonates. These results may be indicative of cochelear immaturity just prior to term birth.

2:05

2pPP3. Evidence of prenatal hormonal effects on the auditory system. Dennis McFadden (Dept. of Psych. and Inst. for Neurosci., Mezes Hall 330, Univ. of Texas, Austin, TX 78712, mcfadden@psy.utexas.edu)

There are marked sex differences in a number of behavioral and physiological measures of the auditory system. Evidence is accumulating that some of these sex differences exist at birth and are attributable to the magnitude of prenatal exposure to androgenic hormones, such as testosterone. A generalization emerging from this evidence is that greater prenatal androgen exposures lead to greater masculinization of these auditory responses. At this point, little is known about the modes of action of androgens on developing cochelear structures, but because the degree of masculinization of otoacoustic emissions (OAEs) appears to parallel the degree of masculinization of other brain structures and behaviors, OAEs appear to have the potential to be a valuable, noninvasive window onto the processes of prenatal development and sexual differentiation. Examples of this potential will be discussed. [Work supported by NIDCD.]

2:35–2:45 Break

2:45

2pPP4. Human auditory system maturation: A neurophysiological comparison between normal-hearing children and children who use a cochlear implant. Curtis W. Ponton, Manuel Don, Betty Kwong, Michael D. Waring (House Ear Inst., 2100 West Third St., Los Angeles, CA 90057), and Jos J. Eggermont (Univ. of Calgary, Calgary, AB T2N 1N4, Canada)

Auditory-evoked responses provide detailed, objective measures of maturational change in the central auditory system from the level of the brainstem to the cortex. Auditory-evoked responses can reflect activity originating from three pathways: the lemniscal and nonlemniscal pathways as well as from a modality nonspecific pathway originating in the reticular activating system (RAS) and its associated thalamic nuclei. Maturational time courses for these pathways were derived from evoked response data recorded from 156 normal-hearing subjects ranging from 5–20 years of age. Analyses of these data indicate that each pathway may have a unique developmental sequence. For normal-hearing children, the lemniscal pathway appears to follow a longer developmental time course than either the nonlemniscal or RAS pathways. For children with implants, the developmental sequences for these pathways are differentially affected depending on onset and duration of deafness. Specifically, maturation of the lemniscal pathway is delayed by an amount that approximates the period of deafness, whereas maturation of the RAS pathway may be related to the age at onset of deafness.

3:15

2pPP5. Psychoacoustic development in humans and the effects of otitis media on psychoacoustic development. Joseph W. Hall III and John H. Grose (Div. of Otolaryngol./Head & Neck Surgery, 610 Burnett-Womack CB#7070, Univ. of North Carolina, Chapel Hill, NC 27599-7070, jwh@med.unc.edu)

Psychoacoustic measures of auditory development in children with normal hearing and fluctuating conductive hearing losses reveal several general trends. First, processing efficiency is usually reduced in young children: they require higher SNRs and are less sensitive to AM. However, underlying mechanisms may be mature: derived measures of frequency selectivity, TMTF time constants, across-frequency processing (CMR and the processing of multiple modulations), and binaural processing (MLD for broadband maskers and sensitivity to off-midline stimuli) are adultlike by age 5. Nevertheless, key differences remain: across-frequency processing in children is less tolerant of asynchronies than in adults and, for narrow-band maskers, young children have poor binaural masking release, a result which is not due simply to perceptual similarities between signal and masker. A second general trend is that deficits in function due to early conductive hearing loss are not permanent, although the duration of abnormal performance appears to be strongly associated with the specific function being measured. Basic measures of CMR recover by approximately six months but the processing of multiple modulations takes longer. The MLD may take several years to recover. Interpretations of these findings in terms of CNS strategies for analyzing peripheral information will be presented. [Work supported by NIH NIDCD.]

3:45

2pPP6. Excessive auditory backward masking and its potential remediation in children with specific language impairment. Beverly A. Wright (Audiol. and Hearing Sci. Prog., 2299 N. Campus Dr., Northwestern Univ., Evanston, IL 60208-3550, b-wright@nwu.edu)

Approximately 5% of children who are otherwise unimpaired have marked difficulties perceiving and producing spoken language. This disorder is typically labeled specific language impairment. Recent psychoacoustic tests indicate that children with specific language impairment have excessive auditory backward masking, as demonstrated by their extreme difficulties detecting brief tonal signals followed by longer noise maskers. These results are consistent with the hypothesis that specific language impairment results from an auditory perceptual deficit that makes difficult the discrimination of brief, rapidly presented sounds, including phonemes in speech. Subsequent studies using adults with normal hearing and language skills indicate that the interference from a backward masker can be reduced with practice. This learning generalizes from trained to untrained backward-masking conditions and appears to result from improvements in temporal rather than spectral processing. Taken together, these data imply that training in auditory backward-masking conditions may be a useful intervention for children with specific language impairment. [Work supported by the McDonnell-Pew Program in Cognitive Neuroscience, NIDCD, the Charles A. Dana Foundation, and March of Dimes.]

4:15

2pPP7. Frontiers and backwaters of research on the biological bases of hearing development. Edwin W. Rubel (Bloedel Hearing Res. Ctr., Box 357923, Univ. of Washington, Seattle, WA 98195, rubel@u.washington.edu)

The goal of this talk will be to summarize recent advances and current challenges in research on the biological bases of hearing development. Three areas will be focused on: functional and structural changes in the cochlea during the early stages of hearing development; the biological bases of critical periods for inner ear and brain development; and the influence of experience on structure and function of central auditory pathways. In each area the dramatic advances in the amount and quality of descriptive information that has emerged over the past decade will be considered. These data provide a firm foundation on which to test hypotheses regarding the mechanisms that govern development. The contributions that the tools of modern cellular and molecular biology can bring toward providing new answers to questions of developmental mechanisms will be discussed.

Session 2pSA

Structural Acoustics and Vibration: Memorial Session for Manfred Heckl

Kenneth A. Cunefare, Chair
G. W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0405

Chair's Introduction—2:15

Invited Papers

2:20

2pSA1. Active control of curve squeal caused by trains. Maria A. Heckl (Dept. of Mathematics, Keele Univ., Staffordshire ST5 5BG, UK)

Squeal noise is generated by a feedback mechanism. It occurs if two bodies are in contact and move relative to each other so that dry friction acts in the contact area. This can lead to an unstable oscillation driven by the feedback between the friction force and the relative velocity of the bodies. Sound is radiated from one or both oscillating bodies. Some well-known types of noise are generated by this phenomenon, for example squeal and the sound of a bowed string. Curve squeal (the noise produced by the wheels of a train that travels fast around a tight bend) is considered in this paper. The frequency spectrum of curve squeal is dominated by one or a few sharp peaks which correspond to certain bending resonances of the free wheel. Squeal noise can be controlled and even suppressed by a secondary feedback system, consisting of a sensor (to pick up the wheel oscillation or the squeal noise), filter, phase-shifter, amplifier, and transducer (to apply the control signal). This type of control amounts to the stabilization of an originally unstable wheel oscillation and is a special form of active control. A mathematical model for this secondary feedback method will be presented. Numerical results for its effectiveness will be shown.

2:35

2pSA2. Realization of the Cremer admittance of silencer linings with passive absorber elements. Friedel P. Mechel (Landhausstr. 12, D-71120 Grafenau, Germany)

L. Cremer has shown that the maximum attenuation in a lined duct is obtained when the solution of the characteristic equation lies in the first branch point of that equation. This condition specifies a surface admittance of the lining, the Cremer admittance. Much effort was undertaken in the Institute of Professor Cremer, later directed by Manfred Heckl, to realize this admittance. It will be shown how a computer program can find the material and dimension parameters of a lining with Cremer admittance, after a suitable principal construction was conceived (e.g., multilayer porous absorber, or Helmholtz resonators, etc.). The potentials and the principal limitations of a lining with a Cremer admittance will become visible.

2:50

2pSA3. The influence of barrier edge treatments. Michael Möser (Inst. für Technische Akustik, Einsteinufer 25, 10587 Berlin, Germany)

The paper deals with the reduction of the sound field in the shadow region behind barriers by means of an attached body at the edge of the screen. The reflecting attributes of the barriers hatpiece are described by a locally reacting impedance. The considerations start with a repetition on the diffraction at ideal soft and hard bodies to demonstrate the governing principle of tangential power transport parallel to their surface: The impedance must be chosen in such a way that the tangential intensity near the edge is lowered, such turning the incoming power to ''undangerous'' directions. Subsequently, the differences due to finite impedances are discussed. The physical principles are demonstrated in a computer animation of the sound fields for the different cases. Theoretical computations give considerable improvement of the level reduction in the shadow zone for not too small angles of diffraction. They are compared with measurement results and practical applications are discussed.

3:05

2pSA4. Induced damping in Heckl's beam experiment. D. Feit and M. Strasberg (Code 7052 and 702, respectively, David Taylor Model Basin, NSWC, 9500 MacArthur Blvd., West Bethesda, MD 20817-5700)

In 1961, Heckl described an experiment showing that the loss factor of a beam in flexural vibration was increased considerably by the attachment of a large number of small cantilevered strips along the span of the beam, their lengths varying so that their frequencies of antiresonance extended over the first five modal resonance frequencies of the beam [J. Acoust. Soc. Am. **33**, 640–651 (1961)]. The increase in loss factor occurred even though the total strip mass was much smaller than the beam mass. More recently, an approximate theoretical estimate of the loss factor of Heckl's beam with attached strips, using a simple expression for induced dissipation based on ''fuzzy'' substructure concepts was presented [J. Acoust. Soc. Am. **99**, 335–344 (1996)]. This estimate agreed well with Heckl's measurements at some frequencies, but disagreed at others for unknown reasons. In an attempt to understand the cause of the partial disagreement, new calculations have been performed using a finite-element model of the beam and strips, the loss factor being estimated from resonance bandwidths in steady-state frequency response and transient decay rate.

3:20

2pSA5. Waveguide absorbers for structural damping. Eric Ungar (Acentech, Inc., 33 Moulton St., Cambridge, MA 02138)

For a series of experimental measurements, Manfred Heckl attached strips of damping tape extending outward from the plate edges to increase the plate's damping. He understood that vibrations induced in the tapes would extract vibrational energy from the plates, but did not explore their effects in detail. The present paper presents an analysis of the damping effects that attached wave-bearing systems can produce and illustrates realizations of waveguide absorbers in the shapes of tapered rods and beams. It also presents related experimental data and describes the general potential of waveguide absorbers.

3:35–3:50 Break

3:50

2pSA6. Heckls influence on SEA development. Richard H. Lyon (RH Lyon Corp, 691 Concord Ave., Cambridge, MA 02138)

Manfred Heckl was at Bolt Beranek and Newman Inc. in the early 1960s during the development of the basic SEA ideas and applications. As various concepts, formulas, and experimental results emerged, he constantly showed how those results related to the earlier work of Westphal, Cremer, and others. Four areas are particularly fresh in my mind—the use of transmissibility notions for connected plates, the use of reciprocity to calculate SEA results, the modal density of cylindrical structures, and the edge damping of plate structures. In every case, Manfred was able to bring fresh insights to the discussions. Although his name is not on many SEA papers, his contributions and influence were strong and are not forgotten.

4:05

2pSA7. The estimation and control of rolling noise from trains. Paul Remington (BBN Technologies, 10 Moulton St., Cambridge, MA 02138)

Throughout his career Manfred Heckl contributed significantly to the body of knowledge that helped to clarify the mechanisms governing the generation of rolling noise from trains. During a brief period at the Technical University in Berlin this author had the opportunity to collaborate with Prof. Heckl and to learn from his insights and expertise in this field. Here an important mechanism for the generation of noise from trains, the interaction force in the zone of contact between wheel and rail produced by the small scale roughness on their running surfaces, will be examined. Analytical techniques for predicting the interaction forces that include the effects of load and wheel and rail profile will be described. The implications for controlling the noise at the source will be examined and the noise reduction achieveable will be estimated using measured wheel and rail roughness data. [Work was supported by the European Rail Research Institute, Utrecht, The Netherlands.]

4:20

2pSA8. The multipole method: Contributions of Manfred Heckl and new developments for high-frequency acoustic scattering. Martin Ochmann (Technische Fachhochschule Berlin, Fachbereich Mathematik und Physik, Luxemburger Strasse 10, D-13353 Berlin, Germany)

In 1988, Cremer and Wang suggested spherical field synthesis (one-point multipole method) for calculating sound radiation from nonspherical bodies. Then, Heckl derived a close connection between Cremer's method and the Helmholtz integral equation. Also, he recommended the use of several multipole locations for applying the method to nonconvex radiators (multipoint multipole method). From these ideas, the multipole method has since been extended in several directions. Recent research results were obtained by investigating the high-frequency scattering from a nonconvex structure. This structure consists of a sphere where the positive octant is cut out. The region of the missing octant is called a ''cat's-eye,'' since it acts like a three-dimensional reflector. Hence, depending on the angles of incidence, multiple reflections can occur. The surface may be rigid or characterized by a local absorbing surface impedance. A boundary element model of this structure consists of nearly 8000 elements for accurate consideration of large values of ka. The results of the multipole method are compared with a newly developed iterative solver of the Helmholtz integral equation for high-frequency scattering. Both methods agree well. However, it is shown that commonly used approximations at high frequencies fail, e.g., the plane wave approximation, if multiple scattering occurs.

4:35

2pSA9. Acoustics teaching and Manfred Heckl. Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115, rossing@physics.niu.edu)

If allowed to have only a dozen acoustics books in my library, one of them would certainly be *Structure-Borne Sound* by Cremer, Heckl, and Ungar. This book, which is the epitome of clarity, could only be written by skilled acousticians who are also great teachers. Although the entire book is the result of a close collaboration between these master teachers, the Foreword tells us that the chapters on Damping, Impedance, and Radiation were written by Manfred Heckl. He also animated several of the figures from the chapter on Radiation for a classroom video he created. His lively lectures, often punctuated with thoughtful demonstrations of acoustical phenomena, serve as examples to teachers of acoustics around the world.

4:50

2pSA10. The use of tuned absorbers in minimizing sound power from vibrating structures. Gary H. Koopman, Eric Constans, and Ashok Belegundu (Ctr. for Acoust. and Vib., Penn State Univ., 157 Hammond Bldg., University Park, PA 16801)

This paper presents a methodology for designing quiet shell structures through the use of multiple tuned resonators. Thin shell structures are targeted for noise reduction because of their ubiquity in industry. They are used to cover and protect noisy devices, such as gearboxes and electrical transformers, and are also used to keep dirt and other contaminants away from moving parts, as in the case of automotive valve covers. The method combines a finite element method vibration prediction code, a boundary element method sound power prediction code, and a combined stochastic/gradient-based optimization algorithm for optimizing the locations and parameters (stiffness, mass, and damping) of the tuned absorbers. A design example has been optimized: a thin shell covering a motor/gearbox/brake combination. It is demonstrated that the addition of a small amount of weight (130 g, or 1.6% of the weight of the shell) can produce substantial reductions in sound power (13 dB in the targeted a $\frac{1}{3}$-octave band centered at 125 Hz).

5:05

2pSA11. Recollections of Professor Manfred Heckl. Gideon Maidanik (Carderock Div., Naval Surface Warfare Ctr., David Taylor Model Basin, Code 7030, 9500 MacArthur Blvd., West Bethesda, MD 20817-5700)

Professor Manfred Heckl was one of the scientific giants in the area of acoustics. He combined theory and experiment so well that both were made simple and transparent. I met Manfred at BBN in 1960. I realized he was one of those individuals that you could approach even when the question was not properly formulated. He helped and taught you how to formulate the question and directed you to the answer. He was an exceptional mentor and generous colleague. When he left BBN to return to Germany to assist in establishing Muller-BBN and to be groomed to succeed Cremer's professorship in Berlin, I lost a major impetus in my early development in acoustics. We tried to maintain close professional contacts, but his personal touch was missed and the efforts were not successful. Periodically, we did have face-to-face professional contacts and these were immensely enjoyable and beneficial, but the day-to-day influence was gone and I was poorer for it. Early in our relationship at BBN we became friends and we remained so to the end. I had many opportunities to test this friendship and it withstood intact the passage of time. It is this friendship that I will cherish forever.

5:20

2pSA12. Concentrated lateral excitation of structures. Bjorn A. T. Petersson (Dept. AAETS, Loughborough Univ., Loughborough, Leicestershire LE11 3TU, England)

Supplementing a recent study, aimed at clarifying theoretical requirements regarding models for prediction and design of transmission characteristics for built-up structures [Petersson and Heckl (1996)], vibration transmission relating to lateral force and moment excitation has been investigated for plates and beams of arbitrary depths. Through the associated mobilities, the power transmission and the cross-coupling between lateral translational and lateral rotational components of excitation and motion are studied and the influence of the volumetric nearfield is examined. Complete descriptions are developed for the two associated point-cross mobility elements. From the theoretical results it can be concluded that reciprocity with respect to the point-cross mobility elements is strictly only valid in the hypothetical case of point excitation. The numerical results demonstrate, however, that reciprocity can be invoked in an approximate sense also for realistic excitation conditions. It is shown that a reducttion of a lateral force excitation, remote from the neutral layer, to a combined force and moment excitation at the neutral layer is not applicable except for very small Helmholtz numbers. It is also concluded that the volumetric nearfield affects all mobility elements. Finally, some manageable estimation procedures for the mobilities will be proposed.

5:35

2pSA13. Measurement of multicomponent point mobilities on thick plates and deep beams. Christ A. F. de Jong (TNO Inst. of Appl. Phys., P.O. Box 155, 2600 AD Delft, The Netherlands)

The accuracy of power-based prediction methods for sound and vibration transmission in built-up structures, such as statistical energy analysis, is crucially dependent on the estimation of the input power as well as the power transmitted at junctions. These transmissions are governed by multi-component mobilities. When the excitation area is small compared with the structural thickness, the mobilities are strongly influenced by the volumetric near fields. Recent theoretical studies of concentrated excitation of plates and beams of arbitrary thicknesses [Petersson and Heckl, JSV 196(3) (1996) and Petersson ICA'98] have lead to some manageable estimation procedures for the mobility components. This paper describes a series of dedicated laboratory experiments undertaken to validate the estimation procedures. The mobility components are extracted from measured frequency response functions via an inverse method. The experimental procedure and the difficulties encountered are described. The results are compared with the estimated mobilities.

Session 2pSC

Speech Communication: Prosody, Production, and Voice (Poster Session)

Stefanie Shattuck-Hufnagel, Chair
Speech Communication Group, Research Laboratory, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

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. To allow for extended viewing time, posters will remain on display until 1:00 p.m. on Wednesday, 24 June. A cash bar will be set-up near the poster area during the Social Hour on Tuesday, 23 June, from 6:15 p.m. to 7:30 p.m.

2pSC1. The domain of phrase-final lengthening in English. Stefanie Shattuck-Hufnagel and Alice Turk (Dept. of Linguist., Adam Ferguson Bldg., Univ. of Edinburgh, Edinburgh EH8 9LL, Scotland, turk@ling.ed.ac.uk)

Intonational phrase final lengthening is well established for the final syllable rhyme in English [Klatt (1976), Wightman (1992), Campbell and Isard (1991)]; there is some evidence for lengthening extending further leftward [Langeveld (1997) for Dutch, Berkowitz (1993) for Hebrew], and for post-boundary lengthening of the initial consonant of a new prosodic constituent [Fougeron and Keating (1997), Fougeron (1997), Gow and Gordon (1995), and A. Cooper (1991)]. It is asked whether boundary-related duration lengthening extends further in both directions from a prosodic constituent boundary, using prosodically disambiguated sentence pairs like ''Please say 'Thomasburg or Terraceville and Watertown will play' '' and ''Please say 'Michigan or Madison and Vatican will play,' '' in which pitch-accent and main-word-stress-related lengthening can be distinguished from boundary-related lengthening. Results will be presented for 40 sentence pairs produced by four speakers (normal and reiterant forms) for individual time segments measured across the full extent of each target phrase. Preliminary results from a single speaker suggest that preboundary lengthening can extend leftward to the primary stress of a polysyllabic word. [Work supported by EPSRC.]

2pSC2. Ambiguity in prominence perception in spoken utterances of American English. Laura C. Dilley and Stefanie Shattuck-Hufnagel (Speech Commun. Group, RLE, 36-511, MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

For most words of English in most spoken contexts, the location of the main stress syllable is clear. However, for words in which (a) the main lexical stress is not obvious from the phoneme string alone (e.g., main-stress-ambiguous words like ''digest'' and ''increase,'' and certain proper names), and (b) the possible positions for the main stress consist of adjacent full-vowel syllables, an ambiguity in the location of the prominence and, consequently, in the identity of the lexical item (for main-stress-ambiguous words) is observed when these words occur in contexts with certain $F0$ configurations extending across the adjacent full-vowel syllables (e.g., a rise followed by a peak, or a peak followed by a fall; $F0$ peaks frequently dissociate from lexically stressed syllables [Ladd (1996)]). In these cases, the syllable affiliation of the main lexical stress is unclear; either the peak or movement (rise or fall) in the $F0$ contour can be heard as the prominence. This suggests that listeners interpret $F0$/syllable alignment in light of their language knowledge and a set of perceptual proclivities [Handel (1989)]. When the language fails to specify enough information or two perceptual proclivities are in conflict, perceptual ambiguity can result.

2pSC3. Intonation of noun phrases in Unangan (Eastern Aleut). Alice Taff (Dept. of Linguist., Box 354340, Univ. of Washington, Seattle, WA 98195-4340) and Jacob Wegelin (Univ. of Washington, Seattle, WA 98195-4322)

This paper reports on the intonation patterns of Unangan. The effect of syntactic structure on intonation was investigated. Data for this study are taken from recordings of 12 fluent speakers each providing around 20 sentences in Unangan elicited by English stimuli. Numerical results of entire pitch tracks generated on the Kay CSL 4300B were subjected to statistical analysis. Exploratory plots of batches of sentences were produced using a smoother. When utterances are grouped by speaker, Unangan words have been shown to have an intonation contour whereby each word attains a maximum in its first half, ending lower than it starts. The combination of highs at each word beginning with the downtrending of sentences creates the overall sentence shape of a cascade, each word a step in the falls [Taff and Wegelin, UCLAWPP (1997)]. Building on that study, this investigation looks at plots of same sentences compared across speakers finding that in the possessor/possessed construction (e.g., cormorant's neck) the second word in the phrase is higher in pitch than the first, contrary to the lowering of each successive word described above. Results indicate that intonation in this language is closely related to its syntax. [Work supported by NSF.]

2pSC4. Vowel devoicing and the loss of lexical accent in Tokyo Japanese. Mafuyu Kitahara (Dept. of Linguist., Indiana Univ., Bloomington, IN 47405)

The fundamental frequency of words with devoiced syllables was measured in order to see how the conflict between devoicing and the pitch accent is resolved in Tokyo Japanese. Previous studies found that the pitch after the devoiced accented syllable is elevated in order to compensate for the loss of the accent. Earlier results [M. Kitahara, J. Acoust. Soc. Am. **101**, 3195 (1997)] show that the pitch elevation occurs only when there is no phrasal high tone before the devoiced region. The present study further eliminates the possibility of accent shifting by controlling the moraic structure of target words. They contain one or two devoiceable syllables and a moraic nasal in front of the devoiceable region. Four speakers of Tokyo Japanese pronounced target words embedded in a carrier phrase. If

the accent shifts away from the devoiced syllable, pitch elevation should not occur because the accent is realized somewhere else. Preliminary analysis of the data shows that the accent on the devoiced syllable is not shifted but simply lost under these conditions because the moraic nasal cannot bear an accent in Tokyo Japanese. This result is discussed in connection with intra- and interspeaker variations of the accentual pattern.

2pSC5. The effect of stress on duration in Aleut. Lorna Rozelle (Dept. of Linguist., Box 354340, Univ. of Washington, Seattle, WA 98195, rozelle@u.washington.edu)

Although duration is often one of the physical correlates of stress, it has been claimed [A. Berinstein, UCLAWPP 47 (1979)] that languages with contrastive vowel length do not use duration as a correlate of stress. This study addresses three questions about the phonetic correlates of stress in Aleut, a language with contrastive vowel length. Is duration a correlate of stress in Aleut? If so, does the change in duration caused by stress neutralize the phonemic contrast between short and long vowels? Finally, does stress assignment precede or follow the optional deletion of the final syllable? Data from four native speakers of Pribilof Aleut were analyzed. Vowel duration was measured from the spectrographs produced by a Kay CSL 4300B speech analysis system. Vowels were categorized as stressed or unstressed according to the Aleut stress rule [A. Taff, MA thesis, University of Washington, 1992] and as long or short according to the words dictionary entry. Statistical analysis shows that duration is a correlate of stress, counter to Berinstein's claim. However, despite the change in duration wrought by stress, contrastive ratios between short and long vowels are maintained. In addition, stress is assigned before final syllable deletion.

2pSC6. Temporal structuring of acoustic segments in speech communication. Rodmonga K. Potapova (Dept. of Appl. and Exp. Linguist., Moscow Linguist. Univ., Ostozenca 38, 199837, Moscow, Russia, mglu@online.ru)

The following problems were studied on the domain of the temporal structuring of speech utterance: the determination of algorithms of synthesis of the rhythmic patterns of the utterance, the problem which is often referred to as preprogramming of the temporal organization of the speech signal being part of a more general problem of synthesis (generation) of prosodic patterns of the utterance; the determination of signals used in the information exchange between linguistic and neuro-physiological levels of speech analysis and synthesis; the rules determining the conversion of transformation of the rhythmic patterns of the utterance into real-time intervals between articulatory positions and movements. The character of a durational correlation which has been revealed in this experiment (with the use of the special correlation coefficient ρ and the test of the H_0-hypothesis) made it possible to define three types of temporal correlation: microcorrelation (for subsound and sound levels), mesocorrelation (for syllabic level), and macrocorrelation (for the level of phrase rhythmic structures—phonetic words). The integral temporal organization of speech utterance is determined on the one hand by a lower level effect—coarticulation, and on the other hand by a higher level effect—the rhythmic patterns of a phrase.

2pSC7. Influence of functional and acoustic parameters of intonation contours on prosody lateralization. Marc D. Pell (McGill Univ., School of Commun. Sci. and Disord., 1266 Pine Ave. West, Montreal, QC H3G 1A8, Canada, czbm@musica.mcgill.ca)

The perception of prosody by adults with unilateral right- (RHD) and left-hemisphere (LHD) damage and subjects without brain injury was assessed through six tasks that varied functional (i.e., linguistic/emotional) and acoustic attributes of a common set of stimuli. Three tasks measured the ability to detect prominence cues in short utterances (focus perception) and three tasks examined the perception of emotional-prosodic features by the same listeners (emotion perception). Within each condition, an initial task tested recognition of each ''type'' of prosody when all naturally occurring acoustic features (but no semantic features) signalled the target response (baseline). Two additional tasks assessed how changes in (a) duration and (b) fundamental frequency parameters of the stimuli influenced performance on the baseline task within each condition. Results indicated that, irrespective of the availability of specific acoustic parameters, recognition of prosody by both RHD and LHD patients was highly influenced by functional attributes of the stimuli (i.e., LHD patients were selectively impaired in detecting linguistic prominence whereas both groups were impaired in deriving the emotional tone). Brain-damaged patients also displayed irregularities when individual acoustic properties of the stimuli were manipulated, indicating that both functional and acoustic attributes of prosody may be determinants of prosody lateralization.

2pSC8. Syntactic ambiguities and their resolution in prosody: Observations in Japanese, Korean, Mongolian, and Turkish. Shigeru Sato, Young-Sook Choi, and Kei Yoshimoto (Grad. School of Intl. Cultural Studies, Tohoku Univ., Sendai, 980-77 Japan, satos@intcul.tohoku.ac.jp)

Acoustic measurement was made of utterances of syntactically ambiguous sentences in Tokyo and Sendai Japanese, Seoul and Kyongsang Korean, Mongolian, and Turkish to observe prosodic strategies for disambiguation. Materials were sentences of the types ADV-VP1-NP-VP2 and ADJ-NP1-NP2-VP, where the former is ambiguous in terms of adverbial modification, ADV modifying either VP1 or VP2, and the latter creates ambiguity for the adjective ADJ modifying either NP1 or NP1+NP2. For these constructions, the four languages show identical constituent structures, and accordingly create same ambiguities. After defining the syntactic depth of a boundary, $F0$ of the phrase before and after the boundary, and duration of the syllable and pause before the boundary were measured. The results for differentiating the syntactic depth: (1) Tokyo Japanese and Kyongsang Korean make use of $F0$ rise after the boundary, (2) Seoul Korean lengthens the duration, (3) Turkish uses $F0$ rise before the boundary followed by a pause, and (4) pause is the indication for Sendai Japanese and Mongolian.

2pSC9. The effects of discourse focus and lexical accent on the prosodic phrasing in Japanese. Young-Sook Choi, Shigeru Sato, and Kei Yoshimoto (Grad. School of Intl. Cultural Studies, Tohoku Univ., Sendai, 980-77 Japan, yschoi@insc.tohoku.ac.jp)

Measurement of $F0$ was made of utterances of Japanese sentences to observe behavior of intonation contours with varied focus assignment and lexical accent specifications. Materials were 80 sentences of right-branching structures of the type NP1-NP2(-NP3)-VP, where 16 have two three-mora NP's and 64 three NP's, including all the permutations of NP's, each with one of the four lexical accent types. The utterances were generated in question–answer discourse contexts so that in a sentence, one NP was always focused. The results are: (1) focus assignment has no significant effect on NP1, (2) NP2 and NP3 are significantly higher when focused than when not focused, and (3) focused/unfocused NP2 preceded by unaccented NP1 is significantly higher than those preceded by accented NP1, and so is NP3 preceded by NP2 with the same conditions. These suggest that focus assignment on NP2/NP3 requires rephrasing there, that

a lexical accent of an NP narrows the pitch range of the following phrase, and that the prosodic status of focus assignment is lower than that of the lexical accent.

2pSC10. Effects of prosodic structure on consonant weakening. Lisa M. Lavoie (Dept. of Linguistics, Cornell Univ., Morrill Hall, Ithaca, NY 14853, LML1@cornell.edu)

While the effects of lexical stress are most obvious on vowels, consonants also show evidence of prosodic structure. In English, stress conditions the flapping of alveolar stops and Turk [''Effects of Position in Syllable and Stress on Consonant Articulation,'' Ph.D. dissertation, Cornell University, 1993] has shown that labial stops also shorten in the flapping environment. This study investigates the strength of intervocalic consonants as a function of word stress, position-in-word, and speech rate. Two stress pairs of closely matched disyllabic words (one for word-initial, one for word-medial) were recorded for 16 English consonants at three rates of speech. Words for [v] are as follows, with the stressed syllable capitalized: VENdor, veNEER; SEver, seVERE. Results for fricatives will be reported, including the following tendencies for fricatives which are the onset to an unstressed syllable: voiced fricatives are more fully voiced, some fricatives are shorter, more formant structure is evident in the noise, and transitions to the following vowel may be longer. Findings for each speech rate will be compared to determine how the rate may impact consonant behavior. Each correlate will be evaluated with respect the hypothesis that consonants in unstressed syllables are more vowellike.

2pSC11. Intergestural coordination adjacent to multiple prosodic boundaries. Dani Byrd (Haskins Lab., 270 Crown St., New Haven, CT 06511-6695)

Recent work has demonstrated that progressively larger prosodic boundaries result in increasing amounts of lengthening for consonant gestures both preceding and following the boundary. This experiment extends these findings by considering kinematic data collected with a magnetometer to determine if: (1) vowel gestures demonstrate lengthening patterns comparable to those of consonant gestures due to adjacent prosodic boundaries, and (2) the relative timing or intergestural coordination between consonant and vowel gestures is affected by adjacent boundaries. Eighty tokens were recorded for each of three speakers using magnetometry for a / . . . amə(#)mi . . . / sequence with five differing boundary conditions: (1) no boundary, (2) a simple word boundary, (3) a boundary between elements in a list, (4) a vocative name boundary, and (5) an utterance boundary. Durations of the bilabial closings and openings, duration of the vocalic tongue raising movement from [a] to [i], and temporal latencies between the pre- and post-boundary [m]'s and the [i] articulation were determined. All subjects demonstrated lengthening of the vocalic tongue raising for larger boundaries, although they differed in whether the lengthening was more pre- and post-boundary. Furthermore, two subjects had no systematic changes in relative /#CV/ timing, but one subject showed relatively later vocalic peak velocities for larger boundary conditions. [Work supported by NIH.]

2pSC12. Nontraditional acoustic features of focus. H. Timothy Bunnell, Stephen R. Hoskins, and Debra Yarrington (The duPont Hospital for Children and Univ. of Delaware, 1600 Rockland Rd., Wilmington, DE 19803)

Natural productions of the sentence ''Bob bought Bogg's box'' in which focus was varied over each of the four words of the sentence were altered to produce prosodic cue neutralized versions. The alterations were applied singly and in all possible combinations to form eight experimental versions of each original sentence (the original and seven cue-neutralized versions). The original sentences as spoken by eight talkers and their cue-neutralized versions were presented to listeners with the task of identifying the focused item in the sentence. Results indicated that (a) overall, *F*0 cues were more important than either amplitude or duration cues in signaling focus, (b) the importance of amplitude and duration cues was greatly enhanced when *F*0 cues were neutralized, and (c) in many cases, identification of focus remained above chance after all three acoustic features were neutralized. The present study reports analyses of the residual acoustic features which continue to convey focus when amplitude, duration, and *F*0 are neutralized. [Work supported by NIDRR and the Nemours Research programs.]

2pSC13. Baby word rhythm preferences of Japanese infants. Akiko Hayashi (Faculty of Education, Tokyo Gakugei Univ., 4-1-1, Nukui-Kita, Koganei-Shi, 184 Japan, hayashi@u-gakugei.ac.jp), Kyoko Yoshioka (Tokyo Gakugei Univ., Koganei-Shi, 184 Japan), and Reiko Mazuka (Duke Univ., Durham, NC 27708-0085)

For infants acquiring English, many studies have shown that the dominant stress pattern of the language (viz., strong–weak syllable sequence) plays a critical role in perception, segmentation, and early speech production. Questionnaires to Japanese mothers revealed that many words they use in child-directed speech take the form of XsX or XsXs (X=regular mora; s=special morae). It was hypothesized that this sequence might count as a dominant rhythmic pattern for Japanese infants. Using a head-turn listening preference paradigm, 8–10-month-old Japanese infants were found to show a preference in listening to lists of pseudowords that conform to the XsX pattern when compared to another list of words with three regular morae (XXX). Four- to six-month-old infants showed no such preference. An additional experiment eliminated the possibility that the result of the first experiment was simply based on the presence or absence of special morae in the lists. The results will be discussed in relation to the increasing body of English data to examine how infants might learn to take advantage of language-specific rhythmic patterns in language development.

2pSC14. The relationship between spectral properties and perceptual evaluation of hypernasality in children with cleft palate—vowel /i/. Ryuta Kataoka, David Zajac, Robert Mayo, Donald Warren (UNC Craniofacial Ctr., Univ. of North Carolina, Chapel Hill, NC 27599-7450), and Kaoru Okabe (Coretex Corp., Tokyo, 192 Japan)

The purpose of this study was to examine the relationship between spectral characteristics and the perceptual evaluation of hypernasality in speakers with cleft palate. Speech samples from 32 cleft palate patients (average age 9.3 years) and 5 normal subjects (average age 9.4 years) were used in this study. The isolated vowel /i/ produced by each subject was filtered using 16 one-third octave bandpass filters from 200 Hz to 6.3 kHz to obtain a 1/3 octave spectrum. These spectra were normalized using the total power level of the 16 1/3 octave bands. Three sophisticated listeners rated the speech samples using a six-point equal-appearing interval scale. Average spectra obtained from the hypernasal and the normal groups were compared. Spectral characteristics of hypernasality in the vowel /i/ were identified as a rise in power level between the first formant (*F*1) and the second formant (*F*2) and a rise in power level of *F*1. Multiple regression analysis revealed a high correlation ($r=0.81$) between the averaged perceptual scores and four levels of 1/3 octave bands: the band containing *F*1, the first, the second, and the fourth bands from *F*1. [Work supported by NIDR Grant Nos. DE07105 and DE06957.]

2pSC15. Effects of levodopa on finger and articulatory movements in Parkinson's disease. Michele Gentil and Pierre Pollak (INSERM, CHU Clinique Neurologique BP 217, Grenoble 38043 Cedex 9, France, michele.gentil@ujf-grenoble.fr)

Efficacy of levodopa therapy on finger and articulatory movements in Parkinson's disease was studied. The production of isometric force was estimated in both conditions, without and with levodopa in 14 patients with a diagnosis of idiopathic Parkinson's disease (the mean duration of symptoms was 9 years). Forces generated by the upper and lower lips, tongue, and right and left forefingers in the presence of visual feedback were measured by means of force transducers. Moreover motor disability of patients was assessed in each condition. Fourteen control subjects also participated in this study. The beneficial effect of levodopa on finger motor disorders of patients was observed as indicated by better motor scores and improved production of finger forces. In contrast, the production of articulatory forces of patients was either not improved or aggravated by levodopa therapy. These results suggest that cerebral nondopaminergic lesions are probably responsible for the impairments of parkinsonian speech.

2pSC16. Stop-consonant production by dysarthric speakers: Use of models to interpret acoustic data. Kelly L. Poort (Res. Lab. of Electron. and Dept. of EECS, MIT, 50 Vassar St., Cambridge, MA 02139)

Acoustic measurements have been made on stop consonants produced by several normal and dysarthric speakers. The acoustic data were previously recorded by H. P. Chang and H. Chen at MIT. In the present study, various aspects of production following release of the oral closure were quantified through the use of acoustic measures such as spectra and durations of noise bursts and aspiration noise, as well as shifts in frequencies of spectral prominences. Through comparison of these measurements from the normal and dysarthric speech, and based upon models of stop-consonant production, inferences were drawn regarding articulator placement, rate of articulator release, tongue-body movements, and vocal-fold state. The dysarthric speakers deviated from normals particularly with respect to alveolar constriction location, rate of release, and tongue-body movement into the following vowel. For example, the lowest front-cavity resonance in the burst spectrum of an alveolar stop is normally in the range 3500–5500 Hz. For three of eight dysarthric speakers, the range was lowered to 1500–2800 Hz, indicating either placement of the tongue tip further back on the palate or formation of the constriction with the tongue body in a location similar to that of a velar stop. [Work supported in part by NIH Grant DC00075.]

2pSC17. Formant trajectory characteristics in persons with Parkinson, cerebellar, and upper motor neuron disease. Gary Weismer and Jill Wildermuth (Dept. of Communicative Disord. and Waisman Ctr., Univ. of Wisconsin-Madison, 1500 Highland Ave., Madison, WI 53705)

Vocalic nuclei appear to have a good deal of importance in understanding the speech motor control deficit in dysarthria. Research in our laboratory over the past ten years has shown that the slope of formant transitions and the size of the acoustic vowel space are both predictive of the speech intelligibility deficit associated with amyotrophic lateral sclerosis (ALS). Moreover, certain aspects of formant trajectories, such as transition duration and transition extent, seem to provide useful inferences about the articulatory disorder in ALS and in Parkinson disease (PD). In the present study the investigation of vocalic nucleus production in persons with dysarthria is broadened, by reporting on formant trajectory characteristics of persons with Parkinson, cerebellar, and upper motor neuron disease due to stroke. The hypothesis of pathology-specific trajectory characteristics is evaluated, and some methodological considerations concerning the parametrization of formant trajectories are discussed. [Work supported by NIDCD DC00319.]

2pSC18. The effects of utterance length on temporal control in aphasia. Shari R. Baum (School of Commun. Sci. and Disord., McGill Univ., 1266 Pine Ave. W., Montral, QC H3G 1A8, Canada, insr@musicb.mcgill.ca)

By comparing the same syllables produced in a variety of utterance-length conditions, the present investigation examined the hypothesis that fluent aphasics exhibit a shorter-than-normal speech planning domain, whereas nonfluent aphasic patients exhibit impaired speech timing at all levels. Subjects included nonfluent aphasics, fluent aphasics, RHD controls, and normal controls. The base stimuli were 10 monosyllabic nouns from which 10 bisyllabic words were derived. Each of the 20 words also appeared medially in the context of a short (4–6 syllable) and a long (8–10 syllable) sentence. Productions were elicited two times each and recorded for acoustic analysis. Durations of the target monosyllable as a proportion of whole utterance durations were computed. Preliminary analyses revealed that normal speakers produced target syllables with shorter durations in 2-syllable relative to 1-syllable words in both long and short utterances. Target syllables were shorter in the long sentences, as expected. A similar pattern was found for the other speaker gruops, but the magnitude of the differences was larger in the fluent and nonfluent aphasic patients productions. Moreover, the nonfluent speakers' proportional durations were aberrant due to abnormally long sentence durations. Results are discussed in rlation to theories of temporal control in brain-damaged patients. [Work supported by MRC.]

2pSC19. Source modeling of severely pathological voices. Bruce R. Gerratt, Jody Kreiman, Norma Antonanzas-Barroso, Brian Gabelman (Div. of Head/Neck Surgery, UCLA School of Medicine, 31-24 Rehab. Ctr., Los Angeles, CA 90095-1794), and Abeer Alwan (UCLA, Los Angeles, CA 90095)

Previous studies using a simplified LF source model [Qi and Bi, J. Acoust. Soc. Am. **96**, 1182–1185 (1994)] to describe pathological voice sources indicated that model fit to the return phase is particularly poor. To examine the perceptual importance of differences in how the return phase is modeled, various functions were fitted to the source waveforms that had been obtained by inverse filtering. Voices were then resynthesized using these different source functions, and the resulting signals were compared perceptually to the original natural voice and to versions synthesized with LF modeled sources. The perceptual importance of changes in closing phase modeling for these severely pathological voices will be discussed, as will the range of functions necessary to obtain perceptually satisfactory synthesis of severely deviant voice sources. [Research supported by NIH.]

2pSC20. Speaking rate-induced acoustic variability in Parkinson's disease. Kris Tjaden (Dept. of Commun. Disord., San Diego State Univ., 5500 Campanile Dr., San Diego, CA 92182-1518, Tjaden@mail.sdsu.edu)

The idea that speech consists of overlapping gestures that slide with respect to one another suggests that measurable kinematic, and hence, acoustic differences should characterize utterances produced with varying degrees of gesture overlap. [Saltzman and Munhall, Ecolog. Psych. **1**, 333–382 (1989)]. Further research suggests that a simple acoustic model of overlapping gestures [Weismer *et al.*, J. Phon. **23**, 149–164 (1995)] accounts for some of the spectral and temporal variability in speech produced by individuals with apraxia of speech. One implication is that articulatory deficits associated with motor speech disorders can be modeled, in part, as disruptions in gesture overlap. The extent to which a model of overlapping gestures accounts for acoustic variability in other motor speech disorders is largely unknown. The present study explored the extent to which a simple acoustic model of overlapping gestures accounted for speaking-rate induced variability in *F*2 trajectories produced in target words by individuals with Parkinson's disease (PD), young adults, and older adults. *F*2 onset frequency was used as an acoustic measure of

overlapping consonantal and vocalic gestures; *F*2 onset frequency also was used to predict across-repetition temporal variability in *F*2 trajectories.

2pSC21. A comparative study of speech motor programming in stutterers and nonstutterers. N. Aravind Kumar and S. R. Savithri (Speech and Hearing, All India Inst. of Speech and Hearing, Mysore 570 006, India)

The present study was undertaken to verify the credibility of theories that attribute the locus of stuttering aetiology to be a central speech planning dysfunction. Ten stutterers and ten nonstutterers participated in this study which utilized a choice reaction time paradigm to determine the effects of three independent variables, viz., word length, phonetic complexity, and linguistic complexity on dependent variable-speech reaction time (SRT). By varying the response complexity in this manner it was intended to manipulate the response preparation in a way that if stuttering depends on a premotor programming, then the response complexity would adversely affect response preparation time and this effect of ''complexity'' would be greater for stutterers (with aberrant speech programming) than for nonstutterers. The results obtained supported the above and significant differences were elicited in both within and across group conditions. However, some peculiar ''effects,'' viz., ''paradoxical complexity effect'' and ''motor ceiling effect'' were also evidenced and are subsequently discussed. The results of the study suggest that the stutterers may be endowed with inefficient motor programming capabilities that are aggravated by increased production demands especially under time stress conditions.

2pSC22. Modeling of articulatory dynamics using cascaded first-order systems. Yorinobu Sonoda and Kohichi Ogata (Faculty of Eng., Kumamoto Univ., 2-39-1 Kurokami, Kumamoto, 860 Japan, ogata@eecs.kumamoto-u.ac.jp)

This paper proposes modeling of articulatory dynamics using cascaded first-order systems. The model, which consists of simple first-order components, can be a good approximation for the behavior of a complex system. In the paper, the approach to modeling the articulator movements was performed on several articulatory organs: tongue, jaw, etc. Parameters describing the model were obtained from observed articulatory data in such a way that time patterns of impulse response of the model were approximated to velocity ones of movements. Good approximations were obtained on time patterns of articulation for vowels and consonants. Therefore, successive movement patterns for continuous speech including consonants were approximated by combining several models with high accuracy. By modeling articulatory dynamics for vowels and consonants based on articulatory measured data, complex articulatory patterns in a form of [-V C V-] were synthesized and the time instant of input commands for articulation of vowels and consonants were estimated in a connected speech sound including consonants. [Work supported by SCAT.]

2pSC23. Testing a normalization procedure for articulatory recovery. Richard S. McGowan (Sensimetrics Corp., 26 Landsdowne St., Cambridge, MA 02139)

In order to recover articulatory movement from human speech in the analysis-by-synthesis procedure used in a current project it is necessary to adapt, or normalize, the anatomy of the vocal tract of an articulatory synthesizer. A normalization procedure was tested in which the vocal tract length and the functions determining the conversion of midsagittal distance to cross-sectional area were adjusted until there was an acoustic match between the synthesizer and the human over a selected set of utterances. Testing of the normalization procedure was performed using vowel data on two subjects from the Wisconsin x-ray microbeam data set to find how many utterances with known midsagittal shape and acoustic information were necessary so that the adjusted synthesizer vocal tract was able to predict output of novel utterances. Another part of the normalization procedure is to find a mapping between the settings of the articulators of the synthesizer and the corresponding midsagittal shapes of the human vocal tract. Neural networks were trained to map from the settings of the articulatory synthesizer to the positions of pellets over limited sets of vowels. The ability of these networks to generalize was tested with novel vowels. [Work supported by grant DC-01247 from the NIH.]

2pSC24. A task-dynamic approach to gestural anticipation in speech: A hybrid recurrent network. Elliot Saltzman (Haskins Labs., 270 Crown St., New Haven, CT 06511, saltzman@haskins.yale.edu) and Suvobrata Mitra (Univ. of Connecticut, Storrs, CT 06269)

Recent simulations of the anticipatory timing of speech gestures within the *task-dynamic* model of gestural patterning are described. To accomplish this, a sequential neural network architecture is used in a simplified dynamical system in which the network's outputs define the activations of gestures within a set of vocal tract constrictions (''tract variables'') whose dynamics are second order. These activations act only to insert target values into the tract-variable dynamics; the stiffness and damping coefficients are fixed. The forward dynamics, from current tract-variable state and target inputs to the next tract-variable state, are then modeled by a second recurrent network whose outputs are fed back to the sequential network. In this manner the simulated articulatory timing patterns are sensitive both to the intended linguistic sequence and to the articulatory dynamics themselves. Simulations will be described suggesting that anticipatory gestural timing is defined along a simple implemented continuum between rigid time-locking and unrestricted lookahead behavior. Temporally elastic anticipatory behavior such as that reported for French lip protrusion falls naturally within this continuum. These results provide an elegant account of interspeaker and interlanguage differences in anticipatory gestural timing. [Work supported by NIH.]

2pSC25. An interactive construction system of three-dimensional vocal tract shapes from tomogram. Tohru Yokoyama, Nobuhiro Miki, and Yoshihiko Ogawa (Faculty of Eng., Hokkaido Univ., Sapporo, 060 Japan, yokoyama@kairo3.hudk.hokudai.ac.jp)

A new interactive system is developed to construct the 3-D vocal tract shape from tomogram, for the purpose of understanding the relationship between the speech production process and the geometrical figure of the human vocal tract. The shaded 3-D vocal tract shape can be seen with this system from any direction, by construction from a set of 2-D images measured by, for example, MRI or x-ray CT, and basic geometrical information is obtained for estimating the area function and transfer function, or creating the mesh model for 3-D FEM. For constructing a 3-D vocal tract shape, the following three operations are conducted for the system. First, the 3-D locations of the tomogram images are interactively determined. Second, the contours of the air tract in the images are extracted after several image processings. Third, the wire frame model of the vocal tract, including a branching case, is constructed computationally from the extracted contours. The system is developed using the C++ language as an object-oriented programming language, Motif toolkit as GUI, and PEX as 3-D graphic libraries. The proposed system provides a convenient tool for speech research on the X-Window environment.

2pSC26. Palate shape effects on characteristic vowel tongue postures. Mark K. Tiede (ATR Human Information Processing Res. Labs., 2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-02 Japan, tiede@hip.atr.co.jp)

In this work MRI vocal tract data obtained during sustained vowel production is used to investigate the interaction between individual speaker differences in hard palate shape and the compensatory adjustments to tongue posture needed to tune language-specific vowel formant resonances. PCA techniques developed in previous work [Tiede and Ye-

hia, J. Acoust. Soc. Am. **100**, 2658(A) (1996)] were used to characterize three-dimensional tongue and palate shapes for vowel inventories of five Japanese and five American English subjects. Parametrized shapes for individual speakers were compared to a normalized mean tract shape pooled across subjects for each language. Preliminary results show that the axis of tongue body displacement is associated with oral cavity depth (OCD, measured from the teeth to the rear pharyngeal wall): Shorter OCD subjects showed greater vertical /a/ versus /i/ differences, and longer OCD subjects used more of a horizontal component. Systematic effects of palate vault shape on tongue constriction location have also been observed. These results are consistent with the speaker-specific articulatory (''A'' space) adaptations suggested in previous work [Honda *et al.*, Proc. ICSLP96, 784–787 (1996)].

2pSC27. Characteristics of phonation offset and onset in normal adults and children. Laura L. Koenig (Long Island Univ., Brooklyn, NY)

A major goal of research into the dynamics of vocal fold vibration is to specify the physical requirements for self-sustained oscillation, i.e., phonation. In recent years, theoretical and modeling studies have sought to define the necessary conditions for initiating phonation in laryngeal models of varying complexity, focusing especially on requisite pressure levels (the phonation threshold pressure), structural configurations (glottal width and the glottal convergence angle), and their interrelations. Conditions surrounding phonation offset have received somewhat less attention but are generally understood to involve different scale values of the same set of variables. This paper investigates voicing behavior around abduction for /h/ in normal English-speaking adults and 5-year-olds. Recordings consist of intraoral air pressure and oral airflow collected using a Rothenberg mask. DFT analyses of airflow signals from speakers whose /h/Us had voicing interruptions reveal that spectral changes accompanying phonation offset differ from those seen at phonation onset. Speakers show various individual patterns but the difference is typically most evident in the amplitude and bandwidth of the first harmonic. These results add to the empirical literature indicating that initiation and cessation of voicing do not occur under identical conditions. [Work supported by NIH.]

2pSC28. Speaker identification using noncontemporary speech samples. Harry Hollien and Reva Schwartz (Inst. for Adv. Study of the Commun. Processes, 50 Dauer Hall, Univ. of Florida, Gainesville, FL 32611)

There are two noncontemporary elements important to speaker identification. In one, listeners are asked to make identifications at various latencies after hearing the speaker (time of identification is noncontemporary) or, in the second, samples of the speaker's voice are obtained at markedly different times (noncontemporary samples). It has been found that memory for voice decays over time in the first case; hence it has been assumed that the noncontemporary samples pose just as difficult a challenge. AN aural–perceptual speaker identification project was carried out to test this second relationship. Noncontemporary samples, with latencies of 4, 8, and 32 weeks, and then 6 and 20 years, were studied. It was found that mean correct identification was initially reduced by 15% and this level was sustained for up to 6 years. It was only after 20 years that levels dropped significantly (to 33%). It is concluded that listener's efficiency in identifying noncontemporary speech will show only modest decay over time and, hence, its use ordinarily should have only a minimal effect on the speaker identification process.

2pSC29. Perception of intoxication effects on speech. Harry Hollien and Gea DeJong (Inst. for Adv. Study of the Commun. Processes, 50 Dauer Hall, Univ. of Florida, Gainesville, FL 32611)

Can speakers, especially actors, effectively mimic the speech of intoxicated individuals and can they volitionally reduce speech degradation resulting from this behavioral state? Auditors were tasked to determine if actors sounded more inebriated when simulating it than when they actually were intoxicated and if they sounded less intoxicated (than they were) when attempting to sound sober. Young, healthy actors, chosen on the basis of a large number of selection criteria were required to produce several types of utterances: (1) during a learning phase, (2) when sober, (3) at three simulated levels of intoxication, (4) during actual, and parallel, levels of controlled intoxication, and (5) at the highest level of intoxication but when attempting to sound sober. Two (counterbalanced, ABX) aural–perceptual procedures were carried out with 40 listeners. It was found that these auditors rated the speakers as being more intoxicated when sober, but simulating drunkenness, 88% more often than when actually intoxicated. They also were judged as sounding less inebriated (than they were), when attempting to sound sober, 61% of the time. The bases for these relationships will be discussed.

2pSC30. Perceptually motivated modeling of noise in pathological voices. Brian Gabelman, Jody Kreiman, Bruce R. Gerratt, Norma Antonanzas-Barroso (Div. of Head/Neck Surgery, UCLA School of Medicine, 31-24 Rehab. Ctr., Los Angeles, CA 90095-1794), and Abeer Alwan (UCLA, Los Angeles, CA 90095)

Research investigating the correlation of acoustic measures of noise and the perception of pathological voice quality has consistently demonstrated a moderate association. However, this correlational approach cannot address basic questions concerning their cause and effect relationship, such as how carefully noise must be modeled to preserve the perceived pathological vocal quality, and what particular aspects of the noise are necessary to preserve natural vocal quality. To address these questions, an analysis by synthesis approach was applied to a large sample of natural pathological vowels varying in severity of pathology. Several synthetic versions of each voice were created using an LF source model and a formant synthesizer. Amplitude and spectral characteristics of the noise were varied as necessary. Stimuli were compared perceptually to the original voice and to each other. A discussion will focus on the procedures necessary to adequately model noise in voices with severe vocal pathology, and on the perceptual importance of differences in how vocal noise is synthesized. [Research supported by NIH.]

2pSC31. A perceptual and acoustic study of the imitated voice. Kirk P. H. Sullivan (Dept. Ling., Umeå Univ., S-901 87 Umeå, Sweden) and Frank Schlichting (Hamburg Univ., D-20144 Hamburg, Germany)

The belief that an individual can discriminate between different voices from memory is central to the concept of the voice line-up, the aural analogue of the police identification parade, and its application in the legal sphere. This paper examines the impact professional and amateur imitations have on the accuracy of the technique and whether differences in impact can be explained by an examination of vowel space, timing, and $F0$ values. Three groups of experimental participants heard a series of different voice line-ups. They were instructed to indicate which voice of each series was the voice they had been asked to identify. Not all of the series contained the voice to be identified. Group one was presented line-ups containing a high-quality professional imitation. This resulted in a significant degree of confusion between the voice to be identified and the imitation. Group two was presented the same series of line-ups but the professional imitation was exchanged for an amateur one. Group three was presented the same line-ups, but containing another amateur imitation. The amateur imitators failed to cause confusion with the voice they were imitating, yet succeeded, to different degrees, in disguising their natural voices such that their own identity was protected.

2pSC32. Acoustic analysis of the effects of alcohol on the human voice. Orla M. Cooney, Kevin G. McGuigan, Peter J. P. Murphy (Dept. of Phys., Royal College of Surgeons in Ireland, St. Stephens Green, Dublin 2, Ireland, kmcguigan@rcsi), and Ronan M. Conroy (Royal College of Surgeons in Ireland, Dublin 2, Ireland)

Experiments to determine the physical effects of alcohol on the human voice are reported. Recordings of sustained phonations and prescribed texts were taken from 12 volunteers at various stages of alcohol consumption, to determine whether a significant change in voice frequencies occurs when alcohol has been consumed. No association between the fundamental frequency of a sustained phonation and level of alcohol consumption is observed (regression coefficient = 0.146, t ratio = 0.28, $df = 71$, $p = 0.781$). Similar results for the higher formant frequencies are also observed. The effect of alcohol on duration of speech is also investigated. Prescribed-text, sentence duration is found to increase at the rate of 6.4% (standard error = 0.5%, $p < 0.0001$) per ml of alcohol consumed per kilogram of body weight. There is no evidence that the sex of the speaker is related to the effect of alcohol on sentence duration. Acoustic analysis of fixed-text voice recordings is therefore not a suitable method of determining whether a person is alcoholically intoxicated. [This research was supported by the Research Committee of the Royal College of Surgeons in Ireland.]

2pSC33. The use of random spliced speech for the recognition of familiar voices. Ricardo M. Figueiredo (Dept. of Legal Medicine, UNICAMP, CX. POSTAL 6153, CEP 13084, Campinas, Brazil, ricfig@turing.unicamp.br)

Various techniques for experimental speech signal manipulation to test recognition of familiar voices have been used, but random spliced speech (RSS) has not been tried. This procedure, which consists of randomly recombining digitally segmented speech signals, offers the advantage of more natural-sounding speech and decreased risk of creating artifacts interfering with perception. Although intelligibility is also destroyed by this technique, RSS preserves the spectral cues related to voice quality, as well as *F*0 level and range. In the present experiment, various 3.5 s RSS samples of 11 males were presented to 34 listeners, each familiar with only one of the voices. In a recognition task with forced yes/no responses, the index of correct identification was well above chance. An analysis of misindentifications revealed that proximity of *F*0 means led to more errors than did similarity of long term average spectra. Since speakers with *F*0 mean less than 2.8 semitones apart were frequently confused, it seems that in such an experimental situation the listeners may be forced to rely on isolated cues, such as *F*0; this may explain the high number of false identifications, even though the number of false rejections is practically nonexistent.

TUESDAY AFTERNOON, 23 JUNE 1998 WEST BALLROOM A (S), 2:15 TO 6:20 P.M.

Session 2pSPa

Signal Processing in Acoustics, Architectural Acoustics and Speech Communication: Multichannel Signal Processing for Acoustical Applications

David J. Evans, Chair
National Institute for Standards and Technology, Building 233, Room A147, Gaithersburg, Maryland 20899

Chair's Introduction—2:15

Invited Papers

2:20

2pSPa1. Reconstruction and projection of interior sound fields using a spherical measurement array. Earl G. Williams (Naval Res. Lab., Code 7137, Washington, DC 20375)

A spherical surface populated with microphones is placed within an aircraft fuselage or automotive cabin, and the sound field is determined throughout the interior of the source. This hypothetical experiment is the goal of the measurement system which will be discussed. The output from this spherical antenna is processed using mathematical and signal processing techniques to solve both a forward (projection) and an inverse (reconstruction) problem when no sources are within the array. The forward projection provides the field inside the measurement surface. The field outside is computed by solving an inverse problem providing a reconstruction of pressure and velocity in a spherical region up to the external enclosing boundaries. The accuracy of a proposed measurement system for both the forward and inverse modes of operation will be discussed. This discussion includes number of sensors and sensor placement schemes, and, most importantly, the effect of ambient noise and SNR on the projections and reconstructions, including the maximum reconstruction distances from the measurement sphere which can be obtained. [Work supported by ONR and NASA.]

2:50

2pSPa2. Multichannel acoustical measurement for sound field mapping. Svend Gade (Brüel & Kjær Skodsborgvej 307, DK-2850 Nærum, Denmark, sgade@bk.dk) and Jørgen Hald (Brüel & Kjær, DK-2850 Nærum, Denmark, jhald@bk.dk)

A newly developed sound field mapping method, which uses time-domain holography, is useful for localization, quantification, and ranking of sound generating regions on engines and gearboxes under run-up (i.e., nonstationary) conditions. The technique, which is called nonstationary STSF, requires simultaneous measurement of time histories on a planar surface covering the sound source. Thus subarray scanning as used with the cross-spectral STSF technique cannot be applied and a two-dimensional microphone array with half wavelength microphone separation must be used. Consequences of the need for a full size fixed array system is that it is not feasible for measurements to cover a whole vehicle, except at low frequencies, and for engine measurements a large number of microphone measurement channels is needed to cover a frequency range up to 3 kHz. The technique maps all sound field descriptors

in the near field, but not only as a function of physical location but also as a function of time for transient events. Nonstationary STSF allows source attenuation simulations similar to the cross-spectral STSF technique to be performed. The development of TDH is part of the Brite–Euram project PIANO, dealing with new pass-by noise optimization methods for quiet and economic heavy road vehicles.

3:20

2pSPa3. Signal processing for sound capture. Daniel V. Rabinkin, Richard J. Renomeron, Atul Sharma, and James L. Flanagan (CAIP Ctr., Rutgers Univ., 96 Frelinghuysen Rd., Piscataway, NJ 08854-8088)

High-quality electret microphones and single-chip processors are economical enough to be used in large numbers. This latitude opens opportunities for dynamic source location and sound capture with spatial selectivity in three dimensions. This report discusses algorithms for matched-filter processing of microphone arrays and for coordinate tracking of moving talkers. A prototype conferencing system is demonstrated in which the automatic source locator steers both a video camera and a beam-forming microphone array to capture image and audio from a moving talker. [Components of this research are supported by NSF Contract 397-26740 and DARPA Contract DABT63-93-C0037, the New Jersey Commission on Science and Technology, and the corporate members of the Rutgers Center for Computer Aids for Industrial Productivity (CAIP).]

3:50

2pSPa4. Modeling in-vehicle engine noise. Filip Deblauwe, Patrick Van de Ponseele, and Geert Lowet (LMS Intl., Interleuvenlaan 68, B-3001 Leuven, Belgium, fdb@lms.be)

Transfer path analysis is accepted as a tool to model or troubleshoot the lower engine orders for in-vehicle noise. The contribution of each force input that is coherent with the engine can be evaluated both for amplitude as well as phase. This allows ranking the individual paths and allows identifying paths that interact (cancel effects). The ability to listen to these contributions, also called partial pressures, is needed for the sound quality evaluation of engine noise, and the next logical step in transfer path analysis. Steady-state applications of time domain transfer path analysis have been reported, but primarily the study of pseudostationary sounds such as engine run-ups was desired. In this document, different methods will be discussed that could enable this type of work. The new developed synthesized order substitution method is presented. This method enables an engine sound engineer to listen to the sound due to a single or a set of noise transfer paths, for a reasonable computation effort.

4:20–4:35 Break

Contributed Papers

4:35

2pSPa5. Localization of acoustic emissions in aerospace structures. Paul J. Wells, Arthur Stephens (Adv. Information Processing Dept., Br. Aerosp. Sowerby Res. Ctr., FPC267, PO Box 5, Filton, Bristol BS12 7QW, UK), and Andy Ibbotson (Br. Aerosp. Airbus, Filton, Bristol, UK)

Health Usage and Monitoring Systems (HUMS) offer the potential to revolutionize the civil aircraft industry's approach to routine scheduled inspections. Developments in acoustic emission measurements allow the direct detection and localization of fatigue damage in complex primary aircraft structures. This information has been used to successfully guide inspections and to locate damage in regions inaccessible to conventional NDT methods. The use of such a system in flight will lead to inspection costs being reduced significantly and will promote confidence in the structural integrity of next generation civil aircraft. This paper will examine the question of the optimum positioning of sensors for a given sample using methods such as genetic algorithms and simulated annealing. The use of dimensional reduction projection techniques is also discussed. In an aerospace structure the propagation of sound will not be isotropic because of the variation in the thickness of the structure. This anisotropy will limit the accuracy of localization unless accounted for. The magnitude of this effect is determined empirically for an aircraft wing panel, and a method for calibrating acoustic emission measurements to minimize the effect of this anisotropy is suggested.

4:50

2pSPa6. Imaging of supersonic surface fields on a cylindrical shell. Charles N. Corrado, Jr., Daniel T. DiPerna, and Matthew Conti (Eng. Tech. Ctr., 240 Oral School Rd., Mystic, CT 06355-1208)

Conventional imaging techniques represent a radiating body as a distribution of equivalent monopole sources. This paper presents a more rigorous approach for imaging the supersonic components of both surface pressure and normal velocity produced by scattering or radiation from a nominally cylindrical body. The technique employs the Helmholtz integral equation to relate monopole and dipole strengths. A modal expansion of the surface pressure on a baffled cylinder forms the basis for determining an array-weighting function used to estimate the surface fields. The technique makes use of acoustic pressure data measured in a horizontal plane at ranges where the evanescent field is insignificant. No far-field constraints are imposed in the derivation of the method, allowing the measurements to be conducted within the geometric near-field of the body. When the array is located in the far field, this method is shown to yield similar array-weighting functions to a method derived by Williams [E. G. Williams, J. Acoust. Soc. Am. **99**, 2022 (1996)]. Mid-frequency, $2 \leqslant ka \leqslant 20$, scattering and radiation data from measurements of a ribbed cylindrical shell are used to illustrate the technique. The intensity images reveal membrane wave and rib scattering processes in both the time and frequency domains. [Work supported by NSWC and ONR.]

5:05

2pSPa7. Holographic reconstruction of active sources in a three-dimensional enclosure. Yang-Hann Kim, Young-Key Kim (Ctr. for Noise and Vib. Control, Dept. of Mech. Eng., Korea Adv. Inst. of Sci. and Technol., Taejon, 305-701, Korea), Byeongsik Ko, and Kangho Ko (Daewoo Motor Co., 199 Chongchon-dong, Pupyong-gu, Inchon, 403-714, Korea)

This study proposes a method to identify active sources. The theory itself does not limit the application of the proposed method, but the main aim has been focused on finding the location of active noise sources and their strengths in an enclosure. The acoustic holography method can be applied to identify active sources, primary sources if one prefers to call them. However, it will depict all sources including imaginary sources, which one may call secondary sources, which are due to reflections on the walls of enclosure. In this study, an attempt has been made to distinguish these two different types of sources. First, boundaries are modeled by the inhomogeneous Robin boundary condition. It makes it possible to represent the boundaries as the combination of general ideal sources and passive boundaries; the ideal sources imply the primary sources, while the

passive boundaries control the secondary sources. Second, this general boundary condition is combined with the holography method to reconstruct the strength of ideal sources. A technique for estimating surface admittance of passive boundaries is also proposed for a three-dimensional sound field by generalizing that for a duct.

5:20

2pSPa8. Estimation of far-field pressures from geometric near-field measurements. Daniel T. DiPerna, Charles N. Corrado, Jr. (Eng. Technol. Ctr., 240 Oral School Rd., Mystic, CT 06355-1208, ccorrado@atinc.com), and David C. Warwick (Naval Surface Warfare Ctr., West Bethesda, MD 20817-5700)

Measurements of radiated noise are often made in the geometric near field of a cylindrical radiator to take advantage of high signal-to-noise ratios. At these ranges the far-field beam patterns are not fully established and the corresponding sound pressure levels cannot be determined via a simple range correction. In this paper, three methods of estimating the far field given measurements primarily made in a horizontal plane are presented. All three methods employ propagator matrices to relate surface field quantities to the near- and far-field pressures. The pressures measured over a segment of the receive array are multiplied by a weighting vector determined from the propagator matrices to minimize the mean square error between the actual and estimated far field. One method represents the surface field as a distribution of monopole sources of unknown strength. This method also employs an image source field representation to attenuate surface bounce contributions. The second method represents the surface field in terms of cylindrical harmonics with unknown modal coefficients employed at discretized axial positions along the cylinder. The third method is formulated using unknown modal coefficients derived from a substitution of the spherical harmonic expression of the free-space Green's function in the Helmholtz integral equation.

5:35

2pSPa9. Feasibility of using imperfect microphone arrays in noise source location. Tarun Bhatt, J. Richard Houghton, and Corinne M. Darvennes (Dept. of Mech. Eng., Tennessee Technolog. Univ., Cookeville, TN 38505, tarun@bigfoot.com)

Microphone arrays are being used in many applications such as noise source location, sound intensity measurement, or acoustic holography. For the purpose of source location using the unwrapped phase method, it is important to use phase-matched microphones. The paper studies the feasibility of replacing phase-matched microphones with an array of unmatched microphones and of finding faulty microphones within the array. For this purpose, two sets of tests were performed. With the first set, the microphone characteristics were measured. An electronic circuit provided time delays, i.e., phase information, between the microphones, and narrow and wideband source signals supplied frequency responses for all the microphones. With the second set of tests, optimization techniques and neural network methods were used to locate a noise source with (1) a matched microphone array, (2) an unmatched array, and (3) an array containing faulty microphones. Comparison of the data from the various microphone arrays is presented in this paper. The feasibility of using unmatched arrays with optimization and/or neural network methods is discussed. It is also demonstrated that a microphone with a significant signal-to-noise ratio and/or a large phase error can be pinpointed within the microphone array.

5:50

2pSPa10. Superdirectivity design for a sphere-buffled microphone array. Kazuhiko Kawahara and Kimitoshi Fukudome (Dept. of Acoust. Design, Kyushu Inst. of Design, 4-9-1 Shiobaru Minami-ku, Fukuoka City, 815-8540 Japan, kawahara@kyushu-id.ac.jp)

For the signal extraction from a diffuse sound field, the superdirecitonal microphone array seems to be more stable and more effective than the noise canceling microphone array system. In this research, for the sphere-baffled microphone array (SBM), the superdirectivity characteristic was designed by the frequency domain beamforming method and the complex least mean square algorithm. The frequency domain beamforming is a method for a broadband beamforming. There is a possibility of improvement of the directivity frequency characteristic. The SBM is an array which was embedded 36 elements on the rigid diffractive sphere of 8.84 cm radius. An advantage of SBM is its isotropy. When the directivity characteristic was fixed for a definite direction, only the shift coefficent can be used to change the mainlobe direction. This idea can be easily generalized into the three-dimensional mainlobe control. In a typical design calculation, for the beamwidth of -3 dB, 30° was achieved at the frequency 2 kHz. This characteristic is comparable to that of a larger size microphone, which is called a shotgun microphone.

6:05

2pSPa11. Beamforming for broadband constant beamwidth through FIR filter and DSP implementation. Yuanliang Ma, Baosong Zhang (Inst. of Acoust. Eng., Northwestern Polytechnical Univ., Xi'an 710072, PROC, ylma@nwpu.edu.cn), and Didier Vray (Creatis, INSA de Lyon, France)

The paper presents a method to design a beamformer for broadband constant beamwidth. This is important in sonar systems for sea bottom observation and classification applications. For a given frequency band, a number of frequency points are chosen to represent the response. For each frequency point, a beamforming weighting vector can be deduced to satisfy a given beamwidth requirement by existing methods, for example, Chebyshev shielding coefficients can be deduced for a linear array. For a broadband array of N elements, a $N*M$ matrix can be obtained if M frequency points are chosen. Each row of the matrix represents a frequency response of an array element, which can be satistied by a FIR filter designed by model-reference adaptive techniques. The sum of outputs from a bank of N FIR filters provides a beamformed output and so it is very handy to be implemented by DSP hardware. A design example shows that the results are satisfactory.

Session 2pSPb

Signal Processing in Acoustics: Implementation Issues of Acoustical Signal Processing in Real Time Systems II

Christian P. de Moustier, Chair
Scripps Institute of Oceanography, Marine Physical Laboratory, 9500 Gilman Drive, La Jolla, California 92093-0205

Contributed Papers

2:15

2pSPb1. FFT beamforming with cylindrical arrays: Application to the Toroidal Volume Search Sonar. Christian P. deMoustier and Timothy C. Gallaudet (Marine Physical Lab., Scripps Inst. of Oceanogr., 9500 Gilman Dr., La Jolla, CA, 92093-0205)

FFT beamforming is typically applied to arrays of coplanar and equally spaced elements. For curved arrays, conventional delay and sum beamforming should be the norm because the elements are not coplanar and the projection of their position on a tangent plane yields unevenly spaced elements. However, with appropriate phase correction coefficients, FFT beamforming can be applied to the elements contained in a 90° sector of a cylindrical array if beams formed within $\pm 35°$ about broadside are kept. Within this sector results are essentially identical to those obtained via conventional beamforming but the FFT technique is more efficient computationally and better suited to real-time applications. The error bounds associated with FFT beamforming on cylindrical arrays are discussed and we present applications of the technique to data collected with the Toroidal Volume Search Sonar by the Coastal System Station, Panama City, FL. [Work sponsored by ONR-NRL (Contract No. N00014-96-1-G913).]

2:35

2pSPb2. Acoustic source estimation based on physical modeling and optimization algorithm. Aggelos J. Amditis, K. S. Nikita, and Nikoloas K. Uzunoglu (Dept. of Elec. and Comput. Eng., Natl. Tech. Univ. of Athens, 15773 Zografou Cambus, 9 Iroon Polytechniou St., Zografou, Athens, Greece)

The propagation of acoustic waves in the sea is analyzed by employing a horizontally stratified model for the sea medium with an arbitrary number of layers. Through this physical model, the acoustic field distribution, originating from an elementary source at an arbitrary location point inside the sea medium, is calculated by implementing a Green's function approach. The proposed optimization algorithm is based on the minimization of the quadratic error function between estimated source signal values and signal values corresponding to known acoustic sources for the frequency content of the measured signal. Initial source estimation is based on the measured signal and the transfer function obtained by the above described propagation model. The minimization of the error function with respect to the model parameters described is carried out independently for each signal, corresponding to a known source, by employing the downhill-simplex minimization method. The measured signal is positively recognized as one corresponding to a known acoustic source if the minimization procedure results in an error value less than a predefined limit.

2:55

2pSPb3. Robust Kalman-based time-varying DOA estimation of impulsive noise. Han-Su Kim and Koeng-Mo Sung (School of Elec. Eng., Seoul Natl. Univ., Seoul 151-742, South Korea)

This paper presents the robust Kalman-based method which is modified with a Huber function for estimating time-varying DOA (direction of arrival) under an impulsive noise environment. The Kalman-based algorithm can estimate the proper parameter in the additive white Gaussian noise. But it is well known that the performance of the Kalman algorithm degrades seriously when the input or the estimation error contain impulsive noise (burst outliers) which can arise frequently in many practical applications. The proposed algorithm has the robustness over the impulsive noise by putting bounds to the derivative of its cost function with respect to the estimation error, using Huber function. It is known in the theory of robust estimation that if the influence function is bounded, the estimator is robust. Simulations show that the performance of the proposed algorithm is less vulnerable to the impulsive noise than the conventional Kalman algorithm.

3:15

2pSPb4. A universal parallel sonar signal processing system. Lifu You and Enfang Sang (Dept. of Underwater Acoust. Eng., Harbin Eng. Univ., Harbin 150001, PROC, euipc@public.hr.hl.cn)

A modern signal processing algorithm has been widely used in the sonar system, which has improved the sonar system features, but has also given the sonar signal processing system more compute burden in the meantime. Parallel signal processing technology is a solution. A universal parallel signal processing module is designed, which is based on the VME bus, so it is convenient to get a more powerful processing system. This module is comprised of four TMS320C40 digital signal processors. As examples, some basic sonar signal processing algorithms such as FFT, beamforming, etc. have been implemented using this system. The efficiency and speedup are given in this paper. This parallel signal processing system is suitable for many kinds of sonar systems.

3:35

2pSPb5. Maximum likelihood track-before-detect matched-field beamforming with SWELLEX data. Kerem Harmancı, Joseph Tabrikian, and Jeffrey L. Krolik (Dept. of Elec. and Comput. Eng., Duke Univ., Durham, NC 27708-0291)

Moving source localization methods are seriously limited at low signal-to-noise ratios because the signal energy becomes spatially spread over long observation times. In this paper, the integration time available to matched-field beamformers is extended by incorporating *a priori* statistical knowledge of the source dynamics. This is accomplished by first showing that the well-known minimum variance beamformer output power is essentially the log-likelihood function with respect to source location under a model consisting of a single target in an unknown but structured noise field. This likelihood function for a single frame of data may be combined with the conditional distribution of the current target position given its previous location to obtain a maximum-likelihood (ML) track-before-detect processor. In cases where the sample covariance matrix is ill-conditioned, derivation of a constrained ML estimate assuming known diffuse noise level is given. Results are presented using vertical hydrophone array data for an underwater acoustic source off the coast of southern California. Comparisons with conventional MV matched-field beamforming demonstrate that incorporating source dynamics into the processor improves its ability to maintain the target track even in the presence of interference.

Session 2pUW

Underwater Acoustics: Scattering and Reverberation (Précis Poster Session)

Dajun Tang, Chair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105-6698

Chair's Introduction—2:15

Contributed Papers

In this session, an oral summary of each paper will be presented followed by a poster session. All posters will be on display, and all contributors will be at their posters, from 4:00 p.m. to 5:30 p.m.

2:20

2pUW1. The effects of bandwidth, dispersion, and correlation length scale on shallow-water reverberation. Kevin D. LePage (SACLANT Undersea Res. Ctr., Viale San Bartolomeo 400, 19138 La Spezia, Italy)

The ensemble average of the short-time average of the monostatic reverberation power expected in a shallow-water waveguide is obtained explicitly in the parameters of modal phase speed, slowness, dispersion, and source bandwidth under the assumption of a Gaussian source envelope and a two-dimensional Gaussian spatial scatterer correlation function. Results show that in the limiting case of very narrow bandwidths, resonant scattering is expected, yielding lower backscattered levels, and significant propagation structure is superimposed on the backscatterer time series, consistent with the two way TL to the scattering patch at the center frequency. In the limit of patch size small compared to the correlation length scale, higher backscattered levels, consistent with the scattering strength of the individual scatterers themselves, are expected, and the superimposed propagation structure is reduced due to the spatial decorrelation of the patches interrogated by the individual modes. Expressions for the time varying vertical coherence are also obtained in this framework.

2:25

2pUW2. High-frequency acoustic scattering from thermally generated turbulence. John Oeschger (Coastal Systems Station, Naval Surface Warfare Ctr., Dahlgren Div., Panama City, FL 32407-7001, oeschger_john@ccmail.nscs.navy.mil)

The results of a laboratory experiment are presented for the case of high-frequency broadband multistatic acoustic scatter from a thermally generated buoyant plume. The dominant scattering mechanism is changes in the relative compressibility, which is related to the temperature difference from ambient. Far-field weak scattering theory is used to describe the scattering process. This results in a simple relationship between the complex acoustic scatter and the one-component, three-dimensional Fourier transform of the temperature field measured at the scattering producing Fourier component K, defined as the Bragg wave number. This wave number is the magnitude of the difference wave vector between the scattered and incident wave vectors. Theory predicts for measurements made in the same scattering direction that spectral estimates for overlapping Bragg wave numbers be identical. Utilizing data collected from simultaneous multiple bistatic configurations with parallel Bragg wave vectors yields a more fully resolved spectral estimate of the scattering producing temperature field. A time series of the spectral estimate of the turbulent temperature field is shown for a three-channel bistatic common Bragg direction measurement. The results indicate agreement between the spectral estimates in the regions of overlapping Bragg wave numbers, validating the theoretical description. The time series also shows strong, yet highly variable nulls in the wave-number spectrum of the turbulence. [Work supported by ONR.]

2:30

2pUW3. Range spread measurement of underwater scatterers and channels using active sonar probing signals and a cross-correlation process. James J. Kisenwether and Dennis W. Ricker (Autonomous Control and Intelligent Systems Div., Appl. Res. Lab., Penn State Univ., State College, PA 16804)

A method for measuring range spread of underwater scatterers and channels based on the cross-correlation processing of the tones of a multicomponent waveform has been developed. The correlation terms which comprise the measure are samples of the time-frequency correlation functions (TFCF) of scatterers and channels. The measure has been used to characterize the time-frequency spread and represents the signal coherence. This paper will provide a review of two frequency correlation function representations and their characterization using the aggregate coherence measure, termed correlation mass. The sensitivity of this measure as a function of waveform parameters, scatterer structure, and measurement platform motion will be defined and its robustness on a multiping basis assessed. It will be shown that this measure can be decomposed into correlated components resulting from interaction with a scatterer and a channel. Essentially the composite correlation is obtained by sampling the product of the scatterer and channel TFCF's. Data from in-water experiments and digital sonar simulations will be presented which validates the linear systems model of the composite scattering process. [Work supported by ONR.]

2:35

2pUW4. Target size measurements via specular echo spectrum. Jerome Mathieu, Patrick Schweitzer, Etienne Tisserand, Serge Weber (L.I.E.N. (Electron. Instrumentation Lab. of Nancy), Univ. Henri Poincare-Nancy 1, BP 239, 54506 Vandoeuvre-les-Nancy Cedex, France, mathieu@lien.u-nancy.fr), and Serge Gauthier (G.E.M.C.E.A., 54500 Vandoeuvre-les-Nancy, France)

Ultrasound principles actually used to determine experimentally the size A of solid targets are based on measurement of scattering attenuation. The principal limit of this method comes from the difficulties in identifying excess attenuation due to scattering. This paper presents a new method to determine the size A of a target by means of specular echo spectrum. When an immerged (in water) target is insonified with a very brief ultrasound pulse, the backscattered echo contains two parts temporally distinguishable. The first part is the specular echo. The specular echo target spectrum St depends on the form function F and A, but also of the measuring chain response. So this method includes a gauging step which consist to calculate the spectrum Sb of a block's echo. It is also showed that Sb/St (which is noted Fmes) is proportional to F in the bandwidth measuring chain B. The comparison of Fmes and F gives the bandwidth B in

terms of KA, where K is a known wave number. Thus A is deduced by comparing Fmes and F. The most interesting advantage of this method is its independence towards attenuation. Some results are presented.

2:40

2pUW5. The extraction of a target scattering response from measurements made in a range-dependent shallow-water environment. Angie Sarkissian (Naval Res. Lab., Washington, DC 20375-5350, angie@aqua.nrl.navy.mil)

The free-field scattering response of a target may be extracted from measurements made over long ranges in a range-independent shallow-water environment by the use of vertical array of sources and vertical array of receivers. The procedure requires the generation of a low-order shallow-water mode by the source array and the extraction from the received echo at the receiver array of a low-order shallow-water mode. The algorithm is generalized to range-dependent environments where additional complications exist due to the mode coupling produced by the range dependence in the shallow-water environment, which exists in addition to the mode coupling produced by the scatterer. Results of numerical simulations are presented for a ribbed cylindrical shell with hemispherical end caps. The target scattering response is computed using the finite elements and infinite elements method, and the propagation is computed using the parabolic equations method. [Work supported by ONR.]

2:45

2pUW6. Formalism for boundary scattering in waveguides. David H. Berman (Dept. of Phys. and Astron., Univ. of Iowa, Iowa City, IA 52242, david-berman@uiowa.edu)

A natural scheme for discussing scattering in a waveguide with rough boundaries involves propagating waves from a source to the rough boundaries, followed by some form of interaction with the boundaries, followed by propagating from the boundaries to a receiver. The interaction with the boundaries is often described by a scattering amplitude whose relationship to the scattering amplitude for the same boundary located in a half-space is not entirely clear. The present work shows what this relationship is and what propagator should be used to carry the waves to and from the boundaries. These results are based on a systematic multiple-scattering analysis and provide non *ad hoc* expressions for both the scattering and the propagation parts of the problem. As an illustration, the problem of computing reverberation time histories is discussed, a problem beyond the realm of parabolic equation methods which are sometimes used to treat multiple scattering.

2:50

2pUW7. Backscattering by layered media: Modeling and comparison with data. Laurent Guillon (Lab. d'Acoust. de l'Univ. du Maine, UMR CNRS 6613, Ave. Olivier-Messiaen, BP 535, 72017 Le Mans Cedex, France, lguillon@ifremer.fr) and Xavier Lurton (Lab. d'Acoust. Sous-Marine, IFREMER, Plouzané, France)

Deep-water multibeam echosounders provide bathymetry data and acoustical images, the latter being sometimes difficult to interpret. Indeed, due to the size of the wavelength used (about 10 cm), the wave's penetration into the seafloor is not negligible and the internal geological structure of the seafloor must be taken into account in the acoustical modeling. A model for backscattering by layered media is proposed here for trying to explain the acoustical response of such geological configurations. It is based on the concept of ''equivalent input backscattering strength'' through a multilayered stratified lossy medium. Each fluid layer may feature a continuous variation in acoustical parameters (sound velocity, density, and attenuation) and generates volume backscattering. The backscattering strengths provided by the sediment finite volume and by an underlying basement are affected by various effects inside the upper layer stack : refraction, attenuation, and interferences. Synthesizing the classical approaches of backscattering (interface and volume) and transmission in layered media, this model allows one to discuss the various components and effects featured in the resultant backscattering strength. It is finally compared to acoustical data recorded in areas where some ground truth is available.

2:55

2pUW8. Measurement at low ultrasonic frequencies of absorption due to suspended particulate matter. Niven R. Brown, Timothy G. Leighton (Inst. of Sound and Vib. Res., Univ. of Southampton, Southampton SO17 1BJ, UK, nrb@isvr.soton.ac.uk), Simon D. Richards, and Tony D. Heathershaw (Defence Evaluation and Res. Agency, UK)

The visco-thermal absorption of sound by suspended particulate matter can be reliably measured using a reverberation technique. This absorption may have an adverse effect on the performance of sonars operating at low ultrasonic frequencies in coastal waters where suspensions are often present in significant concentrations. A series of experiments has been performed to study the viscous absorption by suspensions in the frequency range of 50–150 kHz by taking the difference in reverberation times of a volume of water with and without particles. Measured attenuation agrees reasonably well with that predicted by theory for concentrations above 0.5 g/l. A number of interesting phenomena have also been observed which are related to the measurement system and the turbulent flow required to keep the particulate suspended. For accurate measurement, losses at the container walls must be minimized. Although there are small losses at the surface by transmission, along with losses due to absorption in the acoustic boundary layer, the presence of bubbles and dissolved gas, and turbulence, the relative measurement gives a good assessment of the absorption by the particles. These effects are discussed and results for particulate absorption are presented.

3:00

2pUW9. Turbidity in future high-frequency sonar performance models. Simon D. Richards (Defence Evaluation and Res. Agency, UK), Niven R. Brown, and Timothy G. Leighton (Univ. of Southampton, Southampton, UK)

Current sonar performance models are incapable of accurately predicting the performance of high-frequency sonars in highly variable turbid coastal waters. There is therefore a requirement for improved models incorporating the additional effects in such environments. Turbid coastal waters are characterized by relatively high levels of suspended particulate matter, the presence of which leads to increased attenuation through viscous absorption and scattering, leading to a significant reduction in the detection range of high-frequency active sonars at moderate concentrations. The additional attenuation mechanisms have been incorporated into a propagation-loss model and the effect on detection range has been investigated. The additional attenuation processes are influenced by the ambient temperature, pressure, and salinity, and the effects of these parameters on the total attenuation in seawater containing suspended mineral particles has also been investigated. The results presented demonstrate that the effects of suspended particulate matter should be included in future high-frequency sonar performance prediction models in turbid environments. Temperature is found to be an important factor influencing the attenuation, and the local temperature should therefore be used in performance calculations. [Published with the permission of the Controller of Her Britannic Majesty's Stationery Office.]

3:05

2pUW10. Spectral diffusion of seismo-acoustic waves in shallow water. Robert I. Odom (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105) and Valerie I. Peyton (Yale Univ., P.O. Box 208109, New Haven, CT 06511)

The propagation of seismo-acoustic waves in a strongly forward scattering medium is modeled as an energy diffusion process satisfying a convection-diffusion equation in slowness space [Odom and Mercer, Geophys. J. Int. **127**, 111–124 (1996)]. No energy propagates with a slowness greater than the Stoneley/Scholte wave slowness, and therefore must sat-

isfy a zero flux boundary condition at the maximum slowness. At a slowness equal to 1/(minimum shear speed), energy is lost by radiation to shear waves, and for finite frequencies satisfies a mixed boundary condition in slowness space. Maupin [Geophys. J. Int. **126**, 762–780 (1996)] has recently shown how the branch line integral for the elastic continuum can be written as a dyad, which does not suffer from common problems with other formulations employing leaky modes. Maupin's dyadic representation for the elastic wave continuum is employed to derive the proper mixed boundary condition and illustrate its effect on the diffusion of seismo-acoustic wave energy in a shallow water waveguide. [Work supported by ONR.]

3:10

2pUW11. Experimental verification of the deformed cylinder scattering model and its application for the calculation of backscattering from fish. Kouichi Sawada (Natl. Res. Inst. of Fisheries Eng., Ebidai, Hsaki, Kashima, Japan), Zhen Ye (Natl. Central Univ., Chung-li, Taiwan, R.O.C.), and Masahiko Furusawa (Tokyo Univ. of Fisheries, 4-5-7 Konan, Minato, Tokyo, Japan)

The dorsal aspect target strength pattern of straight cylinders made of polystyrene (EPS) are measured and compared with the deformed cylinder scattering model (DCM) proposed by Ye [Zhen Ye, J. Acoust. Soc. Am. **102**, 877–884 (1997)] in the range of the tilt angle from $-50°$ to $50°$. Both experiments and the theory agreed very well for the wide range of tilt angle. The deformed cylinder model is also applied to predict the target strength pattern of Ayu (*Plecoglossus altivelis*) which is a river fish with an open swimbladder. The dorsal target strength of Ayu is measured by keeping it in alive condition. Then its swimbladder's shape is observed using a soft x-ray system and digitized. Digitized data is used for the calculation on the basis of the deformed cylinder model and also the shape of the swimbladder is reproduced carving a rectangular solid of EPS by hand. The measured TS pattern of alive Ayu, the physical EPS model of swimbladder, and the DCM calculation agreed very well. Expanded polystyrene is a useful material to verify the scattering theory of the gas model and can become an artificial fish model with a swimbladder. Scattering from a complicated shape as a fish is predictable using DCM.

3:15

2pUW12. Comparison between subcritical penetration models and *in situ* data. Alain Maguer, Eric Pouliquen, Edoardo Bovio, Warren L. J. Fox, and Henrik Schmidt (NATO SACLANT Undersea Res. Ctr., Viale S. Bartolomeo 400, 19138 La Spezia, Italy)

This paper presents the recent results of a study whose overall objectives are to determine the mechanism(s) contributing to anomalous high-frequency sound penetration into sediments, and to quantify the results for use in sonar performance prediction for the detection of buried objects. To date, several mechanisms have been hypothesized in order to explain this phenomenon, the most frequently mentioned being: (a) the existence of a Biot slow wave in the sediment, (b) surface roughness, and (c) scattering of the evanescent wave by volume inhomogeneities within the sediment. *In situ* acoustic measurements were performed on a sandy bottom whose geoacoustic properties were carefully identified. A parametric array mounted on a tower moving on a rail was used to insonify hydrophones located above and below the sediment interface. An extensive data set covering a wide range of grazing angles (both above and below the estimated critical angle) and frequencies (2–50 kHz) was acquired and processed. The results are compared to models for the three mechanisms stated in an attempt to explain the anomalously high sound-pressure levels measured. Dependence on frequency, depth into the sediment, and grazing angle is explored.

3:20

2pUW13. Subcritical penetration into a viscoelastic solid due to interface roughness: Laboratory experiment and perturbation theory model. Garfield R. Mellema, Terry E. Ewart, and Kevin L. Williams (Appl. Phys. Lab., College of Ocean and Fishery Sci., and Dept. of Elec. Eng., Univ. of Washington, Seattle, WA 98105, mellema@apl.washington.edu)

The propagation of acoustic energy into sub-bottom sediments at grazing angles below the compressional critical angle has been well documented and various mechanisms to explain this propagation have been proposed. Due to the complexity of the *in-situ* environment the propagated energy might not be the result of a single mechanism, but a sum of a combination of several mechanisms. In order to better consider the validity of one of the more prominent mechanisms, rough surface scattering, a simplified laboratory environment has been constructed in which to make measurements of penetration due to interface roughness. In this environment, the sediment has been replaced by a homogenous viscoelastic solid. To preserve the uniformity of the solid, both the transmitter and the receiver are located in the water column, where their radiation patterns are well characterized. Use of an automated positioning and measurement system and a large interface area ensures that the series of measurements are both repeatable and statistically meaningful. A perturbation theory model has been constructed for comparison with the measurements. In the experiment its accuracy is evaluated as the surface roughness is increased through the limit of perturbation theory. [Work supported by ONR.]

3:25

2pUW14. Aspects of a damped surface wave in the Fourier diamond spaces. New surface wave analysis method (SWAM). Loïc Martinez, Jean Duclos, and Alain Tinel (LAUE, Pl. R. Schuman, BP 4006, 76610 Le Havre Cedex, France)

A damped surface wave (complex wave number $K=K'+jK''$) is described by its space-frequency representation $S(x,\omega)=\exp(jK(\omega)x)$. The wave-number-frequency representation $Ksi(k,\omega)$ is the spatial Fourier transform of $S(x,\omega)$. A Ksi-cut versus k is a k-Breit Wigner form (square modulus maximum for $k=K'$,K'' half-width). A Ksi-cut versus ω is then shown to be an ω-Breit Wigner form (square modulus maximum for $\omega=\Omega'$, with an Ω'' half-width). New interesting properties are found: the phase velocity is $C_p=\Omega'/K'$ and the group velocity is $C_g=\Omega''/K''$. The wave-number-time representation $N(k,t)$ is the inverse frequency Fourier transform of $Ksi(k,\omega)$. $N(k,t)$ is then shown to be $N(k,t)=\exp(j\Omega(k)t)$. The space-time representation $s(x,t)$ is obtained by inverse frequency (respectively, wave number) Fourier transform of $S(x,\omega)$ [respectively of $N(k,t)$]. The four spaces s, S, Ksi, and N are the four Fourier diamond spaces. Used on $s(x,t)$ experimental datas, the two cuts properties are the basis of the surface wave analysis methods (SWAM). Those methods identify the complex K (respectively, Ω) using k-ARMA (respectively ω-ARMA) models. Using SWAM on a cylindrical shell, the A-wave is experimentaly fully characterized. The agreement with theoretical results is excellent.

3:30

2pUW15. Aspects of cylindrical shell resonances in the Fourier diamond spaces. Use of surface wave analysis methods (SWAM) on experimental or numerical datas. Loïc Martinez, Jean Duclos, and Alain Tinel (LAUE, Pl. R. Schuman, BP 4006, 76610 Le Havre Cedex, France)

For an attenuated surface wave (wave number $K=K'+jK''$), the resonant space-frequency $S(x,\omega)$ representation of a cylindrical shell is performed versus the angular position x. In this space, the MIIR properties are demonstrated. The resonant wave number-frequency representation $Ksi(k,\omega)$ is then obtained by spatial Fourier transform of $S(x,\omega)$. This two-dimensional second space clearly separates clockwise and anticlockwise propagating waves. The first ones are observed in the positive k

values of *Ksi*, and the second ones in the negative *k* values of *Ksi*. The modulus of *Ksi* is maximum each time that the resonant condition is encountered [ω so that $K'(\omega)$ is integer]. As new results, the k cut of Ksi reveals the Sommerfeld Watson aspect of the problem and SWAM identify the complex *K*. On an ω cut of *Ksi*, SWAM identify the RST aspect of the target (and a little more): complex frequency resonances $\Omega = \Omega' + j\Omega''$ for real modes *n*. SWAM is performed on experimental datas and on numerical form functions. The agreement with theoretical results obtained by RST is very good.

3:35

2pUW16. New experimental characterization of a resonance: Identification of the mode number using the Argand diagram and the GTD approach. Serge Derible, Jean-Marc Conoir, Jean-Louis Izbicki, and Pascal Rembert (LAUE, Universite du Havre, Pl. R. Schuman, 76610 Le Havre, France)

Until now, the mode n of vibration of an immersed cylindrical shell, associated to the reemission of a peripheral wave, using an incident quasi-monochromatic plane wave is performed by carrying out the method of isolation and identification of resonance. The new method proposed here involves only the FFT of the whole pressure signal backscattered at normal incidence by the shell, normalized using the FFT of the incident wideband signal. At a resonance frequency, corresponding to an unknown value of n, the background phase φ_{exp} is measured from the Argand diagram of the scattered pressure. It corresponds to the argument of the pressure at the resonance frequency. The geometrical theory of diffraction allows us to connect the background phase φ_{GTD} to the mode number n. This mode is then deduced by minimizing the difference between φ_{exp} and φ_{GTD}. Experiments are carried out for *A* and S_0 resonances. Then, with a single backscattered signal, it is possible to obtain the frequency, the width, and the mode of the resonance.

3:40

2pUW17. Scholte wave dispersion by rippled liquid/solid interface topography. Jacques R. Chamuel (Sonoquest Adv. Ultrason. Res., P.O. Box 81153, Wellesley Hills, MA 02181-0001)

The effect of sand ripples on seismoacoustic waves is not yet understood. Quantitative results are presented characterizing the dispersion of Scholte waves propagating along a rippled surface of an immersed ''soft'' solid half-space. One model has a sinusoidal profile with a ratio of corrugation amplitude to wavelength of 0.25. The observed Bragg frequency components of broadband transient received waveforms are shifted by 180° as the receiver is displaced one cycle along the corrugation, however, the other frequency components are delayed by a smaller amount following the dispersion characteristics. Scholte and Rayleigh wave experimental dispersion graphs obtained from ''soft'' solid models are compared with replotted scaled numerical results on Rayleigh wave dispersion along a ''hard'' corrugated solid half-space [Glass *et al.*, ''Propagation of Rayleigh surface waves across a large-amplitude grating,'' Phys. Rev. B **24**, 6843–6861 (1981)]. The findings reveal that a rippled ''soft'' liquid/solid interface can decrease the velocity by more than 70% of high-frequency Scholte wave components propagating normal to the ripples. The accuracy of acoustic inversion and detection of buried objects depends on predicting the effect of seafloor topography. [Work supported by ONR.]

3:45

2pUW18. Scatterer depth estimation using broadband active matched field processing. Brian K. Jennison (Loyola College, Dept. of Elec. Eng., 4501 N. Charles St., Baltimore, MD 21210-2699), C. Allan Boyles, and Kevin J. McCann (Johns Hopkins Univ., Laurel, MD 20723-6099)

The reduction of false alarms caused by boundary reverberation is one of the major challenges confronting active acoustic systems used for anti-submarine warfare. Scatterer depth serves as a powerful discriminant for distinguishing false target echoes reflecting from the surface or the bottom from actual targets located within the water column. Active matched field processing provides a technique for estimating the scatterer depth by comparing the measured acoustic data with a series of modeled replicas generated assuming different scatterer depths. This paper considers active matched field processing to estimate the depth of a scatterer (range is assumed known from the delay time of the echo) using a single receive hydrophone. Since matched field processing is sensitive to environmental mismatch, the search space is expanded to include both the scatterer depth and the unknown environmental parameters. Simulated annealing is used to efficiently search the parameter space for the optimal solution. A novel cost function, based solely on multipath arrival times rather than the relative multipath amplitudes, is introduced to make the search robust to uncertainties in scatter characteristics. Simulation results demonstrate the effectiveness of the technique to accurately estimate scatterer depth with limited environmental data.

3:50

2pUW19. The detection of cylindrical objects of low acoustic contrast buried in the seabed. Ruthven C. Evans and Timothy G. Leighton (Inst. of Sound and Vib. Res., Univ. of Southampton, Highfield, Southampton SO17 1BJ, UK, ecpr@isvr.soton.ac.uk)

It is anticipated that the next generation of submarine fiber-optic telecommunications cables will be invisible to present-day detection systems. In cooperation with Cable and Wireless (Marine) Ltd., research has been conducted into the remote detection of buried objects of low acoustic contrast, with particular emphasis given to the detection of buried cables. Specifically, this has concerned the design of a detection system that can reliably discriminate a small diameter (cm scale), cylindrical target buried to a depth of up to 1 m below the seabed from a range of 1 m above. A purpose-built, laboratory-scale, automated sensing system comprising a bistatic arrangement of adjustable, focused transducers is described. This apparatus has been used successfully in the high-resolution imaging of a range of buried objects, proving an acoustic detection system to be a feasible solution. Recent experimental investigations have focused on two areas: Waveform optimization techniques, to maximize seabed penetration and target interaction, and matched filtering and clutter reduction techniques, to optimize the detection system to the acoustic signature of a particular buried object. Results are presented.

3:55

2pUW20. Characteristic of a semiburied object in shallow water. Shi-e Yang and Xiukun Li (Dept. of Underwater Acoust. Eng., Harbin Eng. Univ., 150001 Harbin, PROC)

The impulse reflection of a semiburied object is strongly complicated due to a multipath channel effect in a shallow-water environment, and usually also mixed with unavoidable bottom reverberation. An efficient method of signal processing in order to obtain necessary characteristics of the object, which can be used for object identification, is discussed, and several experimental examples are given.

Session 2pPLb

Plenary Lecture

William A. Yost, Chair
Loyola University, Parmly Hearing Institute, 6525 West Sheridan Drive, Chicago, Illinois 60626

Chair's Introduction—5:00

Invited Paper

5:05

2pPLb1. Representation of complex spectra in the peripheral auditory system. Eric D. Young, Barbara M. Calhoun, Jane Yu (Dept. Biomed. Eng. Ctr. Hearing Sci., Johns Hopkins Univ., 720 Rutland Ave., Baltimore, MD 21205, eyoung@bme.jhu.edu), and Israel Nelken (Hebrew Univ.—Hadassah Med. School, Jerusalem, Israel)

The representation of acoustic stimuli in the auditory periphery is tonotopic; frequency is mapped onto a place in the array of auditory neurons and the stimulus energy at each frequency is represented by the activity of neurons at the corresponding place in the neural array. Despite this generality, the details of the process can be complex, in part due to nonlinearities in neural processing. In the dorsal cochlear nucleus (DCN), for example, neurons give the same response to an energy peak in the spectrum as they do to an energy notch. Because of the nonlinearities, it may not be possible to predict the responses to arbitrary spectral shapes from responses to tones. In the DCN, a physiological model which accounts for the nonlinearities has been developed, but this method is not generalizable. A method for characterizing receptive fields for broadband stimuli has been developed based on the presentation of ensembles of stimuli with random spectral shapes; a neuron's sensitivity to a particular frequency is obtained from the correlation between the neuron's rate and the stimulus energy at that frequency. Because this method uses complex stimuli, it should better predict responses to complex stimuli of similar bandwidths. [Work supported by NIDCD.]

Session 3aPLa

Plenary Lecture

Patricia K. Kuhl, Chair
University of Washington, Eagleson Hall 204, Box 354875, Seattle, Washington 98195

Chair's Introduction—8:00

8:05

3aPLa1. How listeners find the right words. Anne Cutler (MPI for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands)

Languages contain tens of thousands of words, but these are constructed from a tiny handful of phonetic elements. Consequently, words resemble one another, or can be embedded within one another, acoustics not withstanding. The past few years of psycholinguistic research have seen repeated confirmation that the process of spoken-word recognition by human listeners involves activation of multiple word candidates consistent with the input, and that these candidate words actively compete with one another. Current models of the word-recognition process are computationally explicit and can accurately reproduce the experimental results, as will be demonstrated by simulations with the SHORTLIST model, operating on a realistic vocabulary of 26 000 words. Activation and competition, however, do not exhaust the range of procedures which human listeners can call upon to find words in speech. My colleagues and I have recently established that listeners are sensitive, at an early, prelexical stage of speech processing, to constraints on what could potentially be a word of the language. Furthermore, cross-linguistic comparisons—of English, French, Japanese, and Sesotho—suggest that these constraints are not tailored to the characteristics of a particular vocabulary, but are universal.

Session 3aAAa

Architectural Acoustics and Noise: Structure-Borne Sound in Buildings: Measuring and Modeling Noise Flanking Transmission

Trevor R. T. Nightingale, Chair
National Research Council, Institute for Research in Construction, Building M-27, Montreal Road, Ottawa, ON K1A 0R6, Canada

Chair's Introduction—9:15

Invited Papers

9:20

3aAAa1. Sound transmission through permeable double-leaf membranes. Kimihiro Sakagami (Environ. Acoust. Lab., Faculty of Eng., Kobe Univ., Rokko, Nada, Kobe, 657 Japan), Masakazu Kiyama (Kobe Univ., Kobe, 657 Japan), Toru Uyama, and Masayuki Morimoto (Kobe Univ., Kobe, 657 Japan)

Double-leaf membranes with a permeable interior leaf are one of the typical constructions in membrane-structure buildings. A theory to predict the sound-pressure reflected by, and transmitted through, this type of double-leaf membrane was established by taking into account the coupled motion of the membranes and surrounding sound field [K. Sakagami, M. Kiyama, and M. Morimoto, J. Acoust. Soc. Am. **100**, 2735(A) (1996)]. In this same paper, the theory was experimentally validated. That work focused on absorption characteristics since the lack of absorption is one of the main acoustical problems in this type of building. The present paper investigates the sound transmission through a structure of this type in order to gain insight into its sound insulation performance which, in many cases, needs to be improved. The effects of the parameters of the double-leaf membrane on its transmissibility are demonstrated and discussed through numerical examples: The characteristics drastically change with the permeability of the interior leaf. The effect of the interior permeable leaf's mass is slight, whereas that of the exterior impermeable leaf is significant, particularly at mid-high frequencies. The effect of an absorbent layer in the cavity is also examined.

9:40

3aAAa2. Prediction of noise transmission through commercial profiled metal cladding systems. Y. W. Lam (Dept. of Acoust. Audio Eng., Univ. of Salford, Salford M5 4WT, UK)

Lightweight profiled metal cladding systems generally have poor acoustic insulation characteristics, and traditional theories for evaluating noise transmission through building elements are usually not applicable to double-skin profiled systems because of their orthotropic nature. Based on extensive experimental observations, a new method for predicting the sound reduction index (SRI) of commercial double-skin cladding systems is proposed. The method combines existing theory for orthotropic flat plates with finite element analysis to account for the pronounced ''dips'' in the SRI at midfrequencies, which are caused by the resonances of the profile geometry. The predicted SRI was shown to compare well with measurements on 15 cladding systems used in the experimental work. The method was further extended to cover commercial double-skin cladding constructions. Analytical equations were formulated to account for the orthotropic nature of the cladding sheets, the sound bridging through fixing supports, and the sound reduction through the in-fill insulation. Comparisons of predictions against sound reduction measurements were made on a wide range of commercial cladding products. Good agreement between predictions and measurements was found.

10:00

3aAAa3. European prediction models for building acoustics. Eddy Gerretsen (TNO Inst. of Appl. Phys., P.O. Box 155, NL-2600 AD Delft, The Netherlands, gerretsen@tpd.tno.nl)

About 8 years ago work started in the European standardization organization (CEN) to create prediction models for several acoustic aspects in building: airborne sound transmission, impact sound transmission, facade sound transmission, facade sound radiation, and the noise caused by technical equipment. The main aim is to provide the link between the acoustic performance of products and elements, as measured by standardized laboratory measurement methods, and the performance in the realized buildings. So far the prediction models for the first four aspects are (almost) finished. The work on the part concerning equipment noise has yet to start, handicapped as it is by a lack of available knowledge in this area. The chosen approach for the models on airborne and impact sound transmission applies separate independent transmission paths between rooms, which can be shown to be a first-order estimation of a SEA approach. For these models a new quantity has been introduced to characterize the structure-borne sound transmission at junctions of building elements. Further work is focusing on the description of measurement methods to determine this quantity and on verification of the models in order to indicate more precisely their reliability in predicting the acoustic performance of buildings.

10:20

3aAAa4. Structure-borne transmission in lightweight buildings. Robert J. M. Craik (Dept. of Bldg. Eng. and Surveying, Heriot-Watt Univ., Riccarton, Edinburgh EH14 4AS, UK)

The move away from mass as a means of reducing sound transmission in buildings has led to the development of lightweight buildings. High levels of sound insulation are achieved by introducing discontinuities while retaining structural integrity. This paper will examine structure-borne sound transmission in lightweight framed buildings. For sound transmission through a wall or floor it is the coupling between the frame and the lightweight panels which must be considered. The coupling can be classified as either a line connection or a point connection. Each of these cases will be examined and the theories for sound transmission compared with experimental results. Flanking transmission must also be considered and models for the coupling between the structural elements will also be examined. These models are different from traditional buildings as many of the elements behave as beams rather than as plates. The structure is therefore a collection of beams and plates rather than just plates. Some of the transmission paths for transmission between adjacent rooms will be examined both for the case where all walls and floors are lightweight constructions and for the case where there is a mixed construction with traditional walls but lightweight floors.

10:40

3aAAa5. Flanking paths between walls and lightweight wood frame floors using statistical energy analysis. T. R. T. Nightingale (Inst. for Res. in Construction, Natl. Res. Council Canada, Ottawa, ON K1A 0R6, Canada), and J. A. Steel (Heriot Watt Univ., Riccarton, Edingurgh EH14 4AS, Scotland)

Previous work has examined flanking paths from a floor/ceiling assembly to a load-bearing wall (where the floor joists were oriented normal to the supporting wall). It was shown that beams at the joint (plates and joist header) must be considered when predicting transmission at this type of joint. The dominant flanking path was from the floor deck to the supporting wall which reduced the apparent sound insulation from FSTC 59 to 53. The path from the upper to the lower wall was insignificant. This paper presents the results of a similar study that examined flanking paths from a nominally identical floor/ceiling assembly to a non-load-bearing wall (where the joists were parallel to the joint axis). Measured velocity level differences and apparent sound insulation are compared to the load-bearing wall configuration. Flanking paths involving a non-load-bearing wall are shown to have no appreciable impact. Joint models are presented using both standard wave methods and dynamic stiffness methods. The joint models are incorporated into a statistical energy analysis model of the system. Measured and predicted results for the two joint configurations are compared and show good agreement between 100–3150 Hz.

11:00

3aAAa6. The use of semianalytical calculation models to verify SEA predictions of flanking transmission in building structures. Ivan Bosmans and Gerrit Vermeir (Lab. of Bldg. Phys., Katholieke Univ. Leuven, Celestijnenlaan 131, B-3001 Heverlee, Belgium, ivan.bosmans@bwk.kuleuven.ac.be)

The prediction of structure-borne sound transmission between coupled plates represents a key aspect when studying sound transmission in buildings, since vibrational energy flow between walls and floors leads to the existence of flanking paths. The resulting flanking transmission can be assessed theoretically using statistical energy analysis (SEA). In this method, the coupling between plates is usually estimated based on the wave approach for the corresponding semi-infinite plates. For a specific type of junction, one might consider verifying the predictive performance of SEA experimentally on an isolated plate junction. However, possible discrepancies between measured and predicted data can result either from the failure of the calculation model for the semi-infinite plates, or by the violation of the basic SEA assumptions. In this paper, a semianalytical calculation model based on a modal summation approach for finite-sized plates is proposed to determine the source of the prediction error. For this purpose, calculations are performed for two ensembles of similar corner junctions with randomly distributed dimensions. The analysis emphasizes the influence of the model overlap factor on the accuracy of SEA. [Research funded by the Flemish Institute for the promotion of Scientific Technological Research in Industry (IWT).]

WEDNESDAY MORNING, 24 JUNE 1998 CASCADE BALLROOM II (W), 11:25 TO 12:55 P.M.

Session 3aAAb

Architectural Acoustics and Noise: Impact of Trains on Performance Halls

George P. Wilson, Chair
Wilson Ihrig and Associates, Inc., 5776 Broadway, Oakland, California 94618

Invited Papers

11:25

3aAAb1. Owner/builder perspective of the impact of trains on performance halls for Benaroya Hall, Seattle. Andrew Clapham, Project Manager (Seattle Symphony, 305 Harrison St., Seattle, WA 98109-4645)

The Seattle Symphony began planning for a new concert hall in early 1987 at a location near their existing venue, the Seattle Opera House. During funding negotiations with The City of Seattle in 1992, the City made it clear that they wished to have the new facility located in their downtown core area. They offered a site that had many added challenges. These included being in the flight path of Sea-Tac International Airport; being alongside the Metro transit tunnel which is to accommodate light rail; being located in the heart of a very active city; and finally being located directly above the twin track railroad tunnel that runs north to south through the city center. In order to achieve the desired acoustical quality for the 2 500-seat main auditorium and 540-seat recital hall, not only for performances, but also for digital recording, a number of innovative design and construction techniques were required. These ranged from a massive foundation slab above the tunnel, to construction of a vibration isolated building within a building, which is supported on natural rubber isolation bearings. The challenges of designing and constructing this unique building will be presented.

11:45

3aAAb2. Vibration isolation design for Benaroya Hall. George P. Wilson (Wilson, Ihrig & Assoc., Inc., 5776 Broadway, Oakland, CA 94618)

Location of the new Seattle Concert Hall, Benaroya Hall, directly over an existing Burlington Northern railroad tunnel required development of a very effective vibration and noise isolation system. The first step in the design was to measure the ground-borne vibration and noise levels from trains via boreholes at the site and in the basement levels of nearby buildings located over the tunnel. From the measured vibration, the expected sound level in the performance space was projected for comparison with the criterion established for acceptable level and inaudibility. This comparison determined the amount and frequency range for the reduction required. The degree of reduction of structure-borne noise required determined that the design required two parts, a massive foundation mat for coupling loss and an isolation system with 4-5 Hz effective natural frequency and full reduction over a wide frequency range. Achieving the needed attenuation required a structure with very high mechanical impedance and isolators which provide a low acoustical impedance over a broad frequency range, leading to the use of heavy concrete structure and natural rubber isolation bearings. Details of the design parameters and materials are presented along with results from measurements with the completed isolation system and partially completed hall.

12:05

3aAAb3. Structural design for vibration isolation at Benaroya Hall. Brian H. Glover (Skilling Ward Magnusson Barkshire, 1301 Fifth Ave., Ste. 3200, Seattle, WA 98101-2699)

Vibration and noise isolation criteria, together with an urban site for the Seattle Symphony Hall, created unusual constraints for its structural design. A massive concrete mat foundation bridges and isolates the railroad tunnel running diagonally under the site. Two garage levels and an isolation level further separate the Auditorium from the tunnel, while 30 ft of unexcavated soil remain between the Auditorium and an adjacent underground transit station. Heavy concrete construction is used extensively to add mass for isolation. Concrete topping was added to the Auditorium roof for isolation, and double concrete walls surround the Auditorium. Exterior precast concrete cladding adds further mass and isolation. To speed construction time, the Auditorium balconies, including the inner walls which were cast to odd geometric shapes, are precast concrete. The Auditorium structure is isolated from the surrounding structure by an uninterrupted, 7-in.-wide joint, creating a ''box within a box'' The only structural connections to the Auditorium inner box are two types of special, natural rubber isolation bearings at its base. One bearing type takes gravity loads. The other type, initially prestressed, takes lateral loads. The need for bearing access under the base of the Auditorium created unusually shaped transfer girders.

Contributed Papers

12:25

3aAAb4. Controlling subway noise in the LG Arts Center—Sangnam Hall. D. Stanton Korista and Ahmad Abdelrazaq (Skidmore, Owings & Merrill LLP, 224 South Michigan Ave., Ste. 1000, Chicago, IL 60604)

The LG Arts Center in Seoul, Korea is a multifunction facility including a Hall intended for a variety of performances including western drama and dance, chamber music, traditional Korean folklore plays, opera, oration, and digital recordings. The 1050 seat Hall is designed to guard against both airborne and groundborne noise. Airborne noise is controlled by enclosing the stage house and the audience chamber with 200-mm airspace double walls. Horizontal and vertical acoustical joints at the Hall structure interface with the main building and control groundborne noise generated by the subway trains in a tunnel adjacent to the site. The subway tunnel and the foundation structure of the building are in the same rock strata and the groundborne noise is transmitted as structure-borne noise to the Hall through the lower-level parking structures. Without isolation the noise from the subway would have intrusive effects on the multifunctional use of the performance hall and stage area. Vibration isolation rubber bearings are provided in the interface design to reduce the noise to acceptable levels in the noise critical space. The impact of the proposed isolation bearing pads on the building design and their acoustical benefits are discussed.

12:40

3aAAb5. Train and rail transit noise in cinemas and theaters: Case studies. M. Asselineau (Peutz & Assoc., 103 Bd Magenta, F75010 Paris, France)

Train and subway noise unfortunately often turns out to be part of the show in theaters and cinemas of large cities, as the history of rail lines construction, or simply the urban planning, brought them close together. This paper aims to submit and discuss a few case studies, involving both rehabilitation and new construction projects, that feature cinema and performance facilities located near underground and suburban lines, main lines, or even high-speed lines. Those case studies, which range from real success to total failure, outline the architectural and technical context of the situation and discuss the sound and vibration levels measurements, as well as the technical solutions that were applied to the construction. It turns out that in rehabilitation projects, while the client may be willing to improve the acoustics, structural limitations can often be encountered, while in new projects there often is a need to convince the client of the importance of vibration control.

WEDNESDAY MORNING, 24 JUNE 1998 VASHON ROOM (W), 9:15 A.M. TO 12:40 P.M.

Session 3aAB

Animal Bioacoustics: General Animal Bioacoustics

G. Pavan, Chair

Centro Interdisiplinaire di Bioacustica, e Ricerche Ambientali, Universita degli Studi di Pavia, Via Taramelli 24, 27100 Pavia, Italy

Invited Paper

9:15

3aAB1. Three-dimensional computer modeling of biosonar emission in the common dolphin. James L. Aroyan (Elo TouchSystems, Inc., 41752 Christy St., Fremont, CA)

Several delphinid species are known to possess a highly sophisticated and adaptable biosonar system, yet the exact acoustic mechanisms involved in signal generation, emission, and reception remain poorly understood. Recently, a novel computational approach was utilized in an investigation of the acoustical mechanisms of dolphin biosonar. Three-dimensional acoustic simulation techniques and a new method for mapping acoustic tissue properties from x-ray computed tomographic data were applied to models of the forehead and lower jaw tissues of the common dolphin, *Delphinus delphis.* This approach proved highly effective in studying the detailed processes of biosonar emission and reception in this animal. The results of the investigation which concern the emission

system of this dolphin will be presented. These results include: (1) the common dolphin's skull plays the predominant role in beam formation; (2) the melon contributes significantly to narrowing of the emitted beams; (3) the melon behaves as a waveguide and a lens in the biosonar signal emission process; (4) the results of both inverse simulation and forward extrapolation programs suggest that the biosonar signal source tissues lie within a small volume of the soft tissues of the right nasal passageway.

Contributed Papers

9:45

3aAB2. Acoustic pathways of hearing in the bottlenose dolphin, *Tursiops truncatus*. Whitlow W. L. Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734), Bertel Mohl (Univ. of Aarhus, Aarhus, Denmark), Paul E. Nachtigall (Hawaii Inst. of Marine Biol., Kailua, HI 96734), Jeffrey L. Pawloski, and James L. Aroyan (Elo TouchSystems, Inc., Fremont, CA)

The acoustic pathways for the reception of sound in the Atlantic bottlenose dolphin were studied by measuring auditory brain-stem-evoked potential responses as a specially designed suction cup hydrophone was placed at different areas of the animal's head. The dolphin was trained to come out of the water and beach on a rubber mat. Evoked responses were recorded differentially with suction cup surface electrodes placed between the parietal (noninverting) just posterior to the blowhole and mastoid (inverting). A transient signal was used as the stimulus and the amplitude of the stimulus was lowered in fixed dB steps until the evoked response reached a specific level close to the threshold level. The results indicate that the most sensitive portion of the lower jaw is slightly forward of the pan-bone and about 3/4 of the way toward the front of the lower jaw and along certain spots on the bottom midline of the lower jaw. The relative sensitivity of different areas of the lower jaw was also determined. The measurement results will be compared with results from calculations based on numerically solving the three-dimensional wave equation for wave propagation into the lower jaw of the dolphin. [Work supported in part by ONR, Grant No. N0001496WX30364.]

10:05

3aAB3. Effect of low-frequency seismic exploration signals on the cetaceans of the Gulf of Mexico. Shannon Rankin and William E. Evans (Marine Acoust. Lab, Texas A&M Univ., 5007 Ave. U, Galveston, TX 77550)

The increasing use of seismic exploration has created a need to understand effects that intense low-frequency sounds have on the distribution, abundance, and behavior of cetaceans. The goal of this study is to determine possible effects of seismic exploration on cetaceans in the northern Gulf of Mexico. Low-frequency sounds originating from seismic exploration were compared with the presence/absence of cetaceans as determined by acoustic and visual methods of detection. The recordings were collected via a towed passive hydrophone array during seasonal GulfCet research cruises from 1992–1997. The study area consisted of 14 north–south transects ranging from the 100–2000 m isobath in the northern Gulf of Mexico. For odontocetes, with the exception of sperm whales, no relationship between presence of low- and medium-frequency seismic noise was found on a relatively large spatial scale of hundreds of kilometers. For mysticetes, results were inconclusive due to small sample size. These findings suggest that the repetitive signals generated by seismic exploration may not have a negative impact on the large-scale distribution or overall abundance of delphinids in the Gulf of Mexico. However, additional studies have indicated possible negative impacts on aspects of communication and orientation of behavior in sperm whales (*Physeter macrocephalus*).

10:25

3aAB4. An onboard acoustic data logger to record biosonar of free-ranging bottlenose dolphins. Douglas P. Nowacek, Peter L. Tyack (Dept. of Biol., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), Randall S. Wells (Chicago Zoological Society, c/o Mote Marine Lab., Sarasota, FL 34236), and Mark P. Johnson (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

An acoustic data logger has been developed which utilizes a two-channel DAT recorder housed in aluminum and attached to the dorsal fin. The recorder has a flat frequency response from 10 Hz to 14 kHz, and each tape can store 120 min. The first suction-cup hydrophone (sensitivity −205 dB) was placed 10 cm posterior of the blowhole, and the second 20 cm below the lateral base of the dorsal fin. The anterior ''high-frequency'' hydrophone, designed to record echolocation signals, has unity gain and a one-pole 10-kHz high-pass filter. The ''ambient'' hydrophone located at the base of the dorsal fin has +18-dB gain and a one-pole 1-kHz high-pass filter. To obtain echolocation recordings the ''high-frequency'' hydrophone was filtered through a simple demodulator in one of the deployments. The package includes VHF radio transmitters for tracking the animal and recovering the package after it releases via corrosible magnesium links. The package was attached to temporarily restrained animals which, after release, were followed to record behavioral data while the recorder logged acoustic activity. During the two successful deployments to date the logger recorded animal vocalizations, surfacing events, the sounds of passing boats, and hydrodynamic sounds produced by the animal's fluke strokes. [Work supported by ONR, OVF, SeaGrant-WHOI.]

10:45–11:00 Break

11:00

3aAB5. Results of underwater ambient noise measurements in three large tank exhibits at the Monterey Bay Aquarium. Daniel M. O'Neal and Steven R. Baker (Phys. Dept., Naval Postgrad. School, Monterey, CA 93943-5000)

Measurements have been made of the underwater ambient noise in three large tank exhibits at the Monterey Bay Aquarium: the Kelp Forest Exhibit (335 000 gal), the Monterey Bay Habitats Exhibit (350 000 gal), and the Outer Bay Waters Exhibit (1.4 million gal). A single, calibrated, Navy-type DT-276 was used. The hydrophone output voltage was preamplified and recorded using a 16-bit digital audio tape recorder with a sampling rate of 48 kHz. Measurements were made with various mechanical equipment (motors, fans, pumps, sprinklers, wave machine) turned on and off. On one occasion, the noise was measured in the largest tank, the Outer Bay Waters Exhibit, during a power shutdown. For comparison, measurements were also made at several locations and depths in the inner Monterey Bay. The results of one-third octave band and narrow-band analyses will be presented. Comparisons will be made between the aquarium and bay results and standard deep water ambient noise spectral density curves.

11:20

3aAB6. Identification of dolphin schools by bioacoustical unique features. Konstantin Kebkal (State Oceanarium of Ukraine, 9, Epronovskaya, 335024 Sevastopol, Ukraine, kostia@hotmail.com; kostya@fb10.tu-berlin.d400.de), Rudolf Bannasch (Technische Universitat Berlin, FG Bionik & Evolutionstechnik, Ackerstr. 71-76, D-13355 Berlin, Germany), and Valeriy Kulagin (State Oceanarium of Ukraine, 9, Epronovskaya, 335024 Sevastopol, Ukraine)

Long-term acoustical observations conducted on bottle-nosed dolphins (Tursiups truncatus ponticus) in the State Oceanarium of Ukraine as well as in the open sea revealed the existence of school specific acoustic features. It was found that each dolphin school uses a unique set of whistle contours in conjunction with a preferable frequency band and certain time characteristics. If the latter two characteristics somewhat vary according to particular surrounding and behavioral situations, the composition of normalized whistle contours stays stable. Each single whistle can be described by a set of polynomial coefficients characterizing its shape. Referring to the whole set of whistles emitted, a diagram of signal preferences can be built. Several samples will be presented to demonstrate that such diagrams represent acoustical portraits unique for every dolphin school. Based on those analyses, the development of a system for passive auto-

matic monitoring of dolphin abundance and migration activity is in progress. The system should consist of radio-acoustical buoys placed in the sea to record signals and to transmit them to a coastal station and an automatic signal analyzer in conjunction with an expert-system capable of determining whether the respective signals belong to a school registered earlier, or a new one.

11:40

3aAB7. Matched-field processing of baleen whale vocalizations using calibrated environmental models. Aaron M. Thode, Gerald L. D'Spain, and William A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., Univ. of California at San Diego, La Jolla, CA 92093-0205)

During the 1997 Marine Mammal Vocalization (MMV) experiment conducted by the Scripps Marine Physical Laboratory last September, J15 source tows were conducted in a highly range-dependent environment northwest of San Nicolas Island in the Southern California Bight region. The tows consisted of eight radial tracks ranging from 1 to 6 nautical miles from the research platform FLIP, along all cardinal directions. Among the signals broadcast were low-frequency combs (18–110 Hz), middle-frequency combs (50–200 Hz), FM downsweeps (900 to 50 Hz over 9 s), and an Eastern Pacific type A blue whale recorded during the 1996 MMV experiment. Multiple CTD casts were also taken, and a bathymetric database was obtained from the National Oceanographic Service. All these data are used to calibrate matched-field processing environmental models, which are subsequently applied to find the location in 3D space of a series of blue whale vocalizations. [Work sponsored by ONR.]

12:00

3aAB8. Feasibility of monitoring bowhead whales with a vertical line array. Kevin D. Heaney and Peter N. Mikhalevsky (Sci. Applications Intl. Corp., 10260 Campus Pt. Dr., San Diego, CA 92121)

As part of the ACOUS trans-Arctic acoustic propagation experiment, a vertical line array will be deployed off Pt. Barrow, Alaska in the summer of 1998. The array will be cabled to shore and will be capable of real-time recording of acoustic frequencies from 20 to 500 Hz. This array will provide scientists the opportunity to record bowhead whale songs in real time from hydrophones spanning the water column. Currently, the bowhead whale census is taken during the spring migration from the edge of ice leads. The feasibility of doing a Fall census, when the sea is relatively free of ice, will be presented. Recordings of bowhead whale songs, and simulations of localization using the vertical array will be presented. Using varying bathymetry and sono-buoys, to add to the horizontal resolution of the array, 3D localization of vocalizing whales should be possible.

12:20

3aAB9. Observations of the movements of humpback whales about an operating seismic survey vessel near Exmouth, Western Australia. Robert D. McCauley (Ctr. for Marine Sci. Technol., Curtin Univ. Technol., P.O. Box U 1987, Perth 6854, Western Australia), Micheline-Nicole Jenner, Curt Jenner (Ctr. for Whale Res., Broome 6725, Western Australia), and Douglas H. Cato (Defence Sci. Technol. Organisation, Prymont, NSW 2009, Australia)

As part of a project investigating the interaction of air-gun noise and marine animals, monitoring of migrating humpback whales traversing the track of a seismic vessel was conducted in October 1996. Four separate pods traveling SW in 100–120 m of water were followed as they approached a seismic vessel traveling on an east/west heading and operating a 2678 cui air-gun array every 8 s. The response of pods was to change course at 3–6 km from the vessel so as to pass in front of or behind it. In one instance a single animal was observed moving slowly at 4.2 km NNW of the vessel, then to suddenly accelerate to 10–12 kn on a SW course across the vessel's bow at a closest range of 1.5 km. It then swam S at 5–10 kn until it was 4 km from the vessel. During the period of rapid swimming the animal stayed close to the surface with the tail flukes often breaking the water, maintaining a 40-s blow interval. The animal then swam steadily SW, eventually stopping and commencing singing on encountering a pod of two animals at 15 km from the still operating seismic vessel. [Work supported by the Energy Research and Development Corporation, Australian Petroleum Production and Exploration Association, and WMC Petroleum.]

WEDNESDAY MORNING, 24 JUNE 1998 GRAND BALLROOM A (S), 8:30 TO 10:30 A.M.

Session 3aAO

Acoustical Oceanography and Underwater Acoustics: Acoustic Determination of Ocean Parameters

Jeffrey A. Nystuen, Chair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

Contributed Papers

8:30

3aAO1. Applications of two frequency insonification techniques to oceanic bubble sizing. Andy D. Phelps, Matt D. Simpson, and Timothy G. Leighton (Inst. of Sound and Vib. Res., Univ. of Southampton, Southampton SO17 1BJ, UK, ap@isvr.soton.ac.uk)

This work presents the results from using a combination frequency acoustic technique to measure the near-surface bubble population in both the open sea and surf zone. The benefits of using a combination frequency technique are that the collected data are more accurate and less prone to ambiguities than linear backscatter measurements. The technique monitors the appearance of sum-and-difference signals generated by the nonlinear interaction of two sound fields: one high-frequency signal scattering geometrically from the bubble surface and the other used to excite the bubble into resonant pulsation. The work details the calibration of the apparatus necessary to relate the measured heights of the sum-and-difference terms to actual numbers of bubbles, and describes the experimental procedure for the collection of the oceanic data. The collected data are verified using simultaneous measurements of the dispersive sound speed and attenuation, which can be compared with estimates calculated from the measured bubble population. The results from a number of trials are presented, with both time-averaged and time-variant measures of the local population demonstrated. [Work supported by the Natural Environment Research Council.]

8:45

3aAO2. Possibilities for small-scale emission tomography of bubble plumes in shallow water. Ivan P. Smirnov, Alexander I. Khil'ko (Inst. of Appl. Phys., Nizhny Novgorod, 46, Ulianov, Str., 603600, Russia), Paul A. Elmore, and Jerald W. Caruthers (Naval Res. Lab., Stennis Space Center, MS 39529)

Bubble plumes, created at the sea surface by breaking waves, emit sound in a broad range of frequencies. Based on the spatial and temporal characteristics of the emitted sound, emission tomographic reconstruction of the noise source distribution and evolution may be possible. It is shown that the space/time variables needed for the reconstruction process can be estimated from the coherence function of the radiated noise. The effectiveness of the reconstruction procedure when noise is present is estimated theoretically by the following physical model: (1) the location and time of bubble creation is assumed to be random and statistically independent, (2) the spatial distribution of the noise in the cloud, the spatial structure of the cloud, and the coherence function of the noise are described by a Gaussian distribution, (3) the sizes of the sources are much smaller than the size of the acoustic fields, and (4) the observation distance is much greater than the size of the sources. The requirements for the reconstruction process and the possible limitations of the method are discussed.

9:00

3aAO3. Monitoring air–sea exchange using ambient sound. Jeffrey A. Nystuen (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, nystuen@apl.washington.edu)

Wave breaking and precipitation are important processes contributing to the exchange of momentum, heat, water, and gas at the air–sea interface. These processes are also principal generators of high-frequency (over 1 kHz) underwater ambient sound. Furthermore, the bubbles generated by the breaking waves and raindrop splashes can be stirred downward into the ocean mixed layer forming an effective sound absorption layer. When present this layer modifies the ambient sound field, and can thus be detected using passive acoustics. Using the passive sound field as the exploratory tool, more conventional measurements of salinity, temperature, bubbles, wind, and surface waves will be examined. Anomalies in near-surface salinity are explained by acoustically derived rainfall and wind-speed measurements. Acoustically derived rainfall measurements suggest that rain affects the surface wave field. Enhanced injection of bubbles into the mixed layer by rain in the presence of high wind is indicated by changes to the shape of the underwater sound spectrum. [Work supported by ONR Ocean Acoustics and NSF.]

9:15

3aAO4. Wave scattering and propagation in bubbly liquids. Zhen Ye (Dept. of Phys., Natl. Central Univ., Chung-li, Taiwan, ROC) and Alberto Alvarez (Natl. Central Univ., Chung-li, Taiwan, ROC)

Acoustic wave propagation through bubbly liquids can be modified significantly by its scattering by the bubbles. The effects may be conveniently studied either by diagram methods when the number of bubbles is large, or by numerical simulation when the number is not so large. In this paper, numerical simulation on the acoustic scattering and propagation in a bubbly mixture is reported, with comparison to the effective medium theories. The analysis shows possible Anderson localization of acoustics waves. It is found that collective modes in bubble clouds are possible when the bubble volume fraction reaches a certain value. [Work supported by NSC of ROC and Spanish Ministry of Education.]

9:30

3aAO5. Rough surfaces characterization using a time-reversal mirror. Philippe Roux, Mathias Fink (Lab. Ondes et Acoustique, ESPCI, 10 rue Vauquelin, 75005 Paris, France), and James H. Rose (Iowa State Univ., Ames, IA 50011)

A new method is described for determining the rms height and autocorrelation function of rough surfaces using measurements made with an acoustic time-reversal mirror (TRM). First, the TRM insonifies the surface and the backscattered wave is recorded and time-reversed. Second, the TRM is displaced parallel to the surface and, at each displacement the previously recorded time-reversal wave is rebroadcast. For zero displacement the time-reversed wave is matched to the surface and the original incident pulse is recovered (upon reflection of the rebroadcast wave). However, with increasing displacement the signal ceases to resemble the incident pulse. The decorrelation of the signal allows one to characterize the surface. Measurements are reported for a number of periodic and random rough surfaces.

9:45

3aAO6. The Aharonov–Bohm effect revisited by an acoustic time-reversal mirror. Philippe Roux, Julien De Rosny, and Mathias Fink (Lab. Ondes et Acoustique, ESPCI, 10 rue Vauquelin, 75005 Paris, France)

The violation of time-reversal invariance for acoustic waves to achieve a new way of characterizing a vorticity field with a double acoustic time-reversal mirror (TRM) is taken advantage of. In particular, it is shown experimentally that the double TRM works as a vorticity amplifier. In the case of a vorticity filament, the sound-vorticity interaction is interpreted as the acoustical analog of the Aharonov–Bohm effect. Numerical experiments simulate the acoustical Aharonov–Bohm effect by propagating a plane wave which is scattered by a heterogeneous motionless medium.

10:00

3aAO7. Real-time characterization of unstationary flows using a time-reversal mirror. Agnes Maurel, Julien De Rosny, Sebastien Manneville, Philippe Roux, and Mathias Fink (Lab. Ondes et Acoustique, ESPCI, 10 rue Vauquelin, 75005 Paris, France)

Violation of time-reversal invariance for acoustic waves is a new method for characterizing a vorticity field with a double acoustic time-reversal mirror (TRM). Experiments show that the double TRM works as a vorticity amplifier and, even in the case of small vorticity field, provides the characteristics of the flow. Moreover, the TRM works in real-time and allows one to measure the vorticity field at a 20-Hz frequency rate. The case of a Von Karman vortex street is studied experimentally.

10:15

3aAO8. Results from cross-correlation measurement of estuarine current velocity. Roger W. Bland and Daniel G. Neuman (Dept. of Phys. and Astron., San Francisco State Univ., San Francisco, CA 94132)

A technique for accurate measurement of the average water velocity in estuarine environments has been developed. The technique is based on measuring the travel time of 200-kHz ultrasonic signals. A high degree of precision has been obtained using a cross-correlation technique, and long, detailed time series have been recorded. However, in some geometries the precision is degraded by multipath interference. Measurements of current speed in California's Sacramento-San Joaquin River Delta system will be discussed, with particular reference to the effects of multipath interference. Data were collected throughout a large flood in Dec. 1996–Jan. 1997. An analysis of the variation of the strength of the ultrasonic signal has been carried out to determine the attenuation of the signal due to the load of suspended sediment, and the results will be presented. The current-velocity time series recorded in the San Francisco Bay also show wave trains from boat wakes. These wakes will be discussed.

Session 3aBV

Bioresponse to Vibration/Biomedical Ulttrasound: Ultrasound Assisted Drug Delivery I

Pierre D. Mourad, Chair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

Chair's Introduction—8:00

Invited Papers

8:05

3aBV1. Mechanisms for biological effects of ultrasound. Edwin L. Carstensen (Dept. of Elec. Eng. and Rochester Ctr. for Biomed. Ultrasound, Univ. of Rochester, Rochester, NY 14627)

At least three physical processes have the potential to produce therapeutic benefits through ultrasound exposure. These include heating, cavitation, and a related group of second-order mechanical effects that involve radiation forces and streaming on a macroscopic or microscopic scale. In most applications of ultrasound in physical therapy, deep heating is the postulated rationale for its use. Hyperthermia as an adjunct to cancer therapy requires sophisticated control of ultrasonic heating. At higher levels, ultrasound can be used to destroy precisely selected tissues obviating the need for surgical intervention. In this application, nonlinear propagation can be used to selectively increase the absorption parameters of the target tissues. At low amplitudes the response of bubbles is dominated by the acoustic pressure. At a critical acoustic pressure, however, the inertia of the surrounding medium becomes controlling. At this threshold acoustic pressure a 10% or 20% increase in acoustic pressure leads to an increase in the collapse pressure in the bubble by orders of magnitude. Radiation forces can be detected by auditory and tactile sensors. Macroscopic acoustic streaming may alter transport across biological membranes and microscopic streaming exerts stresses on the membranes and the cytoplasmic and nuclear contents of cells.

8:30

3aBV2. Toxin and gene transfer into cells by extracorporeal shock waves. M. Delius (Inst. for Surgical Res., Univ. of Munich, Klinikum Grosshadern, 81366 Munich, Germany) and U. Lauer (Univ. of Tübingen, 72076 Tübingen, Germany)

Shock wave application to cells *in vitro* causes a transient increase of the permeability of the cell membrane which does not lead to cell death. It was hypothesized that shock waves might be a new method of transferring therapeutic agents directly into cells. To test this, biological effects resulting from the acoustical transfer of proteins and nucleic acids into cells were examined. Protein transfer was examined with the ribosome inactivating proteins gelonin and saporin which inhibit the cellular protein synthesis. Dose response curves were established with three tumor cell lines in the presence and absence of shock waves. Compared to the controls, shock waves enhanced the action of gelonin and saporin from 300 to 40 000 fold. *In vivo* experiments with an animal tumor model established that acoustic transfer of these agents into cells also occurred *in vivo*. Gene transfer by shock waves was examined in a number of cell lines by plasmid vectors carrying standard reporter genes. Transfer succeeded in many cases yet the transduction efficiency was lower than with other established methods of gene transfer. In summary shock wave permeabilization is a new method for tumour therapy and gene transfer.

8:55

3aBV3. Ultrasound-induced transcutaneous transport for drug delivery and diagnostics. Joseph Kost (Dept. of Chemical Eng., Ben-Gurion Univ., Beer-Sheva, 84105, Israel, kost@mit.edu), Samir Mitragotri, and Robert Langer (MIT, Cambridge, MA 02139)

In spite of major research and development efforts in transdermal systems and the many advantages of the transdermal route, impermeability of the human skin is still a major problem that limits the usefulness of the transdermal drug delivery approach. Application of ultrasound induces temporary structural changes in the skin resulting in enhanced transdermal transport of molecules (sonophoresis). Experimental findings suggest that among all the ultrasound-related phenomena evaluated, cavitation plays the dominant role in sonophoresis, suggesting that application of low-frequency ultrasound should enhance transdermal transport more effectively. It was found that at the low ultrasound frequencies high molecular weight molecules including insulin, interferon, and erythropoietin can be transported through the skin. Recently it was shown that the enhanced skin permeability during sonophoresis also facilitates outward diffusion of glucose present in the interstitial fluids beneath the skin. The *in vitro* as well as *in vivo* experiments show that the amounts of glucose extracted by sonophoresis are proportional to the blood glucose concentrations. Furthermore, the amount of glucose extracted by ultrasound can be measured and used to predict temporal changes in the blood glucose levels. These results demonstrate the feasibility of developing a noninvasive method of drug delivery and diagnostics using ultrasound.

9:20–9:35 Break

Contributed Papers

9:35

3aBV4. Ultrasound for the mediation of a therapeutic agent through the blood–brain barrier. Leonard J. Bond (Univ. of Denver Res. Inst., 2050 E. Iliff Ave., Denver, CO 80208) and Lawrence Ng (Univ. of Colorado Health Sci. Ctr., Denver, CO 80262)

There is a critical need to develop safe, noninvasive, nonchemical methods to enhance the delivery of nonlipophilic neuropharmaceuticals to the brain. A novel method for mediating nonlipophilic drug transport across the blood–brain barrier (BBB) will be presented. Preliminary data were obtained using a unit operating at 1.6 MHz with ultrasound coupled through water and an attenuating layer onto an *in vitro* cell-culture model of the BBB. The effects of different intensities on permeation rate and cell integrity were investigated. Various microscopic techniques were used to determine cell condition and viability. At low intensities no effects were seen. At intermediate intensities enhanced permeation up to levels of a factor of 3 increase were measured. At the highest intensity substantial cell disruption was observed. Preliminary observations regarding the mechanism for the interaction of ultrasound with the BBB will be presented. Work is in progress to determine the conditions for effective trans-cranial delivery of ultrasound using phantoms constructed to model *in vivo* brain. By promoting transport of nonlipophilic drugs across the BBB, this approach has the potential to safely and effectively treat a wide range of brain and central nervous system disorders.

9:50

3aBV5. Mechanistic studies of ultrasonically enhanced transdermal drug delivery. J. M. Farrell (Harvard-MIT Div. for Health Sci. and Technol., Boston, MA 02139), S. Mitragotri, S. Theis, J. Kost, and R. Langer (MIT, Boston, MA 02139)

The transdermal route of drug delivery offers a noninvasive and painless alternative to more traditional routes such as injections. However, its applicability is limited to only a handful of drugs due to the low permeability of the skin. It has been demonstrated that low-frequency ultrasound significantly enhances transdermal transport of a variety of drugs (sonophoresis). There is strong experimental evidence to suggest that cavitation plays a key role in this enhancement, although it is not clear which cavitational effects are primarily responsible for the transport. Cavitation can affect transdermal transport in at least two ways: It can induce changes in skin structure, creating pathways or channels for molecules to diffuse across and/or it can induce convective transport across the skin. An experimental system has been developed to assess contributions due to each of these phenomena individually. Although the data are insufficient to eliminate a role for convective processes, these experiments clearly indicate that stuctural changes in the skin play an important, and probably primary, role in enhanced transport.

Invited Paper

10:05

3aBV6. Enhancement of ultrasound-mediated transfection with cavitation micronuclei. James F. Greenleaf (Dept. of Physiol. and Biophys., Mayo Clinic and Foundation, Rochester, MN 55905, jfg@mayo.edu) and Mark E. Bolander (Mayo Clinic and Foundation, Rochester, MN 55905)

Many methods are available for transferring foreign genes into mammalian cells (transfection). The advantages of ultrasound-mediated transfection are that it can be controlled in both space, by focusing, and time, by exposure. It was tested whether ultrasound-mediated gene transfection in cultured immortalized human chondrocytes would be enhanced using artificial cavitation nuclei in the form of Albunex®. Here 1.0 MHz ultrasound transmitted through the bottoms of six-well culture plates containing media with green fluorescent protein reporter gene plasmids at a concentration of 40 μg/ml and Albunex® at 50×10^6 bubbles/ml produced a peak transfection efficiency of about 50% of the living cells when exposure was 4×10^5 P spatial average peak pressure (SAPP) for 20 mins. Using these parameters, transfection efficiency increased linearly with ultrasound exposure pressure with a transfection threshold observed at a SAPP of 10^5 P. Adding fresh Albunex® at 50×10^6 bubbles/ml prior to sequential one second 3.2 or 4.0×10^5 P exposures increased transfection efficiency by 15% for each of three consecutive exposures. Efficient *in vitro* and perhaps *in vivo* transfection appears possible with ultrasound providing both temporal and spatial control over the transfection process.

Contributed Paper

10:30

3aBV7. Defects of bilayers generated by ultrasound. M. S. Malghani, Jie Yang, and Junru Wu (Dept. of Phys., Univ. of Vermont, Burlington, VT 05405)

Transdermal drug delivery is handicapped by the low skin permeability caused by highly ordered structure of lipid bilayers in the outer human skin layer. It has been reported that ultrasound can increase the permeability of human skin [S. Mitragotri, D. Blankschtein, and R. Langer, Science **269**, 850–853 (1995)]. The enhancement was attributed to acoustic cavitation but the underlying physical mechanism is not fully understood. As a model, dipalmitoylphosphatidylcholine (DPPC) lipid bilayers are insonicated by ultrasound of two submegahertz frequencies (168 and 707 kHz). The free-field spatial peak pressure amplitudes of both are measured to be 6×10^5 Pa. Bilayer defects, which have average diameters of tens to hundreds of nanometers and can be detected by an atomic force microscope, are generated within less than 0.5 min. The number of the defects grows with time. The defect growth rate at the 168-kHz frequency is about 3.5 times that at the 707-kHz frequency.

Session 3aEA

Engineering Acoustics: Electroacoustics and Standards

George S. K. Wong, Chair
Institute for National Measurement Standards, National Research Council, Montreal Road, Ottawa, ON K1A 0R6, Canada

Invited Papers

8:00

3aEA1. Selected topics in electroacoustical calibrations, measurements, and standards. Victor Nedzelnitsky (Natl. Inst. of Standards and Technol., Sound Bldg. 233, Rm. A147, Gaithersburg, MD 20899-0001, vnedzel@nist.gov)

Modern instruments including multichannel dynamic signal analyzers capable of FFT and other signal processing, as well as improved concepts for applying these instruments to particular electroacoustical measurement and standardization issues, are essential tools in modern electroacoustical metrology. The overall uncertainty of a measurement, and its suitability for its intended purpose, may involve components of uncertainty that are individually and collectively difficult to evaluate. In such instances, the capacity to perform the same measurement in different ways can be especially valuable. Different measurements exploiting differences in source signals, signal processing implementations, and apparatus can provide valuable insight and can support the validity of measurement results and interpretations. Examples selected from recent and ongoing work at NIST include calibration or characterization of microphones and microphone arrays in a free field, examination of imperfections in anechoic chambers, and measurement of effects of instabilities and nonlinearities in electroacoustical systems. Recent ANSI (American National Standards Institute) and IEC (International Electrotechnical Commission) standardization work involving the coherence function method for examining nonlinearities in the electroacoustical performance of hearing aids has elicited comments proposing an alternative method involving the Hilbert transform. Some significant advantages and disadvantages of these methods are discussed.

8:15

3aEA2. Progress in silicon microphones. Gerhard M. Sessler (Univ. of Technol., Merckstrasse 25, D-64283 Darmstadt, Germany, ses@uet.th-darmstadt.de)

Recently, several advances have been made in the design, technology, and performance of silicon microphones. New or significantly modified transducers are: (1) condenser microphones consisting of a single chip processed with sacrificial-layer technology to form the electrode-air gap system; (2) condenser microphones with small parasitic capacitance to decrease the load on the sensor output; (3) optical-waveguide microphones utilizing novel waveguide designs and operating with intensity modulation, which, to date, has not been used in acoustic silicon sensors. Also, attention has been given to various types of ultrasonic silicon microphones for airborne sound. Of importance in all implementations is the equivalent noise level, which is, in some microphone-preamplifier systems, sufficiently low for certain applications, but still higher than for conventional microphones. The present state of the art with respect to transducer design and performance will be discussed.

8:30

3aEA3. Design and properties of face to face couplers for microphone comparison calibration. Erling Frederiksen (Bruel and Kjaer, Skodsborgvej 307, 2850 Naerum, Denmark. erlingfred@bk.dk)

This paper describes the design and properties of microphone comparison calibration couplers. In 1988 Bruel and Kjaer introduced a sound level calibrator with a built-in reference microphone which, via a feed-back system, controls the sound pressure in its coupler cavity. This was a significant step in the direction of more stable calibrators. The same calibrator presented another new principle. The microphone to be calibrated and the reference are placed face to face in such a short distance that their diaphragms see essentially the same sound pressure which is produced by a rotational symmetrical source. This mounting was found to improve the accuracy and to increase the upper frequency limit of the new calibrator by more than five times in comparison with previous calibrator configurations. Today IEC-TC29 prepares an international standard for comparison calibration of working standard microphones. The method of this standard uses the special symmetrical mounting of the source and the microphones. The paper describes results obtained with some experimental couplers which are made in this connection in order to analyze the calibration accuracy and optimize the frequency range. The couplers are mainly intended for comparison of sensitivity magnitude but may also be used for comparison of phase response characteristics.

8:45

3aEA4. Controlled environment for reciprocity calibration of condenser microphones. George S. K. Wong and Lixue Wu (Acoust. Standards, Inst. for Natl. Measurement Standards, Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

For precision reciprocity calibration of condenser microphones it is necessary to operate the calibration apparatus under a controlled environment such that fluctuations in pressure, temperature, and relative humidity are reduced to a minimum. The design, implementation, and performance of a highly stable environmental chamber for calibration of microphones and instruments is discussed. For reciprocity microphone calibrations, the system enables resolution of better than 0.001 dB.

9:00

3aEA5. Procedures for the precision of sound power determination by S. I. scanning method. Hideki Tachibana (Inst. of Industrial Sci., Univ. of Tokyo, 7-22-1 Roppongi, Minato-ku, Tokyo, 106 Japan) and Hideo Suzuki (Ontec R&D Co., Ltd., Yokohama, 226 Japan)

As the third international standard of ISO 9614 series (Acoustics—Determination of sound power levels of noise sources using sound intensity), the precision method by scanning is now being drafted by ISO/TC43/SC1/WG25. It will be Part 3 (Precision method for measurement by scanning), following Part 1 (Measurement at discrete points: precision, engineering and survey classes) and Part 2 (Measurement by scanning: engineering and survey classes). In this paper, the outline of the draft and discussions are introduced by focusing on the measurement accuracy, the methods for checking the adequacy of the measurement sound field (field indicators), and practical procedures.

9:15

3aEA6. Experimental study on the accuracy of sound power determination of S. I. scanning method. Hiroo Yano (Chiba Inst. of Technol., 2-17-1 Tsudanuma, Narashino-shi, Chiba, 275 Japan), Hideki Tachibana (Univ. of Tokyo, Tokyo, 106 Japan), and Hideo Suzuki (Ontec R&D Co., Ltd., Yokohama, 226 Japan)

In ISO/TC43/SC1/WG25, the precision method of sound power determination by sound intensity scanning method is being drafted as the third standard of ISO 9614 series. For this drafting work, some experimental studies were performed on the measurement accuracy, the methods for checking the measurement sound field (field indicators), and practical procedures of signal processing during the scanning. The sound power determination was made for artificial sound sources (a reference source and a sound source system consisted of two loudspeakers) and actual noise sources (a vacuum cleaner and an air compressor) by changing such measurement conditions as the scanning density, the scanning speed, and the strength of the extraneous noise. From the experimental results, the validities of the specifications in the draft are examined.

9:30

3aEA7. Electroacoustics and standards. Gunnar Rasmussen (G.R.A.S. Sound & Vib. aps., Skelstedet 10B, 2950 Vedbæk, Denmark)

Standardization within electroacoustics has been of great importance for the development of sound measuring instrumentation. The sound level meters are used to build up a homogeneous database, which in turn is used as a foundation for the measurement standards. Based on the database, regulation and legislation of noise exposure, noise control, product labeling, etc., may take place. The development of standards for sound level meters, measurement microphones, and calibration techniques has led to big improvements. The development of standards for intensity instrumentation and measurements will hopefully lead to similar improvements in developing the instrumentation and practical use in the daily work of the practitioners. The revision of the standards is important. They should hopefully lead to improvements for the manufacturer and for the users. The revisions should take into account the technical development and the expanded use of instrumentation in the acoustical and mechanical field. The ability to measure small changes in emitted intensity by critical sources is important for research in noise control. The measurement of sound-pressure level in the outdoor environment is not well treated in the standards, although it is an area where the legislation is most restrictive regarding noise exposure, because it influences the largest parts of the population. Examples and proposals will be given.

9:45–10:00 Break

Contributed Papers

10:00

3aEA8. The study of long-term sensitivity changes in backplate electret type condenser microphones. Kenzo Miura (Matsushita Commun. Eng. Co., Ltd., 4-3-1 Tsunashimahigashi, Kohoku-ku, Yokohama, 223 Japan) and Yoshinobu Yasuno (Matsushita Commun. Industrial Co., Ltd., 600 Saedo-cho, Tsuzuki-ku, Yokohama, 224 Japan)

The estimation of the lifetime of a usual electret condenser microphone (ECM) using an electret foil as a diaphragm has already been reported by authors. On the other hand, backplate electret type condenser microphones (BECMs) used for mobile telephones, recently, and the backplate that is laminated with electret foil has been examined and studied concerning the estimation of sensitivity changes over a long time span. Heat acceleration tests for electret foil on the backplate were taken, and from the results, the surface charge decay of backplate electret was estimated. As the thickness of the diaphragm in a BECM is thinner than the usual ECM, sensitivity increasing because of looseness of diaphragm tension was expected to become larger than the usual ECM. Therefore, heat acceleration tests for the diaphragm only were taken and estimates of the levels of sensitivity increasing by the looseness of the results. The charge decay characteristics on the electret surface and the characteristics of tension looseness in the diaphragm had both results taken together. The long-term sensitivity changes of BECMs were estimated and compared with the sensitivity of BECM samples that were left for about 10 years at room temperature.

10:15

3aEA9. Measurement of microphone random-incidence and pressure-field responses and determination of their uncertainties. Johan Gramtorp and Erling Frederiksen (Bruel & Kjaer, 2850 Naerum, Denmark, jgramtorp@bk.dk)

For many applications it is important to know the frequency response of the measurement microphone. In an earlier paper the determination of the free-field response and the uncertainties of the responses based on actuator calibration for $\frac{1}{2}''$ microphones was discussed [J. Gramtorp and E. Frederiksen, J. Acoust. Soc. Am. **100**, 2672 (1996)]. This paper continues by discussing the determination of the pressure-field and random-incidence responses and the determination of their uncertainties. The pressure-field responses are based on estimated pressure-field corrections obtained using the pressure-field reciprocity technique and electrostatic actuator calibration. The random-incidence responses are based on a large number of free-field measurements with different angles of incidence increased in small steps, and calculated according to IEC 1183, combined with electrostatic actuator calibration. From most manufacturers no infor-

mation on the uncertainties is available, but a complete set of uncertainty values for pressure-field and random-incidence corrections and responses for the new Bruel & Kjaer Falcon Range microphones will be presented and discussed.

10:30

3aEA10. Measuring system for the derivation of thermal loudspeaker parameters. Gottfried K. Behler and Armin Bernhard (Inst. for Tech. Acoust., RWTH Aachen, Templergraben 55, D-52056 Aachen, Germany, gkb@akustik.rwth-aachen.de)

Loudspeakers used in high-power applications are stressed by mechanical and thermal load. Both can be the reason for damage and should be controlled. Whereas the mechanical situation can be derived from the input voltage and the total impedance, which can be measured easily, it is rather difficult to calculate for the actual voice coil (VC) temperature under working conditions. It is possible to derive the VC temperature from the input voltage using linear mathematics if the thermal equivalent circuitry of the loudspeaker including the cabinet is known. Former papers have discussed the thermal equivalent circuitry of the loudspeaker without taking the cabinet into account. In the presented paper an equivalent circuit is presented, which can be used for loudspeakers in different cabinets and a measuring method is shown, that enables long-term measurements on loudspeaker systems with constant input power using a PC-based regulator. The main problem is to calculate the voice coil temperature from the measured impedance data. The solution of this problem is presented. At least the thermal parameters of the equivalent circuit are calculated from the measured data using curve fitting. Different loudspeakers and setups have been tested. The results will be presented.

WEDNESDAY MORNING, 24 JUNE 1998 GRAND CRESCENT (W), 9:15 TO 12:25 P.M.

Session 3aMU

Musical Acoustics: Bowed String Instruments

Thomas D. Rossing, Cochair
Physics Department, Northern Illinois University, DeKalb, Illinois 60115

Carleen M. Hutchins, Cochair
Catgut Acoustical Society, 112 Essex Avenue, Montclair, New Jersey 07042

Invited Papers

9:15

3aMU1. Studies in bowing point friction in bowed strings. Robert. T. Schumacher (Dept. of Phys., Carnegie Mellon Univ., Pittsburgh, PA 15213)

By using force transducers at the terminations of a violin E string, one may reconstruct the force and velocity at the bowing point that produces those termination forces [R. T. Schumacher and S. Garoff, Institute of Acoustics, Proceedings, Vol. 19: Part 5, 43–48 (1997) (ISMA'97)]. That information, combined with examination of the ''friction track'' [R. T. Schumacher and S. Garoff, CAS J. **3**(2), 9–16 (1996)] allows the possibility of determining the frictional properties of the rosin-mediated bow–string interaction. The reconstruction of the force, $f(t)$, and velocity, $v(t)$, allows determination of the force–velocity relation $F(v)$. That reconstruction is a delicate operation requiring careful measurement of the reflection functions of the transverse waves at the terminations. Use of simulations aids in determining the limitations of the reconstructions. Knowledge of $F(v)$, combined with examination of the friction track, variation of the chemical constituents of the rosin, and chemical preparation of the string surface, as well as measurements of the rosin's physical properties, such as softening temperature and x-ray structural examination, will eventually establish the physical basis for the alternating adhesion–lubrication properties of the rosin that produce the slip-stick string oscillation.

9:40

3aMU2. Modal analysis of violins and cellos. Thomas D. Rossing, Mark Roberts, Eric Bynum, and Laura Nickerson (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115)

Normal modes of bowed string instruments are determined by the coupled motion of the top plate, back plate, enclosed air, ribs, neck, fingerboard, etc. Normal mode testing has traditionally been done using sinusoidal excitation, either mechanical or acoustical, although modal testing with impact excitation is a fast, convenient way to determine the normal modes of an instrument. Holographic interferometry provides the greatest detail about mode shapes, but the best understanding of the normal modes of an instrument results from application of more than one testing method. The normal modes in a cello are found to be quite similar to the corresponding modes of a violin, although they may appear in a different order in frequency. Modal frequencies in one cello were found to occur at 0.29 to 0.43 times the corresponding mode frequencies in a violin.

10:05

3aMU3. A normal-mode-based model of violin radiation. George Bissinger (Phys. Dept., E. Carolina Univ., Greenville, NC 27858, bissingerg@mail.ecu.edu)

A simple physical vibration–radiation model based entirely on experimental modal analysis can successfully predict the spatially averaged radiative properties of a bridge-excited violin measured with a rotating microphone array. The energy transfer from bridge excitation to acoustic radiation is modeled using (a) the measured transfer mobility $Y(\omega)$ to characterize energy transfer from the bridge to the corpus, and (b) the radiation efficiency R_{eff}, calculated from a boundary element program for each normal mode to characterize the energy transfer from the corpus to the air. This model provides string-specific radiation predictions and allows one to make quick estimates of the averaged acoustical effects of violin adjustments or modifications which would be expected to change $Y(\omega)$ but not R_{eff}, e.g., bridge trimming. Extensions of the model and future applications and directions for its application will be discussed.

10:30

3aMU4. Influence of the weight of mutes on tones of a violin family. Kenshi Kishi (Dept. of Electro-Commun., The Univ. of Electro-Commun., 1-5-1, Chofugaoka, Chofushi, Tokyo, 182 Japan, kkishi@peacock.ee.uec.ac.jp)

The phrase ''con sordino'' is used for the score of a violin family to indicate the use of a mute. The mute is used to reduce the volume of the violin tones and change their tone quality. This paper examines the influence of the weight of mutes on tones of a violin using the spectral envelopes of the violin tones. Violins with and without mutes were played by a skilled player in an anechoic chamber to produce tone samples which were then recorded on a DAT recorder. Using the tone samples, the time variances of the fundamentals and spectral envelopes obtained by FFT analyses were examined. The force transfer function of a cello bridge with mutes installed, which is used for convenience, was measured. From these examinations and analyses, we made the following observations: (1) the higher the order of harmonics, the lower their relative sound pressure levels compared with that of the fundamental and (2) the level differences between the peaks and dips of the spectral envelopes are reduced as the weight of a mute is increased.

10:55–11:05 Break

11:05

3aMU5. Anomalous low frequencies from a bowed violin string. Roger J. Hanson (Phys. Dept., Univ. of Northern Iowa, Cedar Falls, IA 50614, roger.hanson@uni.edu) and Frederick W. Halgedahl (Univ. of Northern Iowa, Cedar Falls, IA)

With a bow force greater than the Schelleng maximum and careful bow control, it is possible on a violin to produce sounds of definite pitch ranging from approximately a musical third to as much as two octaves below the normal pitch. [Hanson *et al.*, CAS J., **2**, Series II, 1–7 (1994)]. The pitch is in agreement with the fundamental frequency of the observed harmonic series. The fundamental itself is very weak if the sounds are produced on the open violin G string. These anomalous low frequencies (ALF) occur when the bow force is great enough to prevent the Helmholtz kink from triggering the normal release of the string from the bow hair. The bow serves as a quasitermination point of the string so that there are pronounced bow-nut and bow-bridge reflections yielding at the bow a very complex waveform, some portion of which triggers the slipping of the string. The ALF can also be produced on a bowed monochord where the motion at several points is detected optically, and the sound is generated from the resulting signal. Some musical applications of the effect will be discussed.

11:30

3aMU6. Acoustical aspects of chamber music for strings. Juergen Meyer (Physikalisch-Technische Bundesanstalt, D-38116 Braunschweig, Germany, juergen.meyer@ptb.de)

Playing string quartet music under different room acoustics conditions demands adaptation of the playing technique, particularly of the articulation. If not damped by the bow, violin tones have a 60-dB decay time of about 0.5 s in the highest part of the range, increasing to about 1 s for the lowest tones; in the lower stringed instruments, decay times are as much as 3 s. When damped by the bow, the decay times can be shortened to about 0.1 s. Attack times of staccato violin tones are of the order of 30–50 ms; smoothed attacks may be prolonged up to 200 ms. The temporal structure of a fast staccato tone sequence (14 tones per second) is mainly characterized by breaks in the overtones of the order of 25 ms, whereas the fundamentals seem to be blended together or slurred. These values correspond with typical acoustical criteria of historical and contemporary chamber music halls. To determine the reaction of musicians to the boundary conditions as well as the listeners' impressions, a test was performed by playing the same programs of short and different music examples during only 3 h in three halls having different room acoustic properties.

Contributed Papers

11:55

3aMU7. Radiation mechanisms of a violin with known modal characteristics. Lily M. Wang and Courtney B. Burroughs (Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804)

Multiplanar near-field acoustic holography (NAH) involves reconstructing the surrounding radiated sound field from measurements of the complex pressure on at least four intersecting planes enclosing a source [L. M. Wang and C. B. Burroughs, J. Acoust. Soc. Am. **100**, 2654(A) (1996)]. After a brief review of the multiplanar NAH technique, results from applying the algorithm to a bowed violin (Hutchins' SUS295) are shown for each open string. An open-frame mechanical bowing apparatus, which has been previously described, is used to excite the violin [L. M. Wang and C. B. Burroughs, J. Acoust. Soc. Am. **99**, 2502(A) (1996)]. Areas of significant radiation on the violin structure are identified. Measurements are also compared to results from modal analyses previously performed on the same instrument. [Work supported by NSF Graduate Fellowship, Bell Laboratories GRPW Grant, and AAUW Selected Professions Dissertation Fellowship.]

12:10

3aMU8. Behavior of real violin strings, mechanically bowed. Norman C. Pickering and Michael P. Kerr (J. D'Addario and Co., Inc., Farmingdale, NY)

Traditional methods of evaluating various types of strings for bowed instruments have relied almost exclusively on subjective impressions of expert players. In the real world, busy professionals rarely have enough time or objectivity to provide string designers and manufacturers with precise information. Furthermore, such opinions are affected by factors unrelated to the stings themselves, and are subject to change in ways that are often difficult to relate to the design details of the product being tested. The use of a well-designed bowing machine makes it possible to amass quantities of data on specific strings, enabling string makers to discover the characteristics that are favored by professional artists. Data is obtained for strings having cores of gut, synthetic polymers, monofilament, and stranded steel. By varying bow force, velocity, and distance from the bridge an envelope of usable combinations is accurately determined. Time for development of Helmholtz waveform partial by partial, average frequency variation as well as cycle-by-cycle period variation, and overall spectral distribution can all be determined with good repeatability. Driving strings magnetically instead of by bowing permits frequency determination of individual partials, an indication of string flexibility or lack of it. This work has led to new string designs, incorporating new materials and manufacturing processes.

WEDNESDAY MORNING, 24 JUNE 1998

CASCADE BALLROOM I, SECTION A (W), 9:15 TO 11:10 A.M.

Session 3aNSa

Noise and Architectural Acoustics: Low-Frequency Noise

Louis C. Sutherland, Chair
27803 Longhill Drive, Rancho Palos Verdes, California 90275-3908

Chair's Introduction—9:15

Invited Papers

9:20

3aNSa1. A new way of reducing self-sustained flow noise. Reni A. Henry, Anas Sakout, Alex L. Coiret (Lab. L.E.P.T.A.B., La Rochelle Univ., avenue Marillac, 17042 La Rochelle Cedex 1, France), and Aziz Hamdouni (Poitiers Univ., 86960 Futuroscope Cedex, France)

Vortex shedding from shear layers may be driven by the resulting acoustical field. In order to achieve noise reduction, the feedback loop is broken down with the help of auxiliary secondary flows that are created by means of minor geometrical modifications. Experiments have been conducted on a flow through two close orifice plates in a cylindrical duct: they have brought forward the possibility of reaching stable paths of flow bifurcations where an important noise reduction is obtained in a large range of flow rates. By creating a suitable interaction between the main shear layer of the confined jet and thin auxiliary shear layers, noise reduction takes place in a range of 20–40 dB. The latest result is the following: due to the stability of the flow bifurcation path that can be reached, such a level of noise reduction is maintained in a large range of flow conditions. Moreover, by choosing an adequate path, noise reduction takes place for flow conditions where intense tones would be present if these flow conditions were to be reached through conventional means. It may be emphasized that such methods are costless, compared to some sophisticated active control methods.

9:50

3aNSa2. A new approach to low-frequency room noise criteria. Robert D. Hellweg, Jr. and Hsien-sheng (Jason) Pei (Digital Equipment Corp., PKO2-1/J60, 129 Parker St., Maynard, MA 01754, hellweg@mail.dec.com)

There are three different and conflicting room criteria methods that are widely used by consultants today: room criterion (RC), balanced noise criterion (NCB), and noise criteria (NC). The RC procedure is documented in ANSI S12.2-1995 ''Criteria for evaluating room noise'' and the ''1995 ASHRAE Applications Handbook;'' the NCB procedure in ANSI S12.2; and the NC ''tangency'' method in the ''1991 ASHRAE Applications Handbook.'' There are significant differences in the three criteria, especially for low frequencies, which can impact the costs both of new designs and of remedial actions to existing facilities. In order to promote a resolution to these differences, NOISE-CON96 and NOISE-CON97 held special technical sessions on room noise criteria. Based on our analysis of published research and on experiences dealing with other noise sources, the authors suggest an approach to resolve the low-frequency differences: (1) establish basic room criterion curves for noises without abnormal temporal or spectral characteristics, and (2) apply correction factors to the measured data to account for abnormal features before comparison to the basic criterion. This paper presents details of the suggested approach, the status of revisions to ANSI S12.2 and to the ASHRAE handbook, and other ongoing research.

10:20

3aNSa3. Parameters involved in low-frequency noise annoyance. Ulf Landstrom (Natl. Inst. for Working Life, Box 7654, S-907 12 Sweden)

Noise annoyance and related adverse responses evoked during exposures in low-frequency environments are effected by a number of separate and interacting parameters. Laboratory as well as field studies have shown that in most cases the correlations between annoyance and noise levels or linear values are very weak. For descriptions of the fundamental correlations between low-frequency noise exposure, annoyance, and related adverse effects, deeper knowledge about the different interacting factors is needed. One consequence of this deficient knowledge of the critical factors for low-frequency noise annoyance is that well-founded methods for the evaluation of annoyance and related adverse effects are still missing. Noise control, building and rebuilding of new acceptable environments, or other types of measures against the low-frequency noise problems cannot be based only on evaluations of the dBA or other simple rating alternatives. For adequate evaluations and measures, deeper knowledge about interacting factors, e.g., tones, frequency character, bandwidth, exposure time, fluctuations, perception, masking and type of work, is needed. The present paper is a summary of the results and conclusions that could be drawn from studies directed upon some of the most important parameters involved in low-frequency noise annoyance.

Contributed Paper

10:50

3aNSa4. The effects of age on annoyance caused by low-frequency noise. Rukhsana Adams and Bridget M. Shield (School of Eng. Systems and Design, South Bank Univ., London SE1 0AA, UK, shieldbm@sbu.ac.uk)

A laboratory study into the response of subjects to noises of varying intensity and frequency content has been carried out. Eighty subjects ranging in age from approximately 20 to 70 yr were asked to rate sounds according to annoyance. The hearing levels of the subjects were tested using pure-tone audiometry. The annoyance ratings have been analyzed to investigate relationships between annoyance and spectral shape, and also to determine whether the annoyance responses are related to age or hearing acuity. Preliminary results suggest that there is a strong relationship between age and response to low-frequency sound, subjects over the age of 45 finding sounds with high levels of noise below 200 Hz significantly more annoying than younger subjects. However, the annoyance ratings of the older subjects do not appear to be related to hearing levels, suggesting that the response may be due to some other factor such as the association of the noise.

WEDNESDAY MORNING, 24 JUNE 1998

CASCADE BALLROOM I, SECTION B (W), 9:15 TO 11:40 A.M.

Session 3aNSb

Noise: Tire Noise

Paul R. Donavan, Chair

Noise and Vibration, MC 483-324-101, GM Proving Ground, Building 24, Milford, Michigan 48380

Chair's Introduction—9:15

Invited Papers

9:20

3aNSb1. The investigation of a towed trailer test for passenger tire coast-by noise measurement. James K. Thompson (Automated Anal. Corp., 2805 South Industrial, Ste. 100, Ann Arbor, MI 48104, jthompso@autoa.com) and Thomas A. Williams (Hankook Tire, 3535 Forest Lake Dr., Uniontown, OH 44685)

It is difficult to quantify the portion of coasting vehicle noise that is due to tire-pavement interaction alone. There are often contributions from aerodynamic noise of the vehicle, transmission whine, noise from suspension components, and other miscellaneous sources. The towed-trailer method used in the revised SAE J-57 standard has been shown to be an effective means of isolating tire-pavement noise for truck tires. This paper reports the results of a test program conducted by SAE Tire Noise Standards Committee to evaluate the feasibility of towed-trailer coast-by testing of passenger and light truck tires. The results of tests conducted in April 1996 at the Ohio Transportation Research Center are described and they indicate that accurate measurements are possible for towed-trailer testing of passenger tires. It is shown that a key aspect of performing such a test is reducing the noise of the tow vehicle and that sufficient reduction is possible even for quiet test tires. During the period of this work, the ISO/TC 31/WG 3 working group developed a draft trailer test method for evaluating tire noise. This test procedure was also evaluated. The results show the proposed test method to be highly accurate for tire noise evaluation.

9:40

3aNSb2. Assessment of tire/pavement interaction noise under vehicle passby test conditions using sound intensity measurement methods. Paul R. Donavan, Richard F. Schumacher, and Jeffrey R. Stott (Noise and Vib. Ctr., 483-324-101, General Motors Proving Ground, Milford, MI 48380)

Over the past several years, sound intensity measurement methods have become an increasing valuable tool in isolating tire/pavement interaction noise when a vehicle is tested under full throttle acceleration conditions such as the ISO 362 R15 procedure. Several investigations have been conducted and reported which demonstrate the relationship between ''on-board'' sound intensity measured close to a tire contact patch and the sound pressure level measured by a stationary microphone 7.5 m away from the line of travel of the vehicle. Using these relationships, the contribution of tire/pavement noise can be assessed relative to other noise sources associated with a vehicle under acceleration as measured at 7.5 m. In this application, it has been determined that some tires can produce significantly higher noise levels under the torque of acceleration than under cruise conditions. Sound intensity has also been used to separate sound propagation from sound generation effects in the assessment of test surfaces such as those specified by SAE and ISO. This paper reviews the various applications of sound intensity in the assessment of tire/pavement interaction noise issues related to vehicle passby noise, the implication of recent findings on noise reduction strategies, and the potential for standardization of sound intensity techniques.

10:00

3aNSb3. The influence of belt and tread band stiffness on the tire noise generation mechanisms. Wolfgang Kropp, Krister Larsson, and Stephane Barrelet (Dept. of Appl. Acoust., Chalmers Univ. of Technol., S-41296 Gothenburg, Sweden, wk@Ta.chalmers.se)

A brief introduction to a theoretical model is presented which describes a smooth tire rolling on a rough road surface. The model is limited to a two-dimensional contact between tire and road. It includes the vibration properties of the tire, the nonlinear contact, and the sound radiation. It is also taking into account the influence of the acoustic properties of the road surface. The model is applied to determine the sound radiated due to the tire vibrations and the local deformations of the tread band. A parameter study is presented which shows the influence of the bending stiffness of the belt and the stiffness of the tread band on the noise generation and radiation mechanisms. The study also considers the influence of the surface roughness of the road. In this way it is possible to quantify which part of the roughness spectrum is leading to sound generation via tire vibration and which part of the roughness spectrum is involved in the local deformation of the tread band.

10:20

3aNSb4. Identification of sources of tire/pavement interaction noise. Richard J. Ruhala and Courtney B. Burroughs (Grad. Prog. in Acoust., Penn State Univ., State College, PA 16804)

Although most highway noise is due to radiation from rolling tires, the dominant sources of tire/pavement interaction noise remain largely unknown. In this study, near-field acoustical holography (NAH) is used to visualize the acoustic field of a rolling tire on outdoor pavement surfaces. Since NAH measurements are conducted in the near field, areas of maximum radiation from the tire may be localized with a much finer resolution than with conventional measurements. Planar NAH is used to scan planes along the side, front, and rear planes of the tire. A scanning mechanism will be attached to an automobile trailer for measurements of tire noise under realistic conditions and procedure as developed in earlier results [R. J. Ruhala and C. B. Burroughs, SAE 972047 (1997)]. Several types of tires are evaluated, including blank tire, tire with even spaced transverse grooves, tire with only circumferential grooves, and a production passenger car tire. The tires are tested on both smooth and rough pavements, and at several speeds. The results will provide further insight into the sources and mechanisms of tire/pavement interaction noise. [Work supported by Goodyear, DOT, and GM.]

Contributed Papers

10:40

3aNSb5. Effects of the acoustical characteristics of road pavements on long-range sound propagation. Michel C. Bérengier (L.C.P.C. - Ctr. de Nantes, BP 19, 44340 Bouguenais, France), Jean-François Hamet (I.N.R.E.T.S. 25, 69675 Bron cedex, France), and Yves Pichaud (L.C.P.C.-Ctr. de Nantes, 44340 Bouguenais, France)

For many people in industrial countries, traffic noise is considered as one of the most important elements of discomfort. In the past 15 years, the only protections for the neighboring populations against road traffic noise were road barriers. For a short time, the modification of the road pavement can partly contribute to decrease this nuisance. Recent research on the acoustical performances of road pavements have been carried out, and, a first ranking in the close proximity of the road has been established using a standardized method. From these first results, it is possible to predict the road traffic noise far from the road taking into account the acoustical characteristics of various pavements. This prediction procedure requires the theoretical knowledge of all the physical phenomena occurring along the propagation path. This is the reason why the model used for calculation integrates ground effects over mixed surfaces and atmospherical parameters such as vertical sound-speed gradients. The pavement effect is introduced in the model through the sound power spectrum of the tire/road noise source located a few centimeters above the road surface. The predicted results are advantageously compared to *in-situ* measurements carried out over some hundred meters.

10:55

3aNSb6. Acoustic modeling of road vehicles for traffic noise prediction: Determination of the sources heights. Jean-François Hamet, Marie-Agnès Pallas (I.N.R.E.T.S. 25, av. François Mitterand, case 24, 69675 Bron cedex, France), David Gaulin, and Michel C. Bérengier (L.C.P.C.-Ctr. de Nantes, 44340 Bouguenais, France)

For traffic noise predictions road vehicles can be modeled by point sources. When dealing with the evaluation of noise barrier's efficiency or of far-field noise, the knowledge of the sources heights is of primary importance. Their determination can be obtained from two different methods. The first one deals with an inverse fitting procedure performed on experimental excess attenuation results between two receivers 7.50 m apart. Knowing the noise frequency spectrum at the reference microphone and the pavement and close ground impedances, it becomes possible to adjust the source height by minimizing the prediction/measurement differ-

ence. By this procedure, the equivalent noise source is found very close to the surface. Otherwise, sources heights can be obtained directly using microphone arrays. These measuring techniques are now commonly used to localize sound sources of railway vehicles. Their application to road vehicles appears to be more recent. For the present case dealing with the determination of the source height only, a simple vertical array can be implemented. Preliminary measurements made on a light vehicle confirm that the acoustic radiation due to the various sources (including engine noise) comes practically entirely from the car underbody, i.e., from a zone close to the ground.

11:10

3aNSb7. Modeling of tangential contact forces. Krister Larsson, Stephane Barrelet, and Wolfgang Kropp (Dept. of Appl. Acoust., Chalmers Univ. of Technol., S-412 96 Gothenburg, Sweden, kl@ta.chalmers.se)

The main noise source from vehicles at speeds greater than 50 km/h is the tires. To reduce traffic noise it is therefore necessary to understand the noise generation mechanisms in the contact between tire and road. The paper presents an improvement of an earlier developed acoustic rolling model concerning a smooth tire rolling on a rough surface. This model consists of three steps. First, the radial contact forces between road and tire are calculated. These forces give the vibrations of the tire, which finally give the sound radiation from the tire. A method to include the tangential forces into the rolling model is the topic of this paper. To do this, the influence of the tread blocks on the force transmission from the road to the belt of the tire has to be known. A FEM-model of one tread block is made in order to determine its dynamical properties. A simplified model of the block is established, by only taking into account the two lowest modes in the frequency range of interest (i.e., up to 3 kHz). This simplified model is included in the existing contact model to give the response of the tire due to tangential or radial excitations.

11:25

3aNSb8. Computational and subjective procedures for the assessment of sounds with weak tonal components. Wolfgang Ellermeier (Inst. of Psychol., Univ. of Regensburg, Regensburg, Germany) and Matthias Vormann (C.v.O Univ., Oldenburg, Germany)

A discrete tone is defined as ''prominent'' if the level of the tone exceeds the level of the noise contained in its critical band by 6 dB. Tonal or narrow-band noise components of tire sounds often fall short of this level, but are nevertheless rated as prominent by test drivers. Tonality is closely related to the sensation of pitch strength, which depends upon the sensation level, bandwidth, duration, and frequency of the tonal components present. In order to compare a number of computational procedures for calculating the tonality of sounds, a broad range of tire noises, synthesized tire sounds, and artificial tone–noise mixtures in which the level and bandwidth of the tonal components was varied, were subjectively assessed using (a) a paired-comparison technique, and (b) categorical judgments collected from about 60 listeners. They were objectively assessed by tone-to-noise and prominence-ratio measurements, and by a new method based on a model of pitch sensation which accounts for the temporal and frequency resolution as well as for masking properties of the human ear. The advantages and limits of the different methods will be compared and discussed. Especially for weak narrow-band components, the pitch model seems to be superior to the other methods used to quantify the degree of tonality.

WEDNESDAY MORNING, 24 JUNE 1998

CASCADE BALLROOM I, SECTION C (W), 9:15 A.M. TO 12:05 P.M.

Session 3aNSc

Noise and Engineering Acoustics: Occupational Noise Exposure Monitoring: Instrumentation, Methods, and Histories

John P. Seiler, Chair

Mine Safety and Health Administration, Cochran Mill Road, Building 028, P.O. Box 18233, Pittsburgh, Pennsylvania 15236

Chair's Introduction—9:15

Invited Papers

9:20

3aNSc1. Instrumentation for the measurement of occupational noise exposure: Past, present, and future. Theodore J. Kuemmel (Quest Technologies, Inc., 510 S. Worthington St., Oconomowoc, WI 53066)

The instruments used to measure occupational noise exposure have all evolved from the basic sound level meter. Variation of the sound fields and the dynamics of the measurement environment raised questions about the validity and accuracy of occupational noise exposure measurements. Measurement practice and instrumentation have evolved significantly to overcome the limitations inherent in the measurement of sound levels. Modern microprocessor based instruments which automatically compute noise exposures have dramatically improved the validity and accuracy of occupational noise exposure measurements. Today many of the measurements of occupational noise exposure can be reliably replicated. This paper addresses the challenges faced when attempting to assess occupational noise exposure, and their influence on the development of occupational noise exposure measurement instrumentation. The influence of technology, regulation, and standards on instrument development is also addressed.

9:35

3aNSc2. Occupational noise measurement standard: ANSI S12-19. Richard Goodwin (Acoust. Systems, Inc., P.O. Box 146, Vandalia, OH 45377, asi898-3198@worldnet.att.net)

This standard provides procedures for the measurement of occupational noise exposure from all types of noise; e.g., continuous, fluctuating, intermittent, and/or impulse/impact. Measurements may be reported as sound level with corresponding duration, time-weighted average sound level, and/or noise dose. This standard provides for the measurement of the noise exposure of individuals and can be extended to represent groups performing similar activities. It can also be used to measure the noise exposure from a given job or activity. This standard does not provide procedures for the measurement of occupational noise exposures attributable to the rise of earphones or telephone receivers. This standard presents recommended practices for utilizing instrumentation such as a sound level meter, noise dosimeter, integrating sound level meter, and acoustical calibrators. Guidelines are provided for measurement conditions such as the acoustical environment, measurement activities, and operational variations. The parameters surrounding the noise and its impact on the selection of instrumentation is presented. Specific guidelines for using a sound level meter, a noise dosimeter, and an integrating sound level meter are presented. Guidelines for documentation and reporting of the measurement of occupational noise exposure are specified.

9:50

3aNSc3. What are we still doing wrong in assessing occupational noise exposure? Daniel L. Johnson (Larson-Davis, Inc., 1681 West 820 North, Provo, Utah 84601)

Sound level meters/dosimeters and measurement/prediction procedures have improved greatly over the last 20 years. Instrumentation nowadays can be expected to assess the energy of impulsive noise properly, provide a time history of the noise exposure, and have a sufficient dynamic range. Nevertheless, some improvements are still needed. These improvements are to (1) rework the SLM standards so the measurement envelope for all instruments, regardless of type, can be specified to be the same; (2) provide better methodology for determining when infrasound and ultrasound need to be measured; (3) provide a mechanism for properly assessing intermittent exposures, including the inclusion of effective quiet; and (4) finally, as a visionary dream, provide long-term measurement of occupational/nonoccupational noise exposure coupled with environmental factors as well as health factors such as blood pressure and hearing threshold levels. The need for and the details of each of these improvements will be discussed.

10:05

3aNSc4. Types and frequencies of abnormal sound levels when using noise dosimeters to establish worker TWAs. Larry H. Royster (Dept. of Mech. Eng., North Carolina State Univ., Raleigh, NC 27695-7910) and Julia D. Royster (Environ. Noise Consultants, Inc., Raleigh, NC 27622-0698, effective_hcps@compuserve.com)

The potential contamination of noise dosimetry noise exposure findings as impacted by the following types of artifacts was investigated: thumping of the microphone, moving the microphone around during the sample, hollering directly into the microphone, effects of body reflections, and blowing directly at the microphone. The findings from this investigation support the assumption that contamination of noise dosimetry daily noise exposure results is not of significant concern, at least when using a 5-dB exchange rate (OSHA). In addition, the detailed histogram records obtained from conducting noise dosimetry noise exposure surveys at five large industrial facilities were reviewed. Of the over 28 000 histogram samples reviewed, less than 14 samples indicated potential contamination from unusual noise sources. In summary, the results of this study confirm what the authors have observed over many years of conducting industrial noise exposure studies using noise dosimetry: contamination of noise dosimetry findings is not a significant issue (at least when a 5-dB exchange rate is utilized).

10:20

3aNSc5. Sound exposure profiling: A noise exposure assessment alternative. Lee D. Hager (James, Anderson & Assoc., Inc., P.O. Box 23113, Lansing, MI 48909, lee_hager@compuserve.com)

Assessment of worker sound exposure based on incremental task sound levels has a history dating back to the earliest days of recognition of noise as a hearing hazard. Sound exposure profiling (SEP) uses current technology to compile task-based sound measurements to model long-term risk of exposure to hazardous noise. SEP is an alternative to full-shift dosimeter techniques for sound exposure assessment. The methodology offers the advantage of providing additional information about noise sources and how worker interaction with noise sources results in long-term risk of exposure. SEP has been adopted by several major corporations as their corporate strategy for assessment of worker noise exposure, and is supported in ANSI S12.19-1996, ''Measurement of Occupational Noise Exposure.'' This paper will described SEP and its application in industry.

10:35

3aNSc6. Unusual jobs require unusual methods for noise exposure measurement. John P. Barry (U.S. Dept. of Labor—OSHA, 3535 Market St., Philadelphia, PA 19104, john.barry@internetmci.com)

With the diversity of the American workplace, the one size fits all approach cannot accommodate every occupational noise exposure measurement challenge. This is precisely why ANSI Standard S12.19-1996, ''Measurement of Occupational Noise Exposure,'' permits different measurement criteria and a variety of instrumentation and methodologies. The Occupational Safety and Health Administration (OSHA) and the Mine Safety and Health Administration (MSHA) use employee-worn noise dosimeters almost exclusively to document alleged violations of the occupational noise standards. However, employers and consultants may use any method that properly identifies those employees who need to be included in a hearing conservation program and that relates the noise exposure with the risk of occupational hearing loss and the effectiveness of hearing protection devices. Some environmental and

operational hearing loss and the effectiveness of hearing protection devices. Some environmental and operational problems that affect occupational noise measurement are temperature, humidity, rain, wind, radio, magnetic and microwave interference, intrinsic safety, and the presence of hazardous materials. To illustrate effective handling of these measurement problems, the methodologies used in the occupational noise exposure assessment of operating municipal departments will be discussed.

10:50–11:05 Break

Contributed Papers

11:05

3aNSc7. Investigation of intermittency in the noise exposure pattern of longwall coal miners. J. Alton Burks and Roy C. Bartholomae (Pittsburgh Res. Lab., Nat. Inst. for Occupational Safety and Health, P.O. Box 18070, Pittsburgh, PA 15236, aib5@cdc.gov)

It is generally recognized that intermittent exposures to noise are less hazardous than continuous exposures with the same energy. This effect has been accounted for in the USA with the use of a 5-dB exchange rate, instead of the 3-dB rate associated with the equal energy principle. However, the appropriateness of the 5-dB exchange rate for assessing the risk of noise exposures in mining has never been examined closely. The authors have previously reported the results of a worker noise exposure survey conducted at six longwall coal mines [R. C. Bartholomae and J. A. Burks, Proceedings: NOISE-CON '94, 1017–1022 (1994)]; the time-resolved dosimeter data from that study were reanalyzed to determine if the noise exposure pattern of longwall miners can be categorized as intermittent. For this purpose *effective quiet* is defined to be a period of time in which the sound level is less than 75 dBA [W. Melnick, J. Acoust. Soc. Am. **90**, 147–154 (1991)] for 10 s or longer [K. D. Kryter, *The Effects of Noise of Man* (1985), p. 273]. It is found that the miners experience an average of 125 periods of effective quiet and an average total quiet time of 1.7 h on a daily basis.

11:20

3aNSc8. Hearing hazard evaluation and hearing conservation in a very reverberant indoor firing range. Felix Z. Sachs (U.S. Army Ctr. for Health Promotion and Preventive Medicine, 5158 Blackhawk Rd., APG, MD 21010-5422, felix_sachs@chppm-ccmail.apgea.army.mil)

OSHA and Army regulations differentiate steady from impulse noise exposure, but the boundary is fuzzy. A 300-m research indoor firing range produced 10-s reverberation times resulting in noise exposures straddling this boundary. Worker exposure was assessed as both steady and impulsive under both OSHA and Army criteria. The highest peak level was 164.9 dBP, the maximum ''slow'' level was 133 dBA. Hearing protector calculations for steady noise used the NIOSH ''assumed protection values'' method. Army noise criteria use 3-dB exchange between level and time. Any parallel filter analyzer could measure the time-averaged octave band levels. OSHA uses 5-dB exchange rate, a rare feature on parallel filter analyzers. Administrative controls were most stringent for the Army steady noise method and least restrictive for the MIL-STD-1474 impulse noise method. For OSHA steady noise, allowable rounds per day depended on the number of shots fired per burst due to the 5-dB exchange rate. [The opinions or assertions contained herein are the views of the author and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.]

11:35

3aNSc9. Noise from sandblasting. Tonnes Ognedal and Rune Harbak (Sinus AS, Sandvigaa 24, N-4007 Stavanger, Norway)

With respect to the increased focus on the working environment, a project has been carried out in order to investigate the noise exposure for personnel working with sandblasting and UHP-water blasting. For sandblasting, three different aspects that influence the noise exposure have been evaluated: (a) Noise from the nozzle and the blasting object, (b) noise generated by the air supply in the masks/helmets, and (c) sound attenuation of the masks and helmets. Normal A-weighted noise levels outside the protective gear are between 110 and 120 dB while sandblasting. The noise level inside the mask and the inner hood will vary from 95–105 dB from the sandblasting. The noise generated by the air supply can be as high 90–100 dB inside the inner hood. Thus the total average noise exposure for the personnel (without earplugs) is 97–107 dB with today's most common equipment in Norway. Noise from UHP-water blasting can also reach levels up to 110–120 dB, depending on the blasting object and the acoustic environment. Normal working gear gives less protection to the ear than the sandblasting gear. Both the sandblasting and the UHP water blasting may create loss of hearing to the workers.

11:50

3aNSc10. Occupational noise and solvent exposure in the printing industry. Rande Ferris and Beno Groothoff (DTIR, Queensland Govt., Brisbane, Australia 4030, rande.ferris@dtir.qld.gov.au)

The study examined the effects on hearing from occupational noise and solvent exposure in the printing industry and was conducted by measuring a selection of workers' daily noise dose, sampling personal solvent exposures and audiometric testing before and after shifts for the detection of a hearing threshold shift. Audiometric test results were inconclusive due to the inadequate testing areas provided at each workplace. Personal monitoring confirmed that where solvent exposure occurs, high noise levels are usually also present. Fifty-three percent of the noise dose level measurements exceeded the statutory limit $L_{\mathrm{Aeq,8h}}$ of 85 dB and a further 9% exceeded the limit of 140 dB (L_{in}). Twenty-two percent of the analyzed solvent exposure samples revealed benzene, with 5% exceeding the worksafe exposure standard of 16 mg/m^3. The lack of essential information contained in Material Safety Data Sheets (MSDSs) for substances used in the printing industry was highlighted. Accurate MSDSs are required to assist employers in discharging their obligation to provide correct information about hazardous substances to employees under the Workplace Health and Safety Act of 1995, as well as the hazards associated with those substances. Better information is required in an MSDS, from manufacturers and suppliers, as to the specific properties and health hazards of a hazardous substance.

Session 3aPAa

Physical Acoustics: Sonochemistry and Sonoluminescence: SC I

Kenneth S. Suslick, Cochair
Department of Chemistry, University of Illinois, 601 South Goodwin Street, Urbana, Illinois 61801

Jacques L. Reisse, Cochair
Université Libre de Bruxelles, Chimie Organique, CP 165, Avenue F. D. Roosevelt 50, 1050 Brussels, Belgium

Chair's Introduction—7:55

Invited Papers

8:00

3aPAa1. Under which conditions is sonochemistry able to give information about the so-called extreme conditions prevailing inside the collapsing bubbles? Jacques Reisse, Kristin Bartik, and Nicolas Segebarth (Univ. Libre de Bruxelles, Chimie Organique, CP 165, 50, ave. F. D. Roosevelt, 1050 Bruxelles, Belgium)

Sonochemistry is able to give information about the extreme conditions prevailing inside collapsing bubbles. However, this information is obtained only when the experimental data correspond to primary chemical steps taking place in the gas phase inside the bubbles. This is unfortunately often not the case: The experimental data correspond, in the majority of cases, to secondary reactions taking place either in a shell around the collapsing bubble or in the bulk liquid. The short introductory talk to the sonochemistry session will be devoted to trying to answer the following questions: How must reactions, able to give relevant information about the so-called extreme conditions inside the bubbles, be selected? What is the actual reaction zone inside the bubble and what are the temperature limits for this zone? Is it certain that the ''active zone'' for sonochemistry is always the same as the ''active zone'' for sonoluminescence? Why does sonochemistry involve radicals instead of ions?

8:20

3aPAa2. Sonochemical preparation of protein microspheres. Kenneth S. Suslick, Kenneth J. Kolbeck, Gregory S. Kufner, and Gregory W. Szewczyk (Dept. of Chemistry, Univ. of Illinois, 601 S. Goodwin Ave., Urbana, IL 61801)

Aqueous suspensions of protein microspheres with Gaussian-like size distributions in the micron size range can be produced with high-intensity ultrasound by the sonochemical oxidation of cysteines to form interprotein disulfide cross links. The microspheres can be synthesized containing a core of aqueous or organic liquids (i.e., microcapsules) or air (i.e., microbubbles such as Albunex). The micron sized spheres are contained by a 60-nm layer of protein (approximately ten protein molecules.) The small diameter (less than 5 μm) of the microspheres permits the microspheres to circulate through the bloodstream, while sonochemically linking native protein to form microspheres prevents the expected *l* immune response. To date, we have synthesized microspheres that can be monitored with *in vivo* imaging techniques (MRI and x-ray) as well as microscopy. The distribution and circulation lifetime in the bloodstream has been altered by modifying the properties of the microsphere. These microspheres have many potential biomedical applications such as drug delivery vehicles, blood substitutes, and radiological imaging agents.

8:40

3aPAa3. Sonochemistry of aqueous solutions: EPR and spin trapping studies of radical intermediates. Peter Riesz and Vladimir Misik (Natl. Cancer Inst., NIH, Bethesda, MD 20892-1002)

The primary free radical species produced in the sonolysis of aqueous argon-saturated solutions (i.e., H, OH, and O_2-) and in N_2-containing aqueous solutions (H, OH, O_2-, NO) can be identified by electron paramagnetic resonance (EPR) combined with spin trapping. The EPR experiments using cadmium(2+) ions as hydrated electron scavengers show that no detectable level of hydrated electrons is formed in the sonolysis of argon-saturated water at neutral *p*H. The temperature dependence of the semi-classical treatment of the kinetic isotope effect for H and D formation by O–H and O–D bond scission in 1:1 H_2O–D_2O mixtures was used to estimate the effective temperatures (2000–4000 K using phenyl *t*-butyl nitrone spin traps) in collapsing cavitation bubbles. At low concentrations of nonvolatile solutes (e.g., acetate ions, amino acids, and sugars) only radicals formed by H-abstraction were observed. At high concentrations, pyrolysis radicals such as methyl radicals from acetate, amino acids, and sugars could also be spin trapped. For volatile organic solutes, the effect of the vapor pressure of the solute on the effective value of gamma (the ratio of C_p/C_v) in the cavitation bubble is the major determinant of radical yields.

9:00

3aPAa4. Sonoluminescence and sonochemistry in aqueous solutions containing surface active solutes. Franz Grieser, Muthupandian Ashokkumar, and Katrina Barbour (AMPC, School of Chemistry, Univ. of Melbourne, Parkville, 3052 Australia)

The ultrasound induced reduction of gold chloride in aqueous alcohol solutions has been measured and compared with the extent of sonoluminescence (SL) quenching observed in the same solutions, as a function of alcohol concentration. Both ultrasound initiated phenomena were found to be directly related to the Gibbs surface excess concentration of the alcohol at the gas/water interface. The SL quenching occurs in two distinct surface excess regimes. The first region accounts for about 90% quenching of the SL signal. The second quenching region extends over a broader surface excess range and the profile matches the extent of enhancement in the reduction of gold chloride. When the nonvolatile surfactant sodium dodecylsulfate (SDS) was used instead of alcohol, the trends in the SL signal and in the sonochemical reduction process were quite different. At low concentrations (<2 mM) of SDS, the SL signal was enhanced over that obtained from water without surfactant. At higher SDS concentrations, the SL signal declined and leveled off to about 80% of the water signal at about 10 mM. In contrast, the reduction of gold chloride increased to about 8 mM SDS and then remained constant above this value. Possible mechanisms that explain the results will be presented.

9:20–9:40 Break

9:40

3aPAa5. Sonochemical reactor optimization using computational acoustics techniques. Jean-Louis Y. Migeot (LMS Numerical Technologies, 70 Interleuvenlaan, 3001 Leuven, Belgium)

The designer of a sonochemical reactor is confronted by a complex problem characterized by a large number of design parameters: reactor shape and dimensions, number, positions and power of ultrasound sources, frequency, solvant, temperature, etc.? A wide parametric study cannot be conducted experimentally due to time and cost constraints and a numerical model seems the only viable approach. Such a model must reflect the inherent complexity of the system and at the same time be simple enough to provide relevant results in a reasonable amount of time. This paper proposes an iterative modeling methodology based on the following steps: (1) a finite-element model of the fluid in the reactor, (2) a relationship between local acoustic pressure and gas fraction released by cavitation, and (3) equivalent acoustic properties for the cavitating medium. After a few iterations, this model leads to two main results: (1) the pressure field in the reactor and (2) the cavitation zone and the distribution of gas fraction in this cavitation zone. One can further postprocess the pressure field to predict the streaming forces on the fluid and therefore predict the flow pattern using classical CFD techniques.

10:00

3aPAa6. Sonochemistry—A demonstration lecture. Timothy J. Mason (Sonochemistry Ctr., School of NES, Coventry Univ., Coventry CV1 5FB, UK)

The potential uses of sonochemistry for the processing and chemical industries of the future have been established already in the chemical laboratory. The driving force for this technology is acoustic cavitation and a deeper understanding of this phenomenon will be required to help sonochemistry achieve more widespread acceptance and implementation in industry. For this reason a coordinated research effort is needed involving not only chemists but engineers, physicists, and mathematicians. In the past interaction between these disciplines has been limited due to the different scientific circles in which they move, but a few conferences, including this one, have sought to bring such disciplines together to promote cooperation. This presentation incorporates a series of demonstrations which illustrate the types of physical and chemical effects of cavitation and ultrasound which interest the chemist. These will include degassing, the generation of free radicals, the effects of cavitation near a surface, the dispersion of solids, emulsification, and particle manipulation in an acoustic field. Hopefully these will serve not only as an introduction to sonochemistry for those whose main discipline is not chemistry, but also as a focus for the powerful effects of cavitation which need deeper coordinated exploration.

Contributed Papers

10:20

3aPAa7. The effects of organic compound doping in single-bubble sonoluminescence. Muthupandian Ashokkumar, Franz Grieser (AMPC, School of Chemistry, Univ. of Melbourne, Vic 3052, Australia), William B. McNamara, Kenneth S. Suslick (Univ. of Illinois—Urbana, IL 61801), Thomas J. Matula, C. Allen Frensley, and Lawrence A. Crum (Univ. of Washington, Seattle, WA 98105)

Controlled experiments have been conducted to ascertain the effects of doping with organic compounds on single-bubble sonoluminescence (SBSL) in water. Previous experiments with multi-bubble sonoluminescence (MBSL) indicate that millimolar additions of alcohols quench the sonoluminescence intensity [Ashokkumar *et al.*, J. Phys. Chem. (in press)]. For a given concentration of alcohol, the degree of quenching apparently depends on the hydrophobicity of the alcohols. Similar results are observed in SBSL, using Ethanol, 1-Propanol, and 1-Butanol [intensity quenching measurements of 1-Butanol agrees with those of Weninger *et al.*, J. Phys. Chem. **99** (1995)]. Spectra and radius-time measurements are obtained for these systems. Further experiments with weak organic acids and bases, which change their neutrality with *p*H, are also being conducted. The results and possible interpretations will be discussed. [This research supported by DOE, NSF, and the ARC.]

10:35

3aPAa8. An interpretation of the MBSL temperatures deduced from metal atom emission. Lawrence S. Bernstein (Spectral Sci., Inc., 99 South Bedford St., Burlington, MA 01803, larry@spectral.com)

Recent MBSL observations [McNamara *et al.*, Univ. of Illinois (private communication)] of metal atom emission from a variety of metal carbonyls seeded into Ar display similar peak emission temperatures of approximately 5000 K. Additionally, metal ion emission lines have not been detected. These observations can be interpreted to mean that the peak emission temperature is also the peak collapse temperature. The same conclusion was previously reported based on the same 5000 K temperature observed for C2 MBSL. However, it was argued [Bernstein *et al.*, J. Phys. Chem. **100**, 6612 (1996)] that the chemistry giving rise to C2 emission

turns on and off in a narrow temperature window around 5000 K. Thus it was concluded that C2 emission is not a useful thermometer for higher collapse temperatures which may be attained in MBSL. It is argued that a similar situation exists for the metal atom emission chemistry. Detailed calculations based on a coupled chemical–hydrodynamic collapse model support this hypothesis. The lack of observed metal atom ion lines is attributed to the absence of sufficiently strong lines in the observational spectral window of the experiments.

Session 3aPAb

Physical Acoustics: Nonlinear Acoustics I: 1. History; 2. Solids; Rocks; 3. Surface Waves, Part 1

Lev A. Ostrovsky, Cochair
NOAA/ERL/ETL, 325 Broadway, R/E/ET-1, Boulder, Colorado 80303

Mack A. Breazeale, Cochair
National Center for Physical Acoustics, Coliseum Drive, University, Mississippi 38677

Invited Papers

7:45

3aPAb1. Development of solid state nonlinearity. M. A. Breazeale (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677)

The nonlinear acoustics of solids developed from the nonlinear acoustics of fluids. Originally, the problem was to undo correctly the Hooke's Law assumption without losing nonlinear terms. At present, one problem is to find a use for nonlinear acoustics of solids. The first problem has been solved; the second one is an ongoing one. It is the reason for special sessions. To develop geological exploration based on nonlinear acoustics, at first it was thought that nonlinear acoustics of crystalline solids would do. Now it is known that other nonlinearities must be considered. The purpose of the present paper is to show how solid state nonlinear acoustics developed, how nonlinear acoustics of crystals requires only one nonlinear term, and to give evidence of the fact that the subject needs to be expanded.

8:05

3aPAb2. Acoustic nonlinearities in earth solids. Lev A. Ostrovsky (Univ. of Colorado, CIRES/NOAA Environ. Technol. Lab., 325 Broadway, R/E/ET1, Boulder, CO 80303, lostrovsky@etl.noaa.gov) and Paul Johnson (Los Alamos Natl. Lab., Los Alamos, NM 87545)

Nonlinear elastic response in earth solids is a robust and representative physical characteristic. In this lecture the evidence leading to this conclusion is presented by providing an overview of theoretical and experimental developments in the domain. Illustrated measurements include those of nonlinear response in rock from a variety of dynamical wave experiments. The evidence leads to a pattern of unifying behavior. Nonlinear elasticity in earth solids is large relative to most materials; hysteresis and ''discrete'' memory play an important role in nonlinear properties of earth solids; nonlinear response is evident over a large frequency interval (dc to several MHz at least); and nonlinear response is significant, as is commonly appreciated, at large static and dynamic strain levels, but also at small strains where this behavior and the manifestations of this behavior are commonly disregarded. Recently, the methodology for extracting the nonlinear coefficients for a material have been proposed. Some theoretical models of structural nonlinearity of solid media and of wave propagation in such media are also described.

8:25

3aPAb3. Model equations for nonlinear surface waves. M. F. Hamilton (Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78712-1063), Yu. A. Il'inskii, and E. A. Zabolotskaya (MacroSonix Corp., Richmond, VA 23228)

The theoretical framework developed originally for modeling nonlinear Rayleigh waves in isotropic elastic half-spaces [Zabolotskaya, J. Acoust. Soc. Am. **91**, 2569 (1992); see also Knight *et al.*, *ibid.* **102**, 1402 (1997)] has been extended to encompass a broad class of nonlinear surface wave problems. The theory is based on Hamiltonian mechanics, in which spatial Fourier amplitudes are used as generalized coordinates, and coupled spectral equations are obtained for the harmonic amplitudes in the wave. The nonlinearity coefficient matrices are expressed explicitly in terms of the second- and third-order elastic moduli of the materials. Solutions of the equations describe harmonic generation, waveform distortion, and shock formation not only at the interface but also in the interior of the solid. Cylindrical spreading and beam diffraction may be taken into account, as can transient effects in pulses. Model equations, and simulations of nonlinear surface waves in real materials, shall be presented for Rayleigh, Stoneley, Scholte, and Lamb waves, and surface waves in crystals and piezoelectric materials. The validity of the theory has been demonstrated via quantitative agreement with measurements of laser-generated Rayleigh wave pulses with shocks in fused quartz [Lomonosov *et al.*, J. Acoust. Soc. Am. **101**, 3080(A) (1997)]. [Work supported by NSF and ONR.]

8:45

3aPAb4. Principles and applications of nonlinear acoustic nondestructive testing. Alexander M. Sutin and Dimitri M. Donskoy (Stevens Inst. of Technol., Davidson Lab., 711 Hudson St., Hoboken, NJ 07030)

Nonlinear acoustic techniques have recently been introduced as a new tool for nondestructive inspection and evaluation of fatigued, defective, and fractured materials. Various defects such as cracks, debonding, fatigue, etc., lead to anomalously high levels of nonlinearity as compared with flawless structures. One of the acoustic manifestations of such nonlinearity is the modulation of ultrasound by vibration. This effect is explained by using a model of the defect as contact between two rough surfaces. The vibration varies the contact area modulating the ultrasonic probing wave passing through the defect. A theoretical model considering interaction of different modes of vibrations and probing waves is discussed. Two methods employing the nonlinear interaction of ultrasound and vibration were developed, namely vibro-modulation (VM) and impact-modulation (IM) methods. The VM method employs forced harmonic vibration of the tested structure, while the IM method uses impact excitation of the structures natural modes of vibration. The crack detection tests were carried out for different objects made out of steel, plastic, glass, concrete, and carbon. These nonlinear techniques demonstrated certain advantages as compared with the conventional linear acoustic technique, specifically high sensitivity and applicability to highly inhomogeneous structures. [First author was supported in part by RFBR, Grant No. 96-05-64459.]

Contributed Papers

9:05

3aPAb5. Measurements of surface-wave harmonic generation in nonpiezoelectric materials. Donna C. Hurley (Natl. Inst. of Standards and Technol., Div. 853, 325 Broadway, Boulder, CO 80303, hurley@boulder.nist.gov)

Methods for exciting surface waves suitable for harmonic generation experiments in nonpiezoelectric materials have been investigated. Such methods are needed for improved characterization of engineered surfaces. Unlike conventional ultrasonic techniques, nonlinear experiments require the ability to produce finite-amplitude, spectrally pure surface waves. Comb structures were selected from the methods evaluated, because they appear to offer the greatest potential for harmonic generation experiments on arbitrary substrates. The out-of-plane component of the surface waves was measured using a Michelson interferometer. The apparatus enabled direct measurement of absolute displacements with a spatial resolution of approximately 50 μm. Using this combination of excitation and detection techniques, the evolution of narrowband surface waves in the 1–10 MHz range was studied. Displacement amplitudes as large as 35 nm at a fundamental frequency of 10 MHz were observed. The spatial behavior of the fundamental and second-harmonic displacements indicated that, for correct interpretation, it was necessary to account for diffraction and other geometrical effects. Quantitative results will be presented for a variety of experimental configurations, including different substrates and comb shapes. Experimental results will also be compared to predictions of available theories for nonlinear surface wave propagation.

9:20–9:30 Break

9:30

3aPAb6. Pulsed nonlinear surface acoustic waves in crystals. R. E. Kumon, M. F. Hamilton (Dept. of Mech. Eng., The Univ. of Texas at Austin, Austin, TX 78712-1063), Yu. A. Il'inskii, E. A. Zabolotskaya (MacroSonix Corp., Richmond, VA 23228), P. Hess (Inst. of Physical Chemistry, Univ. of Heidelberg, 69120 Heidelberg, Germany), A. M. Lomonosov, and V. G. Mikhalevich (General Phys. Inst., Russian Acad. of Sci., 117942 Moscow, Russia)

A theoretical model developed recently for the propagation of nonlinear surface acoustic waves in crystals [Hamilton *et al.*, *Nonlinear Acoustics in Perspective*, edited by R. J. Wei (Nanjing U.P., Nanjing, 1996), pp. 64–69] is used to investigate transient effects associated with pulses. The present work extends an earlier theoretical investigation of pulsed nonlinear Rayleigh waves in isotropic solids [Knight *et al.*, J. Acoust. Soc. Am. **96**, 3322(A) (1994)]. The latter was verified via comparison with measurements of laser-generated nonlinear Rayleigh waves with shocks in fused quartz [Lomonosov *et al.*, J. Acoust. Soc. Am. **101**, 3080(A) (1997)]. Here simulations of waveform distortion and shock formation are presented for pulses in real crystals, including Si, KCl, and Ni, and for a variety of surface cuts and propagation directions. The second- and third-order elastic moduli in the nonlinearity matrix are taken from measurements reported in the literature. Attention is focused on waveforms corresponding to recent measurements of laser-generated nonlinear surface waves propagating in Si along the $\langle 112 \rangle$ direction in the (111) plane [Lomonosov and Hess, *Nonlinear Acoustics in Perspective* (Nanjing U.P., Nanjing, 1996), pp. 106–111]. Preliminary comparisons of theory with measurements obtained in related experiments are in good agreement. [Work supported by ONR and the RFBR Foundation.]

9:45

3aPAb7. Rayleigh wave solitons in layered media. Robert I. Odom (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

The problem of weakly dispersive long waves in a thin nonlinear solid layer over a linear isotropic solid half-space is formulated as a nonlinear thin plate problem with one free surface and one surface bonded to the underlying linear isotropic half-space. If terms cubic in the displacement gradient are retained in the nonsymmetric Piola–Kirchoff stress tensor, a multiple-scale perturbation analysis of the equations of motion leads to the Benjamin–Ono equation $u_t + \alpha u u_x + \beta \mathscr{H}[u_{xx}] = 0$, where u is a particle displacement, $\mathscr{H}[\cdot]$ indicates the Hilbert transform, and α and β are constants that depend on the material properties of the layer and half-space. The Benjamin–Ono equation admits soliton solutions, and can be solved by the inverse scattering transform. This work may provide a mathematical model for small amplitude nonlinear seismic waves propagating in a thin sediment layer over a much harder basement. It may also have application to surface acoustic wave (SAW) device physics.

10:00

3aPAb8. Nondestructive flaw detection in a solid medium using nonlinear acoustic effects. C. M. Song, K. I. Jung, and S. W. Yoon (Acoust. Res. Lab., Dept. of Phys., Sung Kyun Kwan Univ., Suwon, 440-746, Rep. of Korea, swyoon@yurim.skku.ac.kr)

In a solid material it is very difficult to find tiny defects or cracks with the conventional linear acoustic techniques as nondestructive evaluation methods. When defects exist in a material, nonlinear acoustic responses become dominant in a propagating medium and can make it possible to detect such defects. As an application of such nonlinear acoustics, a nonlinear acoustic modulation technique for nondestructive flaw detection described in previous work [Kim *et al.*, J. Acoust. Soc. Am. **101**, 3029 (1997)] was applied to several different sample materials to confirm its

feasibility. The second harmonic measurements can easily indicate existence of tiny defects or cracks in the samples. The time-delay measurements of the modulated sum and difference frequency signals can accurately locate defects in aluminum, steel, and brass solid rods, respectively.

10:15

3aPAb9. Slow nonlinear dynamics in rock. R. A. Guyer (Dept. of Phys. and Astron., Univ. of Massachusetts, Amherst, MA 01003), K. R. McCall (Univ. of Nevada, Reno, NV 89557), and K. E-A. Van Den Abeele (Los Alamos Natl. Lab., Los Alamos, NM 87545)

Recent experiments on the elastic response of Berea sandstone show that nonlinear excitation of a resonant bar is accompanied by a slow dynamics, a dynamics involving time scales many orders of magnitude longer than the excitation period, $2\pi/\omega$. A lumped-element description of a resonant bar is developed. A complementary phenomenological theory of slow nonlinear dynamics is postulated. This phenomenological theory is used in the lumped-element description to explain the experiments. The results illustrate the possibility of experimental exploration of slow nonlinear dynamics using the acoustic analogue of a NMR experiment. [Work partially supported by OBES, Engineering and Geosciences Grant No. W7405-ENG-36, the National Science Foundation Grant No. EAR-9528965, and the Institute of Geophysics and Planetary Physics at Los Alamos National Laboratory.]

10:30

3aPAb10. Nonlinear behavior in rock: Experiments show two distinct behaviors. James A. TenCate, Koen E. A. VanDenAbeele,[a] and Thomas J. Shankland[b] (Los Alamos Natl. Lab., Los Alamos, NM 87545)

The compliant features of rocks—cracks, pores, and the fluids filling them—give rise to a variety of nonlinear elastic effects. As a means of achieving large dynamic strains, longitudinal resonance experiments on thin bars of sandstones were performed. Resonance curves were obtained by measuring acceleration at the end of the bar while sweeping frequency at a fixed drive level. As drive levels increased, resonance curves showed peak bending toward lower frequencies (softening modulus). Two very distinct nonlinear behaviors were observed for several different rock samples. At moderate strain amplitudes ($\approx 10^{-6}$), the effects of nonlinearity became quite pronounced: rapid peak shifting and resonance curve shape changes were accompanied by slow dynamics—shapes of the nonlinear resonance curves depend on sweep direction and rate [Geophys. Res. Lett. **23**, 3019–3022]. However, even at lower strain amplitudes, nonlinearity remains. Small frequency shifts are observed, even down to the measurement noise of our apparatus. The exact nature of the transition between the two nonlinear behaviors (e.g., effects of temperature or composition) and an explanation for the nonlinearity at very low strain levels are being pursued. [Work supported through the Office of Basic Energy Sciences—Geosciences of the Department of Energy.] [a]Currently at Laboratorium Bouwfysica, K.U. Leuven, Belgium. [b]Currently at Bayerisches Geoinstitut, Bayreuth, Germany.

WEDNESDAY MORNING, 24 JUNE 1998 FIFTH AVENUE ROOM (W), 9:15 TO 11:40 A.M.

3a WED. AM

Session 3aPP

Psychological and Physiological Acoustics: Auditory Attention I

Donna L. Neff, Cochair
Boys Town National Research Hospital, 555 North 30th Street, Omaha, Nebraska 68131

Ervin R. Hafter, Cochair
Department of Psychology, University of California, 3210 Tolman Hall, #1650, Berkeley, California 94720

Chair's Introduction—9:15

Invited Papers

9:20

3aPP1. Shared attention to simultaneous stimulation. Ervin R. Hafter, Anne-Marie Bonnel, and Erick Gallun (Dept. of Psych., Univ. of California, Berkeley, CA 94720, hafter@socrates.berkeley.edu)

Shared attention is studied using a dual task, where observers respond independently to two simultaneously presented signals. Cost is defined by a loss in performance relative to one signal tested alone. Standard stimuli are followed by signals that can be either incremental or decremental changes in level, frequency, or spatial location, relative to the standard. In a ''detection'' paradigm, the subject says ''change'' for signals of either polarity, but in an ''identification'' paradigm where all trials present signals, the subject responds with the polarities. Comparing results with fixed standards to those with a standard that roves from trial to trial shows that a major determinant of the cost of shared attention is the type of memory used in the task. Fixed standards, which show reliance on absolute memory, yield relatively high performance, but there is a cost of shared attention. Conversely, with roved standards, where absolute memory cannot be used, performance is relatively low and there is no cost. Thus a dissociation is found in which the additional process used in the context-coding mode is costly for attention but not performance. [Research supported by the Natl. Inst. of Health (NIDCD 07787), USA, Univ. of California, USA, and CNRS, France.]

9:45

3aPP2. Cross-modal links between auditory and visual attention. Charles Spence (Dept. of Psych., Oxford Univ., South Parks Rd., Oxford OX1 3UD, England, charles.spence@psy.ox.ac.uk) and Jon Driver (Univ. College London, London WC1E 6BT, England)

Using the orthogonal spatial-cuing paradigm, it has been shown that people make elevation discriminations concerning target sounds more rapidly and accurately when auditory targets are presented at covertly attended azimuths location [C. Spence and J. Driver, J. Exp. Psychol. Hum. Percept. Perform. **20**, 555–574 (1994)]. These shifts of auditory spatial attention could either be elicited endogenously, by informative cues that predicted the likely target location, or exogenously by peripherally presenting spatially uninformative auditory cues. More recently, extensive cross-modal links between audition and vision in the control of both exogenous and endogenous orienting have been reported, such that, for example, a shift of auditory attention typically leads to a shift of visual attention to the same location [C. Spence and J. Driver, J. Exp. Psychol. Hum. Percept. Perform. **22**, 1005–1030 (1996)]. Participants in all previous behavioral studies of crossmodal attention had their eyes and head in spatial alignment. New data is reported showing a partial remapping of audiovisual space when the eyes and head are misaligned (i.e., when people look out of the corner of their eyes). This remapping ensures that attention is oriented to the correct distal location, irrespective of receptor misalignment. [Work supported by MRC (UK).]

10:10

3aPP3. Auditory objects of attention. C. J. Darwin and R. W. Hukin (Exp. Psych., Univ. of Sussex, Brighton BN1 9QG, UK, cjd@biols.susx.ac.uk)

Although interaural time differences (ITDs) are the dominant cue for the lateralization of complex sounds, they are remarkably weak at grouping together simultaneous sounds. Experiments are reviewed which explore the relationship between this observation and auditory attention. A small difference in ITD between two sentences spoken in the same voice is sufficient to allow a listener to say which of two target words occurred in the attended sentence, even when the two sentences are resynthesised to have the same fundamental frequency. By contrast, a large difference in ITD does not allow a listener to exclude a single harmonic from a target steady-state vowel within an attended sentence. The results can be explained by assuming that listeners are not able to attend directly to frequency components that share a common ITD. Rather, listeners attend to the direction of auditory objects, whose frequency composition is determined by grouping cues such as onset time and harmonicity, and whose lateral position is determined from the lateralization cues of the component frequencies. [Work supported by UK MRC.]

10:35–10:50 Break

10:50

3aPP4. Perceiving talking faces: Insights into attention. Dominic W. Massaro (Dept. of Psych., Univ. of California, Santa Cruz, CA 95064)

Perceivers naturally integrate auditory and visual information in the perception of speech and emotion. Although both modalities contribute to perception, instructions and intention appear to modulate the impact of these modalities. Although there is some influence of the to-be-ignored modality, perceivers can attenuate its influence so that some degree of control is possible. The fuzzy logical model of perception provides a good account of performance under the different instruction conditions. For the estimation of the free parameters corresponding to the parameters that correspond to the evaluation of the auditory and visual sources, the FLMP assumed different parameter values in the different instruction conditions. Different parameter values were sufficient to capture the results. Intention can thus be accounted for in terms of information with no qualitative changes in information processing. Within the framework of the FLMP, the fundamental pattern recognition algorithm does not change. Bearing out this assumption, intention can enhance or attenuate the information contribution of sources that are available, but intention does not seem capable of enforcing another type of pattern recognition algorithm. Future research using expanded factorial designs should be able to provide a more definitive understanding of the role of intention and attentional set in pattern recognition. [Work supported by NIDCD.]

11:15

3aPP5. Electrophysiological studies of auditory spatial attention. Steven A. Hillyard and Wolfgang A. Teder-Salejarvi (Dept. of Neurosci., U.C. San Diego, La Jolla, CA 92093, shillyard@ucsd.edu)

Recordings of event-related potentials (ERPs) from the scalp can track the processing of auditory signals through the afferent pathways from brainstem to cortex. ERP components elicited in auditory cortex during the time range 50–200 ms are highly sensitive to selective attention, with attended sounds eliciting greater amplitudes than unattended sounds. Using combined ERP and behavioral measures, the allocation of auditory attention was studied in a free-field array of seven or eight speakers. Brief noise bursts were presented in a random, fast-paced sequence from all the speakers, and subjects were required to attend to the sounds at one location while ignoring the others. Both behavioral detection responses and ERP amplitudes were highest for stimuli at the attended location, with a progressive falloff for responses to increasingly distant sound sources. These data revealed the shape of spatial tuning curves for auditory attention in free-field and support gradient models of attentional allocation. An analysis of ERP waveforms indicated that spatial selectivity is established by consecutive stages of progressive fine-tuning over the time range 50–350 ms poststimulus. [Work supported by Grants from NIMH and ONR.]

Session 3aSA

Structural Acoustics and Vibration: Damping and Absorption

Jean R. Nicolas, Chair
Department Genie Mecanique, Universite de Sherbrooke, Uds 2500 Boulevard de l'Université, Sherbrooke, QC J1K 2R1, Canada

Contributed Papers

8:00

3aSA1. An algorithm for calculation of material loss factors using single degree of freedom method and digital data measured by FFT analyzer. Toshiyuki Maeda (Dept. of Mech. Eng., Nishinippon Inst. of Technol., 1633, Aratsu, Kanda, Miyako, Fukuoka, 800-03 Japan, t.maeda@nishitech,ac,ip)

The half-power band-width method that is conventionally used to calculate material loss factors cooperated with one single degree of freedom method using data of each resonance response of a material bring out estimation errors caused by approximation of resonance frequency due to discrete frequency when data are measured with FFT analyzer. The proposed new algorithm is able to determine the actual resonance frequency, and its amplitude directly from digital data and loss factors without approximation.

8:15

3aSA2. Dynamic and acoustic properties of beams of composite material. Shiva Sander Tavallaey (MWL, Dept. of Vehicle Eng., KTH, 100 44 Stockholm, Sweden)

Modern engineering techniques require the use of more sophisticated and optimized structural design. One way towards this aim is the use of materials which in some way optimize their inherent properties. A concept which suits this purpose very well is composite/sandwich material. Sandwich/composite materials find numerous applications where high strength and low weight are important criteria. The dynamic behavior of a sandwich beam with different boundary conditions has been investigated. Measurements have been carried out. Free-free and clamped-clamped boundary conditions were included. Modal analysis has been used to estimate resonance frequencies and modal parameters. Further, acoustical parameters of a composite plate are measured. An analytical model for prediction of wave propagation constants in sandwich/composite structure is presented. Measured and predicted results are compared.

8:30

3aSA3. Full-spectrum-based acoustic spectroscopy for determining elastic and damping constants in forced-free bars. Qiushuang Guo and David A. Brown (Dept. of Elec. and Comput. Eng., and Ctr. for Marine Sci. and Technol., Univ. of Massachusetts—Dartmouth, N. Dartmouth, MA 02747)

Elastic moduli and damping have been previously determined by measuring the resonant frequencies and quality factor of a virtually free–free bar. The development of a complete analytical solution to the longitudinal wave equation including damping effects for a bar freely supported, driven at one end, and under the influence of transducer mass and stiffness at both ends was previously reported on [Qiushuang Guo and David Brown, ''Effects of transducer mass on resonance frequency and quality factor of free-free bar,'' J. Acoust. Soc. Am. **101**, 3062 (1997)]. The analytical solution to the torsional and flexural waves in a bar including damping effects will be reported on here. When the entire spectrum of data is used, a more accurate determination of elastic moduli and internal damping coefficient is obtained. [Work supported by the University of Massachusetts—Dartmouth.]

8:45

3aSA4. Numerical method using filter banks for vibroacoustic analysis in the medium frequency domain. Sylvie Gorog and Philippe Micheau (Université de Sherbrooke, Sherbrooke, QC J1K 2R1, Canada)

Efficient numerical tools are needed to predict the vibroacoustic response of a viscoelastic structure in the medium frequency range. The presented method is formulated within the framework of nonuniform modulated filter banks, a signal processing tool. The structure is described using a variational approach. The method consists of dividing the frequency domain of interest into several subdomains and hence solving a time domain equation of motion for each subdomain. The complete displacement field is then reconstructed from the different time domain solutions. This time-frequency computations rely on the design of analysis and synthesis filter banks. Flexural vibrations of plates *in vacuo* are studied to illustrate the numerical efficiency of the method. According to the variational approach that is used, results show that the proposed method allows the computation of the solution in frequency domains where classical modal methods and direct methods are limited from a numerical point of view. This numerical tool is of great interest for the study of structures having frequency-dependent properties.

9:00

3aSA5. Improvement of damping via analysis of viscoelastic composites. Ibrahima Sow, Jean Nicolas, and Toan Vu-Khanh (Dept. Genie Mecanique, Universite de Sherbrooke, UdS 2500 Blvd. de l'Universite Sherbrooke, QC J1K 2R1, Canada)

High damping materials allow vibration and noise to be suppressed or at least reduced. Unfortunately, materials with high loss tangents also exhibit low modulus of elasticity. A new composite damping material is investigated, which consists of specially designed fibers embedded in a viscoelastic matrix. This new approach not only enhances damping through improvement of matrix loss tangent by the addition of elastomer, but also through selection of adequate ply orientation, ply stacking, geometry of fibers, and fraction of fibers. Theoretical analysis of the viscoelastic properties of this composite is presented in order to identify the optimum material with high stiffness and high loss tangent according to all parameters. To address validation theoretical developments are compared to experimental results.

9:15

3aSA6. Characteristics for vibration damping of the concrete included secondary subject. Noboru Ishikawa and Hidemaro Shimoda (Inst. of Technol., Shimizu Corp., 3-4-17 Etchujima Koto-ku, Tokyo, 135 Japan)

A concrete structural material with damping performance for vibration has been investigated. It is proposed that in order to reduce the sound propagation in the solid material, the concrete material disperses many

small cells in which many easy moving particles are included. It is expected that the vibration energy which is applied to the material is absorbed by the impaction and friction of the particles alternatively and between particles and cell's wall. The preliminary experiment to confirm the capability of the high damping material on the basis of this concept has been carried out. It was shown that damping performance of the specimen depended on the size, the density, and the volume fraction of the particles by the impact vibration testing for the coupon-sized test specimen. The effect of the easy moving particles for reduction of the sound propagation in the solid was confirmed by the sound propagation testing for the middle-sized panel specimen.

9:30

3aSA7. Finite-element modeling of the vibroacoustic behavior of poroelastic materials. Hamid Bouhioui and Murray Hodgson (UBC, Dept. of Mech. Eng., 2324 Main mall, Vancouver, BC V6T 1Z4, Canada, bouhioui@mech.ubc.ca)

Porous materials are widely used nowadays as part of the insulation systems in the automotive industry or as effective sound absorbers in other areas, such as architecture. Their complex dynamic behavior is due to the various interaction phenomena within the elastic porous material. Such materials have been effectively modeled thanks to the original work of Biot. In the work reported here, the classical Biot theory for elastic porous materials is used for the development of a new three-dimensional finite element. While the finite-element development remains classical, a new numerical implementation is proposed. This new approach allows very fast computation times compared to the approaches existing in the literature. The finite element developed is general and suitable for any type of finite-size, extended-reaction porous material, with any boundary conditions. Comparisons are made using solutions found in the literature, along with a fine analysis of different aspects of the vibroacoustic behavior of porous materials.

9:45

3aSA8. Modeling of porous material added damping on a vibrating plate. Nicolas Dauchez, Sohbi Sahraoui (Laboratoire d'acoustique, Université du Maine, UMR CNRS 6613, Av O. Messiaen, 72085 Le Mans Cedex 9, France, dauchez@laum.univ-lemans.fr), and Noureddine Atalla (Université de Sherbrooke, Sherbrooke, QC J1K 2R1, Canada)

Porous materials like plastic foam are well known for their ability to absorb sound. When bonded onto a vibrating structure they may add damping. Their damping performance is investigated here. The generic configuration is a free aluminum plate (28 cm$\times$22 cm$\times$1 mm) mechanically excited (20–500 Hz), damped by a 1- to 5-cm-thick foam layer. The analysis of porous material behavior, based on Biot–Allard theory [J.-F. Allard (Chapman and Hall, 1993)] shows that damping is mainly related to the frame viscoelasticity. A formulation of an equivalent plate, substituting the porous layer by a monophasic viscoelastic material, is presented. This formulation gives behavior indicators, like amount of added stiffness and damping. It also accelerates the computation of the structure vibration, compared with a 3-D finite element code, including poroviscoelastic elements [R. Panneton and N. Atalla, J. Acoust. Soc. Am. **100**, 346–354 (1996)]. Good agreement is found with both the complete discretized formulation and experimental results. Maximum damping available from biphasic phenomena, like air flow induced by the layer flexure, is also evaluated and compared with structural damping. This shows that more sophisticated mounting conditions should be used to take advantage of the biphasic property.

10:00

3aSA9. On the use of assumed spatial distributions for the vibration analysis of built-up plate structures. Ross A. Fulford and Bjorn A. T. Petersson (Loughborough Univ., Loughborough, Leicestershire LE11 3TU, UK)

For the vibratory analysis of built-up structures, traditional pointlike connections cannot be applied where the interface is large and the wavelength small. Instead, the spatially distributed wave field has to be accounted for whereby the interface has to be considered as either a surface or, for a one-dimensional contact, a line. To analyze such, the traditional approach is to rely upon numerical methods such as the finite element method. These approaches are often cumbersome, however, in that they reveal little about the physics of the systems. As an alternative, this paper considers the development of a method whereby the spatially distributed properties of a line contact can, with the salient physics retained, be included in simplified mathematical models. The basis of the method is to describe the spatial properties with respect to Fourier components and hypothesize that the first order, i.e., uniform, component is most influential. Via an assumed uniform force distribution, the approach is applied, with some promise, to the problem of predicting the vibratory power transmission between a box and a plate.

10:15

3aSA10. Statistical analysis of boundary sound absorption in a panel-cavity system. K. S. Sum and J. Pan (Dept. of Mech. and Mater. Eng., Univ. of Western Australia, Nedlands, Western Australia 6907)

Sound absorption due to acoustic-structural coupling has a significant effect on the acoustical behavior of an enclosed sound field in vibroacoustic systems. In the high-frequency range, this absorption is related to the average behavior of acoustic and structural modes and a statistical description (e.g., statistical energy analysis) of the coupling and its effect on the sound field response is well known. In both the low- and medium-frequency ranges, acoustic-structural coupling of low-order modes tends to dominate and there are large variations in acoustical behavior between acoustic modes. Therefore accurate modal description of the sound field boundary conditions is now important. However, in the medium-frequency range, the complicated boundary conditions are difficult to describe due to the presence of a large number of dominant acoustic and/or structural modes. Thus acoustic-structural coupling is not well understood and the physical interpretation of sound absorption in this frequency range is not clear. The aim of this paper is to obtain some physical insights in the behavior of sound absorption due to acoustic-structural coupling at medium frequencies. The coupling is statistically analyzed for a panel-cavity system and its physical relationship to the boundary sound absorption is illustrated.

10:30

3aSA11. Sound transmission through multilayer structures with isotropic elastic porous materials. Noureddine Atalla, Raymond Panneton, Michel Tournour, and Celse K. Amedin (G.A.U.S., Dépt. Génie Mécanique, Univ. de Sherbrooke, Sherbrooke, QC, Canada)

This paper discusses the prediction of transmission loss through multilayer structures with porous materials. First, a brief review of recent formulations for multilayer structures with poroelastic materials is presented. Second, the transmission loss problem is described and formulated using a coupled finite element and boundary element procedure. Issues such as poroelastic-elastic coupling, poroelastic-fluid coupling, and radiation from a poroelastic material will be addressed. The developed model is validated through numerical examples. Its advantages and limitations are discussed. Finally, typical results showing the vibroacoustic effects of several parameters such as the multilayer configuration, the types of porous materials, and the mounting conditions are presented. [Work supported by Bombardier, Inc., Canadair and N.S.E.R.C.]

Session 3aSC

Speech Communication: Language Specific Influences During Infancy

Peter W. Jusczyk, Chair
Psychology Department, Johns Hopkins University, Ames Hall, Charles Street at 34th, Baltimore, Maryland 21218-2686

Chair's Introduction—9:15

Invited Papers

9:20

3aSC1. Bilingual exposure and some consequences on native-language recognition processes at four months. Laura Bosch (Dept. Psicologia Basica, Universitat de Barcelona, Passeig de la Vall d'Hebron, 171, 08035 Barcelona, Spain)

Evidence of language-specific influences on infants' speech perception presupposes the existence of a basic level of native-language representation. Findings from a language discrimination task, indicating that two-month-old infants perform successfully only when one of the two languages is the native one, but not when both are unfamiliar languages, have been interpreted as evidence that a primary representation of the structure of the native language has been reached, and is probably based on the prosodic/rhythmic properties of the input language (Mehler *et al.*, 1996). The capacity to discriminate languages seems also to be dependent on their prosodic distance. Recent research in this laboratory has addressed the issue of language discrimination for a prosodically close pair of Romance languages, i.e., Spanish and Catalan. While four-month-old infants from monolingual environments seem able to recognize their native language, either Spanish or Catalan, infants from bilingual families did not show a preference for either language. In a second series of experiments, infants from bilingual environments were tested with the same procedure, this time selecting one of their familiar languages to be contrasted with a completely unfamiliar one (English and also Italian). When compared to infants from monolingual families, they showed slower orientation latencies toward the familiar language.

9:45

3aSC2. Effects of language experience on speech perception. Patricia K. Kuhl (Dept. of Speech and Hearing Sci., Univ. of Washington, Seattle, WA 98195)

Language experience during the first year of life has a dramatic effect on infants' perception and production of speech. In this presentation, (a) studies of cross-language effects of language experience on consonant and vowel perception, (b) studies of the effects of speech experience on infants' vowel production, and (c) infant Mismatched Negativity (MMN) responses for speech will be used to show that language experience causes infants to ignore physical differences between speech stimuli that are irrelevant in their native language. A theory is offered to explain how early language experience alters speech perception. The theory holds that listening to native-language speech causes infants to develop ''perceptual maps'' of the incoming speech information. These maps functionally alter the perceived distances between physical stimuli. [Work supported by NIH.]

10:10

3aSC3. The development of phonetic perception in the first year of life. Linda Polka (School of Commun. Sci. and Disord., McGill Univ., Montreal, QC H3G 1A8, Canada, cztg@musica.mcgill.ca)

There is considerable evidence that language-specific influences on speech perception begin early in life with attunement to different aspects of language structure emerging at different ages in the first year. With respect to segmental units, cross-language research has focused on the perception of syllables differing in a single consonant. This talk will focus on research investigating developmental patterns in cross-language vowel perception in the first year of life. A series of experiments exploring vowel perception in English-learning and German-learning infants as well as English-speaking and German-speaking adults will be presented. The results provide further insights into the emergence of language-specific influences and also reveal language-universal perceptual biases that infants bring to the task of vowel perception. Implications of these findings with respect to the early development of phonetic perception and vocal production will be addressed. [Work supported by NSERC.]

10:35–10:45 Break

10:45

3aSC4. Developing sensitivity to native language sound patterns. Peter W. Jusczyk (Depts. of Psych. and Cognit. Sci., Johns Hopkins Univ., Baltimore, MD 21218, jusczyk@jhu.edu)

During the second half of their first year, English-learning infants begin to show sensitivity to the structure and organization of native language sound patterns. In particular, they give evidence of recognizing the phonetic, phonotactic, and prosodic characteristics of native language words. Moreover, their sensitivity to these features of native language sound organization appears to play a critical role in infants' ability to segment words from fluent speech. Evidence for the use of this language-specific information in word

segmentation by infants will be presented. In particular, we will focus on changes in the development of word segmentation abilities between 7.5 and 10.5 months of age. In addition, findings will be considered that suggest infants' representations of the sound patterns of words are detailed both in terms of the phonetic properties and indexical features that they include. [Work supported by NICHD and NIMH.]

11:10

3aSC5. Updates on becoming a native listener. Janet F. Werker, Judith E. Pegg, Rushen Shi, and Christine Stager (Dept. of Psych., Univ. of British Columbia, 2136 West Mall, Vancouver, BC V6T 1Z4, Canada)

Infants' sensitivity to the sound structure of the native language becomes finely honed during the first year. New studies confirm and extend this fact in the domains of phonetic, metrical, and grammatical processing. First, infants were tested on their ability to discriminate phonetic, but nonmeaningful differences in the native language. Although most 6- to 8-month olds discriminated [d] versus [t], 10- to 12-month olds did not. Thus, by 10–12 months, infants listen to only those phonetic differences that distinguish acceptable native-language syllabic shapes. Next, English-learning infants' discrimination of single syllable words such as ''clone'' from their WS counterparts, e.g., ''cologne'' was examined. Testing infants from the newborn period through the first year of life revealed that this ability is not robustly evident until around 10 months. Finally, the authors asked whether infants are sensitive to the acoustic and phonological cues distinguishing content from function words. Using lists of content versus function words excised from infant-directed utterances, it was found that both neonates and 6-month olds distinguished these broad grammatical categories. These studies confirm and extend our understanding of initial biases and listening experiences on infant speech processing. A critical question the authors are now addressing is how these changes in speech perception prepare the child for language learning.

11:35

3aSC6. Rapid gains in the speed and efficiency of word recognition by infants in the second year. Anne Fernald, John P. Pinto, and Daniel Swingley (Dept. of Psych., Stanford Univ., Stanford, CA 94305-2031)

Toward the end of the second year, during the period of rapid development in productive vocabulary known as the ''vocabulary explosion,'' infants also make dramatic gains in the speed and efficiency with which they are able to recognize familiar words in fluent speech. Here several studies using an on-line measure of speech processing by infants from 15 to 24 months of age are presented. In this research, an auditory-visual matching procedure, in which infants looked at two familiar pictures while hearing a target word that matched one of the pictures, was used. Monitoring infants' eye movements from the onset of the target word, word recognition was assessed by measuring infants' latency to look to the matching picture in response to the spoken word. These studies show that by the end of the second year, children are progressing toward the highly efficient performance of adults, using incomplete acoustic information to make rapid and accurate decisions about the identity of spoken words. [Work supported by NIMH.]

12:00–12:10 Discussion

WEDNESDAY MORNING, 24 JUNE 1998 METROPOLITAN BALLROOM (S), 7:45 TO 10:45 A.M.

Session 3aUW

Underwater Acoustics and Acoustical Oceanography: Bottom Geoacoustic Characterization and Inversion I (Précis Poster Session)

Subramaniam D. Rajan, Chair
Scientific Solutions, 5907 106th Avenue, NE, Kirkland, Washington 98033

In this session, an oral summary of each paper will be presented followed by a poster session. All posters will be on display, and all contributors will be at their posters, from 9:30 a.m. to 10:45 a.m.

Chair's Introduction—7:45

Invited Paper

7:50

3aUW1. Measuring parameters that control acoustic propagation in granular sediments near the seafloor. Robert D. Stoll (Lamont-Doherty Earth Observatory of Columbia Univ., Palisades, NY 10964, stoll@ldeo.columbia.edu)

Propagation of acoustic waves in granular sediments near the water-sediment interface is controlled by a number of parameters such as fluid and grain compressibility, permeability, and the compressibility and shear stiffness of the skeletal frame in a water environment. A new experimental technique for evaluating several of these properties is described wherein a cylindrical, water-saturated sample is subjected to vertical, steady-state fluid flow. By controlling the hydraulic gradient, any desired vertical, effective

stress may be generated and acoustic measurements then made across the diameter. When the effective stress is reduced to zero, acoustic measurements yield the compressibility of the grains, whereas at nonzero effective stress levels, it is possible to assess the effects of depth of embedment which causes changes in the compressibility and stiffness of the frame. Finally, by measuring the flow rate at each gradient, a direct evaluation of permeability is made. Preliminary measurements agree with most historical evaluations of these parameters suggesting that some recent work which challenges the traditional values is in error. [Work supported by ONR, code 321OA.]

Contributed Papers

8:10

3aUW2. Matched-field inversion using the down hill simplex algorithm. Martin Musil (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 2Y2, Canada), John M. Ozard (Defence Res. Establishment Atlantic, Victoria, BC V9A 7N2, Canada), and Michael J. Wilmut (Royal Military College, Kingston, ON K7K 7B4, Canada)

In recent years inversion techniques have been proposed to effectively and efficiently determine geophysical parameters. The focus of this matched-field inversion (MFI) study is to determine the geoacoustic and geometric parameters for a North-East Pacific shallow water data set. MFI correlates modeled data, called replicas, with the measured data and uses a search algorithm to find model values which maximize the correlation function. In our MFI the parabolic equation (PE) is used to compute the replicas to account for mode coupling in the range-dependent environment. The first stage of the optimization algorithm employs a random search to determine $N+1$ parameter sets with the best correlations, where N is the number of parameters being determined. Next the $N+1$ sets are the input to the downhill simplex algorithm. The algorithm is shown to perform well for simulated vertical line array data for search parameter ranges representative of the environment in the shallow North-East Pacific Ocean. The inversion technique was then applied to measured data obtained in the North-East Pacific on the continental shelf using towed CW and broadband sources and a vertical array of hydrophones. The geometric parameters of the experiment and dominant geoacoustic parameters were successfully recovered.

8:14

3aUW3. Numerical analysis on inversed bottom geoacoustic models for predicting a shallow-water acoustic field. Guoliang Jin and Renhe Zhang (Shanghai Acoust. Lab., Academia Sinica, Shanghai 200032, PROC)

The bottom reflection coefficient as a function of the grazing angle is calculated for a liquid bottom with depth-dependent sound velocity, density, and attenuation. The bottom geoacoustic parameters of a simplified model, which describes only the depth dependence of sound velocity and treats the density and attenuation as the depth independent, are derived from the bottom reflection coefficients by the least-squares fit. The inversion accuracy in reproducing an acoustic field in the water column is tested by the performance of matched field processing, the sound propagation losses, modal group velocities, and modal eigenvalues for numerical examples. In general, the test results are good for single frequencies, but degrade for broader frequency bands. The possible distortion of the depth dependence of sound velocity from the inversion under the simplified model is discussed.

8:18

3aUW4. A multifrequency inversion method for geoacoustic parameters. Renhe Zhang, Fenghua Li, and Wenyu Luo (Natl. Lab. of Acoust., Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC, zrh@public.east.cn.net)

The effects of the bottom parameters on the sound field are discussed theoretically and two characteristic angles relevant with geoacoustic parameters are put forward. Based on the theoretical analysis of the characteristic angles, a multifrequency inversion scheme is given as follows: High-frequency and far-field data are used to invert the upper bottom parameters first, and then low-frequency and near-field data are used to invert the lower bottom parameters. Simulated geoacoustic inversions based on simulated annealing method and the beam-displacement ray-mode (BDRM) sound propagation model are applied to the test cases from the 1997 Geoacoustic Inversion Workshop. The simulated inversions show that the inversion results of multi-frequency inversions are more accurate and reliable than those of the single-frequency inversions. The multi-frequency inversion scheme is also applied to the experimental data from '96 China–US Far Yellow Sea Experiment. [Work supported by NSF of China and ONR of USA.]

8:22

3aUW5. Geoacoustic parameter inversion from waveform structure. Fenghua Li and Renhe Zhang (Natl. Lab. of Acoust., Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC, zrh@public.east.cn.net)

Based on the beam-displacement ray-mode (BDRM) sound propagation model, the effect of the bottom reflection coefficient on the pulsed waveform structure in shallow water with a thermocline is discussed. Simulated numerical calculations show that in shallow water with a thermocline, the bottom reflection amplitude and the bottom reflection phase can be inverted by the envelope of the pulsed waveform and the pulsed waveform structure, respectively, when the source and/or the receiver are above the thermocline. Using the experimental data from '96 China–US Far Yellow Sea Experiment, inversions for bottom reflection amplitude and reflection phase are performed. The waveform correlation between the experimental waveform and the inversion waveform is about 0.88 while the envelope correlation between them is about 0.95. The inversion of the experimental data shows that the inversion for the bottom reflection coefficient is effective and the inversion results are stable when the source and/or the receiver are above the thermocline. [Work supported by NSF of China and ONR of USA.]

8:26

3aUW6. Geophysical measurements on sediment cores and geoacoustic modeling. Philip P. Thomson, John I. Dunlop (Dept. of Appl. Phys., UNSW, Sydney 2052, Australia), and Frank Neissen (Alfred Wegner Inst., Bremerhaven, D-27568, Germany)

Geophysical measurements on marine sediments cores have considerable interest and importance in underwater acoustics and marine geophysics. A series of studies on geoacoustic properties of marine sediments have been conducted and modeled on the samples from Australian continental shelves by the technique developed by UNSW. More works were carried out on cores from North Atlantic ocean basins at the Alfred Wegner Institute in Germany by using a multisensor core logger (MSCL). MSCL is a logging device which enables a number of geophysical measurements such as p-wave velocity, gamma ray attenuation, and magnet susceptibility on marine sediment cores encased in cylindrical plastic core liners. A pair of compressional wave transducers at 250 kHz is used to measure dilatational velocity and a gamma ray source and detector are used to measure attenuation in the cores. A comparative study of physical properties measured by a MSCL system and UNSW technique has been applied to a wide range of frequencies and modeled with modified Biot–Stoll theory.

8:30

3aUW7. Building an acoustic library for categorizing bottom substrate type. John Hedgepeth, Colleen Sullivan (BioSonics, Inc., 4027 Leary Way NW, Seattle, WA 98107, jhedgepeth@biosonicsinc.com), Robert Sullivan (Parametrix, Inc., Seattle, WA), and Patrick Schneider (Instituto de Ciencias del Mar, Barcelona, Spain)

A hypothesis to categorize bottom substrate type is that information about the bottom sediments is encoded in the echo signal. Bottom typing means that to decode this information, however, for data verification, the user of automated typing programs should establish libraries where data acquisition of acoustic data is accompanied by physical core samples of the bottom. Separate verification should be done for different types of bottoms (e.g., %rocks, %sand, %silt, and %clay). Data from 126 observations of sediments in the vicinity of Commencement Bay near Tacoma, WA was used to establish libraries and to test the capability of the Visual Bottom Typer software to categorize correctly. Four different algorithms were used: division of the first bottom echo, ratio of the first and second bottom echo, cumulative intensity of the first bottom echo, and fractal dimension of the first bottom echo. The echo signal parameters depended not only on the type of bottom sampled but also on the equipment parameters. Two different frequencies and two different beam widths were compared. Verification results were valid for particular combinations of equipment parameters.

8:34

3aUW8. A measuring method for shear wave velocity of marine sediment using radiation impedance. Masao Kimura (Dept. of Ocean Eng., Tokai Univ., 3-20-1 Orido, Shimizu, Shizuoka, 424 Japan, mkimura@scc.u-tokai.ac.jp)

Radiation impedance of a circular vibrating plate on the surface of viscoelastic medium such as marine sediment varies with the viscoelastic properties of the medium. Therefore it is possible to predict the viscoelastic properties using the characteristics of the radiation impedance. It was demonstrated that the shear wave velocity of viscoelastic media can be determined using the value of the frequency at which the imaginary part of radiation impedance becomes zero, and the measured results of shear wave velocities for dry and wet sands were shown. In this study, the radiation impedance is measured for some kinds of marine sediment models. Then, the values of shear wave velocities obtained from the measurements of the radiation impedance, and the measured values using a bimorph type piezoelectic transducer, are compared. Moreover, the effect of the size of the radius of the circular vibrating plate on the measured values of shear wave velocities are considered.

8:38

3aUW9. Measurements of high-frequency acoustic scattering from seabed vegetation. Anthony P. Lyons and Eric Pouliquen (SACLANT Undersea Res. Ctr., APO AE 09613-5000, lyons@saclantc.nato.int)

Knowledge of the acoustic properties of seabed vegetation is required for accurate seafloor characterization in many shallow-water sites. As the study and modeling of acoustic scattering from seabed vegetation have been very limited to date, characteristics such as the dependence of scattering strength on grazing angle or frequency are still largely unknown. In order to quantify these properties, acoustic scattering experiments were conducted in beds of *Posidonia Oceanica* at several sites near the islands of Elba and Sardinia, Italy and in Saros Bay, Turkey with side-scan, single beam, and parametric transducers. Normal incidence broadband measurements were made at both the primary (40 kHz) and secondary (8 kHz) frequencies of a parametric sonar. Oblique incidence measurements were made with linear sources at frequencies from 29 to 385 kHz and with coverage over a large range of grazing angles. Results of these measurements are presented in terms of scattering strength versus grazing angle as well as scattered amplitude statistics. The mean scattering strength of the posidonia beds was found to be quite high (≈ -22 dB) and also to be fairly constant over most of the range of grazing angles, indicative of a volume scattering mechanism. Ultimately this effort will aid in the development of theoretical models of acoustic backscatter in seagrass covered sites.

8:42

3aUW10. Use of Scholte waves to determine the shear wave velocity structure in marine sediments. Sayuri Kawashima and Masao Kimura (Dept. of Ocean Eng., Tokai Univ., 3-20-1 Orido, Shimizu, Shizuoka, 424 Japan, 60oid001@scc.u-tokai.ac.jp)

Scholte waves propagating in marine sediments exhibit velocity dispersion, if the shear wave velocity of the sediment under the seabed surface depends on the depth. Therefore, it is possible to obtain the shear wave velocity profile of marine sediment, using velocity dispersion characteristics of Scholte waves. In this study, *in situ* experiments of Scholte waves in marine sediments are conducted off the coast of Shizuoka. A weight-drop system, lifting a weight and releasing it, was used as the source. Propagated signals are detected by gimbaled geophones. Group velocity dispersion characteristics are obtained from the recorded signals using the multiple filter method. Next, the depth dependence of shear wave velocity is obtained from the group velocity dispersion curves of Scholte waves using an inversion method. Furthermore, the estimated depth dependence of shear wave velocity is investigated.

8:46

3aUW11. Simplex simulated annealing: A hybrid approach to geoacoustic inversion. Mark R. Fallat and Stan E. Dosso (School of Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada)

The problem of matched-field inversion of acoustic fields for ocean-bottom geoacoustic properties has received considerable attention in recent years. In order to avoid the large number of local minima typical of multidimensional geoacoustic parameter spaces, many inversions apply global optimization methods, such as simulated annealing (SA), to minimize the mismatch between measured and modeled fields. However, the random manner in which SA searches the parameter space can be quite inefficient, particularly near convergence and/or in cases where the minima represent narrow valleys not aligned with the parameter axes. In addition, standard SA has no memory, so a good solution may be discarded at the early stages and never revisited. In order to improve the efficiency and effectiveness of SA, a local inversion component based on the downhill simplex method within a global approach to geoacoustic inversion has been incorporated. The simplex method is sensitive to local gradients, and provides efficient perturbations at all stages which are independent of the coordinate axes. Also, the simplex of models carried through the inversion retains good models and provides a memory for the algorithm. The hybrid approach of simplex simulated annealing will be described and illustrated for geoacoustic inversion. [For Underwater Acoustics Best Student Paper Award.]

8:50

3aUW12. Frequency dependence of acoustic response due to randomness and uncertainties in physical properties of porous media. Yongke Mu, Mohsen Badiey, and Alexander H.-D. Cheng (Univ. of Delaware, Newark, DE 19716)

Models of acoustic waves in porous media are complicated because of the multiple physical parameters involved. The physical structure of these parameters can be random, and the accuracy toward their direct measurement or indirect estimate is generally uncertain. In order to better understand the extent by which different parameters can affect the interaction between the acoustic waves and the porous media, these randomness and uncertainty issues need to be addressed. In this paper, the investigation is based on the Biot model. First, the model is improved based on effects arising from the statistical distribution of pore/grain size. Next, effects of

randomness in sediment stratigraphy are examined. Also, we take into consideration the uncertainty resulting from data scarcity. Numerical tools used include propagator matrix, PE, and Monte Carlo simulation. From the variation of different parameters a different result is found. But frequency is found to always play an important role. For example, at lower frequencies the effects of pore/grain size can be neglected, but at higher frequencies they have significant impact. The material frequency response for the microscopic and macroscopic randomness is examined here. The results offer a better understanding of the interaction between the acoustic waves and the porous media.

8:54

3aUW13. Inversion of explosive shallow-water transmission loss data to obtain acoustic seabed models. Marshall V. Hall (DSTO Maritime Operations Div., P.O. Box 44, Pyrmont, NSW 2009, Australia, marshall.hall@dsto.defence.gov.au)

Acoustic transmission losses (TL) measured with explosive charges along a track in shallow water have been analyzed at frequencies in octave steps starting from 16 Hz. The results have been inverted to obtain the acoustic properties of the seabed, which is modeled as a uniform half-space. There are five unknown parameters: sound and shear speeds (C_p and C_s), sound and shear absorption coefficients (A_p and A_s), and density. The number of unknowns is reduced to three by using a regression equation for density in terms of C_p, while A_s is set to its maximum physical value. By holding C_s fixed for subsets of the process, the number of unknowns is reduced to two (C_p and A_p). For a given seabed, the theoretical TL is computed as a function of range using TL models valid for stratified media. The rms residual between the measured and theoretical TL is computed over rectangles in the C_p–A_p plane. The position at which the rms residual is a minimum is noted, and expressed as a function of C_s. Criteria for selecting the optimum result are discussed. The half-space properties contain useful information, but since they vary with frequency, they do not describe the actual seabed.

8:58

3aUW14. The development of the multibeam echosounder: An historical account. Ståle Vilming (Norwegian Univ. of Sci. and Tech., Fac. of Elec. Eng. and Telecom., Trondheim, N-7034, Norway, vilming@tele.ntnu.no)

The first paper on narrow-beam echosounding was published in 1960. Since then, the advantages of the principle of multibeam swath bathymetry for seabed mapping has been recognized by many institutions working in the field of area mapping underwater. We intend to review the development the multibeam echosounder has undergone from the ''beginning'' until the present in order to better understand the challenges facing the field today. It can safely be said that the field of area mapping has diversified into many directions since the 1960s. At that time most of the development was related to the mechanical and electronic system components, whereas present research focuses on: better algorithms for more precise mapping, seabed classification, image processing, acoustical modeling of the backscattering process from seafloor and volume, and high-resolution methods for direction finding of plane waves. Multibeam echosounding thus encompasses many scientific disciplines, and we consider it important to summarize the state of the art in these disciplines as far as they are related to seabed characterization. This could help scientists to better understand the requirements for further research.

9:02

3aUW15. A comparison of Hankel transform algorithms' performance in shallow-water waveguides and their effect on the determination of seabottom plane-wave reflection coefficients. Ben Cox and Phillip Joseph (Inst. of Sound and Vib. Res., Univ. of Southampton, Southampton SO17 1BJ, UK, btc@isvr.soton.ac.uk)

Complex acoustic pressure data as a function of range may be decomposed into its constituent wave-number components by taking the Hankel transform. This paper investigates the accuracy of a number of discrete Hankel transform algorithms applied to computer simulated pressure data in a shallow water channel for a variety of seabed bottom types. The results are compared against the exact solution and the most accurate algorithm identified for shallow-water applications. The differing behavior of these algorithms is illustrated by their use in the Hankel transform method proposed by Frisk *et al.* [J. Acoust. Soc. Am. **68**, 602–612 (1980)] for the determination of seabottom plane-wave reflection coefficients. The results from computer simulations and scale model tank experiments are compared.

9:06

3aUW16. Comparison of the performance of global optimization methods with perturbative methods in the estimation of ocean bottom properties from multifrequency data. Subramaniam D. Rajan (Sci. Solutions, Inc., 18 Clinton Dr., Hollis, NH 03049)

During 1988 a series of acoustic propagation experiments was conducted in the Hudson Canyon area off the coast of New Jersey. It included synthetic aperture experiments in which a source transmitting multiple tones was towed toward/away from a vertical array of receivers. By this means, a two-dimensional structure of the acoustic field was obtained at eight different frequencies. This data set is used to estimate the bottom acoustic properties using global optimization approaches such as simulated annealing and genetic algorithm and to compare their performance with that of perturbative methods, which because of their simplicity have some advantages over the global methods.

9:10

3aUW17. Performance evaluation of horizontal and vertical vector sensor arrays in shallow water environments. Peter Gerstoft (Marine Physical Lab., UCSD, La Jolla, CA 92093-0704) and Joo Thiam Goh (DSO Natl. Labs., 118230 Singapore)

In planning an experiment it is essential to have as diverse information as possible in order to obtain a good estimate of the relevant parameters. Traditionally, the inversion for waveguide parameters as well as the source location is done exclusively using pressure measurements. Recently, it has been suggested by Nehorai and Paldi [IEEE Trans. Signal Process. **42**, 2481–2491 (1994)] that arrays of vector sensors (sensors that measure both pressure and velocities) could be used. Experiments have also been conducted in deep water by D'Spain *et al.* [IEEE Oceans (1992)]. In a realistic shallow-water environment the use of either vector or pressure sensors on a linear horizontal/vertical array with amplitude-only or complex-valued data at a number of discrete frequencies are discussed. Based on developed maximum-likelihood functions, the performance for the above type of data will be classified using either a global or a deterministic approach. For the deterministic approach, the Cramer Rao lower bound (CRLB), based on derivatives of the likelihood function, is used. For the global approach the *a posteriori* distributions are examined. These are based on Monte Carlo integration of the likelihood function. [Work supported by DSO National Laboratories, Singapore.]

9:14

3aUW18. Range-dependent environmental mismatch in matched-field tomography (MFT). Ronald T. Kessel (School of Earth and Ocean Sci., Univ. of Victoria, P.O. Box 3055, Victoria, BC V8W 3P6, Canada)

The goal of MFT is to determine unknown geoacoustic properties of the ocean by fitting a model to data recorded when the source–receiver configuration is known. The best-fit model presumably gives the best estimate of the actual geoacoustic properties along the propagation path. But the best-fit model, like the family of trials from which it is drawn, is necessarily constrained and idealistic. For simplicity, the horizontal variation of the sediments and bathymetry might have been ignored, and all trials might have had the same number of sediment layers, fixed arbitrarily perhaps, with each trial simply assigning different geoacoustic properties to those layers. If the best-fit model does not match the actual environment in these and other respects, then in what way does its properties represent those of the actual environment? When is the best-fit narrowly functional, fitting one data set well enough, but failing given other source–receiver configurations or frequency bands? These fundamental questions are addressed using the adiabatic theory of mode propagation, following a general perturbative approach to mismatch in any geoacoustic parameter in weakly range-dependent environments. Systematic and random mismatch are considered, and the implications for MFT are explored.

9:18

3aUW19. Application of coordinate rotation to inversion for sediment parameters. John S. Perkins, Michael D. Collins, and Laurie T. Fialkowski (Naval Res. Lab., Washington, DC 20375)

A global optimization technique (e.g., simulated annealing), combined with an accurate acoustic simulation capability, provides the means to invert measured acoustic data for unknown environmental parameters (e.g., sediment properties). Two problems are often encountered: (1) Which parameters are actually important to the solution? and (2) What are the parameter couplings? These problems can be regarded as difficulties in navigating through a multidimensional landscape to find the global minimum of the cost function (the error between the measured data and the simulated acoustic field corresponding to a given set of parameter values). A coordinate rotation technique based on the eigenvectors of the covariance matrix of the gradient of the cost function has been developed [M. D. Collins and L. Fishman, J. Acoust. Soc. Am. **98**, 1637–1644 (1995)]. In applying this technique to low-frequency (<200 Hz) SWellEX-96 data involving a drifting source of known position and a 32-element vertical receiver, a 12-parameter representation of the sediment structure and properties leads to a frequency-dependent parameter hierarchy and parameter couplings. The resulting environmental inversions are compared with previous characterizations and are used in matched-field processing to locate the source at other times during the experiment. [Work supported by ONR.]

9:22

3aUW20. Offshore geoacoustic inversions using sounds from land vehicle activity. Gerald L. D'Spain, William A. Kuperman, LeRoy M. Dorman, Lewis P. Berger, and William S. Hodgkiss (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0704)

Land-based vehicle activity can be clearly detected and tracked by underwater acoustic sensors located outside the surf zone. This capability has been demonstrated with nearshore seismoacoustic data collected during the Marine Physical Lab's Adaptive Beach Monitoring (ABM) program. The purpose of this paper is to demonstrate how the sounds from these land vehicles can be used to perform inversions for the offshore geoacoustic properties, somewhat akin to the use of Vibroseis in geophysical exploration. The event of interest here is where four tracked vehicles traveled down the beach at a speed of 10 m/s, creating signal beam levels that exceeded background levels by 35 dB in the 30- to 70-Hz band. The signal-to-background noise ratio in single element spectra also is sufficiently high that these vehicles can be used as sources of opportunity. The datum for the inversions is the strong frequency dependence of the phase velocity of the arrivals in this frequency band. Forward modeling is performed with a parabolic equation (PE) code written by Mike Collins. Results are compared to those obtained from the offshore arrivals of broadband detonations conducted on the beach. [Work supported by ONR, Code 32.]

9:26

3aUW21. The multiple scattering interactions of cobbles and pebbles lying on the ocean floor. C. Feuillade and R. W. Meredith (Naval Res. Lab., Stennis Space Center, MS 39529-5004)

A new model for predicting the sonar target strengths of ensembles of rocks lying on the ocean floor has recently been presented. The model represents rocks by either rigid movable spheres or elastic spheres of variable size, density, and numerical/geometrical configuration. When the rocks are densely packed, and therefore in close proximity to each other, it is important to consider the contribution that multiple scattering interactions make to the collective scattering response. To investigate this question, a T-supermatrix formalism, a self-consistent procedure which includes all orders of multiple scattering [see Lim and Hackman, J. Acoust. Soc. Am. **91**, 613–638 (1992)], has been developed to describe the radiatively coupled ensemble. The method allows the coherently summed scattered pressure field in any direction to be determined, as well as the total power scattered by the system. Application to rocks of cobble and pebble size classes indicates that the effect of multiple interactions is dependent upon the acoustic frequency, the rock spacing relative to the radius, and the grazing angle. The results obtained provide a basis for estimating the relative importance of these interactions in determining target strength. [This work was supported by ONR/NRL.]

Session 3aPLb

Plenary Lecture

Terry Ewart, Chair
Applied Physics Laboratory, University of Washington, MS HW-10, Seattle, Washington 98105

Chair's Introduction—11:00

Invited Paper

11:05

3aPLb1. Stochastic effects in ocean acoustics. Advances in theory and experiment. Barry J. Uscinski (Dept. of Appl. Mathematics and Theoretical Phys., Univ. of Cambridge, Cambridge, UK)

As the need to operate in the kilohertz range increased in the 1960s, existing approaches to ocean acoustics, based on sound-speed profiles and ray-tracing, became quite inadequate. The raggedness of acoustic transmissions, previously regarded simply as noise, was so pronounced at the higher frequencies that the stochastic properties themselves became an urgent area of research. Several major ocean transmission trials were carried out to obtain fluctuation data. The trial at Cobb Seamount produced results that could not be explained by existing random-wave propagation theory. The research that followed led to significant advances in several areas. New propagation theory emerged based on the parabolic equations for the acoustic field moments. In particular it was the construction of analytical solutions of the equation for the fourth moment that provided the first satisfactory explanation of the Cobb intensity fluctuation spectra. Advances in numerical simulation methods and capability have allowed pictures of the full stochastic acoustic field in an ocean transmission trial to be produced. These have not only confirmed experimental variances and cross correlations but have also revealed the spatial structure of the field, including interesing features such as the elongated sound ribbons that arise in some cases.

Meeting of Accredited Standards Committee S2 on Mechanical Vibration and Shock

to be held jointly with the

U.S. Technical Advisory Group (TAG) Meeting for ISO/TC 108 Mechanical Vibration and Shock (and Subcommittees ISO/TC 108/SC1, SC2, SC3, SC5, and SC6)

D. J. Evans, Chair S2 and Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108, Mechanical Vibration and Shock
National Institute of Standards and Technology (NIST), Acoustics, Mass and Vibrations Group, Building 233, Room A147, Gaithersburg, Maryland 20899

Accredited Standards Committee S2 on Mechanical Vibration and Shock. Working group chairs will present reports of their recent progress on writing and processing various shock and vibration standards. There will be a report on the interface of S2 activities with those of ISO/TC 108 (the Technical Advisory Group for ISO/TC 108 consists of members of S2, S3, and other persons not necessarily members of those committees), including plans for 1998 meetings of ISO/TC 108 and/or its Subcommittees.

SCOPE OF S2: Standards, specifications, methods of measurement and test, and terminology in the fields of mechanical vibration and shock, and condition monitoring and diagnostics of machines, but excluding those aspects which pertain to biological safety, tolerance and comfort.

Session 3pPL

Plenary Lecture

Whitlow W. L. Au, Chair
Hawaii Institute of Marine Biology, P.O. Box 1106, Kailua, Hawaii 96734

Chair's Introduction—1:00

1:05

3pPL1. Marine mammal ears: An anatomical perspective on underwater hearing. Darlene R. Ketten (Biol. Dept., WHOI, Woods Hole, MA 02543 and Dept. of Otolaryngol., Harvard Med. School, Boston, MA 02114, dketten@whoi.edu)

Analyzing structure and function in specialized ears can produce new insights into fundamental hearing mechanisms and lead to technological advances. Research into dolphin echolocation is a classic example. Recently, however, concerns over anthropogenic sounds in the oceans pushed us to develop a broader knowledge of marine mammal hearing, and, in the last five years, hearing research on marine mammals expanded considerably. The resulting data on their hearing, ear anatomy, and vocalizations suggest that marine mammal ears are more diverse and complex than previously expected, with acoustic capabilities spanning infra to ultrasonic ranges. Seals are amphibious hearers with middle and inner ears similar to land carnivores, while the ears of whales are strikingly different and are adapted exclusively to hearing underwater. Consistent with high sound speeds in water, specialized fats, not air-filled canals, conduct sound to the ear, and both middle and inner ears are located well outside the skull. Vestibular components are reduced, consistent with cervical fusion related to hydrodynamic body shapes; but their cochlear components, particularly the auditory fibers, are hypertrophied. Neural hypertrophy may be adaptive for high background noise, but may also be related to exceptional signal processing mechanisms in both infra and ultrasonic whales.

Session 3pAA

Architectural Acoustics: The Technical Committee on Architectural Acoustics Vern O. Knudsen Distinguished Lecture

A. Harold Marshall, Chair
Marshall Day Associates, Wellesley Street, P.O. Box 5811, Auckland 01, New Zealand

Chair's Introduction—2:15

Invited Paper

2:20

3pAA1. Concert hall research: Balancing complexity and practicality. John Bradley (Inst. for Res. in Construction, Natl. Res. Council, Montreal Rd., Ottawa, ON K1A 0R6, Canada)

An approach that has carefully balanced complexity and practicality has led to several new developments in concert hall research. These new developments are reviewed and their practical implications are explored. The two separate dimensions of spatial impression, apparent source width and listener envelopment, are each related to different physical parameters. The relationship between these acoustical parameters and the design of halls will be explored. For example, listener envelopment is related to the relatively late arriving lateral sound energy. Measurements in a number of halls will be used to explore how this and other acoustical measures of spatial impression are related to hall design. This same new understanding of spatial impression can also further enhance our understanding of sound fields under balconies and help us to design improved electroacoustic enhancement systems. Spatial impression is also influenced by sound level and especially by lower frequency levels. The importance of levels and the balance between high- and low-frequency levels has often been neglected. New work relating the perception of bass and treble sound to different factors will be reviewed. Previous studies and new analyses will be used to explore how hall design can influence perceptions of tonal balance in concert halls.

Session 3pAO

Acoustical Oceanography and Animal Bioacoustics: Acoustics of Fisheries and Plankton I

Christopher Feuillade, Chair
Naval Research Laboratory, Stennis Space Center, Mississippi 39529-5004

Invited Paper

1:30

3pAO1. An intermediate-range sonar for fish detection. David Farmer, Mark Trevorrow (Inst. of Ocean Sci., 9860 W. Saanich Rd., Sidney, BC V8L 4B2, Canada, dmf@ios.bc.ca), and Bjarke Pederson (LIC Eng., D-2900 Hellerup, Denmark)

A primary challenge in acoustic fish stock assessment is the limited sampling volume of vertically oriented sonars. While long-range imaging with powerful shore-based systems has demonstrated interesting possibilities, practical and environmental constraints motivate an intermediate technology with the goal of mapping fish distributions to ranges of a few km. This work describes a 40-element 12-kHz array deployed both for subsurface towing and also for mounting on the sea floor with motor drive allowing azimuthal rotation. Field tests in the towed configuration have been carried out in the Strait of Georgia, with the array maintained on the sound channel axis at a depth of 40 m. While there was some sensitivity to towing depth, fish could be detected out to the 7-km sampling limit at the same time as fishing activity showed the presence of migrating sockeye salmon. Sea floor deployment in a shallow coastal area demonstrated the sonar's ability to track herring and provides experience in the use of a rotating mount for sector imaging of fish schools. Based on experience gained with this implementation, both the potential applications and limitations of this approach for intermediate range fish detection are identified.

Contributed Papers

1:50

3pAO2. Shallow-water salmon and herring detection using 100-kHz sidescan sonars. Mark V. Trevorrow, David M. Farmer (Inst. of Ocean Sci., 9860 W. Saanich Rd., Sidney, BC V8L 4B2, Canada), and Bjarke Pedersen (LIC Eng. A/S, Hellerup, Denmark)

Acoustic results of fish detection using fixed 100-kHz sidescan sonar installations in the Fraser River (British Columbia) and in the Oresund (near Copenhagen, Denmark) are reviewed. Looking transverse to the river flow, echo traces due to migratory salmon were observed at ranges up to 250 m in water 6 to 13 m deep. The salmon trajectories were readily identified and counted against the background of reverberation from the river surface and bottom sediments. Similarly in the Oresund, long-term monitoring detected migrating herring schools above background reverberation at ranges up to 500 m in water 10 to 13 m deep. In both cases, occasional strong backscatter interference from vessel wakes and wind-induced surface bubble layers was observed. In the Oresund, changes in acoustic propagation conditions due to bottom inflow of more saline oceanic water drastically reduced fish detectability. For salmon in the river, acoustic multipaths due to boundary reflections were distinguishable from direct echoes at ranges up to 150 m. The general problems of shallow-water boundary reverberation and multipath reflection focusing will be reviewed, with discussion on how to compensate for these waveguide effects.

2:05

3pAO3. Multibeam sonar image processing and three-dimensional analysis of fish schools. Chafiaa Hamitouche, Valerie Fracasso, and Carla Scalabrin (ENST-Bretagne, Technopole de Brest-Iroise, B.P. 832, 29285, Brest Cedex, France)

Digital data from a high-resolution multibeam sonar were processed to provide 3-D fish school descriptors. The multibeam sonar operates at a carrier frequency of 455 kHz, with 60 beams of 1.5° each, allowing an observation angle of 90°. The envelope of the signal is sampled every 5 cm. The usual operating range is 100 m with a ping rate of 7 pings/s. This huge amount of data provides larger biovolume samples with higher data resolution than standard fisheries research vertical echosounders used until now to describe fish school structure [C. Scalabrin *et al.*, ICES J. Mar. Sci. **53**, 181–188 (1996)]. Furthermore, the use of a multibeam sonar allows construction of 3-D images of fish schools. Data were processed by specially designed image algorithms in order to account for anisotropy in the range and vessel movement direction planes. Three-dimensional analyses such as smoothing, segmentation, visualization of the segmented volume, and its quantification were performed. Three-dimensional patterns description tools (invariant with respect to translation, rotation, and scaling) were studied. These tools might lead to the classification of the detected structures using pattern recognition methods.

2:20

3pAO4. Near-resonance scattering from schooling fish. R. W. Nero, C. H. Thompson, C. Feuillade, and R. H. Love (Naval Res. Lab., Stennis Space Center, MS 39529-5004)

Most fish have air-filled swimbladders which provide some buoyancy to aid in swimming. These swimbladders resonate like damped air bubbles and are very efficient scatterers. A mathematical formalism has recently been developed to describe near-resonance backscattering from schooling fish. [C. Feuillade, R. W. Nero, and R. H. Love, J. Acoust. Soc. Am. **99**, 196–208 (1996)]. This model predicts that coupled resonance and interference effects will occur when fish are closely spaced. These effects cause the frequency response of a tightly packed school to differ significantly from that of a single fish. As fish spacings increase beyond several body lengths, these effects rapidly diminish. When this model was developed, only one small set of data was available for comparison to model predictions; these data gave indications that the model was valid [Holliday, J. Acoust. Soc. Am. **51**, 1322–1332 (1972)]. Experiments off the west coast of the United States and in the Gulf of Mexico have recently been conducted on several different species of schooling fish in order to confirm the validity of the model and determine its robustness. Measurements on numerous schools confirm that scattering from schooling fish is accurately represented by the model. [Work supported by ONR/NRL.]

2:35

3pAO5. Sound scattering from a few scatterers: Application to swimbladder fish. Zhen Ye and Alberto Alvarez (Dept. of Phys., Natl. Central Univ., Chung-li, Taiwan, ROC)

This paper presents a general method for describing sound scattering by an ensemble of a few scatterers. A scheme is developed for calculating the effective scattering function of such an ensemble, including all orders of multiple scattering between targets. The method is then applied to low-frequency sound scattering by small fish schools. It is shown that in the backscattering, certain regular patterns in scattering appear as a result of interference rather than multiple scattering, when the fish form an ordered structure. The interference peaks are due to in-phase scattering, and can be explained by the Bragg reflection mechanism. Such regular interference patterns are gradually degraded, as the random deviation in each individual fish location from its mean position increases. For a completely random distribution of fish, the model recovers the results from the random-phase approximation. For the forward-scattering case, it lacks either the regular or the noisy interference. The scattering strength seems rather smooth, and may be well approximated by the incoherent summation of forward-scattering function of each individual fish. The scattering by partially or perfectly ordered fish schools is hardly distinguishable from that by totally random distributions of fish in the forward direction. [Work supported by the National Science Council and the Spanish Ministry of Education.]

WEDNESDAY AFTERNOON, 24 JUNE 1998 EAST BALLROOM B (S), 1:00 TO 2:55 P.M.

Session 3pBV

Bioresponse to Vibration/Biomedical Ultrasound: Ultrasound Assisted Drug Delivery II

Pierre D. Mourad, Chair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

Invited Papers

1:00

3pBV1. Chemopotential by low-level ultrasound. George H. Harrison and Elizabeth K. Balcer-Kubiczek (Radiation Res. Div., Univ. of Maryland, 655 W. Baltimore St., 6-015BRB, Baltimore, MD 21201, gharriso@umabnet.ab.umd.edu)

Recent laboratory studies have identified a new approach to cancer therapy: the potential of drug action by ultrasound. Both standard chemotherapeutic cytotoxins and relatively nontoxic compounds related to those used in photodynamic tumor therapy are being investigated. Common to most reported investigations has been the use of continuous wave ultrasound at spatial peak, temporal average intensities exceeding 0.5 W/cm^2, implying the likelihood of temperature elevation in insonated tissue, and usually involving cytotoxicity from the action of ultrasound alone. On the other hand, for potentiation of chemotherapy using tone-burst or pulsed ultrasound at spatial average, time average intensities less than 0.3 W/cm^2 has been demonstrated [G. H. Harrison, E. K. Balcer-Kubiczek, and P. L. Gutierrez, Ultrasound Med. Biol. **22**, 355–362 (1996)]. Chemopotentiation data will be presented for several biological test systems with emphasis on human prostate tumor cell lines. Mechanistic evidence will be analyzed; fluid streaming and shear stress are likely important determinants of chemopotentiation and the cell membrane can be expected to be an important site in this process. Ultrasound-induced modification of drug resistance may have important apoptotic links. These factors suggest the importance of treatment-induced differential expression of membrane-related genes.

1:25

3pBV2. Sonochemistry of sonodynamic therapy. Peter Riesz and Vladimir Misik (Radiation Biol. Branch, Natl. Cancer Inst., N.I.H., Bethesda, MD 20892)

Sonodynamic therapy is a promising modality for cancer treatment based on the synergistic effect of cell killing by a combination of drugs (sonosensitizers) and ultrasound. The effectiveness of sonodynamic therapy was demonstrated in cell studies and in tumor-bearing animals. The mechanism of drug-dependent sonosensitization is unknown, but it seems likely that various mechanisms operate for different classes of sonosensitizers. Ultrasound-mediated sonodynamic activation of porphyrins is particularly poorly understood and hypotheses ranging from singlet oxygen production to peroxyl radical formation have been advanced. Evidence has been obtained against some of the arguments which were proposed in favor of the singlet oxygen hypothesis and the requirement of extracellular localization of gallium porphyrin ATX-70 molecules for sonosensitization in HL-525 cells has been established. Short-lived toxic intermediates produced from sonosensitizers by ultrasound are implicated in the mechanism of sonodynamic activation of some sensitizers. EPR spectroscopy was used to identify proposed intermediates in the activation of certain sonosensitizers (e.g., *N*,*N*-dimethylformamide, DMSO) by ultrasound. Certain water-soluble azo compounds, which are stable at hyperthermia temperatures and can be decomposed by ultrasound to give peroxyl radicals capable of initiating peroxidation of lipids and damaging other cellular sites, appear to be promising sonosensitizers.

1:50

3pBV3. Gene activation and gene delivery with ultrasound. Evan C. Unger (Dept. of Radiol., The Univ. of Arizona, Tucson, AZ 85724-5067, eunger@vms.ccit.arizona.edu), Thomas McCreery, Robert Sweitzer, DeKang Shen, GuanLi Wu, and Veronica Caldwell (ImaRx Pharmaceutical Corp., Tucson, AZ 85719)

Transfection experiments were performed *in vitro* in cell culture and *in vivo* in fish and mice using cationic liposomes with several different reporter genes. The effect of continuous-wave 1-MHz ultrasound on cell viability, cell permeability, and gene expression was assessed. RT-PCR was used to monitor gene expression of cell repair genes in insonated cells. Acoustically active halocarbon-containing liposomes were prepared for gene delivery. Ultrasound increased cell permeability and gene expression following liposomal transfection. RT-PCR showed upregulation of repair of a number of cell repair genes after ultrasound exposure. The acoustically active gene delivery vehicles provided highly efficient transfection both with and without ultrasound *in vitro* and *in vivo*. Following IV administration of halocarbon-containing liposomes/gene complexes, preferential gene expression was observed in the insonated tissue. Ultrasound has potentially useful applications for targeting and enhancing gene therapy.

2:15

3pBV4. Ultrasound activation of new drugs for killing cancer cells. Katsuro Tachibana and Toshiki Uchida (First Dept. of Internal Medicine, Fukuoka Univ. School of Medicine, 7-45-1 Nanakuma, Jonan, Fukuoka, 814-80 Japan)

The study of the destructive action of ultrasound in conjunction with drugs upon cancer has become an exciting area. New substances that are chemically activated by low level ultrasound energy have recently been discovered [Tachibana *et al.*, Lancet **349**, 325 (1997)]. Three nontoxic red stains frequently used for cosmetics and food, Rose Bengal, Eosin Y, and Erythrosine B were evaluated if cytotoxic effects can be induced by ultrasound energy *in vitro*. Gastric and leukemic cancer cell line suspensions were exposed to ultrasound (1 MHz) with or without the drugs. The survival rate of the cells was measured immediately after treatment. The temperature increase was less than 2 °C. Treatment with ultrasound plus drugs resulted in a significant reduction of the survival rate compared to ultrasound alone. Whereas the drug alone showed no change in survival rate. Although the mechanism is unclear, it is postulated that ultrasound energy, possibly cavitation, activated these drugs. Further experiments should be carried out to evaluate its use in the clinical situation. The three stains could be applied as new booster agents for treatment of cancer by ultrasound.

Contributed Paper

2:40

3pBV5. Control of ultrasound-mediated reversible membrane permeabilization for targeted drug delivery. Mark R. Prausnitz, Keyvan Keyhani, Aimee Parsons (School of Chemical Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0100), and Thomas N. Lewis (Georgia Inst. of Technol., Atlanta, GA 30332-0405)

Ultrasound-mediated permeabilization of cell membranes is a potentially useful method to noninvasively enhance delivery of therapeutic compounds into targeted cells. Unlike previous studies concerned only with cell death, in this study the effects of ultrasound exposure parameters on reversible membrane permeabilization of prostate cancer cells were determined. Cells were suspended in a solution containing a membrane-impermeant marker (calcein) and exposed to 24-kHz ultrasound. Viability and molecular uptake were quantified using flow cytometry. Average uptake increased with incident ultrasound pressure above a threshold and reached a plateau (20%–25% of external calcein concentration); increased with total exposure time; was weakly dependent on pulse length (3.3 ms to 1.0 s) with a small maximum at 10–33 ms and a sharp decline at 1.0 ms; and was independent of duty cycle (0.5%–50%). The fraction of viable cells decreased continuously with increasing pressure, decreased as uptake increased for most other examined condition was approximately 25% for maximum uptake under optimal conditions. Uptake decreased approximately exponentially with the time at which calcein was added to the cell suspension after exposure, with a time constant of about 1 min. The membrane disruptions allows uptake of albumin (MW=66 000) but not R-phycoerythrin (MW=240 000).

WEDNESDAY AFTERNOON, 24 JUNE 1998 EAST BALLROOM A (S), 1:00 TO 3:00 P.M.

Session 3pEA

Engineering Acoustics: Advanced Transduction Materials

Elizabeth A. McLaughlin, Chair
Naval Undersea Warfare Center, Code 2131, 1176 Howell Street, B1170, Newport, Rhode Island 02841-1708

Contributed Papers

1:00

3pEA1. Efficiency and electromechanical resonance in magnetostrictive transducers. Frederick Calkins and Alison Flatau (Iowa State Univ., Dept. of Aerosp. Eng. and Eng. Mech., 2019 Black Eng., Ames, IA 50011)

Analysis of the transduction phenomenon of a Tonpilz Terfenol-D magnetostrictive transducer is presented. Transducer efficiency is discussed from electroacoustic and scalar energy perspectives. Equations are presented for identifying the frequency of most efficient transducer operation based on electroacoustic theory [F. V. Hunt, in *Electroacoustics: The Analysis of Transduction, and Its Historical Background* (Acoustical Society of America, New York, 1982)]. Experimental results are presented that show this analysis to be accurate. However, two nonintuitive results are observed: most efficient transducer operation occurs at a frequency above mechanical resonance; and changes in operating conditions which result in an increase in transducer efficiency cause the frequency of highest efficiency to increase even further above the frequency of mechanical

resonance. This motivated a comparison with a scalar energy analysis for development of a better, or at least more intuitive, understanding of the mechanisms which might explain this result. Energy analysis aids in recognizing the distinction between the transducer's mechanical and electromechanical resonances and that the magnitude of achievable efficiency will be enhanced by separation of electromechanical and mechanical resonances. Coupling the energy and electroacoustic analyses, one arrives at persuasive arguments for explaining the experimental observations regarding magnetostrictive transducer efficiency.

1:15

3pEA2. A study of 1-3 piezocomposites high drive limits. Thomas R. Howarth (Naval Res. Lab., Washington, DC, howarth@nrl.navy.mil), David Van Tol, Charles Allen, and Jack Hughes (Penn State Univ., State College, PA)

A preliminary study of the electrical drive limits of 1-3 piezocomposites has been completed. Two injection molded 1-3 piezocomposite samples were designed, fabricated, and procured from Material Systems, Inc. of Littleton, MA. Each of the composites had a 25% volume fraction of piezoceramic, a hard epoxylike matrix, a silver epoxy electrode pattern, 6-mm thickness, 100×100-mm total cross-sectional area with an active cross-section of 20×20 mm. One of the samples featured type II (PZT-5H) piezoceramic while the other featured type I (PZT-4). Each of the transducers was fitted with thermocouples to monitor temperature effects. Transducer efficiency, voltage response (TVR), current response (TCR), and power responses (TPR) were measured over the frequency range of 25 Hz–500 kHz in 200-Hz steps where the transducer thickness resonance frequency was approximately 240 kHz. This presentation shall outline the sample fabrication, the tests, and the final results. Photographs will be shown on the units upon failure. A comparison between the 1-3 piezocomposite materials limits shall be made with traditional piezoceramics. [Work supported by Dr. Kam Ng of ONR 334.]

1:30

3pEA3. The development of affordable constant beamwidth transducers using injection molded 1-3 piezoelectric composite. Kim C. Benjamin, A. L. Van Buren, and Sheridan Petrie (Naval Undersea Warfare Ctr., Newport, RI 02481)

The design, fabrication, and acoustic calibration results for a new class of constant beamwidth transducers (CBTs) is presented. This experimental study extends previously reported work by Van Buren *et al.* [J. Acoust. Soc. Am. **73** (1983)] in which a CBT was constructed using hundreds of individual end-capped PZT ceramic cylinders. Each ceramic element assembly was wrapped in corprene to provide mechanical isolation from the massive metal housing and to provide a pressure release boundary condition. This presentation describes a new design approach which employs injection molded piezocomposite which is curved to form a spherical cap. Design equations will be presented which allow one to dimension the active spherical section for a given application. Of key importance to a successful CBT design is the proper implementation of velocity amplitude shading. The current approach uses an area shading technique that is achieved through selective electroplating of copper. The design, construction, and measured results for three different geometries, including a 0.76-m (30″) diameter unit will be discussed.

1:45

3pEA4. Lamb wave generation in composite plates with a thin linear array of piezoelectric elements. Thierry Demol, Pierrick Blanquet, Emmanuel Moulin, and Christophe Delebarre (I.E.M.N. (UMR CNRS 9929), Dept. O.A.E., U.V.H.C., Le Mont Houy, B.P. 311, 59304 Valenciennes Cedex, France, demol@univ-valenciennes.fr)

Since Lamb waves are very efficient in the nondestructive testing of composite plates, new transducers are developed nowadays. They may be movable, stuck onto the surface, or integrated into the plates (''smart structures''). In order to detect defaults with few transducers, a multi-element linear array transducer has been developed to generate single-mode Lamb waves able to propagate over long distances. Movable prototypes were designed in order to evaluate their ability to generate the appropriate waves. Then, composite plates have been equipped with the same kind of transducers consisting of several thin PZT plates. High response levels have been obtained using a unidirectional composite with propagation along the fiber direction. Results for cross-ply laminates were inferior but still acceptable. The main result obtained during this experimental work was the understanding of the generation process for this kind of transducer. Unlike comb transducers, each element generates a separate wave and suitable time delays or spacing between the elements can improve the global response in accordance with linear rules. Furthermore, the delays are easily determined from the responses of each element. This behavior is very interesting as default detection is possible using the properties of the generation process.

2:00

3pEA5. Built-in transducer design for Lamb wave generation in composite structures. Pierrick Blanquet, Thierry Demol, Emmanuel Moulin, and Christophe Delebarre (I.E.M.N. (UMR CNRS 9929), Dept. O.A.E., U.V.H.C., Le Mont Houy, B.P. 311, 59304 Valenciennes Cedex, France, lamb@univ-valenciennes.fr)

Conventional ultrasonic inspection of large composite structures is very time consuming due to the necessity of surface scanning. Because they can propagate over long distances, Lamb waves are an attractive solution to this problem. Two methods are available for the generation of Lamb waves in composite plates. The first method, and the most useful one, consists of using bonded transducers. The second solution is to use a built-in transducer. In both cases, the transducers must be designed to generate efficiently the signal and choose the mode of propagation. In the following, a method to design a built-in transducer will be discussed. First, the case of a unidirectional composite plate is considered. An analytical method is developed, then the results are compared with those obtained using finite element analysis. In the case of a square element transducer embedded in the composite plate, the computed results are compared with experimental data. Then, the method is extended to the case of a multilayered plate. An analysis is performed to determine the optimal design of a built-in multi-element transducer.

2:15

3pEA6. Optimization of the piezoelectric circular bimorph. Andrzej Dobrucki (Wroclaw Inst. of Technol., Inst. of Telecom. and Acoust., Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland)

The combined finite-element and boundary-integral method was applied for modeling of vibration and sound radiation of piezoelectric circular bimorphs. The bimorph consists of two circular piezoelectric disks, which are separated by a metal plate. When an electric voltage is supplied to the bimorph, a bending moment develops which actuates the transducer to transversal deflection. During the design process of the transducer the following limitations must be taken into account: (1) piezoelectric material and thickness of the disks are usually predetermined, (2) the outer diameter of the transducer, i.e., of the metal plate is given. The influence of the metal plate material and its thickness on the frequency response of the transducer is discussed. Young's modulus, Poisson's ratio, density, and the loss factor characterize the properties of the materials. The influence of the ratio of a piezoelectric disk's diameter to metal plate diameter is also discussed. Another important property which was considered during the

optimization process, is the mounting of the transducer, i.e., applied boundary conditions. As an example, the design of a small transducer is presented. The results of calculations are in good accordance with the result of the measurement of realized construction.

2:30

3pEA7. Research on inhomogeneous 1-3 PZT/polymer composite transducer. DaLun Xu (Dept. of Modern Appl. Phys., TsingHua Univ. Beijing, 100084, PROC) and TieYing Zhou (TsingHua Univ., Beijing, 100084, PROC)

An improved equal stress physical model [W. A. Smith and A. Shaulov, Ultrasonics Symposium, 642–647 (1985)] for the inhomogeneous 1-3 PZT/polymer composite material is presented. This model calculates the properties that determine the thickness-mode oscillations in thin plates. Based on this model, together with the constitutive relations for polymer and PZT, expressions are derived for the effective material parameters such as electromechanical coupling constant, acoustic impedance, etc., in terms of the properties of constituents. A relation between the vibration amplitude of the transducer's surface and the volume fraction of the piezoceramic is also deduced. The distribution of the vibration of the transducer's surface, hence the distribution of the acoustic field, can be precisely controlled by adjusting the distribution of the volume fraction of the PZT phase in the transducer. To test our model experimentally, a number of special transducers were made. The measurements of the acoustic field that is produced by the transducers are in good agreement with the theory. The research provides guidelines for the transducer engineer to design inhomogeneous 1-3 PZT/polymer composite transducer, such as transducers of Gauss type, Bessel type, or other special types.

2:45

3pEA8. The ring-type all-fiber Fabry–Perot interferometer hydrophone system. Donglin Li (Harbin Eng. Univ., Dept. of Underwater Acoust. Eng., Harbin 15001, PROC)

In the paper the theory of the ring-type all-fiber Fabry–Perot interferometer hydrophone is given. Such a hydrophone system has been constructed. Experimental results have shown the system has a free-field open-circuit voltage sensitivity level of about -130 dB *re*: 1 V/μPa.

WEDNESDAY AFTERNOON, 24 JUNE 1998 GRAND CRESCENT (W), 2:15 TO 3:00 P.M.

Session 3pED

Education in Acoustics: Undergraduate Research Poster Session

Murray S. Korman, Chair
Department of Physics, U. S. Naval Academy, Annapolis, Maryland 21402

Contributed Papers

All posters will be on display, and all contributors will be at their posters, from 2:15 p.m. to 3:15 p.m.

3pED1. Nonlinear effects in the paraxial region of diffracting sound beams radiated by cylindrical sources. Svetlana V. Babenkova, Vera A. Khokhlova (Dept. of Acoust., Phys. Faculty, Moscow State Univ., Moscow 119899, Russia, vera@na.phys.msu.su), and Steven G. Kargl (Univ. of Washington, Seattle, WA 98105)

Intense cylindrical sources are widely used in underwater acoustics and medical ultrasound; however, nonlinear-diffraction phenomena in acoustic fields of this geometry are not studied yet as well as the fields produced by unfocused and weakly focused circular piston sources. The present work is based on a special analytical method that has been developed recently [M. F. Hamilton, V. A. Khokhlova, and O. V. Rudenko, J. Acoust. Soc. Am. **101**, 1298–1308 (1997)] to model nonlinear and diffraction effects near the axis of an acoustic beam radiated by apodized circular source. This method combines the advantages of the parabolic approximation with nonlinear geometrical acoustics, and it permits an exact solution of the nonlinear problem. The approach is extended to acoustic waves radiated by a finite-length cylinder. A system of nonlinear coupled equations describing waveform distortion in the paraxial region of the beam is derived from the modified KZ equation. Analytical solutions are obtained in the time and frequency domains for an initially sinusoidal wave radiated by the cylinder with Gaussian shading of the amplitude along the axis. Nonlinear waveform distortion, shock formation distance, harmonic propagation curves, directivity pattern, and effect of steering are investigated for various parameters of the source. [Work supported by ONR, RFBR, and CRDF.]

3pED2. Sentential prosodic contour facilitates adults' short-term memory for nonsense syllable strings. Diana L. Schenck (Dept. of Psych., and Cornell Phonet. Lab., Dept. of Modern Lang., Cornell Univ., Ithaca, NY 14853, dls14@cornell.edu)

Recent linguistic theories propose that prosodic contour functions to facilitate encoding, processing, and storing of information in the speech signal. To date, most evidence for these theories comes from studies of infant perception and processing of prosodic speech. The present study employed a short-term memory CV nonsense syllable recall task in order to attempt to replicate with adults the findings for infants that have supported these theories. Ten adult subjects listened to nonsense CVs in two sets, one presented with sentential prosodic contour, the other presented in list form, and recalled as many syllables, in order, from each string as they could. Results were significant at the $p<0.05$ level and supported the hypothesis that subjects would better recall those syllable strings presented in prosodic contour, as the number of errors was fewer for the prosodic set than for the list set. Small familiarity/order of presentation effects and individual differences were found for all subjects. These results suggest that (1) adults use processing and coding strategies similar to those employed by infants in response to linguistic stimuli; and (2) sentential prosody does, in fact, facilitate short-term memory for these stimuli, possibly by enhancing the perceptual saliency of the linguistic input.

3pED3. Design and development of PC-IMAT: Teaching strategies for acoustical oceanography, part II. Lee Anne Hurley, Kevin M. Thomas, Murray S. Korman (Dept. of Phys., U.S. Naval Acad., Annapolis, MD 21402), John W. Schuler (Navy Personnel Res. and Development Ctr., Brooks AFB, TX 78235), and Eleanor Holmes (Integrated Performance Decisions, Arlington, VA 22202)

The PC-IMAT (Personalized Curriculum for Interactive Multisensor Analysis Training) project is proving to be a flexible and effectively evolving computer-based training/educational platform needed to help tackle ASW and other tasks which require extensive analysis, classification, and interpretational skills. Midshipman taking SP411 (Underwater Acoustics and Sonar) are currently using PC-IMAT [J. Acoust. Soc. Am. **101**, 3096A (1997)] to help investigate effective instructional strategies which convey understanding of a complex multivariate domain (like ray trace or propagation loss models). Classroom demonstration lectures and out-of-class projects allow students to successfully interact with experimental apparatus and make actual measurements (e.g., Fourier analysis, detection theory, beam pattern functions, sound speed versus temperature, computer ray tracing, reflection and transmission at an interface, and target strength versus angle of a scale model sub). However, a link between textbook theories, the actual physics experiments, and the real world setting of tactical sonar has been missing. Usage of PC-IMAT on actual submarine training exercises shows the limitations and enhancements of tactical maneuvering when sonar systems are subject to complex surroundings. The research and development of ''scientific visualization'' on PC-IMAT should ultimately incorporate a strong link between both learning environments. [Work supported by ONR.]

3pED4. Sonic gas analyzers. Matthew V. Golden, Robert M. Keolian, and Steven L. Garrett (Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804, golden@sabine.acs.psu.edu)

Whistles were used to determine the presence of hydrogen and/or methane in mines before 1900 in Germany. Since early in the 20th century, the measurement of gas thermal conductivity using a hot wire technique has been the method of choice for nonspecific detection of ''contaminating'' gases in air. At that time, the choice of thermal conductivity detectors was motivated by the popularity of the Wheatstone Bridge circuit for precision measurement [P. E. Palmer and E. R. Weaver, ''Thermal-conductivity method for the analysis of gases,'' Tech. Papers of the Bureau of Standards, No. 249, Vol. 18, Jan. 7 (1924)]. At the present time, the precision-to-cost ratio for frequency/time measurements exceeds that for voltage/current measurement by more than an order of magnitude. This presentation will examine the advantages which an electronic measurement of sound speed using inexpensive components and novel resonator geometries can bring to the detection of gaseous contaminants in air. Prototype sensor/electronics results will be presented. A discussion of temperature compensation techniques as well as the sensitivity to changes in pressure and humidity will be included. [Work supported by ONR.]

3pED5. Use of pseudo-random codes in pulse compression. Paul Shoning and Antal Sarkady (Dept. of Elec. Eng., U.S. Naval Acad., Annapolis, MD 21402)

Pseudo-random, or pseudo-noise (PN), codes are numeric sequences which appear random to the uninitiated observer yet are easily predictable to those who know the generation algorithm. Traditionally, PN codes have been generated by a hand wired sequence of binary shift registers, but are now created using programmable logic chips wired from user defined netlists. A properly implemented logic chip can generate PN codes of maximum length $L=2^n-1$ where n is the number of shift register segments. The code's period is limited only by the number of shift registers available. Bipolar PN codes have a normalized autocorrelation function that is unity at zero shift and $1/L$ at all other shifts. Thus PN codes can be used to amplitude modulate (AM) the ''main bang'' of an active RADAR or SONAR system. This modulation produces compression which vastly improves the range resolution and the sensitivity of the system. The properties of PN codes and their use in a simulated active SONAR system will be demonstrated using MATLAB programs.

WEDNESDAY AFTERNOON, 24 JUNE 1998

CASCADE BALLROOM I, SECTION A (W), 2:15 TO 3:15 P.M.

Session 3pNS

Noise and Architectural Acoustics: Progress Report and Discussion on the Continuing Activity on ASA's Role in Noise and Its Control

David Lubman, Cochair
14301 Middletown Lane, Westminster, California 92683

T. James duBois, Cochair
9424 Crystal View Drive, Tujunga, California 91042

Chair's Introduction—2:15

A meeting is being sponsored by the Technical Committees on Noise and Architectural Acoustics to review current progress and invite further discussion on activity to increase the role of the ASA in noise, noise control and related architectural acoustics issues. This outreach effort has included activity on increasing public awareness about noise and noise control, on public hearing screening testing and development of self-testing techniques, on seminars on industrial noise, on meeting room acoustic environments and, currently of special interest, on classroom acoustics. Discussions about these and/or new related activity, including encouraging joint activity in these areas with other professional organizations will be encouraged.

Session 3pPAa

Physical Acoustics: Sonochemistry and Sonoluminescence: SC II

Jacques L. Reisse, Cochair
Universite Libre de Bruxelles, Chimie Organique, CP 165, 50 avenue F. D. Roosevelt, 1050 Bruxelles, Belgium

Kenneth S. Suslick, Cochair
University of Illinois, A420 CLS Laboratory, Box 15-6, 601 South Goodwin, Urbana, Illinois 61801

Chair's Introduction—12:55

Contributed Papers

1:00

3pPAa1. Hydrodynamical perturbation effects in multiple-bubble sonoluminescence. Alex Robinson, David W. Kuhns, Anatol M. Brodsky, and Lloyd W. Burgess (Dept. of Chemistry, Univ. of Washington, Seattle, WA 98195-1700)

This presentation will describe the observation of new dynamical effects in multibubble sonoluminescence (MBSL). The experiments show a specific time-dependent nonmonotonic response of MBSL in aqueous and organic solutions to hydrodynamical perturbations caused by sudden pressure changes. Through a fast pressure jump, the bubble field is forced into an excited metastable state, subsequently relaxing over a period of seconds to minutes. The period and intensity of the system excitation is highly dependent upon physical parameters and small concentrations of surface active species. By monitoring the system via sonoluminescence intensity, low level quantitation of alcohols and earth metals in water as well as percent concentrations of water in alcohols has been performed. [Work supported by the Center for Process Analytical Chemistry.]

1:15

3pPAa2. Dependence of multibubble sonoluminescence intensity on the sound field in a sonochemical reactor. Hideto Mitome, Teruyuki Kozuka, Toru Tuziuti (Natl. Industrial Res. Inst. of Nagoya, AIST, 1 Hitare-cho, Kita-ku, Nagoya, 462 Japan, mitome@nirin.go.jp), and Liming Wang (Japan Small Business Corp., Tokyo, 105 Japan)

Since sonochemical reaction is related to the generation of acoustic cavitation bubbles, sonoluminescence induced by the collapse of bubbles can be a measure to specify the reaction field. The present paper discusses dependence of multibubble sonoluminescence intensity on the sound field in a sonochemical reactor to specify the cavitation field quantitatively. Deionized and distilled water was filled in a square reactor cell $50\times50\times140$ mm, which was settled in a temperature-controlled water bath, and was irradiated by cw ultrasound changing the frequency from 20 to 250 kHz for various values of water depth and temperature. The intensity of sonoluminescence was measured by a photomultiplier tube with a converging lens ensuring all the sonoluminescing position was detectable. The experimental results show that the luminescence intensity is very sensitive to the water depth. A change as small as 1 mm alters the vibration mode in the cell. The intensity is also strongly dependent on the amount of dissolved air and temperature. Correlation between the luminescence intensity and the harmonics in the ultrasonic waveform generated by nonlinear vibration of bubbles is shown. [Work supported by Japan Small Business Corporation.]

1:30

3pPAa3. Problems of ultrasonic energy transferring into liquid-phase load. Oleg V. Abramov, Vladimir O. Abramov, Vitaly N. Bulgakov, and Oleg M. Gradov (Lab. of Ultrasound, N. S. Kurnakov Inst., Leninsky prosp. 31, Moscow 117907, Russia, abramov@ionchran.msk.ru)

The rational circuits of construction and modes of operations for a system: electric generator, electroacoustic transducer, sonotrode system, irradiator, and liquid phase are considered from the point of view of transferring vibrations of maximum power into the load. The analysis has shown that for realization of vibration transfer of the greatest possible acoustic power created by the electroacoustic transducer, it is necessary to use the sonotrode system with developed surface. It can be realized by using an irradiator of the tubular form, in which the longitudinal fluctuations are transformed into radial ones. Calculations of tubular type sonotrode systems, based on use of the analytical approaches and of computer modeling, are conducted. A number of sonotrode systems that transform longitudinal vibrations into ones, extending in a radial direction in relation to axes' tubular irradiator, is made and tested. For improving the technological efficiency of the device, the tubular irradiator form was optimized to increase the number of the cavitational bubbles beneficially worked in the liquid and intensity of hydroacoustic streaming. So, for increasing the number of the cavitational bubbles on the external surface of the tubular irradiator at the high-amplitude zones for the longitudinal vibrations, the cross lugs are formed. The stand for valuation of acoustic capacity, entered in liquid load, is assembled, and experimental research at the choice of optimum modes of ultrasonic processing of liquids with various physical properties (viscosity, surface tension, density) is conducted.

1:45

3pPAa4. Enhancement of ultrasonic cavitation yield by a bifrequency irradiation and its frequency effect. Ruo Feng, Channgping Zhu, Huamao Li, and Zhouhua Chen (State Key Lab. of Modern Acoust., Nanjing Univ., Nanjing, 210093, PROC)

The ultrasonic cavitation yield given by a combining orthogonal irradiation at frequencies of 28 kHz and around 1 MHz is studied experimentally by using electric detection method. It is shown that the cavitation yield produced by the bifrequency irradiation is much larger than the sum of the yields produced separately by the two irradiations. For instance, when the intensity of 28-kHz ultrasound is fixed but that of 0.87-MHz ultrasound changes in the range of 4–7 w/cm^2, the cavitation yield of the combining irradiation is equal to 1.9–3.4 times of the sum of the yields given separately by the two irradiations. The ultrasound beams of 28 kHz combined separately with 1.0 and 1.7 MHz are also used. The results show that the enhancement of cavitation yield by bifrequency irradiation is notably dependent on the used frequencies. When the intensity of 0.87-, 1.0-, and 1.7-MHz ultrasound is taken 6.5 w/cm^2, the combining irradiation

gives 4.4, 3.2, and 1.8 times the cavitation yield of 28-kHz irradiation, respectively, which is similar to that obtained previously by using single-frequency ultrasound irradiation. The authors would like to acknowledge Professor T. J. Mason for his valuable suggestion.

2:00

3pPAa5. Two-phase theory of acoustic wave propagation in magnetic liquids. Tomasz Hornowski (Inst. of Acoust., Adam Mickiewicz Univ., Matejki 48/49, 60-769 Poznań, Poland, hornaku@phys.amu.edu.pl), Mariusz Kaczmarek (Inst. of Fundamental Technolog. Res., Mielżyńskiego 27/29, 61-725 Poznań, Poland), Mikołaj Łabowski, and Andrzej Skumiel (Inst. of Acoust., Adam Mickiewicz Univ., Matejki 48/49, 60-769 Poznań, Poland)

A new theoretical approach to ultrasonic wave propagation in magnetic liquids is proposed on the basis of a two-phase model in which the solid phase (magnetic particles coagulating to clusters forming a rigid skeleton) and the liquid phase (carrier liquid) are described by separate sets of equations for the balance of mass, momentum, and angular momentum. Interaction between clusters is described by components of the stress and strain tensors in the solid phase whereas interaction between the clusters and the liquid is accounted for as Stokes, Basset, and added mass forces and the respective angular momenta. Expressions are derived for the velocity and attenuation as functions of the wave frequency and the angle between the propagation direction of the ultrasonic wave and the magnetic field. Moreover, the expressions obtained are checked experimentally for several ultrasonic wave frequencies *versus* the strength and direction of the magnetic field under various conditions. [Work supported by KBN Research Grant No. 2PO3B 07912.]

2:15

3pPAa6. Sonofluorescence image in an analogous sonochemical reactor. Huamao Li, Andong Xie, Fan Zhong, Huijun Wan, Lijun Liu (Sci. Res. Managerial Office, Ji'an Teachers' College, Ji'an, 343009, PROC), and Ruo Feng (Nanjing Univ., 210093, PROC)

According to the sonofluorescence principle of ultrasonic cavitation oxidation suggested by the authors [Huamao Li *et al.*, Acta Acust. **22** (1997)], the cavitation field distribution images in both analogous symmetrical and nonsymmetrical sonochemical reactors filled with aqueous Luminol–KOH alkaline solution have been taken when using an ultrasound with an operating frequency of 1.45 MHz and an ultrasonic electric power of 25 W. The images seem to show that the ultrasonic cavitation field in the analogous symmetrical sonochemical reactor appear symmetrical with uniform distribution, and the ultrasonic cavitation fields in the analogous nonsymmetrical sonochemical reactors appear nonsymmetrical with nonuniform distribution, but the nonsymmetrical cavitation field distribution can be guided along some curved way. The authors think that the sonofluorescence image method can be available for use in a sonochemical reactor designed for getting uniform or positioning cavitation field. [Work supported by the Natural Science Foundation of Jiangxi province and the Science Foundation of Ji'an Teachers' College.]

WEDNESDAY AFTERNOON, 24 JUNE 1998 CEDAR ROOM (S), 1:00 TO 2:15 P.M.

Session 3pPAb

Physical Acoustics: Nonlinear Acoustics I: 1. History; 2. Solids; Rocks; 3. Surface Waves, Part 2

Akira Nakamura, Cochair
3-6-1 Gakuen, Fukui Institute of Technology, Fukui, 910 Japan

Mack A. Breazeale, Cochair
National Center for Physical Acoustics, Coliseum Drive, University, Mississippi 38677

Contributed Papers

1:00

3pPAb1. Interaction of counterpropagating acoustic waves in materials with nonlinear dissipation and in hysteretic media. Hélène Bailliet, Vitalyi Gusev, Pierrick Lotton, and Michel Bruneau (Lab. d'Acoustique, IAM–UMR 6613, Univ. du Maine, av. O. Messiaen, 72085 Le Mans Cedex, France, helene@laum.univ-lemans.fr)

Current interest in the interaction of sound waves traveling in opposite directions is due to progress in thermoacoustics (where the acoustic resonator is a part of the refrigerator) and achievements in diagnostics of solids via the evaluation of standing waves in rods. In the present analysis, the second-order nonlinear processes are described by the equation $\partial^2 V_+/\partial t^2 - c_0^2(\partial^2 V_+/\partial x^2) = -(\partial/\partial t)\{(\partial/\partial x)[(\epsilon-1)V_-^2+V_+^2] + RV_+|V_+| - h[V_-^A+V_-\mathrm{sgn}(\partial V_-/\partial t)](\partial V_-/\partial x)\}$. Here $V_\pm = v_+(\mu x, t-x/c_0) \pm v_-(\mu x, t+x/c_0)$, $\mu \ll 1$, $v_\pm$ are particles velocities, and superscript A denotes amplitude of a periodic wave. Elastic nonlinearity [the first group of terms on the right-hand side (rhs)] does not provide effective interaction of ''$\pm$'' waves. Nonlinear dissipation (the term proportional to R on the rhs) is responsible for the additional attenuation of the ''$-$'' wave induced by the ''$+$'' wave. The chosen form of dissipative nonlinearity is typical of gas-filled porous structures used in thermoacoustics. Hysteretic nonlinearity typical of micro-inhomogeneous materials (the last group of terms on the rhs) leads to amplification of the weak ''$-$'' wave in the field of the strong ''$+$'' wave. Predicted stimulated backscattering of sound in hysteretic media can influence the operation of parametric underground antennas. [Work supported by DGA.]

1:15

3pPAb2. Pseudo-Rayleigh and pseudo-Scholte surface acoustic solitons. Vitalyi Gusev (Lab. d'Acoustique, Univ. du Maine, av. O. Messiaen, 72085 Le Mans, France), Christ Glorieux, and Jan Thoen (Lab. Akoest. Therm. Fys., Katholieke Univ., Leuven, Celestijnenlaan 200D, B 3001 Leuven, Belgium)

Soliton solutions for the equation $s^2(\partial v/\partial x) = \alpha(\partial/\partial\tau)v^2 + \beta(\partial/\partial\tau) \times (H[v])^2 - D(\partial^2/\partial\tau^2)H[v]$, describing wave propagation on a fluid–solid interface covered by elastically linear film, have been found analytically. In this equation x is a ''slow'' spatial coordinate, τ is a ''fast'' time, α and β are nondimensional nonlinear parameters, S is the wave velocity in the absence of a film, and H denotes the Hilbert transform with respect to τ. This equation generalizes the recently proposed model for nonlinear Rayleigh waves [J. K. Hunter, Contemporary Math. **100**, 185–202 (1989)] and the equation derived for Scholte waves [Gusev *et al.*, IEEE UFFC (1998)], where $D=0$, by accounting for linear variation of the wave velocity with frequency $D \equiv \partial s_p/\partial\omega(\omega=0+0)$. Velocity dispersion $s_p = s_p(\omega)$ $[s \equiv s_p(0)]$ is due to interface loading by an acoustically thin

film and described by the last term in the right-hand side of the equation. The problem of the nonlinear surface/interface wave amplitude distribution in the interior of the medium is also addressed.

1:30

3pPAb3. Nonlinear acoustics of solids: History and modern trends. Vladimir A. Krasilnikov (Dept. of Phys., Moscow State Univ., Moscow, 119899 Russia, snk@snk.phys.msu.su)

The author had the pleasure to be involved in the pioneering experiments on the nonlinear acoustic wave propagation in crystals and metals (1960). It was found that even for small wave amplitudes acoustic harmonics generation took place for longitudinal waves. The results obtained were interpreted from the point of view of Hooke's law nonlinearity due to nonparabolic interaction potential taking into account accumulation effects in nondispersive media. Besides that, transversal acoustic waves under some external effects like pressure, heating, etc., also demonstrated nonlinear behavior despite that it was forbidden in a framework of crystal lattice theory. It was accounted for by the development of structural disturbations in a material and seemed to be the first attempt to demonstrate the importance of acoustic nonlinear nondestructive testing of different materials. It was shown later that some construction materials and rocks have much higher nonlinear parameter values than those for crystals and metals. The paper will also present an historical review of recent developments in other areas of nonlinear acoustics in solids: measurements of nonlinear acoustic parameters, surface acoustic waves nonlinearity, ''clapping'' nonlinearity, defective solids and rocks, etc.

1:45

3pPAb4. Measurement of third-order elasticity in isotropic solids. Z. W. Qian (Natl. Lab. of Acoust., Inst. of Acoust., Acad. Sinica, Beijing 100080, PROC, qzw@canna.ioa.ac.cn) and Wenhua Jiang (Nanjing Univ., Nanjing 210093, PROC)

The responsibility of second-order harmonics to the SV incidence was investigated in an isotropic solid such as aluminum, and each sample to be measured was designed as a semicircular disk of which the cross section is a semicircular polygon in order to provide many incidental angles for measuring. The displacements of both the linear and second-order waves were observed in the area near the center of the semicircle, where the surface is stress free. According to the theory proposed by one of the authors (Qian), the third-order elastic constants can be obtained acoustically by the inverse techniques.

2:00

3pPAb5. Nonlinear acoustics of zero-volume nonbonds and cracks. Igor Yu. Solodov (Dept. of Phys., Moscow State Univ., Moscow, 119899 Russia, snk@snk.phys.msu.su)

The results of experimental investigation of nonlinear acoustic phenomena for surface (SAW) and bulk waves (BAW) in materials with nonbonded interfaces and closed cracks are reported. Several mechanisms of acoustic nonlinearity specific for such materials were found and analyzed. In a propagation mode SAW harmonics generation and wave interaction effects revealed two to three orders of magnitude increase in nonlinear parameter values for a nonbonded interface. Unconventional higher harmonics ratio was measured and manifested in a distinctive nonlinear SAW waveform distortion. Efficient generation of harmonics forbidden in homogeneous isotropic solids was observed for shear BAW in the presence of nonbonds and cracks. Peculiar rectification effects were found for SAW propagation and BAW reflection at an interface and shown to result in both acoustic broadening and creeping of a nonbonded area. Extremely high efficiency was measured for nonlinear BAW–SAW transformations accompanying combined three-wave interactions at a nonbonded interface. In a reflection mode such an interface demonstrated a unique feature of nonlinear reflection with inverse spectral distribution of the harmonic amplitudes. Some prospective applications of the effects observed were demonstrated in signal processing, tribology, and nonlinear NDE of zero-volume nonbonds and closed cracks.

WEDNESDAY AFTERNOON, 24 JUNE 1998 DOUGLAS ROOM (S), 12:25 TO 3:00 P.M.

Session 3pPAc

Physical Acoustics: Memorial Session for Isadore Rudnick

Richard Stern, Cochair
Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804-0030

Robert M. Keolian, Cochair
Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804-0030

Isadore Rudnick was born on May 8, 1917 and passed away in Los Angeles on August 22, 1997. He was considered America's foremost acoustical physicist during the second half of the twentieth century, having made seminal contributions to both classical and quantum acoustics and having trained scores of graduate students and several postdoctoral fellows. He received his Ph.D. in Physics at the University of California, Los Angeles (UCLA) and, after work at Duke University and Penn State, he returned to UCLA where he was a professor of Physics for the past 39 years. He held visiting academic appointments in Paris, Haifa, Rome, Tokyo, Copenhagen, and Nanjing. Professor Rudnick was a member of the National Academy of Science and the Governing Board of the American Institute of Physics and received the prestigious Fritz London Award for his research in Low Temperature Physics from the International Union of Pure and Applied Physics. He was a recipient of the Acoustical Society's Biennial Award (1948), its first Silver Medal in Physical Acoustics (1975), and the Gold Medal (1982). He served as the Society's vice president (1962) and president (1969).

Friends, family, colleagues and former students will present reminisces, photos, video clips, and jokes that we hope will remind attendees at this special session of how fortunate we all were to have benefited from the life of this great and wise gentleman and scientist. Attendees will have the opportunity to make informal comments at the conclusion of the scheduled presentations.

Chair's Introduction—12:25

Invited Papers

12:30

3pPAc1. Isadore Rudnick's contributions to nonlinear acoustics. Robert T. Beyer (Dept. of Physics, Brown Univ., Providence, RI 02912)

12:45

3pPAc2. Izzy Rudnick's educational demonstrations and videos. Robert M. Keolian (Applied Res. Lab. and Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804-0030)

1:00

3pPAc3. Metals, superconductors, phase transitions and Izzy. Moises Levy (MJL Assoc., Monaco Beach Club 807, Naples, FL 34103)

1:15

3pPAc4. Dr. Mudnick, rough-on-rates Rudnick, etc. Steven L. Garrett (Grad. Prog. in Acoust. and Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804)

1:30

3pPAc5. My happy years with my mentor—Professor Isadore Rudnick. Junru Wu (Dept. of Physics, Univ. of Vermont, Burlington, VT 05405)

1:45

3pPAc6. Isadore Rudnick: Making waves in superfluid helium. J. D. Maynard (Dept. of Physics, Penn State Univ., University Park, PA 16802)

2:00

3pPAc7. Isadore Rudnick: Making waves in liquid helium. Seth J. Putterman (Physics Dept., Univ. of California, Los Angeles, CA 90095)

2:15

3pPAc8. Twelve-year study on hydrodynamic forced standing soliton of Faraday wave experiment—In memory of the late Isadore Rudnick. Rongjue Wei (Inst. of Acoust., Nanjing Univ., Nanjing, China)

2:30–3:00 Open Microphone

WEDNESDAY AFTERNOON, 24 JUNE 1998 WEST BALLROOM B (S), 1:00 TO 3:00 P.M.

Session 3pSA

Structural Acoustics and Vibration: Nonlinear Dynamics and Chaos

David Feit, Chair

Carderock Division, Naval Surface Warfare Center, 9500 MacArthur Boulevard, West Bethesda, Maryland 20817-5000

Contributed Papers

1:00

3pSA1. Vector fields and quadratic maps. Huw G. Davies and Konstantinos Karagiosis (Dept. of Mech. Eng., Univ. of New Brunswick, Fredericton, NB E3B 5A3, Canada)

Vector fields describing responses of nonlinear systems are often investigated by sampling on a suitable Poincaré section. For example, period-1 limit cycles yield one fixed point, period-2 two points, and so on. More information could be obtained if the full return map on the section rather than just the fixed points were known. Such a map can be obtained numerically from transient analysis of the system. Instead, an analytic approximation of the map is discussed here, in the general quadratic form $x_i(t)=a_{ij}(t)x_j(0)+b_{ijk}(t)x_j(0)x_k(0)$. $x_i(t)$ are perturbations away from the limit cycle, $t=0$ being chosen on the Poincaré section. This equation with $t=T$, the fundamental period, is the required map. $a_{ij}(T)$ is the Floquet monodromy matrix. Differential equations for a_{ij} and b_{ijk} are obtained from a Taylor expansion of the system equations and solved recursively. The three-dimensional Rössler attractor which is known to yield a smooth, one-dimensional unimodal map is used as an example. [Work supported by NSERC.]

1:20

3pSA2. Experimental and theoretical analysis of the nonlinear dynamics and acoustics of a rattling plate. Karen J. Fegelman and Karl Grosh (Dept. of Mech. Eng. and Appl. Mech., Univ. of Michigan, 2350 Hayward St., Ann Arbor, MI 48109-2125)

Rattle has become an important issue in the motive and aerospace industries. Design for the prevention and reduction of rattle noise requires that the underlying mechanisms be understood and powerful, flexible numerical tools be developed. In this paper, the focus is on the former, developing a basic theoretical and experimental foundation for determining the vibroacoustic behavior arising from the nonlinear dynamics associated with the rattle process. In order to understand the fundamental mechanics of rattle, a model problem was formulated involving a hinged plate rattling against a stiff contact point excited by base motion. The plate was modeled as a flexible beam and the resulting equations of motion were solved explicitly. It was found that the calculated accelerations at the tip of the beam quantitatively agree with those measured from an experimental test stand. In addition, the predicted and measured sound pressure levels (SPL) at various points were found to agree qualitatively. Finally, the closed form flexible body solutions were analyzed with regard to stability to predict transitions from periodic to chaotic rattling behavior. In particular, the sensitivity of the transition region to chaos on the base motion amplitude and frequency are investigated experimentally and theoretically.

1:40

3pSA3. Effects of disorder and nonlinearity on the propagation of classical waves. Olivier Richoux, Claude Depollier, Jean Hardy, and Abderrahmane Brezini (Inst. d'Acoustique Mécanique, UMR CNRS 6613, Univ. du Maine, 72085 Le Mans Cedex 9, France, bal@laum.univ.-lemans.fr)

The interplay between the effects of disorder and nonlinearity on the propagation of classical waves is examined by means of an analytical model and numerical simulations. A quasi-one-dimensional string is loaded by N masses, each one fixed to a spring. Disorder is introduced onto the system by considering masses and/or lattice spacing as random variables. Nonlinearity is assumed to arise from the dynamical potential associated to each spring. The wave propagation is formulated in terms of the transfer matrix. The Lyapunov exponent is computed for various system size, yielding the spectrum and the localization length. The numerical simulations have been performed for the linear ordered case which serves as support. The purely disorder linear system offers a direct realization for the observation of the classical Anderson localization. Disorder in mass perturbs the allowed band from the upper edge in contrast to the disorder in position from both edges, the latter appears more efficient in inducing the Anderson localization. The nonlinear system in presence of both order and disorder exhibits bandwidths greater than their respective similar linear case. These findings suggests strongly that nonlinear waves are robust enough to propagate through the system even in the presence of disorder.

2:00

3pSA4. Nonlinear vibroacoustical free oscillation method for crack detection and evaluation. Leonid M. Gelman and Sergey V. Gorpinich (Dept. of Nondestructive Testing, Natl. Tech. Univ. of Ukraine, 37, Peremogy pr, Kiev, 252056, Ukraine, ania@gelman.pp.kiev.ua)

For detection and evaluation of cracks, the vibroacoustical method of free oscillations, based on nonlinear dynamics of a tested object, is widely used. However, in spite of wide practical application, the theoretical basis of this method is insufficiently investigated. The theoretical research of the decrement of the damped nonlinear free oscillations is given in the presence of a crack in a tested object. New generalizations of the decrement for nonlinear oscillations are proposed and investigated. New analytical expressions between the decrement and relative size of a crack, factors of damping, and the natural frequency of a tested object without a crack are received, on the basis of which a new testing feature is proposed and investigated. The cases of small and large damping are considered. The results obtained allow one to explain the phenomena that appeared at the method usage. [Work supported in part by the John D. and Catherine T. MacArthur Foundation.]

2:20

3pSA5. Nonlinear dynamic behavior of a piezoelectric 1-3 composite. Thomas J. Royston (Dept. of Mech. Eng., Univ. of Illinois, Chicago, IL 60607-7022, troyston@uic.edu) and Brian H. Houston (Naval Res. Lab., Washington, DC 20375-5000)

The nonlinear vibratory behavior of a smart material component, a 1-3 piezoceramic composite, is characterized theoretically and experimentally. The developed theoretical model for the electroelastic behavior of the 1-3 composite follows conventional assumptions made by prior investigators but includes higher-order electrical, mechanical, and electromechanical terms in the constitutive equations for the embedded PZT-5 phase. Dynamic experimental measurements of the 1-3's mechanical response to harmonic electrical excitation over a range of excitation frequencies and mechanical loading conditions quantify the level of higher harmonic content (harmonic distortion) in the device and illustrate the dependence of this on the mechanical coupling conditions of the component to its surrounding. While this study focuses on a particular piezoelectric device, the techniques developed are applicable to the wide range of smart materials and structures which exhibit significant nonlinearity, including those based on piezoelectric, electrostrictive, magnetostrictive, and electrorheological technologies, to name a few. [Work supported by ASEE/NRL Summer Faculty Program.]

2:40

3pSA6. Theoretical bases of the vibroacoustical nonlinear forced oscillation diagnostics method. Nadezhda I. Bouraou (Dept. of Orientation and Navigation, Natl. Tech. Univ. of Ukraine, 37, Peremogy pr., Kiev, 252056, Ukraine), Leonid M. Gelman, and Nataliya Yu. Ossokina (Natl. Tech. Univ. of Ukraine, Kiev, 252056, Ukraine)

A new passive vibroacoustical forced oscillation diagnostics method, based on nonlinear dynamics of tested objects, is considered. The method uses nonstationary vibration excitation with changed frequency. For the first time, theoretical research of the nonlinear oscillation system is given with mentioned excitation and the presence of the crack in the tested object. New analytical expressions between the estimates of spectral density of signals, radiated at tested object forced oscillations, and the relative size of the crack, duration of diagnostics, and speed of frequency excitation changing are received. On the basis of mentioned expressions, testing data are proposed and investigated. Comparison with the case of stationary excitation is given.

Session 3pSP

Signal Processing in Acoustics: General Topics

Edmund J. Sullivan, Chair
Naval Undersea Warfare Center, Code 82101, Newport, Rhode Island 02841

Contributed Papers

1:00

3pSP1. Using acoustic impulses to detect buried objects. Charles G. Don, David E. Lawrence, and Andrew J. Rogers (Dept. of Phys., Monash Univ., Vic 3168, Australia)

A detector of nonmetallic buried objects, such as land mines, has been developed using acoustic impulses with a peak energy around 1 kHz. The initial proposal was to compare the difference in signals from two microphones spaced equally on either side of the sound source held a few centimeters above the ground. Ideally, over a uniform surface the difference signal is zero, however, when one microphone is over an object an additional reflected pulse remains after subtraction. In practice, this pulse is small and often obscured by noise. Better results have been obtained if data from a single microphone is subtracted from a ''reference'' waveform. Some problems which have to be overcome are optimizing the alignment of the two pulse trains and deciding whether or not to normalize the signals prior to subtraction. If care is not exercised, either process may distort or mask the required object reflection. This small reflected signal can be further enhanced, compared to the background noise, by correlating with an appropriate known waveform. Mention will also be made of the effects of different media and surface contours.

1:20

3pSP2. Automatic noise source recognition. Didier Dufournet, Philippe Jouenne, and Adam Rozwadowski (01dB, Inc., 1583 E. Genesee Street, P.O. Box 796, Skaneateles, NY 13152)

MADRAS is an acoustic monitoring system that features the automatic recognition of noise sources. It is based on a tree classifier able to use different classification techniques like signal processing, morphology, fractional analysis, and neural networking. A very high accuracy of recognition is achieved due to the optimization of the choice of real-time technique applied and to the intrinsic performance of each algorithm. MADRAS can learn the characteristics of new sources and also use a database of previously processed cases and typical events. The paper presents the results of multiple source recognition in a complex acoustic environment. Implemented on a real time PC-based measuring analyzer called Symphonie, MADRAS has been able to distinguish all noise sources. Acquired data may then be processed to obtain the energy contribution of each source and a list of the frequency of each appearance. Simulations may be viewed and modified.

1:40

3pSP3. A novel octave graphic equalizer. Marija F. Hribsek and Dejan V. Tosic (Faculty of Elec. Eng., Bul. Revolucije 73, 11000 Belgrade, Yugoslavia)

An octave constant Q graphic equalizer using a new class of active bandpass filters is presented. It covers all 11 standard octaves in a frequency range from 16 Hz to 20 kHz. The level in each octave can be changed within ± 20 dB with respect to the reference level which corresponds to the zero dB gain of the equalizer. The advantages of the proposed bandpass biquads are the use of only one operational amplifier per octave, a very small resistors spread ratio, and equal values of the resistors in all octaves. The computer simulation has been used to verify the design of the novel circuit. The new design has been compared to the existing solutions reported in the literature. The comparison shows excellent agreement with the published results. The effects of the finite gain and bandwidth of the real operational amplifier have also been investigated. The proposed new equalizer shows to be robust against these amplifier imperfections.

2:00

3pSP4. Tomographic scanning acoustic microscope system. Daesik Ko, Daekun Song, and Kyesuk Jun (24, mok-dong, chung-ku, Taejon, 301-729, South Korea)

In this paper, a new type of acoustic microscope system using the tomographic image processing technique of STAM (scanning tomographic acoustic microscope) and operation principles of the transmission-mode and detection technique of SAM (scanning acoustic microscope system) is proposed. Since SAM and scanning laser acoustic microscope system (SLAM) used in the past were operated at the reflection mode, these systems could not obtain high-resolution tomographic images. Although STAM could process tomographic images, due to the detection-using laser, problems such as complex detection hardware, alignment, and low S/N ratio occur. If the tomographic image processing technique is used for ultrasonic microscope systems, the axial resolution can be enhanced and it can help and work together with the existing systems to detect cracks that are difficult to find in the reflection mode. These experimental results will be used for various ultrasonic microscope tomographic image processing and ultrasonic measurements.

2:20

3pSP5. Motion compensated high-frequency synthetic aperture sonar. James T. Christoff (Coastal Systems Station, Dahlgren Div. Naval Surface Warfare Ctr., Code R21, Panama City, FL 32407-7001, christoff_jim@ccmail.ncsc.navy.mil)

The Coastal Systems Station has recently completed a set of sea tests using a high-frequency synthetic aperture sonar (HFSAS) mounted on a towed underwater vehicle. The processed results from these tests demonstrate that 2.5-cm resolution can be achieved in shallow water from a small physical array. This system had a Doppler unit to control the ping rate, but had no inertial navigation system and depended entirely upon acoustic backscatter from the seabed to correct for vehicle deviations in the line of sight. In these sea tests the dual overlapped phase center technique [Sheriff, Symposium on Autonomous Underwater Vehicle Technology, pp. 236–245, 1992] was used as the motion compensation algorithm. The tests were conducted over a target field in the Gulf of Mexico that contained a variety of targets, both small sphere and minelike targets. This paper will present results from the HFSAS in the form of a point response function and an acoustic image. The HFSAS achieved between 2.5- and 3-cm resolution on the smallest sphere and produced excellent images of the remaining targets out to a range of approximately 70 m. [Work supported by ONR.]

2:40

3pSP6. A study of extracting characteristic parameters of a tubular target. Renqian Wang and Yingchun Li (Dept. of Electron., Peking Univ., Beijing 100871, PROC)

An inverse problem of sound scattered by a tubular target is studied for extracting its characteristic parameters. A geometrical acoustic method is used to describe the scattered field of the tubular target. A physical model is suggested: the backscattering of a tube shell is simplified into reflection of a plane wave on layered media under conditions: $ka \gg 1$, $c/c2 \gg h/b$, here a, b, and h are outer radius, inner radius, and thickness of the tube shell, respectively, and c and $c2$ are sound velocity of the external medium of the tube shell and (longitudinal wave) sound velocity of the tube shell. Reconstruction formulas of the characteristic parameters of the tubular target are given. The characteristic parameters of an elastic tube shell, such as a metallic tube shell, a metallic tube covered with sound-absorbing materials, a bone tube of a pig, and soft tube shell such as a blood vessel, and a trachea of a pig, are reconstructed by using computer simulation and experiments. The theoretical results obtained by this method are supported by experiments.

WEDNESDAY AFTERNOON, 24 JUNE 1998 METROPOLITAN BALLROOM (S), 1:00 TO 3:00 P.M.

Session 3pUW

Underwater Acoustics and Acoustical Oceanography: Bottom Geoacoustic Characterization and Inversion II (Précis Poster Session)

Subramaniam D. Rajan, Chair
Scientific Solutions, 5907 106th Avenue, Northeast, Kirkland, Washington 98033

In this session, an oral summary of each paper will be presented followed by a poster session. All posters will be on display, and all contributors will be at their posters, from 2:00 p.m. to 3:00 p.m. To allow for extended viewing, posters will remain on display until 5:30 p.m.

Chair's Introduction—1:00

Contributed Papers

1:05

3pUW1. Head wave data inversion for geoacoustic parameters of the ocean bottom off Vancouver Island. Mark C. A. Laidlaw, Oleg A. Godin, and N. Ross Chapman (SEOS, Univ. of Victoria, P.O. Box 1700, Victoria, BC V8W 2Y2, Canada)

As a part of the Pacific Shelf-93 experiment, small explosive charges were used as underwater sound sources on the continental shelf off Vancouver Island. Broadband acoustic signals received by a vertical line array show the rich structure of precursors arriving prior to the direct water wave. Kinematic properties of the precursors are found to be consistent with those of head waves with different numbers of reflections at the ocean surface and bottom. Power spectra of the precursors reveal complicated interference structure, including deep nulls at relatively low frequencies. Positions of the nulls in the frequency domain and their depth dependence are shown to be sensitive to the variation of compressional wave speed in the bottom. A combination of beamforming, spectral analysis, and travel-time considerations is applied to separate and identify individual head wave arrivals in the sequence of precursor signals. Range- and depth-dependence of the travel times for individual head waves are inverted to determine compressional wave speeds in two distinct subbottom layers and to estimate the thickness of the sediment layers. The frequency- and range-dependence of the head waves amplitudes are analyzed to estimate the compressional wave attenuation. [Work supported by NSERC.]

1:12

3pUW2. Acoustic propagation in gassy sediments. Michael D. Richardson, Warren T. Wood (Marine Geosciences Div., Naval Res. Lab., Stennis Space Center, MS 39529, mike.richardson@nrlssc.navy.mil), Aubrey L. Anderson (Texas A&M Univ., College Station, TX 77843), and Roy H. Wilkens (Hawaii Inst. of Geophys. and Planetology, Univ. of Hawaii at Manoa, Honolulu, HI 96822)

Measurements of values of sediment physical properties, bubble volume, and bubble size distribution are used to predict sound speed and attenuation in the fine-grained, gassy sediments of Eckernförde Bay, Baltic Sea. The predicted results are compared to *in situ* and laboratory measurements of sound speed and attenuation over the frequency range of 5–400 kHz. Dispersion of sound speed is used to determine the upper limit of methane bubble resonance near 20 kHz. These data combined with bubble size distribution measured using CT scans of sediments held under ambient conditions yield an estimate of effective bubble size of 0.3–5.0-mm equivalent radii. Histograms of bubble size distribution are then used to predict frequency-dependent sound speed and attenuation based on the model of Anderson and Hampton [J. Acoust. Soc. Am. **67**, 1865–1903 (1980)]. Given the highly variable spatial distribution of bubble volume and size, the agreement between theory and measurement is remarkably good.

1:19

3pUW3. An optimization technique applied to the recovery of acoustic impedance profiles. Marcelo Magalhaes and Roberto Tenenbaum (Acoust. and Vib. Lab., Federal Univ. of Rio de Janeiro, Rio de Janeiro, Brazil)

A method based on a global optimization technique is used to solve the inverse (identification) problem of one-dimensional wave propagation in transversely infinite inhomogenous media. The scheme consists of minimizing a cost function given by the square distance (according to the Euclidean norm) between a measured and a synthetic signal, the latter produced by an arbitrary impedance profile. This cost function imposes several difficulties, since the model for the synthetic signal causes it to be polynomial whose degree is at least equal to the number of layers in which the profile is discretized. The optimization method used is a secant one, which, instead of making use of information about the second derivatives of the cost function, which is too time consuming, generates sucessive approximations of its Hessian matrix, based on information gathered as the procedure progresses. Moreover, in order to guarantee that the method follows a descent direction, that is, the cost function is always reduced, a BFGS (Broyden–Fletcher–Goldfarb–Shanno) updating rule is used, together with a Wolfe line search procedure. Results are shown for some continuous as well as stratified media, showing that the method works well with both smooth and nonsmooth profiles.

1:26

3pUW4. Sequential versus iterative methods for recovering acoustical impedance profiles of inhomogeneous media. Roberto Tenenbaum and Marcelo Magalhaes (Acoust. and Vib. Lab., Federal Univ. of Rio de Janeiro, Rio de Janeiro, Brazil)

Optimization methods for identification of nonhomogeneous media are known to be very time consuming. On the other hand, sequential methods (layer stripping methods) are computationally very efficient and rather unstable under noisy data. In this work, some recent optimization techniques are applied to recover the impedance profile from the reflected data. A careful approach has to be used though, since the cost function is highly nonlinear and its evaluation is quite costly in a computational point of view. The optimization method chosen is a secant one, which, instead of using information about the second derivatives of the cost function, generates successive approximations of its Hessian matrix, based on information gathered as the procedure progresses. The sequential method lies in a new algorithm for computing the reflection response of a truly layered media for an arbitrary input, which can easily be inverted. The results of both methods are compared for some synthetic echoes. The behavior of the methods for random and zero-average noise added to the response is studied. The advantage of using each of the methods as a function of the signal-to-noise ratio is then discussed.

1:33

3pUW5. An experiment on inversion of sound profile through ray modeling. Yuanliang Ma, Xiaoquan Jiang, Yuanhai Sheng (Inst. of Acoust. Eng., Northwestern Polytechnical Univ., Xi'an 710072, PROC, ylma@nwpu.edu.cn), Shixing Yang, and Qingzuo Cheng (CSSC, PROC)

An experiment was conducted in a lake of 150-m depth. A very short hydrophone array (3-m long or 1/50 of the water depth) was used for data acqisition, which consists of 13 hydrophones with 25-cm spacing. Receiving signals are modeled as sums of eigenrays determined mainly by the sound profile, while the sound profile is expressed by six measured samples at different depths in addition to a known sound speed at the bottom. The inversion is based on a simulated annealing algorithm off-line using a 486/100 microcomputer and completed within 2 min. The resultant rms error of sound profile is about 1.6 m/s in comparison with the measured data. An important feature of this approach is that instead of matching the measured to the modeled data vector, arriving structure expressed by the DOA angles and relative time delays are matched for the inversion. To estimate the arriving structure, high-resolution spectrum analysis techniques such as OWSF and MUSIC are adopted. The experiment has shown that inversion of the sound profile can be successfully achieved through ray modeling and using a very short vertical array.

1:40

3pUW6. Waveguide characterization estimation in shallow water using a vertical array. Dejun Jiang (Inst. of Acoust., Academia Sinica, Beijing 100080, PROC)

Waveguide characterization estimation in shallow water using the complex mode approximation of an acoustic field is presented. The advantages of the method are that (1) the dependence on the environmental parameters will obviously decrease. (2) The waveguide characterization parameters (k_m, β_m) can be obtained at the same time, even when the fine structure in the bottom medium, such as the number and thickness of layers, the velocity of compressive and shear wave, and attenuation, is less known or unknown. (3) It can extract the bottom reflection coefficient $V(k)$. Numerical examples illustrate that (1) the accuracy of k_m is at least 10^{-5}, the related error of β_m is mostly less than 20%, and the smaller the mode number, the smaller the related error is. (2) The reflection coefficient in the bottom at low grazing angle can be obtained using a vertical array where the total number of samples is $N=18$ and the length $L=90$ m. The experimental data of 500 Hz are emphatically discussed here. The results for an $f=500$-Hz-wide impulse acoustic signal source which is located at $z_s=11$ m and $r_s=21$ km show that the extracted accuracy of k_m is between 10^{-4} and 10^{-5} and the related error of β_m is between 9% and 26%.

1:47

3pUW7. The error analysis of inverse algorithms for reconstructing the acoustical parameter profile of layered media. Jun-xuan Lin and Qing-hua Bao (Dept. of Phys., Ocean Univ. of Qingdao, Qingdao 266003, PROC)

Reconstructing the parameter profile of layered media is studied with the telemeter of tiny structure in hydrology, oil exploration, and the position selection of petroleum platform in the ocean. It is a much studied topic in the field of ocean acoustics. An inverse scattering experiment of thermocline was completed by us in 1992. The inverse results by the sample iteration algorithm and DC algorithm agreed well with the data measuring by the CTD instrument. It is still an open question. Further discussion in the aspect of reliability is presented in this paper. Generally speaking, inverse errors are caused by: (1) the inverse algorithm itself, (2) signal processing for the stability of computation, and (3) the improper problem. The result of a model inverse experiment shows that the errors caused by the algorithm itself and signal processings are not so critical if the improper problem was not paying enough attention.

ASA Medals and Awards Plenary Session

Lawrence A. Crum, Chair
President, Acoustical Society of America

Presentation of Certificates to New Fellows

Presentation of Awards

R. Bruce Lindsay Award to Robert L. Clark

Honorary Fellowship to Carleen M. Hutchins

Helmholtz-Rayleigh Interdisciplinary Silver Medal to David E. Weston

Gold Medal to Floyd Dunn

Session 4aPLa

Plenary Lecture

David T. Blackstock, Chair
Applied Research Laboratories, University of Texas, P.O. Box 8029, Austin, Texas 78713-8029

Chair's Introduction—8:00

8:05

4aPLa1. Fuzzier but simpler analytic models for physical acoustics and structural acoustics. Allan D. Pierce (Dept. of Aero. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02315, jones@bu.edu)

Acousticians have long been forced to analyze physical systems for which they apparently do not have enough information to make detailed analytical or computational predictions. One could say that their knowledge of such systems is ''fuzzy.'' Alternately, the means of measuring the details or of describing such a system may be in principle possible, but the answers one really desires, those which one can ''deal with,'' are of seemingly far less complexity than the modeled system in all its detail. The trick, of course, is to hypothesize a small number of simple and relevant descriptors of the system and to build a predictive model that uses only these descriptors. The rich and venerable literature of acoustics abounds with successful attempts in this vein, and a few such, taken from diverse subfields such as physical acoustics, architectural acoustics, underwater acoustics, and structural acoustics, are briefly reviewed in this talk. New problems and new questions lead to new searches for appropriate descriptors and for simple models which employ these descriptors. Theoretical tools which help to fill in the gaps caused by the incompleteness of these descriptors are general conservation principles and Jayne's theory of maximum-likelihood based on Shannon's uncertainty function.

Session 4aAA

Architectural Acoustics: Case Study on the New Tokyo Performing Arts Center

Leo L. Beranek, Chair
975 Memorial Drive, Suite 804, Cambridge, Massachusetts 02138-5755

Chair's Introduction—7:45

Invited Papers

7:50

4aAA1. Japan's new National Performing Arts Center. Takahiko Yanagisawa (TAK Associated Architects, Inc., 1-7 Kanda Nishiki-cho, Chiyoda-ku, Tokyo, 101 Japan)

Japan's first National Performing Arts Center, a ''theater city,'' was opened in the Shinjuku district of Tokyo in the fall of 1997. The center is composed of the New National Theatre, which was the winning design in an international architectural competition in 1986, and Tokyo Opera City and covers a total area of 11 acres. The New National Theatre includes an opera house, a drama theater, and an experimental theater, while Tokyo Opera City comprises a concert hall, a skyscraper, and shops. The center promises to become the focal point of theatrical and musical activity in Japan. The architectural goal was to create spaces which would fully engage the human senses, becoming a ''sanctuary of memorable experiences.'' Special focus was placed on the enhancement of the live sound produced by the human voice as well as instrumental sound. The acoustical consultant and associated laboratory were intimately involved from the beginning, and it was therefore possible to integrate their ideas into the design at an early stage. The team also conducted measurements of many of the world's finest halls. Each of the halls boasts its own distinct character, and they are acclaimed by musicians, audiences, and critics alike.

8:05

4aAA2. Acoustical design of the opera house of the New National Theater, Tokyo, Japan. Leo L. Beranek (975 Memorial Dr., Ste. 804, Cambridge, MA 02138), Takayuki Hidaka, and Sadahiro Masuda (TAK Associated Architects, Inc., Tokyo, 101 Japan)

The New National Theater (NNT) project containing three halls started with an open international design competition in 1986 and ended with a grand opening October 10, 1997. The architect is Takahiko Yanagisawa. The opera house, the core of the NNT facility, has a seating capacity of 1810, a volume of 14 500 m^3, and a reverberation time with full audience of 1.5 s (stage curtain open).

Various model experiments, using a CAD model, 1:10 scale model, and full-sized materials samples, were conducted over a 7-year period. As a result, the main floor has an almost rectangular shape, three-layered balconies have a modest fan shape, and the balcony fronts at each level create a rectangular shape. The overhang of each balcony is minimized to take visibility requirements into consideration. The unique design has a large curved reflector in front of the proscenium and over the orchestra pit and curved reflecting surfaces at the fronts of each of the side balconies to reflect the singer's voices uniformly from a large portion of the tremendous stage. The balcony fronts, the side walls, and the ceiling also augment the reflections from those surfaces. The unoccupied values are: EDT=1.6 s, $C_{80}(3)$=2.4 dB, $[I-IACC_{E3}]$=0.66m and bass ratio=1.1.

8:55

4aAA3. Acoustical design of the medium theater of the New National Theater, Tokyo, Japan. Leo L. Beranek (975 Memorial Dr., Ste. 804, Cambridge, MA 02138), Takayuki Hidaka, and Sadahiro Masuda (TAK Associated Architects, Inc., Tokyo, 101 Japan)

The medium theater is mainly intended for the performances of musicals, dramas, dances, and small-size operas. It features use of two types of stage arrangement: a proscenium type with modest fan plan (1038 seats with V=7150 m^3 and occupied RT=1.0 s), and an open stage type with wide fan plan (1010 seats with V=9250 m^3 and occupied RT=1.3 s). The stage types are changeable from one to the other by operation of a movable floor, which is the front half of the main floor, and of slidable side walls. To cope with a wide variety of performing types, two stage types, two scenery lofts, and extensive lighting and loudspeaker requirements, minimal effective initial reflecting surfaces to enhance raw voice were studied through experiments and studies with CAD models and 10:1 scale model tests. The rear wall is a combination of a Schroeder QRD diffuser and absorbing material to eliminate echo. Satisfactory sound strength without amplification is possible only because of the short depth of the hall.

9:10–9:20 Break

9:20

4aAA4. Acoustical design of the Tokyo Opera City (TOC) Concert Hall. Takayuki Hidaka (Takenaka R & D Inst., 1-5-1 Otsuka, Inzai, Chiba, 270-13 Japan), Leo L. Beranek (Cambridge, MA 02138), Sadahiro Masuda, Noriko Nishihara, and Toshiyuki Okano (Takenaka R & D Inst., Chiba, 270-13 Japan)

The TOC Concert Hall, called ''Takemitsu Memorial'' in honor of the late composer Toru Takemitsu, was opened September 10, 1997. With a seating capacity of 1636 and a volume of 15 300 m^3, the hall is sized to cover the musical range from recitals to orchestral concerts. The plan is rectangular in shape but, by request of the architect Takahiko Yanagisawa, the ceiling is a distorted pyramid, with its peak nearer the stage than the rear of the wall. This unique shape had to be analyzed using a CAD model and a 1:10 scale model so that all interior surfaces would be adjusted in shape and absorption to yield optimum values for RT, EDT, $IACC_{E3}$, surface diffusion, initial time delay gap, and loudness [Beranek, *Concert and Opera Halls* (ASA, New York, 1996)]. To provide a better ensemble condition for the musicians on stage and to provide early reflections to several other regions, a square canopy, almost 10 m on a side, is suspended above the stage. The pyramidal ceiling has diffusing elements added to simulate coffers. Schroeder QRDs on the ceiling surface facing the orchestra are used to control echo and to add sound diffusion. With audience, RT=1.95 s, G=6.4 dB, and $[1-IACC_{E3}]$=0.72.

10:10

4aAA5. Relation of acoustical parameters with and without audiences in concert halls and a simple method for simulating the occupied state. Takayuki Hidaka (Takenaka R & D Inst., 1-5-1 Otsuka, Inzai, Chiba, 270-13 Japan), Leo L. Beranek (Cambridge, MA 02138), and Noriko Nishihara (Takenaka R & D Inst., Chiba, 270-13 Japan)

Six acoustical parameters, RT, EDT, C_{80}, D, G, and $IACC_{E3}$, were measured with and without audiences at six concert and opera halls. The range in volume was 3576 to 15 300 m^3. The upholstering of the chairs ranged from light to heavy. Comparisons were made of these parameters before and after addition of the audience. The conclusions are: (1) Predictions of the occupied reverberation times were found inaccurate when the measured unoccupied values were adjusted by using unoccupied versus occupied data for the chairs obtained in a reverberation chamber. If diffuse sound conditions are assumed, it would be possible to predict EDT, C_{80}, D, and G from the occupied RT values whether measured or calculated as above. Such predictions are also inaccurate. (2) Plots of the six parameters, i.e., unoccupied versus occupied values, are nearly straight lines with correlations generally in excess of 0.9. These high correlations permit the use of empirical equations to predict these parameters for occupied halls from measurements made in unoccupied halls. (3) It was also found that $IACC_{E3}$ is almost the same whether measured with or without audience. This supports a previous report that $IACC_{E3}$, for unoccupied halls, is of importance in assessing their subjective quality when occupied [Hidaka *et al.*, J. Acoust. Soc. Am. 989–1007 (1995)]. This paper also reports on the successful use of a cloth covering over unoccupied seats to simulate occupied conditions.

10:30

4aAA6. Objective and subjective measurement of 15 opera houses in Europe and the USA. Takayuki Hidaka (Takenaka R & D Inst., 1-5-1 Otsuka, Inzai, Chiba, 270-13 Japan) and Leo L. Beranek (Cambridge, MA 02138)

Acoustical measurements were executed at 15 major opera houses in eight countries under unoccupied conditions. The objective parameters determined are RT, EDT, C_{80}, D, G, $IACC_{E3}$, and $IACC_A$, which are common attributes needed for concert hall evaluation. The measured unoccupied RT's range from 1.26 to 2.15s and occupied RTs, usually predicted using the empirical equations of the preceding presentation, range from 1.19 to 1.74 s. The range of the other six parameters are also examined. In particular, $[1-IACC_{E3}]$ has a range from 0.39 to 0.72, which is similar to that for concert halls [Hidaka *et al.*, J. Acoust. Soc. Am. 998–1007 (1995)]. Questionnaires have been returned from about 20 opera conductors which enable the authors to present subjective evaluations of the halls. Finally, the subjective evaluations are compared with the objective data.

Session 4aAB

Animal Bioacoustics: Temporal Patterns and Rhythm

Christopher W. Clark, Chair
Section of Neurobiology and Behavior, Bioacoustic Research Program, Cornell University, Ithaca, New York 14850

Invited Papers

8:30

4aAB1. It's all a matter of timing: Temporal mechanisms for production and encoding of acoustic signals. Andrew H. Bass and Deana A. Bodnar (Neurobiology and Behavior, Cornell Univ., Mudd Hall, Ithaca, NY 14853, ahb3@cornell.edu)

Studies of acoustic communication in teleost fish have identified basic principles underlying both the motor control of acoustic signaling and the sensory encoding of the physical attributes of those signals. Parental male midshipman fish excavate denlike nests under rocky shelters in the intertidal and subtidal zones along the western coast of the United States and Canada. Nesting males generate long duration, multi-harmonic signals with a sinusoidallike appearance that are known as ''hums.'' Observations of courtship behavior together with phonotaxis experiments support the hypothesis that the hum functions as a mate call. Early studies showed that a rhythmically active, pacemaker-motoneuron circuit in the hindbrain establishes the temporal features of hums [Bass and Baker, J. Neurobiol. **21**, 1155 (1990)]. Recent studies suggest that temporal coding in the auditory midbrain contributes to the segregation of, and discrimination between, concurrent acoustic signals which form acoustic beats like those generated by the overlapping hums of nesting male midshipman during the breeding season [Bodnar and Bass, J. Neurosci. **17**, 7553 (1997)]. Hence, the rhythmic timing of action potentials is important for both acoustic signaling and reception. [Work supported by NIH and NSF.]

8:55

4aAB2. Temporal rhythms in the signals of insects. T. G. Forrest (Dept. of Biol., Univ. of North Carolina at Asheville, Asheville, NC 28804, tforrest@unca.edu)

The signals generated by acoustic insects exhibit temporal patterns on several time scales. Annual or seasonal rhythms provide information about temporal changes in insect population size. Knowledge of such changes might be useful in ecological monitoring programs. Circadian or daily rhythms commonly seen in insect signaling are generated by endogenous ''clocks'' that are reset by environmental cues. Understanding these daily patterns is also important for monitoring insect populations. Within insect populations, the signals of neighbors sometimes show temporal patterns of synchrony or alternation with timing on the order of seconds. The phase relationships among individuals are maintained by acoustic resetting, whereby an individual's temporal rhythm will change depending on the phase that neighboring signals are received. On still smaller time scales, temporal patterns of insect songs are diverse and generally differ among species. These small-scale patterns are often essential components of mating signals and are used in mate recognition and mate choice.

9:20

4aAB3. Stochastic resonance in the amphibian auditory system? It's just a matter of time. P. M. Narins (Dept. of Physiological Sci., UCLA, 405 Hilgard Ave., Los Angeles, CA 90095)

Noise has traditionally been a factor to minimize or eliminate for optimum performance of a communication system. Recently, it has been proposed that in some cases, noise may *benefit* information transfer and that its existence in sensory systems may be an adaptation to *enhance* detection of weak signals. One possible mechanism that could take advantage of noise to enhance signal transmission is a physical phenomenon known as *stochastic resonance* (SR). In order to study such effects in auditory neurons, use was made of the ability of amphibians to function over a range of ambient temperatures, and thus over a range of internal noise levels. The goal of the present study was to determine the effect of temperature (and internal noise) on information transmission in the frog. To this end, core temperature shifts were induced experimentally and the resulting changes in signal-to-noise ratio (S/N) were quantified. Although our results do not demonstrate SR in the sense that the S/N passes through a maximum at a particular internal noise intensity, we do demonstrate the profound influence of the internal noise on the S/N derived from the neural spike trains. Moreover, recent field and behavioral data will be presented and interpreted in the framework of stochastic resonance. [Work supported by NIH Grant DC-00222 to PMN.]

9:45–10:00 Break

10:00

4aAB4. Whale voices from the deep: Temporal patterns and signal structures as adaptations for living in an acoustic medium. Christopher W. Clark (Section of Neurobiology and Behavior, Bioacoustics Res. Prog., Cornell Univ., Ithaca, NY 14850, cwc2@cornell.edu)

Whales produce long, rhythmic patterns of sounds and some sounds travel many hundreds of miles underwater. Species can be distinguished by temporal features, but there has been a history of infatuation with melodic qualities as the primary features of measurement. Temporal rates and time-bandwidth products are generally related to bathymetry and transmission properties, suggesting that signal features are adapted for communication and navigation. Pelagic species such as blue and fin whales rely on signals in the 10–30-Hz band, presumably to take advantage of the excellent low-frequency propagation properties of the deep ocean, with 10–200-s patterns of sound delivery. Examples will be presented illustrating the remarkable cadence of signal delivery and that animals retain a rhythm after minutes of silence. Shallow water species produce mid-frequency signals (50–1000 Hz) with temporal patterns on the order of seconds. These species with faster rhythms have greater signal variability covering a greater frequency range. At what level is this relationship between temporal pattern and spectral bandwidth indicative of adaptation to optimize for communication and navigation? Is temporal pattern a retained, conservative feature and spectral variability more of an embellishment by individuals to adjust to local conditions?

Contributed Paper

10:25

4aAB5. Time pattern of Sperm whale codas recorded in the Mediterranean Sea 1985–1996. G. Pavan (Centro Interdisciplinare di Bioacustica e Ricerche Ambientali, Università degli Studi di Pavia, Via Taramelli 24, 27100 Pavia, Italy), T. Hayward (Naval Res. Lab., Washington, DC 20375), J. F. Borsani, M. Priano, M. Manghi, and C. Fossati (Università degli Studi di Pavia, Via Taramelli 24, 27100 Pavia, Italy)

Sperm whales, *Physeter macrocephalus*(*catodon*), emit short click sequences, called codas, with regular time patterns [Watkins and Schevill, J. Acoust. Soc. Am. **62**, 1485 (1977)]. Since codas recorded in different geographical areas have different and stable patterns, they possibly serve to convey regional information. More than 120 codas were recorded in the Central Mediterranean Sea in the years 1985–1996 by several research groups using a number of different detection instruments, including stationary and towed hydrophones, military sonobuoys and passive sonars. All of the recorded codas share the same time pattern 3+1 (///-/) with an overall duration ranging from 456 to 1280 ms and an average value of 910 ms. Even if the coda duration varies, the click pattern remains significantly stable. In the present work, the repetition rate and time pattern of Mediterranean codas are characterized, also taking into consideration their possible biological and geographical role. [Work supported by the Italian Ministry of the Environment, Inspectorate for Sea Protection, and by the Italian Navy. Work of the second author supported by ONR Base funding at NRL.]

THURSDAY MORNING, 25 JUNE 1998 WEST BALLROOM B (S), 9:15 TO 11:55 A.M.

Session 4aAO

Acoustical Oceanography and Animal Bioacoustics: Acoustics of Fisheries and Plankton II

D. Vance Holliday, Chair
Tracor, Inc., 4664 Murphy Canyon Road, #102, San Diego, California 92123-4333

Chair's Introduction—9:15

Invited Paper

9:20

4aAO1. Acoustic scattering models of zooplankton from several major anatomical groups—theory and experiment. Timothy K. Stanton, Dezhang Chu (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, tstanton@whoi.edu), and Peter H. Wiebe (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

In order to estimate zooplankton abundance with active acoustics methods, it is crucial to understand the acoustic scattering properties of the zooplankton. The major challenges in formulating scattering models include the facts that there are no exact solutions for these complex bounded bodies and that the boundary conditions are difficult to determine due to animal size and heterogeneities. A series of laboratory scattering measurements and associated model development has been conducted involving zooplankton from three major anatomical groups: (1) fluidlike (euphausiids), (2) elastic shelled (marine snails, both benthic and pelagic), and (3) gas-bearing (siphonophores). The measurements include backscatter versus acoustic frequency (part or all of the range 24 kHz to 1 MHz) and angle of orientation (up to one-degree precision). Dominant acoustic scattering mechanisms are identified through both

spectral and time-domain (pulse compression) analyses and scattering models are developed accordingly. Data and modeling results from the past several years are reviewed and the following new results are presented: (1) a new formulation for predicting scattering at end-on incidence by euphausiids and (2) observations and modeling of a strong peak in scattering levels of the shelled bodies in the upper Rayleigh scattering region. [Work supported by ONR and NSF.]

Contributed Papers

9:40

4aAO2. Covariance Mean Variance Classification (CMVC) techniques: Application to the acoustic classification of zooplankton. Linda V. Martin Traykovski (Dept. of Biol., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, lmartin@whoi.edu), Timothy K. Stanton, Peter H. Wiebe, and James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Accurate acoustic classification of zooplankton species has the potential to significantly improve estimates of zooplankton biomass made from ocean acoustic backscatter measurements. Theoretical models have been developed for three zooplankton scattering classes (hard elastic shelled, e.g., pteropods; fluidlike, e.g., euphausiids; gas bearing, e.g., siphonophores), providing a sound basis for a model-based classification approach. The Covariance Mean Variance Classification (CMVC) techniques classify broadband echoes from individual zooplankton based on comparisons of observed echo spectra to model space realizations. Three different CMVC algorithms were developed: the Integrated Score Classifier, the Pairwise Score Classifier, and the Bayesian Probability Classifier; these classifiers assign observations to a class based on similarities in covariance, mean, and variance, while accounting for model space ambiguity and validity. The CMVC techniques were applied to several hundred broadband (~350–750 kHz) echoes acquired from 24 different zooplankton to invert for scatterer class. All three classification algorithms had a high rate of success with high-quality, low SNR data. The CMVC approach was also applied to several thousand echoes from fluidlike zooplankton to invert for angle of orientation using both theoretical and empirical model spaces; excellent success rates were achieved with the empirical model spaces.

9:55

4aAO3. A DWBA-based representation of the extinction cross section of weakly scattering objects: Application to zooplankton. Dezhang Chu (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543) and Zhen Ye (Natl. Central Univ., Chung-li, Taiwan, ROC)

Because of the weakly scattering nature of the marine organisms, the Distorted Wave Born Approximation (DWBA) has been successfully used in modeling the acoustic backscattering by zooplankton [Stanton *et al.*, J. Acoust. Soc. Am. **94**, 3463–3472 (1993); Wiebe *et al.*, IEEE J. Ocean. Eng. (in press)]. However, in the cases where the scattering-induced attenuation is noticeable, the previously used DWBA-based solutions fail to predict the attenuation since they ignore the imaginary part of the scattering form function and result in a zero extinction cross section. In this paper, a revised DWBA-based approximation is presented to include the imaginary part of the scattering form function. In deriving the revised solution, phase compensation based on ray paths is used. The solutions are compared with the exact solutions, and the agreements between the approximate solution and the exact ones are reasonably good. The results from this study can be applied to bioacoustic applications where the attenuations due to scattering or/and multiple scattering by zooplankton are significant.

10:10

4aAO4. An *in situ* target strength model for Atlantic redfish. Stephane Gauthier and George A. Rose (Fisheries Conservation Chair, Memorial Univ., P.O. Box 4920, St. John's, NF A1C 5R3, Canada, sgauthie@caribou.ifmt.nf.ca)

An *in situ* target strength (TS) model for Atlantic redfish is proposed based on a series of acoustic-trawl experiments conducted at the outer edge of the Green and Grand Banks in Newfoundland waters (NAFO div. 3Ps) in July 1996 and January 1997. Acoustic data were collected using a SIMRAD EK500 echosounder with a hull-mounted 38-kHz split beam transducer and a deep-tow system with a 38-kHz dual beam transducer that can be towed down to 400-m depth. Bottom and midwater trawling sets were performed to account for fish species composition and length distributions. Study sites were selected on the basis of low variance in individual fish size and the absence of other fish species. The model tests the dependence of TS on diel migratory behavior, TS methodology (split and dual beam), distance from transducer, fish density and size, and the presence and relative abundance of zooplanktonic organisms (krill).

10:25

4aAO5. Sensitivity of acoustic scattering models to fish morphometry. J. Michael Jech and John K. Horne (Cooperative Inst. for Limnology and Ecosystems Res., 2205 Commonwealth Blvd., Ann Arbor, MI 48105, jech@glerl.noaa.gov)

Current efforst to model fish backscatter (e.g., Kirchoff ray-mode model [C. S. Clay and J. K. Horne, J. Acoust. Soc. Am. **96**, 1661–1668 (1994)] use digitized images of fish anatomy for a more realistic representation of swimbladder shape, volume, and aspect. However, the degree of image resolution necessary for acoustic characterization and improved predictions, and the variability associated with resolution has not been quantified. Does a more accurate and higher resolution image result in improved correlation between theory and empirical measurements? Backscatter amplitudes display strong relationships with carrier frequency, swimbladder shape, and swimbladder aspect at rather high frequencies. Are predictions of these relationships influenced by image resolution? X-ray images of Atlantic cod (*Gadus morhua*) and brook trout (*Salvelinus fontinalis*) were digitized at fine resolution, and sensitivity analyses were performed to quantify the effect of varying image resolution on predicted backscatter as a function of carrier frequency and swimbladder shape and aspect. Quantifying image resolution is a step towards determining morphometric effects on backscatter variability and estimating fish abundance using multi-frequency data and the inverse approach. [Work supported by ONR.]

10:40–10:55 Break

10:55

4aAO6. Quantifying intraspecies variation in acoustic backscatter models. John K. Horne and J. Michael Jech (Cooperative Inst. for Limnology and Ecosystems Res., 2205 Commonwealth Ave., Ann Arbor, MI 48105, horne@glerl.noaa.gov)

Fisheries researches are increasing the use of frequency-, length-, and aspect-dependent backscatter models in population abundance estimates of fish and zooplankton. Model predictions are often based on measurements from a single or a limited number of organisms. The resulting model parameters are used to predict backscatter for organisms of any length in the population. At rather high frequencies (i.e., length/wavelength ratios of 1 to 20), choice of carrier frequency, swimbladder shape, and swimbladder aspect influence amplitudes of returned echoes. Morphological differences among individuals will therefore determine the variance observed in amplitudes of echo ensembles. To quantify variance in predicted backscatter within species as a function of carrier frequency, fish length, and swimbladder aspect, we combined data from digitized x rays with Kirchhoff ray-mode backscattering models of brook trout (*Salvelinus fontinalis*) and Atlantic code (*Gadus morhua*). Quantifying backscatter variances within and between species allow us to examine the applicability of multi-frequency data and the inverse approach to estimate fish length abundances. [Work supported by ONR.]

11:10

4aAO7. Estimation of the target strength of juvenile salmonids at any aspect. Thomas J. Carlson (USCE—Waterways Experiment Station, CENPP-PE-E, P.O. Box 2946, Portland, OR 97206-2946)

Many tasks, such as assessment of the state of stocks and observation of migrant behavior, elemental for recovery of Columbia River basin salmonid stocks, are currently performed using active acoustic methods. Detectability models are required to configure acquisition systems, design elements of studies, and process and analyze data. Accurate and precise estimates of juvenile salmonid target strength are necessary input to detectability models. The target strength at 420 kHz of 87 juvenile salmonids was measured within the length range of 45 to 300 mm. The fish were rotated through pitch, roll, and yaw planes. Forty target strength measurements were obtained for each degree of rotation in each plane for each fish. Mean target strengths were computed for 1° and 5° increments of rotation. The mean target strength estimates for each 5° increment of rotation for all fish were fit to the equation $\sigma/\lambda^2 = a(L/\lambda)^b$. Estimates of mean target strength for regions of the sphere encompassing the fish between the planes of rotation were estimated by interpolation to complete a description of the target strength of juvenile salmon at any aspect.

11:25

4aAO8. Target strength measurements of walleye pollock (*Theragra chalcogramma*). Jimmie J. Traynor (Natl. Marine Fisheries Service, NOAA, Seattle, WA, jimt@afsc.noaa.gov)

The results of recent target strength measurements of walleye pollock from the North Pacific are presented. The measurements, made with a lowered transducer system are compared with historical results. The system allows the transducer to be moved closer to the target fish thus reducing the well-known bias of *in situ* target strength measurements due to range-dependent noise thresholds. Results using a conventional system with the transducer mounted on the research vessel and this system are discussed. The AFSC currently uses a target strength-to-length relationship [TS=20 log(L)−66.0, L, length in centimeters] to scale echo integration information to estimates of fish density. The appropriateness of this practice is discussed in light of the recent target strength measurements. Caveats regarding the limitations of *in situ* target strength measurement techniques are presented and suggestions for appropriate conditions for such measurements are provided.

11:40

4aAO9. Three-dimensional acoustic measurements of zooplankton swimming behavior in the Red Sea. Duncan E. McGehee (Tracor Appl. Sci., 4669 Murphy Canyon Rd., San Diego, CA 92123, dmcgehee@galileo.tracor.com), Amatzia Genin (H. Steinitz Marine Biol. Lab., 88103 Eilat, Israel), and Jules S. Jaffe (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0238)

Three-dimensional swimming trajectories of several hundred thousand individual zooplankters were measured using the 445-kHz acoustical imaging system FishTV during a three night moored deployment in the Gulf of Eilat, Israel. The sonar examined a 5-m^3 volume at 27 m depth, in water 300-m deep. The image rate was 4 images/s. Targets were tracked from image to image using an automatic three-dimensional tracking algorithm. Over 14 000 targets remained in view for over 5 s, and their trajectories were used in subsequent analysis. Data from net tows indicated that most targets were euphausiids. The horizontal speeds of targets −75 dB and below were highly correlated with flow measurements from an S4 current meter. These targets also exhibited strong vertical motions, apparently due to internal waves. Estimated mean flow was subtracted from each trajectory to compute the swimming speeds of the animals themselves. These were generally much lower than the mean flow. During the night, variance in the vertical component of the flow-removed tracks was much greater than variance in the horizontal. However, the variance became more isotropic as dawn approached. A hop-and-sink foraging behavior offers one possible explanation for this.

THURSDAY MORNING, 25 JUNE 1998 EAST BALLROOM B (S), 9:15 A.M. TO 12:00 NOON

Session 4aBVa

Bioresponse to Vibration/Biomedical Ultrasound: Applications of Microbubble Based Echo Contrast Agents I

Inder Raj S. Makin, Cochair
Ethicon Endo-Surgery, 4545 Creek Road, Cincinnati, Ohio 45242

Nico De Jong, Cochair
Erasmus University, Room Ee2302, 3000 DR Rotterdam, The Netherlands

Invited Papers

9:15

4aBVa1. Specific characteristics of ultrasound contrast agents. Nico De Jong and Peter Frinking (Erasmus Univ., Rm. Ee 2302, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands)

Ultrasound contrast agents possess acoustical characteristics which differ from the surrounding medium. Specific signatures of the agent, consisting of free or encapsulated gas bubbles, will improve the discrimination between the blood containing the agent and the tissue. One of the specific signatures is the initiation of a volume pulsation when the ultrasound wave hits a free or encapsulated gas bubble. Depending on the magnitude of the ultrasound wave the pulsation will be linear to the applied pressure or nonlinear to the applied pressure. Nonlinear vibration can be split into stationary (harmonic) and transient (power) scattering. Besides the nonlinear vibration (harmonics), specific characteristics include enhanced scattering as a function of the acoustic amplitude, transient scattering which is very useful for measuring flow, LOC (loss of correlation) imaging, and the persistence as a function of ambient parameters like pressure and temperature. Recently, a 2-D echomachine has been adapted for imaging of the nonlinear vibration of contrast agents and further developments in both hardware and software will be advantageous for the many specific characteristics of these agents.

9:35

4aBVa2. Nonlinear properties of microbubbles and applications to medical ultrasound imaging. Volkmar Uhlendorf, Thomas Fritzsch, Michael Reinhardt, and Frank-Detlef Scholle (Res. Labs. of Schering AG, 13342 Berlin, Germany, volkmar.uhlendorf@schering.de)

Gas bubbles smaller than 10 μm dissolve within milliseconds even in gas-saturated liquids, but coatings, etc. can prevent dissolution. These stabilized microbubbles serve as transpulmonary contrast agents. When only their linear properties are employed, physical laws limit the diagnostic potential to detection in large diameter vessels. Nonlinear acoustic properties of contrast agents permit very important additional applications. Nonlinearity arises from bubble pulsations, shell properties, finite amplitude waves, and electronic hardware. The first two sources can dominate, allowing new diagnostic imaging modes sufficiently sensitive to detect isolated microbubbles *in vivo*. Harmonic imaging modes detect 2nd harmonics of the transmit frequency, mainly from microbubbles. Acoustic Emission modes destroy bubble shells by one pulse of moderate amplitude. Consequently, the free bubbles respond strongly to this and other pulses before dissolving. Observed lifetimes of 1–20 ms are enough for harmonic or conventional Doppler detection. Typically, bubbles are replenished too slowly to get similar signal intensity for more than one frame at normal rates. Optimal pulse timing should allow new microbubbles to move close to the beam axis before the next pulse arrives. Applications may be counting methods, flow tagging, and quantification of tissue perfusion, possibly combined with single-shot three-dimensional acquisition and HI modes.

9:55

4aBVa3. Harmonic imaging for microbubble contrast agent detection. P. N. Burns (Imaging Res., Sunnybrook Health Sci. Ctr. S660, 2075 Bayview Ave., Toronto, ON M4N 3M5, Canada, burns@srcl.sunnybrook.utoronto.ca), C. T. Chin, D. Hope Simpson (Medical Biophys., Univ. of Toronto), and J. Powers (ATL Ultrasound, Bothell, WA 98041-3003)

The detection of microbubble contrast agents in the vessels of the microcirculation is currently limited by the relatively low concentrations attainable without causing acoustic attenuation, set against the high clutter signal imposed by surrounding soft tissue. By inducing nonlinear resonant oscillation in a population of microbubbles in transit in the circulation, echoes at the second harmonic can be detected preferentially, thus segmenting the blood signal from that due to tissue. Real-time harmonic imaging and Doppler have been implemented using a clinical array imaging system, and it is shown that with bubble agents it can detect flow in 40 μm vessels of the kidney. At higher incident peak pressures (above about 0.5 Mpa at about 3 MHz), some microbubble contrast agents are irreversibly disrupted by ultrasound. They then produce transient echoes which are high in amplitude and rich in harmonics. Furthermore, transient echoes can be obtained repeatedly over a period of milliseconds, allowing correlation imaging. Such ''transient Doppler'' imaging can demonstrate microcirculatory blood in the capillaries of the moving myocardium, a new development with significant clinical potential. [Work supported by Medical Research Council of Canada.]

10:15

4aBVa4. Acoustic characterization of contrast agents for medical ultrasound imaging. Lars Hoff (Dept. of Telecommunications, The Norwegian Univ. of Sci. and Technol., N-7034 Trondheim, Norway, hoff@tele.ntnu.no) and Per C. Sontum (Nycomed Imaging AS, N-0401 Oslo, Norway)

Nycomed's ultrasound contrast agent *NC100100* has been investigated by *in vitro* acoustic measurements. Acoustic attenuation spectra were used to determine resonance frequencies of the particles. The spectra were correlated with size distributions, and it was found that the shell-encapsulated gas bubbles can be described as viscoelastic particles with bulk modulus 700 kPa. When exposed to hydrostatic overpressures mimicking those found *in vivo* during the systolic heart cycle, the resonance frequency increased, as expected by the particles' increased stiffness. This effect was reversible: After the pressure was released, the particles went back to giving the original attenuation spectrum. This shows that the particles are not destroyed or otherwise changed by the pressure. Acoustic backscatter measured as a function of distance through a contrast agent was used to estimate the backscatter efficiency of the particles, that is, the ratio between scattered and absorbed ultrasound. Results from these measurements agree with theoretical estimates based on the attenuation spectra. Measurements on *NC100100* were compared with earlier results from measurements on *Albunex*® and measurements on an experimental polymer-encapsulated contrast agent, showing how different shell materials cause differences in particle stability and stiffness.

10:35–10:45 Break

Contributed Papers

10:45

4aBVa5. Nonlinear response of microbubbles to pulsed diagnostic ultrasound. Sascha Hilgenfeldt, Detlef Lohse (Fachbereich Physik, Univ. of Marburg, Renthof 6, D-35032 Marburg, Germany), and Michael Zomack (Schering AG, Berlin, Germany)

The sound radiation from micrometer-size bubbles driven by short ultrasound pulses emitted by diagnostic ultrasonography devices is investigated. The frequency spectrum is analyzed in order to predict useful parameter ranges for the application of the second harmonic method.

11:00

4aBVa6. Acoustical nonlinearity parameter of liquids with microbubbles. X. F. Gong, S. G. Ye, D. Zhang (Inst. Acoust., State Key Lab. Modern Acoust., Nanjing Univ., Nanjing, 210093, PROC), S. S. Feng, R. Q. Zhang, R. T. Wang (Acad. Sinica, PROC), Z. Z. Xu, L. M. Liu (Shanghai Zhong Shan Hospital, China), and K. L. Ha (Pukyoung Natl. Univ., Pusan, Korea)

The gas bubbles in liquid are the strong scatters of the sound propagation in such liquid. A liquid containing microbubbles is used as an effective ultrasound contrast agent in medical diagnosis to improve the

contrast of the ultrasonic image. However, the existence of bubbles in liquid may enhance its nonlinearity parameter. In our previous paper, the preliminary results of some ultrasound contrast agents were reported and their large values of the nonlinearity parameter were obtained. This paper is devoted to an experimental demonstration of the influence of microbubbles on the nonlinearity parameter B/A values. These include measuring the dependence of B/A values of Echovist 300 on time. Results show that B/A values decrease with time (from 2678 to 16 over 45 min). B/A values of human blood with a different portion of Echovist were also studied. Larger B/A values are obtained with more Echovist. B/A values of a kind aqua, such as 76% Injectio Meglumini Diatrizoatis Composita (MDC), were studied before and after sonication with different bubble contents. Results indicate that the values of nonlinearity parameter depend on the presence of microbubbles obviously. Some explanation of this effect is discussed. [Work supported by NSF of China.]

11:15

4aBVa7. Bubble dynamics of ultrasound contrast agents. Michalakis A. Averkiou, Matthew F. Bruce, and Jeffry E. Powers (ATL Ultrasound, P.O. Box 30003, Bothell, WA 98041)

The bubble dynamics involved with microbubble contrast agents under insonification is investigated. The acoustic field of an ATL HDI-3000 diagnostic ultrasound system in a contrast specific harmonic imaging mode is reviewed first, and its major features that are related with microbubble behavior are discussed. Issues relating to sound attenuation, mechanical index, and bubble destruction are addressed. The nonlinear oscillatory behavior of contrast microbubbles is modeled with the Gilmore equation. The acoustic pressure field of a short pulse utilized in harmonic imaging is measured with a hydrophone and used as the driving pressure of the Gilmore model. Radius-time $[R(t)]$ and bubble wall velocity-time $[U(t)]$ curves are shown. Frequency domain analysis of $U(t)$ indicates transient resonance characteristics in both the fundamental and second harmonic components that are somewhat different from what one would expect with a continuous-wave steady-state response. The times for complete solution of microbubbles in water are calculated and correlated to observations seen in ultrasound images with contrast agents. Radio frequency (rf) data of scattered pulses from contrast agent microbubbles in an *in-vitro* experiment were collected with a phased array. This data is used to support and explain the contrast microbubble behavior.

11:30

4aBVa8. Estimation of contrast agent concentration using spectrum analysis. Cheri X. Deng (Riverside Res. Inst., 330 W 42nd St., New York, NY 10036, cheri@rrinyc.org), Frederic L. Lizzi (Riverside Res. Inst., New York, NY 10036), Ronald H. Silverman, and D. Jackson Coleman (Cornell Univ. Medical College, New York, NY 10021)

A theoretical scattering model was formulated to calculate the calibrated power spectrum of ultrasound contrast agents. The analysis incorporates the scattering coefficient from an encapsulated contrast agent bubble and includes the effects of realistic focused beams. The results relate specific spectral features to the characteristics of contrast agents, including their concentration and size distribution. Analytical calculations and experimental studies were conducted to analyze the calibrated power spectra of backscatter data using Albunex® and Aerosomes™ over frequency bands employed in typical medical ultrasound (3–12 MHz) and very-high-frequency ultrasound (VHFU, 15–50 MHz). The results showed that contrast agent concentration can be estimated from the spectral features such as the spectral slope and intercept values, which are measured using linear regression analysis. These studies also demonstrated the feasibility of employing contrast agent with VHFU, using wide-band transducers with center frequencies near 40 MHz. The VHFU scans of *in vivo* rabbit eyes showed the time history of contrast agent activity and revealed increased backscatter (as much as 8 dB) within subsegments of the ciliary body. Estimation of contrast agent concentration as a function of time promises to provide key information about perfusion in a contrast agent perfused organ.

11:45

4aBVa9. Detection and estimation of microbubble size distribution in blood. Klaus V. Jenderka, Georg Dietrich, Ulrich Cobet (Inst. of Medical Phys. and Biophys., Medical Faculty, Martin Luther Univ., D-06097 Halle, Germany), and Petr Urbanek (HP-medica, Bahnhofstrasse 30, D-86150 Augsburg, Germany)

The characterization of the microbubble size distribution in blood is of great practical interest for the design and the use of heart–lung machines and for clinical research. Another application is the continuous estimation of size distribution of ultrasound contrast agents. The measurement system is based on a 2-MHz pulse Doppler device and a special transducer, designed to connect to a 3/8-in. tube. The computer controlled system consists of two simultaneously working channels. The system is self-calibrating. The amplitude of a reference echo from a reflector behind the tube is used for the compensation of different acoustical properties of the tube. The minimal size of detectable gas bubbles and also the resolution of the size distribution is 1 μm in diameter. The detected size distribution will be measured continuously and saved every 5 s. Therefore it is possible to study the time course of the number and the size of gas bubbles, for example, during a heart operation. Showing the effectiveness of an arterial blood filter is done by measurements before and after the filter. Examples of the size distribution of contrast agents in different suspensions versus time after injection will be presented.

Session 4aBVb

Bioresponse to Vibration/Biomedical Ultrasound: Medical UItrasound I—Transduction and Propagation

Junru Wu, Chair
Department of Physics, Box 2-11200, University of Vermont, Burlington, Vermont 05405

Contributed Papers

9:15

4aBVb1. Ultrasonic phased array design for reduced crosstalk. John Dodson and Karl Grosh (Dept. of Mech. Eng. and Appl. Mech., Univ. of Michigan, Ann Arbor, MI 48109-2125)

In recent years, therapeutic ultrasound has received increased attention as a treatment modality for cancer therapy (hyperthermia) and heart arrythmia (ablation surgery). Ultrasonic phased arrays offer the benefit of noninvasive treatment with the flexibility of variable focus and beam steering. In designing and building these arrays, electromechanical isolation is extremely important as interelement crosstalk will degrade array performance. However, manufacturability and cost, constraints which may conflict with the goal of reduced crosstalk, must also be a consideration. Efficient numerical methods present a useful design tool for array design and allow the investigation of structural acoustic phenomena in general. In this paper, a two-dimensional finite-element model for an ultrasonic phased array in contact with the human body is used to study array design. The effects of material selection and geometry for matching layer(s) and interelement structural matrix are examined for their impact on radiation pattern, power input to the fluid domain, and deleterious effect on the electromechanical behavior of the transducers. Finally, design guidelines are formulated and, if possible, compared to experimental results.

9:30

4aBVb2. A new method of ultrasonic hydrophone calibration using KZK wave modeling. Hendrik J. Bleeker (ATL Ultrasound, 22100 Bothell Everett Hwy., P.O. Box 3003, Bothell, WA 98041) and Peter A. Lewin (Drexel Univ., Philadelphia, PA 19104)

Most commercially available ultrasonic transducers exhibit finite amplitude distortion in water during hydrophone measurements needed to comply with regulatory requirements. The frequencies observed due to finite amplitude distortion can easily exceed ten times the transducer center frequency and 100 MHz. Typically, hydrophone calibrations are supplied only up to 15 or 20 MHz and do not exhibit a flat response. The frequency response above 15 MHz should be known to accurately represent the acoustic information, especially for high-frequency transducers ranging between 7.5 and 15 MHz. A new hydrophone calibration technique has successfully predicted the frequency response of hydrophones up to 100 MHz. A circular source transducer was first characterized and then modeled using the KZK wave propagation model. This model accounts for diffraction, absorption, and nonlinearity. The transducer frequency response was measured with a hydrophone and compared to the simulation. This difference characterized the frequency response of the hydrophone and was used to estimate the hydrophone calibration. The estimated calibration at 20 MHz was checked and provided good agreement with the manufacturer calibration supplied. Acoustic measurement accuracy will be improved if the hydrophone frequency response is deconvolved from the actual acoustic transducer response.

9:45

4aBVb3. High-frequency backscatter measurements of bovine tissues with unfocused and focused transducers. Subha Maruvada and Kirk K. Shung (Grad. Prog. in Acoust. and Bioengineering Prog., Penn State Univ., State College, PA 16801, sxm104@psu.edu)

In order to improve resolution of ultrasonic imaging, high-frequency scanners which are employed in the range 20–50 MHz are needed. As a result, it is critical to obtain data on ultrasonic scattering and attenuation in this frequency range. At these high frequencies, it is not feasible to make scattering measurements with unfocused transducers, therefore focused transducers are needed. Using the standard substitution method to calculate the backscatter coefficient, as is used with unfocused transducers, yields erroneous results for focused transducers. The assumption that the reflected echo from a perfect reflector in the far field can be calculated as though the transducer acts like a point source does not hold for focused transducers. Therefore a method is presented for focused transducers where the flat reflector is substituted by a particulate reference medium whose backscatter coefficient is well known and documented. Unfocused transducers in the range 10–20 MHz will be used in comparison with focused transducers to measure backscatter. For 15–35 MHz, focused transducers will be used. Results for the backscatter coefficient will be presented for various bovine tissues.

10:00

4aBVb4. The acoustic properties of microcalcifications in the context of breast ultrasound. Martin E. Anderson (Dept. of Biomed. Eng., Duke Univ., Durham, NC 27708), Mary S. C. Soo (Duke Univ. Medical Ctr., Durham, NC 27708), and Gregg E. Trahey (Duke Univ., Durham, NC 27708)

Breast microcalcifications are small crystals of calcium phosphates which are important diagnostic indicators under mammography. Despite continuing advances in the resolution and sensitivity of modern scanners, the role of ultrasound in the breast clinic is limited by the inability of current ultrasound practice to visualize microcalcifications reliably. While mammography is currently the gold standard for their detection, approximately 70% of breast biopsies prompted by the presentation of microcalcifications result in negative biopsy. The improved visualization of microcalcifications under ultrasound could provide additional diagnostic information in these cases, potentially reducing the number of negative biopsies. A review is presented of the theoretical and empirical results to date of this group's studies of the acoustic properties of microcalcifications as characterized through the application of classical theoretical models, including the Faran model, and through clinical studies of suspected *in vivo* microcalcifications and breast tissue backscatter. The similarities found between theoretical and clinical results strongly suggest that microcalcifications have elastic scattering properties. The findings are presented in an analysis of the relative impact of phase aberration, spatial resolution, and tissue backscatter on the visualization of microcalcifications with ultrasound. [Work supported by NIH Grant No. RO1-CA43334.]

10:15

4aBVb5. Acousto-mechanical imaging for cancer detection. Armen Sarvazyan (Artann Labs., 22 Landsdowne Rd., East Brunswick, NJ 08816)

One of the hot areas of ultrasonic medical diagnostics is elasticity imaging (EI). Recently, an alternative method of imaging tissue structures in terms of their elasticity, the method of mechanical imaging (MI), has emerged. The essence of MI is the solution to an inverse problem using the data of stress patterns on the surface of tissue compressed by a force sensor array [A. P. Sarvazyan, Proc. Tenth IEEE Symposium ''Computer based medical systems,'' Maribor, Slovenia, 120–125 (1997)]. A prototype of the device for mechanical imaging of the prostate has been developed and feasibility of MI technology has been proven in recent clinical studies. The present paper describes a modality of medical imaging, acousto-mechanical imaging (AMI) comprising features of both EI and MI. The data on the stress pattern measured by the force sensor array of MI complement the strain data obtained by ultrasonic EI device. An embodiment of the AMI device comprises a force-sensing array made of sound transparent film inserted between an ultrasonic scanner probe and the tissue. Synergy of combining these two complementary methods results in the fact that for certain applications the AMI technology has a diagnostic potential much superior to that of EI and MI separately.

10:30–10:45 Break

10:45

4aBVb6. Shear wave excitation in a rubberlike medium by focused shock pulse. Yury A. Pishchal'nikov, Valery G. Andreev, Oleg V. Rudenko, Oleg A. Sapozhnikov (Dept. of Acoust., Phys. Faculty, Moscow State Univ., Moscow 119899, Russia, yura@naexp.phys.msu.su), and Armen P. Sarvazyan (Artann Labs., East Brunswick, NJ 08816)

Ultrasonic excitation of shear waves is important for elasticity imaging of biological tissue. In the shear wave elasticity imaging shear strain is remotely induced in tissue by the radiation force of focused modulated ultrasound. Previous study has shown that tone bursts of millisecond duration at megahertz frequencies with intensities and exposures typical for diagnostic ultrasound may induce bulk radiation force of the order of 0.01 N/cm^3 and shear displacements of the order of 10 μ. In this work, new nonlinear mechanisms are discussed that provide detectable shear displacements by means of single submicrosecond pulses. Two new possibilities are discussed. Significantly enhanced radiation force and greater shear displacement can be produced in tissue by single acoustic videopulse with nonzero momentum. Another possibility is based on the enhanced nonlinear absorption which occurs after the shock is formed in the temporal profile of the pulse. Nonlinear mathematical models describing the process of generation of shear waves using single pulses are developed. Experimentally, single focused pulses are excited by high-power laser radiation inside an optoacoustic cell. Shear displacements are detected in a sample of gel using optical scattering. Results of modeling and experiments are in good agreement. [Work supported by CRDF and RFBR.]

11:00

4aBVb7. A virtual source method for evaluating ultrasound propagation through tissues mapped from medical images. Jie Sun and Kullervo Hynynen (Div. of MRI, Dept. of Radiology, Brigham and Women's Hospital, Harvard Med. School, 221 Longwood Ave., Boston, MA 02115)

A numerical model which uses digitized layer interfaces was developed for calculating multilayer ultrasound propagation, reflection, and refraction. This model was used to project the acoustic field from a spherical phased array (diameter 10 cm, F number 1, frequency 1 MHz) into multilayer tissues, and was verified by performing phantom measurements. A section on each layer interface in the propagation path was used as a virtual source. Digitized profiles of the tissue layer interfaces of arbitrary shapes can be obtained from MRI, CT, or ultrasound. The phase delay for each element of the phased array was first calculated, and then applied to each element for phase correction. With no phase correction, the focus of the array was shifted and defocused. By using the phased array, the shifted focal point was corrected and the side lobes were reduced. The effects of the array element size (number) were investigated for a deep sonication where the layer thickness information was obtained from a series of MRI scans of a volunteer. For this particular array, it is shown that the element size of 0.39 cm^2 is small enough to produce a near-optimum focus. [Work supported by NCI Grant No. CA 46627.]

11:15

4aBVb8. An A-mode ultrasound technique for tracking the advance of coagulation front in laser irradiated tissue. Zhigang Sun, Hao Ying, Brent Bell, Massoud Motamedi (Biomed. Eng. Ctr., The Univ. of Texas Medical Branch, Galveston, TX 77555-0456), and Jialiang Lu (The Univ. of Houston, Houston, TX 77204)

This communication presents an A-mode ultrasound technique for automatically determining the advance of coagulation damage front in laser irradiated tissue. The basic assumption underlying our technique is that the waveform of the echo signal scattered from a tissue region undergoing coagulation should be changing more rapidly than tissue regions where no coagulation is taking place. In this technique, we first track rf echo signals scattered from many small tissue regions during heating by computing the cross-correlation coefficient between two consecutively acquired echo signals. We then use the resulting coefficients as a measure of waveform change to determine the position of coagulation front via an automatic procedure. To test our technique, we carried out ten *in-vitro* experiments in which pig liver samples were irradiated using an Nd:YAG laser with fluence in the range of 32.4–112.0 W/cm^2 for 290 s. A 10-MHz broadband single-element spherical focused ultrasound transducer was used to detect the thermal coagulation front. The mean-square difference between ultrasonically and histologically determined coagulation depths was 1.0 mm, whereas the mean coagulation depth was 7.2 mm. This good agreement between ultrasonically and histologically determined results shows the potential of our technique for monitoring coagulative tissue damage during thermal therapy.

11:30

4aBVb9. Two-dimensional nonlinear propagation of pulsed ultrasound through a tissue-like material. Ibrahim M. Hallaj, Steven G. Kargl (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105), Ronald A. Roy, and Robin O. Cleveland (Boston Univ., Boston, MA 02215)

During a lithotripsy operation or ultrasound surgery, high-intensity acoustic waves propagate through tissue. Damage and thermal dose delivered to the target site and to tissue neighboring the target site is of significant interest in such operations. Results of a study of the propagation of intense ultrasonic pulses through a tissuelike material are presented. Two-dimensional simulations showing the effects of nonlinearity in inhomogeneous absorbing materials are used to calculate the acoustic pressure field, from which input is obtained for a bioheat equation calculation. Results show the mean and peak acoustic pressures within a target volume as a function of time, and the thermal dose delivered to a target volume due to a pulsed source in the range of frequencies used for ultrasound surgery. The degradation of the focusing ability of an ultrasonic array in nonlinear inhomogeneous media is illustrated, and conclusions regarding therapeutic ultrasound are made. [Work supported by ONR and DARPA.]

Session 4aEA

Engineering Acoustics: Sensors for Smart Systems

Robert D. Corsaro, Chair
Naval Research Laboratory, Code 7135, 4555 Overlook Avenue, SW, Washington, DC 20375-5000

Chair's Introduction—9:15

Invited Papers

9:20

4aEA1. Sensor engineering at the microscale. Thomas B. Gabrielson (Appl. Res. Lab., Penn State Univ., P. O. Box 30, State College, PA 16804, tbg3@psu.edu)

Modern techniques for microfabrication have led to visions of significant new sensor applications. However, the translation of devices from centimeter scales to micrometer scales can cause effects that are unimportant at the larger scales to dominate the behavior at the smaller scales. When sensor structures approach material grain sizes or ferroelectric or ferromagnetic domain sizes, design philosophies need to be adjusted; when clearances are very small, viscous and thermal effects can dominate the dynamic response. While microfabrication has made dimensional downscaling of several orders of magnitude possible, intuition and ''engineering practice'' developed on the macroscale may lead to unsound devices on the microscale. Successful miniaturization of devices and structures depends on an understanding of the changing roles of various forces, energies, and interactions on the appropriate scales. Even with this level of understanding, optimum performance may not be achieved by simple shrinking of a macroscale device. Nowhere is this more evident than in biological organisms where markedly different solutions to sensing, searching, navigation, locomotion, and thermal regulation are used by organisms of different sizes.

9:45

4aEA2. Area averaging sensors for vibro-acoustic control. Brian H. Houston and Robert D. Corsaro (Naval Res. Lab., Code 7130, Washington, DC 20375)

NRL has been conducting acoustic control research for underwater and in-air sound and vibration applications. In many of these studies, the success of the control system depends substantially on the ability of the sensors to measure the physical quantities of interest accurately while ignoring extraneous contributions. For example, the sensors developed for use in the Acoustic Boundary Control (ABC) research were required to monitor accurately the normal component of the average acoustic field on the tile surface, while being insensitive to transverse and high wave number components. The extremely low noise figures required in this study forced the development of special-purpose acoustic pressure and velocity sensors which now are finding use in other applications, including vibration isolation mounts.

Contributed Papers

10:10

4aEA3. Sound power emission measurement and control on a 300-t hydraulic press. Fabrizio Bronuzzi, Caterina Cigna, Mario Patrucco (Dip. Georisorse e Territorio, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, patrucco@vdiget.polito.it), and Maurizio Sassone (''E. Sassone'' s.r.l., Monale (AT) Italy)

The paper summarizes the results of research work carried out to evaluate properly the noise emission characteristics of a 300-t hydraulic press in a complex noise propagation environment. On the basis of the collected data, some control measures were designed, and the results are discussed. As well known, noise is widely affected by the acoustical environment the sound propagates through, and an effective description of the source emission properties is required for the control design. Traditional measurement techniques are not the best for this purpose, because the information on acoustic pressure depends on the environment. To overcome this problem, the emitted acoustic power was then evaluated by means of acoustic intensity measurements; in such a way it is possible to qualify real and complex sound sources, locating the contribution of single parts that participate in sound generation. Resulting maps describing the contribution of every single part, an effective individuation of improvement strategies was possible, since the priorities for intervention were available. A new campaign of sound intensity measurements confirmed the expected control measures' results, and the direct measurements of the operators' exposure confirmed the success of the intervention.

10:25

4aEA4. Active control of sound transmission through an industrial sound encapsulation. Paul Sas, Wouter Dehandschutter, Rene Boonen, and Antonio Vecchio (Dept. of Mech. Eng., Div. PMA, Katholieke Universiteit Leuven, Celestijnenlaan 300B, B-3001 Heverlee, Belgium, wouter.dehandschutter@mech.kuleuven.ac.be)

In this paper, an active control system is developed to enhance the sound transmission loss of a sound encapsulation at low frequencies. The sound encapsulation is used to shield a noise source, such as an air compressor. The control systems comprise two control loudspeakers and two error microphones. The Filtered-X LMS adaptive feedforward control algorithm is used to drive the signals at the error microphones to a minimum. The reference signal for the control algorithm is taken from the noise source, i.e., the tacho signal of the compressor. This study deals with the optimization of the location of control sources and error sensors in view of the achieved control performance. From this perspective it is important to note that minimizing the sound-pressure level at the error microphones, which are constrained to be located inside the sound encapsulation, does not necessarily reduce the acoustic intensity measured in the

far field of the encapsulation (or at least, not to the same degree). Relevant issues which are considered are: the coupling of control sources with the acoustic modes of the enclosure, changes to the interior sound field due to active control, etc.

10:40

4aEA5. Structural vibration mode imaging using photorefractive holography. K. L. Telschow and V. A. Deason (Idaho Natl. Eng. and Environ. Lab., Lockheed Martin Idaho Technologies Co., MS 2209, Idaho Falls, ID 83415-2209, telsch@inel.gov)

Photorefractive processing of optical interference offers a noncontacting optical method for vibration detection that forms an optical ''lock-in'' amplifier. Multiwave mixing with synchronous detection allows measurement of both the vibration amplitude and phase of a vibrating surface directly as a function of the excitation frequency. Narrow bandwidth detection with flat frequency response can be achieved at frequencies above the photorefractive response ($\sim$100 Hz). A minimum detectable displacement amplitude of a few picometers has been demonstrated for a point measurement, with the possibility of further improvement. The method can also be configured to provide full-field imaging of resonant vibrational modes of materials with various microstructures and shapes. Both specularly and diffusely reflecting surfaces can be accommodated. Results will be presented showing the noncontacting imaging capabilities for vibrational spectral analysis of structures and characterization of material properties that these sensors provide for smart systems. [Work suported by the U.S. Department of Energy.]

10:55

4aEA6. Piezoelectric resonant sensor for sound velocity of liquids. Helmut Nowotny (Institut für Theoretische Physik, Vienna Univ. of Technol., Wiedner Hauptstr. 8/136, A-1040 Wien, Austria), Ewald Benes, Branka Devcic-Kuhar, Martin Gröschl, Dagmar Harrer, Rudolf Thalhammer, and Felix Trampler (Institut für Allgemeine Physik, Vienna Univ. of Technol., Wiedner Hauptstr. 8/134, A-1040 Wien, Austria, benes@iap.tuwien.ac.at)

The sensor consists of two circular piezoelectric plates with the liquid in between. Each piezoelectric plate is a bulk resonator with two electrodes driven in thickness extensional modes. The two resonator plates are electrically connected in parallel and the total admittance is measured in the vicinity of two quasiharmonic series resonance frequencies by a PC-controlled electric admittance measurement system. The extensional thickness mode purity was checked directly by corresponding mode pattern measurements using a laser speckle vibration amplitude measurement system. The sensor element is mounted in a thermostated aluminum housing to avoid temperature effects and to obtain the sound velocity of the liquid at a defined temperature. Two quasiharmonic resonance frequencies are used to allow the elimination of the unknown density of the liquid. The sound velocity is calculated out of a rigorous one-dimensional theoretical model of the three-layer sensor arrangement. In addition to the determination of the sound velocity of the liquid, the mass density of the liquid can be obtained from the same primary measurements. However, at present the accuracy of the density value is one order of magnitude lower than that of the velocity (typically 0.1% for water, acetone, glycerin).

11:10

4aEA7. Transient behavior of acoustic gyrometers. Philippa Dupire and Michel Bruneau (Lab. d'Acoustique, IAM—UMR 6613, Univ. du Maine, av. O. Messiaen, 72085 Le Mans Cedex, France, amb@laum.univ-lemans.fr)

Acoustic gyrometry was developed during the past 15 years as a new miniaturized or low-cost technology. The operation of acoustic gyrometers employs acoustic fields inside fluid-filled resonant cavities to determine angular velocities. Until now, research efforts and design methodology have concentrated both on trapezoidal miniaturized gyros (etched on silicon chips) as well as on cylindrical gyros designed using classical techniques. These approaches are restricted to the frequency domain, which involves only the steady Coriolis effect. Nowadays, the need for a time domain solution for inertial effects on acoustic fields clearly arises when dealing with applications involving strong variations of the rotation rates. This paper aims at providing such advances in ''inertial-acoustic'' theory and modeling. The discussion covers cylindrical gyros because the presence of an unsteady rotational velocity gradient (with respect to the radial coordinate) of a gas in a cylindrical cavity adds one of the most important features both in the basic physics underlying ''inertial-acoustic'' transient processes and in the behavior of the gyros. In conclusion, only some new requirements, which have to be taken into account in the design of acoustic gyros, will be addressed as a discussion of the transient behavior in detail is beyond the scope of this paper. [Work supported by DRET-DGA.]

THURSDAY MORNING, 25 JUNE 1998 GRAND CRESCENT (W), 8:00 TO 10:45 A.M.

Session 4aMU

Musical Acoustics: Timbre of Musical Sound I

James W. Beauchamp, Chair
University of Illinois, 2136 Music Building, 1114 West Nevada, Urbana, Illinois 61801

Invited Papers

8:00

4aMU1. Spectral versus harmonic information for timbre: Pilot and experimental results. Einar Mencl (Haskins Labs., 270 Crown St., New Haven, CT 06511, einar@tom.haskins.yale.edu)

This work summarizes results from six experiments on steady-state timbre perception. The primary variable of interest was the physical similarity of tones within each trial: tones were either similar by virtue of their spectral envelope (i.e., they shared resonance characteristics) or their harmonic envelope (i.e., they shared overtone characteristics). These two continua were chosen as candidates for basic perceptual dimensions because (1) they each contain specific types of information about the sound source/event, and (2) a cursory computational analysis suggests that very different processes are needed to extract information about resonances versus overtones from the auditory input. It was reasoned that if different subsystems subserve the processing of timbre along each of these two dimensions, then performance on perceptual tasks would dissociate under appropriate experimental manipulations. These included

(1) attentional focus and response type (timbre similarity judgment, pitch direction judgment, or pitch contour judgment), (2) perceptual priming, and (3) auditory masking. Results demonstrate that in general both dimensions of similarity can support perceptual priming, but also show specific dissociations which inform the understanding of steady-state timbre perception. [Portions of this research were supported by NSF Grant No. DBS-9222358 to Jamshed Bharucha.]

8:25

4aMU2. Macrotimbre: Contribution of attack, steady state, and verbal attributes. Gregory J. Sandell (Parmly Hearing Inst., Loyola Univ., 6525 N. Sheridan, Chicago, IL 60626, sandell@sparky.parmly.luc.edu)

The timbre attributes of a source that are contained across variations in pitch, dynamic, and articulation can be referred to as its ''macrotimbre.'' The timbre of a musical instrument changes across pitch, and thus learning its macrotimbre can be dependent upon the ability to make comparisons among individual pitches. Using 20 pitches shared by two easily confused instruments ($C4$ to $G5$, oboe and English horn), listeners performed better at categorizing novel notes (from a pitch range different from those heard in training) when training facilitated comparison of notes (versus isolated presentations) [Sandell and Chronopoulos, Proc. Third ESCOM Conf., Uppsala, Sweden, 222–227 (1997)]. This indicated that information contained across pitches was useful in generalizing timbre knowledge to previously unheard notes. The present study used (1) multiple versions of the tones (natural, shortened, attack + synthetic steady-state, synthetic steady-state alone), (2) a sorting task under conditions where comparisons were facilitated or hindered, and (3) ratings of verbal attributes (nasal, breathy, muffled, etc.) to isolate the relevant across-pitch timbre properties, and appraise the relative contributions of the attack and steady state.

8:50

4aMU3. Methods for measurement and manipulation of timbral physical correlates. James W. Beauchamp (School of Music and Dept. of Elec. and Comput. Eng., Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801)

Multidimensional scaling calculations based on the results of listening tests comparing musical sounds almost always identify one perceptual dimension which is strongly correlated with a measure of the spectral centroid, whereas other dimensions are less obviously correlated to particular physical parameters. Some parameters which have been suggested as good correlates are attack time, spectral flux, and spectral irregularity [J. Krimphoff *et al.*, J. Phys. IV (C5), 625–628 (1994)]. Also, spectral rolloff, preattack noise, general noisiness, and inharmonicity have been proposed as important factors. Based on time-variant spectral analysis with custom software [J. Beauchamp, Audio Eng. Soc. Preprint No. 3479], these parameters and others can be quantified and manipulated, leading to various applications. For example, synthesis using simplified parameter contours provides stimuli for discrimination testing to reveal listeners' sensitivities to these parameters. Measurement of spectral differences can be correlated to discrimination results. Equalization of parameters common to two instrument sounds allows listeners to focus on those timbral aspects which are different. Finally, hybrid musical sounds created by interchanging parameter contours between instruments can shed light on timbral perception.

9:15–9:30 Break

9:30

4aMU4. Perceptual independence of excitor and resonator properties in percussive instrument sounds. Stephen McAdams (Inst. de Recherche et de Coordination Acoustique/Musique (IRCAM), 1 place Igor Stravinsky, F-75004 Paris, France), Koei Kudo, and Holle Kirchner (IRCAM, F-75004 Paris, France)

The perceptual independence of various sensory dimensions has been studied using a speeded classification paradigm developed by Garner. In this paradigm the speed of classification of items varied along one dimension is compared to the same classification when random variation along a second dimension is present. Perceptual interaction is inferred when performance in the latter is slower or worse compared to the former. Such interactions have been shown for simple dimensions such as pitch, loudness, and a spectral dimension of timbre. The present study tested the independence of complex dimensions with overlapping spectral and temporal characteristics. One dimension consisted of the hardness of a percussion mallet (excitor property) and the other of either xylophone bar density or tympani size (resonator properties). These complex source-related dimensions show evidence of interaction in both performance scores and reaction times. A greater independence is found when unrelated variation is introduced along a third dimension (striking force), although performance scores decrease globally. The results are discussed in terms of previously reported interactions among simple dimensions and interactions among complex source properties.

9:55

4aMU5. The effect of amplitude and centroid trajectories on the timbre of percussive and nonpercussive orchestral instruments. John M. Hajda (Prog. in Systematic Musicology, Dept. of Ethnomusicology, UCLA, Box 951657, Los Angeles, CA 90095-1657)

This study explores the salient acoustical attributes of the associative aspects of Western orchestral instrument timbres. Association refers to a subject's ability to connect or link a stimulus with its intended comparison; this occurs in tasks such as identification, matching, or categorization. Unlike most previous research in instrument association, this study considers both continuant (nonpercussive) and impulse (percussive) classes of instruments. Certain acoustical transformations, such as the partitioning of a signal into attack, steady state, and decay, are not possible for both classes. Findings which have been made regarding the salience of signal segments must be questioned in terms of their application to a more general and inclusive theory of musical timbre. Previous perceptual and timbre synthesis research has demonstrated the importance of global rms amplitude and spectral centroid (the

amplitude-weighted mean frequency). Here, both amplitude and centroid are considered in terms of individual trajectories. The evolutions of and relationships between these trajectories are markedly distinct for continuant and impulse classes of instruments. The perceptual significance of these trajectories is measured through the experimental control of their magnitude and direction. A convergence method is employed to ascertain the necessary and sufficient conditions for instrument association.

10:20

4aMU6. Computer identification of wind instruments using cepstral coefficients. Judith C. Brown (Phys. Dept., Wellesley College, Wellesley, MA 02181, brown@media.mit.edu)

Earlier results on computer recognition of the oboe and saxophone [J. C. Brown, J. Acoust. Soc. Am. **101**, 3167(A) (1997); ''Musical instrument identification using cepstral coefficients as features'' submitted to J. Acoust. Soc. Am.] have been extended to include the clarinet and the flute. Approximately 30 samples of duration 2–10 s of each of the four instrument types comprise the test set. The training set consists of longer segments of approximately 1 min duration. Eighteen mel-based cepstral coefficients were calculated for each of the sounds. The training data were summarized by a cluster analysis with Gaussian probability density functions formed from the mean and variance for each of the clusters. The probability of belonging to each of the four classes was then calculated for the test sounds, and a Bayes decision rule was invoked to assign them to one of the classes. Results indicate that this is a very promising method for automatic recognition of musical instruments.

THURSDAY MORNING, 25 JUNE 1998

CASCADE BALLROOM I, SECTION A (W), 8:40 TO 10:45 A.M.

Session 4aNS

Noise: Acoustical Materials and Propagation

Trevor R. T. Nightingale, Chair

National Research Council, IRC Acoustics, Building M27, Montreal Road, Ottawa, ON K1A 0R6, Canada

Chair's Introduction—8:40

Contributed Papers

8:45

4aNS1. Effect of the flat panel resonance on the absorption characteristics of a microperforated-panel absorber. Ke Liu (Inst. of Acoust., Academia Sinica., P.O. Box 2712, Beijing 100080, PROC), Fenglei Jiao (Beijng Inst. of Light Industry, Beijing 100037, PROC), and Hui Ding (Beijing Municipal Inst. of Labour Protection, Beijing 100054, PROC)

The flat panel resonance can affect the absorption characteristics of a microperforated-panel absorber. This effect was not considered in the existing theory. In this paper, the method of measuring the no-hole flat panel resonance is applied to analyze this effect. The sample is a stainless steel microperforated panel, with the diameter of perforations $d=0.2$ mm, the separation of holes $b=2.5$ mm, the thickness of panel $t=0.2$ mm, and the thickness of cavity behind the microperforated-panel $D=95$ mm. Theoretically, its resonance frequency is 508 Hz, and the maximum absorption coefficient is 0.9. However, the experimental result shows that there are two absorption peaks at 280 and 680 Hz, respectively. Two resonance peaks also appear as a sample of a no-hole flat panel is measured. When the thickness of cavity is changed, the result is still similar. Moreover, when sound intensity is enhanced, two resonance peaks also exist, and move towards low frequency. Therefore, it is possible that the absorption frequency band can be widened, or the absorption coefficient within absorption frequency band can be enlarged, by using the resonance frequency of different materials. [Work supported by the National Natural Science Foundation of China.]

9:00

4aNS2. Prediction system for rotor–stator interaction noise generation, propagation, and optimization of acoustic liners. Yadong Lu (Inst. of Acoust., Acad. of Sci., P.O. Box 2712, Beijing, 100080, PROC), Zongan Hu, and Jiya Cui (Beijing Univ. of Aeronautics and Astronautics, Beijing, 100083, PROC)

Based upon the acoustic mode matches between the source modes and propagating modes, this paper applies the flexible tolerance optimization method to optimize the acoustic parameters (impedance), geometric structure parameters, such as open area ratio, cavity depth, and hole diameter, and operating condition parameters, such as blade passing frequency. The optimum values of the design variables are determined when the in-duct sound suppression approaches the maximum. It can be derived from the optimum results that the emphasis of the engineering optimization design of the perforated plate honeycomb structure should be placed on the optimum choice of the open area ratio and cavity depth. Some other referential criteria for the engineering design of the multi-linings are also provided. Thus, the theoretical prediction system for rotor–stator interaction noise generation and in-duct propagation and optimization of acoustic liners has been developed in this paper. By means of this prediction system, the acoustically multi-sectioned treatments can be theoretically designed for the suppression of rotor–stator interaction noise with discrete frequency, in advance of the beginning of the practical engineering design of acoustic liners.

9:15

4aNS3. The Wiener–Hopf technique and scattering of acoustic waves in ducts. Fredrik Albertson (MWL, Dept. of Vehicle Eng., Royal Inst. of Technol., 100 44 Stockholm, Sweden, fal@fkt.kth.se)

The reactive silencer is the standard product for attenuating noise in ducts with flue gases. A model problem is solved for such a silencer by calculating the scattering of acoustic waves at sharp edges in a two-dimensional waveguide. In this waveguide the walls are considered acoustically hard and to ensure the existence of a unique solution an edge condition is applied. An exact analytical solution is found using the Wiener–Hopf technique, which relies heavily on the Fourier transform with assumptions on regularity of the solution. Control methods are employed to check the theory and the accuracy of the numerical results. One of the control methods is derived from the energy conservation principle. Symmetries, so-called reciprocity relations, are found for the reflection and transmission matrices and are first used to check the analytic calculation and later to reduce the numerical work. A quasistationary model is also used for control purposes. It is valid for low frequencies and checks the phase of the transmission and reflection coefficients, while the energy method only covers the corresponding absolute values. Physical results as the behavior of the reflection and transmission coefficients are also presented. These coefficients are analyzed as functions of both frequency and geometry.

9:30

4aNS4. Acoustic response of a thin poroelastic plate. Kirill V. Horoshenkov (Dept. of Civil and Environ. Eng., Univ. of Bradford, Bradford BD7 1DP, England) and Kimihiro Sakagami (Kobe Univ., Rokko, Nada, Kobe, 657 Japan)

The expression for the viscosity correction function in the Attenborough model for the acoustic properties of rigid frame porous media is modified to account for the viscothermal effects in the oscillatory flow in the vibrating frame. The Helmholtz integral equation formulation is used to produce the solution for the sound field reflected from an infinite, thin, porous, elastic plate. The effect of an air cavity behind the plate is considered. A parametric study is performed to predict the effect of variations in the microscopic parameters of the poroelastic plate.

9:45–10:00 Break

10:00

4aNS5. A nondestructive method for measuring the airflow resistance of locally reacting jet engine nacelle components. T. R. T. Nightingale (Inst. for Res. in Construction, Natl. Res. Council Canada, Ottawa, ON K1A 0R6, Canada) and Brandon Tinianow (Johns Manville Tech. Ctr., Littleton, CO 80127)

As a noise control measure to reduce inlet fan noise of jet engines, the nacelles are often equipped with a locally reacting absorber formed by a thin resistive wire mesh placed over a honeycomb structure. Historically there has been a problem to measure, in nondestructive fashion, the airflow resistance of repaired or refurbished nacelle parts for the purpose of demonstrating OEM compliance. ASTM-C522 is the traditional airflow resistance test method and is destructive, requiring a sample be cut from the specimen and fit into an apparatus where the flow velocity due to a constant pressure drop can be measured. This paper presents the theory and practical considerations of a nondestructive test method to measure the airflow resistance using an ASTM-E1050 impedance tube placed against the wire mesh of the complete specimen. Transfer matrix theory is used to describe the combined impedance of the wire mesh and air cavity. It is shown that at the air cavity resonance frequency the measured system impedance will be that of the wire mesh with the real part being the airflow resistance. Airflow resistance estimates obtained using the presented method and the ASTM-C522 test method are shown to be in excellent agreement.

10:15

4aNS6. A measurement of sound power lever of noise source—Vibration velocity method. Weicheng Yang (Res. Inst. of Household Elec. Appliance of China, No. 6 Yuetan Beixiaojie, Beijing 100037, PROC) and Wuzhi Qiao (Beijing Inst. of Light Industry, Beijing 100037, PROC)

Sometimes it is necessary to measure on-line the noise of some products, such as the household electrical appliances and their parts (motors, compressors, and contacts). Nowadays, several measurements (sound intensity measurement and vibration flow measurement) have been applied, but are not good enough. In this paper, another on-line measurement of sound power lever of noise source vibration, velocity measurement, is proposed. This measurement has two advantages, viz., without the influence of the surrounding area and with less measuring time. After the theoretical analysis and the experiment to the small-type compressor of the refrigerator in a semi-anechoic chamber, the sound-radiating efficiency curve is attained. Based on the vibration velocity measurement, the sound power lever of the compressor is predicated through this curve. The error of the predicated result is less than 2 dB, compared with that of the standard ISO 3745.

10:30

4aNS7. The limit of absorbing characteristics of folded resonators silencers and ways of its realization. Roudolf Starobinski (Lab. Acoustique, Univ. Le Maine, B.P. 535, 72017 Le Mans, France) and Jean Kergomard (Univ. Le Maine, 72017 Le Mans, France)

The folded resonance silencer is a version of 1/4-wave resonator having a very compact shape. They are usually used for suppressing of narrow-band and discrete frequency noise. The sets of resonators tuned to adjacent frequencies can be successfully used for a wide frequency band as well. The general theorems governing the behavior of these resonators are represented in this report. The limits of the muffling characteristics of the resonators and the ways of its realization are investigated on their base. It is shown that the potential muffling ability of a folded linear resonator is 1.2 dB less than for a Helmholtz resonator. However, taking into consideration the possibility of increasing the neck's cross-sectional area, they are more effective at high sound levels because of the higher quality factors. It is shown also that with the help of coupling folded resonators the resonance effects can be created in the frequency band wider than two octaves. These peculiarities make the use of the folded resonators very interesting in noise control systems.

Session 4aPAa

Physical Acoustics and Bioresponse to Vibration/Biomedical Ultrasound: Cavitation Dynamics: In Memoriam Hugh Flynn I

Charles C. Church, Cochair
Acusphere, Inc., 38 Sidney Street, Cambridge, Massachusetts 02139

Ronald A. Roy, Cochair
Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Chair's Introduction—9:15

Invited Papers

9:20

4aPAa1. Hugh Guthrie Flynn, 1912–1997. David T. Blackstock (Appl. Res. Labs. and Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78712-1063, dtb@mail.utexas.edu)

Hugh Flynn spent most of his first three decades in Columbus, Ohio. During the 1930s he was a newspaper crime reporter and also studied physics at Ohio State University. After graduating from Ohio State in 1939, he took a job with the Navy's Bureau of Ordnance and was a Naval officer during WWII. While working for the NAS Committee on Undersea Warfare after the war, he met F. V. Hunt, who encouraged him to study acoustics at Harvard. There Hugh became interested in cavitation, the beginning of a lifelong love affair with bubbles. After his doctoral degree in 1956, he stayed on as one of Hunt's postdocs. The year 1960 was a propitious one. He married Prudence Turgeon and joined the faculty of the new Electrical Engineering Department at the University of Rochester. He spent the rest of his life in Rochester, where he solidified his international reputation as a respected authority on cavitation and fathered three children, Kitty, Molly, and Nathaniel. Despite retiring as Professor Emeritus in 1978, he stayed active in research for the rest of his life. His projects during retirement included bubble phenomena in medical ultrasonics and nuclear fusion from cavitation.

9:40

4aPAa2. Sonically induced growth of gas bubbles by diffusion. Anthony I. Eller (Sci. Applications Intl. Corp., 1710 Goodridge Dr., McLean, VA 22102, eller@osg.saic.com)

Small gas bubbles that might normally dissolve and disappear may instead grow if set into pulsations of sufficient amplitude by an applied sound field. The mechanism involved is a fundamental asymmetry related to the spherical geometry of the conditions which favors diffusion of gas into the bubble during its expanded state over diffusion out of the bubble during a contracted state. Values of the threshold acoustic pressure for growth were computed by a theory that accounts for spherical bubble dynamics as well as convection terms in the diffusion equation and shows general agreement with measured counterparts. Rates of bubble growth for above-threshold conditions can be unexpectedly high if nonspherically symmetric acoustic streaming occurs. A closely related phenomenon—the onset of nonspherical bubble vibration—occurs within the same general range of acoustic pressures observed for bubble growth and will also be described.

10:00

4aPAa3. Reminiscences on bubble dynamics and cavitation. M. Strasberg (Code 702, David Taylor Model Basin, NSWC, 9500 Macarthur Blvd., West Bethesda, MD 20817-5700)

A review of various topics in cavitation and bubble dynamics discussed with Hugh Flynn, most of them still containing unsolved details. Among these are rectified diffusion of gas into pulsating bubbles and the effect of surface films; thermal rectified diffusion and the associated temperature rise; parametric shape oscillations of bubbles and the failure to observe resonances; the threshold conditions for cavitation, the cause for the persistence of the required nuclei, and the effect of a temporary static pressure increase on the nuclei; the dynamics of an individual cavity growing and collapsing in a compressible liquid and the associated internal pressure and radiated sound; and the cavitation noise radiated by cavitating propellers.

10:20

4aPAa4. Physics of acoustic cavitation in liquids: H. G. Flynn's review 35 years later. A. Prosperetti (Dept. of Mech. Eng., The Johns Hopkins Univ., Baltimore, MD 21218)

In 1964 Hugh G. Flynn published in Vol. I, Part B of the series *Physical Acoustics*, edited by W. P. Mason (Academic Press), a long review article by the title ''Physics of Acoustic Cavitation in Liquids.'' This was the first comprehensive review of the topic and has had a enormous influence on the field, serving as an introduction to cavitation and bubble dynamics for at least two generations of researchers (including the present author). A summary of that work will be presented in the light of subsequent developments and the present understanding of the cavitation phenomena. [Work supported by NASA.]

10:40

4aPAa5. Cavitation terminology. Wesley L. Nyborg (Phys. Dept., Univ. of Vermont, Burlington, VT 05405)

Among the important contributions Flynn made to the subject of acoustic cavitation was his introduction of concepts and language which have greatly helped our thinking and our communications. In 1964, he defined two acceleration functions and used them for classifying the behavior of gas-filled cavities in a sound field. The more violent kind of behavior, involving ''collapse,'' occurs when one of the functions, the inertial function, dominates at a critical time during cavity contraction. This type of cavitation was initially called ''transient,'' partly because the collapse events are of short duration, and partly because the cavities themselves are often short-lived, disappearing through fragmentation and dissolution after collapse events. The other class of cavitation is called ''stable,'' and represented oscillatory motion which might be complicated but which persists indefinitely. This classification has become very helpful in discussing cavitation phenomena and is widely used in the literature. In recent years, however, the term ''inertial'' has come into use, replacing ''transient,'' for cavitation in which Flynn's inertial function dominates. Examples will be discussed in which Flynn's classification is applied to experimental situations.

11:00–11:15 Break

Contributed Papers

11:15

4aPAa6. Influence of nonradial experimental conditions on the acoustic scattering behavior of a single cavitation bubble. Torsten Niederdraenk (Institut für Technische Akustik, RWTH Aachen, 52056 Aachen, Germany, tni@akustik.rwth-aachen.de)

The dynamic behavior of a single cavitation bubble in a sound field can be evaluated using the acoustic scattering technique. A high-frequency image sound is scattered by the bubble, which pulsates in the primary driving sound field. The scattered sound contains information about the bubble oscillation. The shape of the bubble radius can be observed by evaluating the amplitudes of the scattered sound or by determination of frequency shifts caused by the motion of the bubble wall. If the wavelength of the high-frequency image sound and the bubble size are of the same order, or if the boundary and excitation conditions in the experimental setup do not fulfill the demand of a spatially independent primary driving field, the bubble is exposed to nonradial driving conditions. Modes of bubble oscillation may develop that can be decomposed into spherical harmonic components of different order. These modes affect the acoustic scattering behavior, because sound emissions of the nonradial contributions of bubble oscillation are characterized by different directivities. Moreover, the amplitudes of sound emitted by higher-order spherical harmonics depend more strongly on the receivers' distance to the bubble.

11:30

4aPAa7. Susceptibility of silicon wafers to acoustic microcavitation damage. Sameer I. Madanshetty and Jogesh B. Chandran (Mech. and Nuclear Eng., Kansas State Univ., 339 Durland Hall, Manhattan, KS 66506-5106)

Acoustic methods are being increasingly used in semiconductor processing. Given the increase in circuit densities being implemented on semiconductor chips the requirements for particulate contamination tolerance are becoming more stringent. The semiconductor industries roadmapped specifications for 1998, 0.25 μm feature size require control of all particulate contaminants larger than 0.08 μm. Fine cleaning of post CMP wafers is currently being done by using the megasonics process. The cleaning effectiveness depends on the strength of the acoustic fields used. In using high-intensity, high-frequency acoustic fields it is crucial to ensure that the surface of the silicon wafer does not suffer any damage due to cavitation. This paper assesses the damage potential of semiconductor wafers to acoustic microcavitation. Acoustic microcavitation is brought about by low megahertz acoustic fields giving rise to micron size bubbles that live a few microseconds. In exposing a surface to continuous waves one could obtain cavitation effects in an average, overall sense; the details of nucleation, evolution of inertial events, however, get glossed over. Both, pulsed and cw insonification were studied. Experimental measurements of the thresholds for cavitation damage of semiconductor wafers will be presented.

11:45

4aPAa8. Study of the correlations between the cavitation noise power and bubble volume rate in an acoustic cavitation bubble field. Stéphane Labouret, Jacques Frohly, Nathalie Poulain, Roger Torguet, and Francis Haine (IEMN-DOAE, Université de Valenciennes, B.P. 311, Le Mont Houy, 59 300 Valenciennes Cedex, France, labouret@univ-valenciennes.fr)

The effects of ultrasound cavitation depend on both the insonification conditions and the liquid characteristics, by means of their influence on the bubble cloud. For this reason, the study of the cavitation bubble field is of interest. The implemented hyperfrequency method allows the monitoring of the bubble volume rate, called void rate a, during the insonification, versus acoustic pressure or dissolved air concentration in water. On the other hand, the oscillating behavior of bubbles found expression in a characteristic acoustic spectrum. The cavitation noise power, defined as the integral of the cavitation acoustic spectrum, permits the estimation of the importance of the dynamic activity of the bubble field. The aim of this paper is to set the relations between the void rate evolution and the cavitation noise power. The evolutions of the cavitation noise power versus the void rate will be examined in two situations, whether the void rate evolution is due to the acoustic pressure variation or to a change of the dissolved air content in water. A link between the cavitation noise power and the increased speed of the void rate will be set.

Session 4aPAb

Physical Acoustics: Radiation and Diffraction

Andrew A. Piacsek, Chair
Department of Physics, Central Washington University, MS 7422, Ellensburg, Washington 98926

Chair's Introduction—9:15

Contributed Papers

9:20

4aPAb1. Acoustical helicoidal waves and Laguerre–Gaussian beams: Applications to scattering and to angular momentum transport. Brian T. Hefner and Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814)

Traveling waves having helicoidal wavefronts are cylindrically symmetric except for an azimuthal angular dependence of $\exp(im\phi - i\omega t)$, where m is an integer. The amplitude vanishes at the screw-phase dislocation [J. F. Nye and M. V. Berry, Proc. R. Soc. London Ser. A **336**, 165–190 (1974)] on the z propagation axis. Examples of paraxial helicoidal waves are the Laguerre–Gaussian beam solutions of the parabolic wave equation. Some potentially useful properties of acoustical helicoidal waves are analyzed. Just as for linearly polarized optical Laguerre–Gaussian beams [L. Allen *et al.*, Phys. Rev. A **45**, 8185–8189 (1992)], the ratio of the axial angular momentum flux to the acoustical beam power is found to be m/ω so that axial radiation torques are generated when the acoustical energy is absorbed. Helicoidal waves are predicted to backscatter in a way which reveals the axisymmetry of the scatterer when combined with a helicoidal mode-selective detector. Some strategies for acoustical helicoidal wave generation and detection will also be described. [Work supported by the Office of Naval Research.]

9:35

4aPAb2. Vibrations of a film with arrays of point masses. Takato Handa, Andy Piacsek, and Roger Yu (Dept. of Phys., Central Washington Univ., Ellensburg, WA 98926)

As a result of the Bloch theory, the eigenstates of an electron in a strictly periodic potential are extended and the eigenenergy levels form allowed energy bands separated by forbidden gaps. Very recently there has been a great deal of excitement in the physics community concerning the extension of the idea to electromagnetic and acoustic waves. The three-dimensional (3-D) photonic band material with a periodically varying dielectric constant has been found. The frequency band structure of a rectangle thin film with arrays of point masses on it interests us. It is assumed that the film is supported by a frame so that at the boundaries, the vibration vanishes. By taking advantage of this zero-boundary condition, a double sine wave expansion was used to transform the 2-D wave equation into a matrix equation. Up to 12 000 eigenvalues and eigenfunctions can be obtained accurately. A number of point mass configurations (or arrangements) were studied and their frequency structure provides rich insight into the physics of the system. In this report, the calculated frequency band structure of the system and the wave functions will be discussed.

9:50

4aPAb3. Absorbing boundary conditions for acoustic waves and Huygens principle. Gérard Mangiante (Lab. Mécanique d'Acoust., Ctr. Natl. Recherche Sci., 31 Chemin Joseph-Aiguier, 13402 Marseille Cedex 20, France, mangiante@lma.cnrs-mrs.fr) and Sylvestre Charles (Schlumberger Geco-Prakla, Inversion Methods, Houston, TX 77077)

When numerically simulating acoustic wave propagation, artificial boundaries must be introduced into the model to limit the area of computation. To eliminate the spurious reflections produced by these artificial boundaries, absorbing boundary conditions are needed. Such boundary conditions are usually approximated according to two main strategies: paraxial wave equation and the use of viscous media. This paper describes a new approach of the boundary conditions problem using Jessel's formulation of Huygens principle [Jessel (1991), Mangiante (1994), Charles (1996)]. A suitable distribution of Huygens secondary sources was arranged in a transition zone surrounding the model, and adjusted to produce destructive interference to cancel the wave field outside this transition zone. Perfect absorption can be theoretically obtained and can be numerically illustrated. However, due to practical difficulties, only an approximation of this solution has been derived for a finite difference scheme. This approximation consists in splitting the history of the wave field within user defined time windows. The main difficulties are related either to the space and time behavior of the Green functions, or to the solution of the wave equation problem via a finite difference scheme.

10:05

4aPAb4. Diffraction by a hard half-plane: Useful and simple frequency forms to an exact time-domain solution. Djamel Ouis (Eng. Acoust., P.O. Box 118, S-22100, Lund, Sweden, djamel.ouis@kstr.lth.se)

The general case of diffraction of a spherical wave by a hard half-plane is considered. The calculations are based on Biot and Tolstoy's exact and explicit form of the solution to the diffraction problem by a hard wedge. This solution is given in the time domain but its Fourier transform is not available. It is the aim of this paper to present some useful and simple approximations for the early time diffracted field which when transformed into the frequency domain is shown to be satisfactory in comparison to numerical integration. This could give some valuable contributions to calculations on scattering by noise barriers and the like. The results of some experiments are also presented.

10:20–10:35 Break

10:35

4aPAb5. A discussion on nonradiating sources. Ricardo E. Musafir (Acoust. & Vib. Lab/Dept. of Mech. Eng. and Hydraulics Dept., Univ. Federal do Rio de Janeiro-CP 68503, 21945-970, RJ, Brazil, rem@serv.com.ufrj.br)

The fact that a given source sound field can be generated by more than one source distribution is quite well known but is frequently given an overexaggerated importance. This paper discusses the structure of sources of silence, the name given by Doak [Proc. I.O.A., 10-2, 693–700 (1988)] to the reactive part of a source distribution, i.e., to the part with no far field and which is responsible for the ambiguity. It is shown that, since all ambiguities can always be traced back to the impossibility of distinction between the sound fields of a point monopole and of an isotropic point quadrupole, a simple criteria can be established for identifying a general source distribution as active or reactive. The particular case of the velocity field due to a point dipole (i.e., to a point force) in an homogeneous medium at rest is worked out in detail, since it contains a source of silence (associated with the vorticity at the source point) which is frequently mis-

handled. The concepts presented are applied to aeroacoustic wave equations, to discuss optimization of the choice of dependent variable as well as some case of identification of nonradiating sources found in the literature.

10:50

4aPAb6. To lend an ear to the dimensionality of space. Jean Hardy (Lab. acoustique, Univ. Le Mans BP535 72017 Le Mans CNRS UMR 6613, France)

Since the laws of physics are essentially based on our feeling of the real word, it seems natural to consider the problem of the space dimensionality. The usual conception of the space follows from the vision and leads to the conviction that dimensionality of the real word is 3; a similar approach but using the ears is done in this work. An acoustical source emits an impulse in a homogeneous and isotropic N-dimensional space. The Green's function obtained shows that odd and even dimensionalities have different behaviors. For odd dimensionalities the Green's function is a combination of a step, a delta function and its derivatives; the step function is responsible for reverberation, the other gives the structure of the impulse front which travels at the wave velocity. The even dimensionality uses the half derivatives for expressing propagation, the wave propagation is subject to reverberation which decreases with the time, the wave velocity is not well defined, and the signal transmission and consequently the communication are difficult if not impossible. An explicit demonstration with 1, 2, 3, 4, and 5 dimensions can be done using musical and speech samples.

11:05

4aPAb7. Resolution of the convected Helmholtz's equation by integral equations. M. Beldi (France Acoust., Res. and Development Dept., 11, rue du 8 Mai 1945, 60350 Berneuil sur Aisne, France) and F. S. Monastir (Univ. of Center, 5000 Monastir, Tunisia)

Using the free-field convected Green's function, the integral formulation for the acoustic velocity potential and its normal derivatives associated with Helmholtz's equation in a uniform flow is established. An original development for the calculation of the Cauchy principal value in the singular integrals is presented. As far as the finite part in the hypersingular integral is concerned, a new transformation of the convected double-layer potential normal derivative is proposed. The result of this transformation avoids the explicit calculation of the finite part. Moreover, it generalizes the famous works [M. P. Stallybrass, J. Math. Mech. **16**, 247–1286 (1967)] [A. W. Maue, Zeit. für Phys. **126**, 601 (1949)] on the resolution of the Helmholtz's equation, by integral equations, in a fluid with null velocity. This new variational method by integral equations can easily be coupled with the variational formulation for studying the convection and refraction effects of the acoustic waves in a nonuniform flow [M. Beldi, Third International Conference on Theoretical and Computational Acoustics, Newark (14–18 Aug. 1997)]. The numerical results shows the importance of these effects.

11:20

4aPAb8. Enhanced ray theory. Nikolai E. Maltsev (Instrumar Ltd., P.O. Box 13246, Stn 'A' 39 Pippy Pl., St. John's, NF A1B 4A5, Canada)

The numerical implementation of standard ray theory faces two intrinsic difficulties, the infinite value of the field at caustics and the nonlinear boundary value problem for computation of eigenrays—rays which connect source and receiver. In an earlier work [N. E. Maltsev, J. Math. Phys. **35**, 1387–1389 (1994)], a nontraditional approach was developed to treat the caustics problem. This report describes a procedure of embedding a nonlinear boundary value problem for eigenrays into the space of a higher number of dimensions, where it becomes an initial value problem for a new system of equations, differential equations for eigenrays. These two enhancements of the ray theory produce a regular method of computation of the field anywhere in the media, including shadows, for harmonic and pulse sources with all physical attributes of rays, such as arrival times and angles.

11:35

4aPAb9. Experimental study on the vortex-shedding sound from a yawed circular cylinder. Jong-Soo Choi and Hoon-Bin Hong (Dept. of Aerosp. Eng., Chungnam Nat'l Univ., 220 Kung-Dong, Taejon, Korea)

For a cylinder in a uniform flow stream, sound is generated by the fluctuating pressure on the surface of a cylinder due to the vortex shedding behind the cylinder. It is known that the major parameters in predicting the acoustic pressure at the far field are the flow velocity, the correlation length, and the fluctuating lift coefficient. In this experimental study, the correlation length of the wake is measured with hot-wire sensors and compared to the one obtained with pressure sensors on the cylinder surface. A new measurement technique for the unsteady lift estimation by using a pressure sensor rotating circumferentially is used to show reasonably good results. The surface pressure and the radiated sound are also measured simultaneously for different yaw angles and showed that the reduced normal velocity component to the cylinder axis reduces the unsteady lift fluctuation which results in lowered sound-pressure level. However, experimental result shows that ''the cosine law'' which uses the normal velocity component as a characteristic velocity for noise generation from a yawed cylinder needs to be carefully reviewed.

Session 4aPAc

Physical Acoustics: General Topics (Poster Session)

James P. Chambers, Chair
National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677

Contributed Papers

All posters will be on display from 9:15 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 9:15 a.m. to 10:45 a.m. and contributors of even-numbered papers will be at their posters from 10:45 a.m. to 12:00 noon. To allow for extended viewing time, posters will remain on display until 12:00 noon on Friday, 26 June.

4aPAc1. Theory on mass transport generated by ultrasonic Lamb waves. Gonghuan Du, Zhemin Zhu, Xiaoliang Zhao (Inst. of Acoust. and State Key Lab of Modern Acoust., Nanjing Univ., Nanjing 210093, PROC), and Junru Wu (Univ. of Vermont, Burlington, VT 05405)

It was reported that acoustic streaming generated by a Lamb-wave sensor (LWS) can be used to pump fluids and transport solids in a small-scale system [Appl. Phys. Lett. **59** (1991)]. The speed, observed by attracting small polystyrene spheres suspended in water, is proportional to the square of the wave amplitude and is about 100 μm/s for 6.5-nm surface displacement of the plate. In this paper, a fairly practical model, i.e., a thin isotropic plate in contact with a viscous liquid layer on its side, is presented. Based on the rigorous propagation theory of Lamb waves and Nyborg's [Phys. Acoust. B **2**, 265–295] and Bradley's [J. Acoust. Soc. Am. **100**, 1399–1408 (1996)] theories of acoustic streaming, the first-order ultrasonic field in the viscous liquid layer and the second-order mass-transport velocity distribution of the liquid are calculated. Comparison between our theoretical values of the Ao-mode Lamb wave and Moroney *et al.*'s experimental result is given. It shows good agreement. [Work supported by NSF of China.]

4aPAc2. Photoacoustic study of the electron–hole transport in a piezoelectric semiconductor. Weimin Gao, Viltali Gusev, Christ Glorieux, Kris Van Rostyne, Ward Van Looy, Jan Thoen (Lab. Akoestiek en Thermische Fysica, Dept. Natuurkunde, Katholieke Universiteit Leuven, Celestijnenlaan 200D, B-3001 Leuven, Belgium), and Gustaaf Borghs (IMEC, B-3001 Leuven, Belgium)

In a piezoelectric semiconductor acoustic waves can be excited via the inverse piezoelectric effect. This excitation process is the result of a screening of the applied electric field by the expanding space distribution of electrons and holes which are generated by pulsed laser radiation. The time and space characteristics of the electron–hole transport in this process are contained in the corresponding acoustic signals. The piezoelectric direct-gap semiconductors cadmium sulfur selenide and gallium arsenide are investigated experimentally at room temperature. A pulsed nanosecond UV excimer laser is used to excite electrons and holes on these samples. The generated longitudinal acoustic waves are detected by a laser interferometer. The obtained acoustic pulses of cadmium sulfur selenide show that both the expansion of the screened region in space and the electron–hole plasma expansion are supersonic at the time scale of laser action. The observations in this optoacoustic experiment indicate that the fast carrier transport is governed by photon recycling, i.e., the reabsorption of photons released by radiative recombination of electron–hole pairs. Experiments for gallium arsenide are in progress and will also be reported.

4aPAc3. Ultrasonic investigation of glass transition dynamics of polyurethane systems. Pierre-Yves Baillif, J. Mohamed Tabellout, and Jacques Emery (Gr. ultrason, UMR 6515 Chimie et Phys., des Materiaux Polymeres, Univ. du Maine, Av. O. Messiaen, 72035 Le Mans Cedex, France, ultrason@aviion.univ-lemans.fr)

Ultrasonic technique is an interesting tool to probe high-frequency dynamic properties of materials by determination of the mechanical modulus. Associated with other methods (rheology, DLS, etc.), it allows the investigation of molecular dynamics associated with sol–gel and glass transitions in polymers on a wide frequency range. Polypropylene glycol triol crosslinked with diisocyanate is a good model system for this purpose in polymer networks. By varying the amount of diisocyanate, stable systems are formed which correspond to different stages of polyurethane gel formation. Three triols with different molar masses (260, 700, and 6000 g/mole) are used as precursors in order to apprehend the influence of the macromolecular chain size on the systems dynamics. The temperature dependence of the ultrasonic absorption and velocity shows a relaxation process, the so-called alpha-relaxation associated with the glass transition. The influence of the stoichiometric ratio $r=[\mathrm{NCO}]/[\mathrm{OH}]$ on this transition is much less important in a large molecular mass precursor system, indicating that entanglement effects become predominant. The ultrasonic relaxation times associated with those obtained by other techniques are represented in an Arrhenius diagram and follow the same William–Landel–Ferry (WLF) law, showing the different systems dynamics' similarity. The WLF parameters seem, in that case, to have a character of universality.

4aPAc4. Longitudinal structural relaxation of glass-forming epoxy oligomer. Mami Matsukawa (Dept. of Electron., Doshisha Univ., Kyotanabe, Kyoto, 610-03 Japan) and Norikazu Ohtori (Niigata Univ., Niigata, 950-21 Japan)

The glass transition process of a common epoxy oligomer (diglycidyl ether of bisphenol A) has been characterized by Brillouin scattering (GHz range) and ultrasonic pulse spectroscopy (MHz range). As the temperature decreases, longitudinal wave velocity rapidly increased. It showed a large frequency dispersion at temperatures higher than the thermodynamical glass transition temperature (T_g). At T_g, velocity changes showed a clear kink. Wave attenuation also showed a peak of the main (alfa) relaxation at temperatures higher than T_g. The peak temperature shifted to lower temperatures in the MHz range, following the VFT type relaxation behavior, which is very common in the main relaxation. The velocities, however, still showed a clear frequency dispersion at temperatures lower than T_g. Because the main relaxation usually results from a large-scale conformational rearrangement of the polymer chain backbone, the effect of this relaxation seems to be neglected at these temperatures. Taking into ac-

count the frequency changes of attenuation peak widths, the effect of other subrelaxations at lower temperatures is discussed. These results are not in contradiction with the acoust-optic dispersion measurements of this oligomer [Kruger *et al.*, Phase Transition (to be published)].

4aPAc5. Hot-spot model of single-bubble sonoluminescence. Kyuichi Yasui (Dept. of Phys., Waseda Univ., 3-4-1 Ohkubo, Shinjuku, Tokyo, Japan)

The hot-spot model of single-bubble sonoluminescence (SBSL) is studied. The temperature is assumed to be spatially uniform inside the bubble except at the thermal boundary layer near the bubble wall even at the strong collapse based on the theoretical results by Kwak *et al.* [Phys. Rev. Lett. **77**, 4454 (1996)]. The effect of the kinetic energy of gases inside the bubble is taken into account, which heats up the whole bubble when gases stop their motions at the end of the strong collapse. A bubble in water containing air is assumed to consist mainly of argon based on the hypothesis by D. Lohse [D. Lohse *et al.*, Phys. Rev. Lett. **78**, 1359 (1997)]. Numerical calculations under an SBSL condition reveal that the kinetic energy of gases heats up the whole bubble considerably. It is also clarified that vapor molecules (H_2O) undergo chemical reactions in the heated interior of the bubble at the collapse and that chemical reactions decrease the temperature inside the bubble considerably. It is suggested that SBSL originates in quasithermal radiation from the whole bubble rather than a local point (the bubble center) heated by a converging spherical shock wave widely suggested in the previous theories of SBSL.

4aPAc6. Nonlinear reflection of acoustic waves at the dissipative isotropic solid–solid interfaces. W. H. Jiang, J. J. Chen, and Y. A. Shui (Inst. of Acoust. and State Key Lab. of Modern Acoust., Nanjing Univ., Nanjing 210093, PROC, whjiang@nju.edu.cn)

The nonlinear reflection at the solid–solid interfaces is investigated under the condition that a SV wave is incident from a nondissipative solid medium to a dissipative solid medium. The amplitude of the freely propagating reflected second harmonic SV wave is calculated. Comparing with previous calculation without consideration of dissipation, it is observed that the peaks of the reflected second harmonic SV wave in the vicinity of critical angles for refractive P and SV waves become smooth and the influence of nonlinear properties of the refractive medium on the amplitude of the reflected second harmonic wave is more notable. A experiment is carried out for glass–copper or rock interface. The SV wave of 8 MHz is incident from glass in which the attenuation of the acoustic wave is negligible upon glass–copper or rock interface. The reflected second harmonic SV wave is detected by a transducer with the working frequency of 16 MHz. The amplitudes of the second harmonic wave at the different incident angles are measured. The resultant angular dependence is in basic agreement with the calculation. [Work supported by NSF of China.]

4aPAc7. Estimation of forward-propagated ultrasonic fields in layered fluid media. K. L. Ha, M. J. Kim (Dept. of Phys., Pukyong Natl. Univ., 599-1 Daeyon-dong, Nam-gu, Pusan 608-737, Korea), W. Y. Zhang, S. G. Ye, and X. F. Gong (Inst. of Acoust., Nanjing Univ., Nanjing 210093, PROC)

The forward propagation of two-dimensional ultrasonic fields through layered fluid media was investigated theoretically by the angular spectrum method combined with the Rayleigh–Sommerfeld diffraction theory and experimentally by measurement of the fields using a precise scanning system with a needle-point hydrophone. Two types of transmitting transducers, one of which is a plane and the other of which is a concave type with the same frequency and the same aperture, were used. In theoretical analysis fields of incidence on a boundary of layer were calculated by the Rayleigh–Sommerfeld diffraction theory and those of transmission throughout the boundary were obtained by applying the angular spectrum method to the fields of incidence and considering the spatial-frequency-dependent transmission coefficients on each boundary. The optimum angular range and the spatial sampling interval could be obtained by supposing the virtual boundaries in the homogeneous medium of water and its appropriateness was confirmed experimentally. It was shown that the forward-propagated fields which could be estimated theoretically agreed reasonably with the measured data in two kinds of simple layered media with one or two boundaries.

4aPAc8. High temporal resolution of the acoustic transients associated with optical cavitation in water using a large area piezoelectric transducer. David C. Emmony and Robin D. Alcock (Dept. of Phys., Loughborough Univ., Loughborough, Leicestershire LE11 3TU, UK, d.c.emmony@lboro.ac.uk)

A PVDF piezoelectric transducer has been used to study the shock wave associated with a laser generated cavitation bubble in water. The measurements depend upon the accurate recording of the voltage developed in a large (3 mm) diameter but thin (9* m) piezoelectric film transducer placed close to a Nd YAG laser-generated cavity in water. The spherical shock transients excite the film over the duration of the passage of the wave through the transducer, and an invertion algorithm of the voltage trace gives the pressure profile along a radius. Pressure profiles due to initial breakdown and first collapse are in excellent agreement with those obtained using full field high-speed optical interferometry of the laser–liquid interaction.

4aPAc9. Particle interactions in coupled phase theory for sound propagation in concentrated emulsions. J. M. Evans and K. Attenborough (Faculty of Technol., The Open Univ., Milton Keynes MK7 6AA, UK, j.m.evans@open.ac.uk)

Theoretical studies of the acoustically induced interphase heat transfer in concentrated emulsions are reviewed. Interactions between the thermal fields around the particles influence the heat transfer and hence the complex wave number of propagating sound. Two modeling approaches are discussed: a hydrodynamic treatment and thermal wave scattering. These methods are used to calculate the interphase heat transfer at high concentrations. Modified heat transfer terms including the effect of particle interactions are obtained, which may be used in coupled phase theory. Predictions of the complex wave number using a revised coupled phase model [J. M. Evans and K. Attenborough, J. Acoust. Soc. Am. **102**, 278–282 (1997)] are compared with experimental data on ultrasound propagation in oil in water emulsions. The comparisons show the nondimensional frequency and volume fraction regions where interactions have a significant effect on the wave number.

4aPAc10. Particle trajectories in a drifting resonance field separation device. Bernhard Handl, Martin Gröschl, Felix Trampler, Ewald Benes (Institut für Allgemeine Physik, Vienna Univ. of Technol., Wiedner Hauptstr. 8/134, A-1040 Wien, Austria, benes@iap.tuwien.ac.at), Steven Woodside, and James Piret (Univ. of British Columbia, Vancouver, BC V6T 1Z3, Canada)

Drifting resonance field (DRF), a new method for the separation of particle/fluid suspensions is introduced. Unlike particle filters using ultrasound-induced coagulation, the DRF method only utilizes the primary acoustic radiation force to achieve the separation effect. This is done by cyclically switching the driving frequency of the DRF resonator between consecutive resonances, causing a directed motion of the particles and thus concentrating the particles in a certain region of the separation volume. Trajectories of polystyrene spheres (approximately 12 microns in diameter) suspended in distilled water have been measured in a prototype DRF resonator, varying the number of the utilized resonance frequencies and the acoustic power input. The measured particle trajectories are compared to computer simulations based on the same process parameters as used in the experimental setup, showing a high correspondence between experi-

mental and simulated data. Both experiments and simulations prove the high potential of the DRF method for separation technology. A promising field of application is the up-scaling of acoustic cell retention systems recently introduced in animal cell culture, potentially yielding higher flow-through rates. The presented results are a valuable basis for design and process parameter optimization of future DRF separation devices.

4aPAc11. Methods of processing laser Doppler anemometry signals to extract sound field information. John S. Cullen, David B. Hann, Clive A. Greated, and D. Murray Campbell (Dept. of Phys. and Astron., Univ. of Edinburgh, The Kings' Bldg., Mayfield Rd., Edinburgh EH9 3JZ, UK)

For several years laser Doppler anemometry (LDA) has been used to measure mean flow rates in gaseous and liquid flows. It is only recently that LDA has been developed as a technique for recording instantaneous particle velocities, enabling nonintrusive measurements to be made in sound fields. The paper describes a typical experimental setup and discusses various methods of digitally processing the LDA photodetector signals to extract the required velocity information. Particular emphasis is placed on the zero counting and Hilbert transformation techniques. The problem of signal dropout is also addressed. LDA measurements of acoustic particle velocity amplitudes are compared with acoustic velocity amplitudes deduced from probe microphone and impedance measurements. It is found that LDA measurements could be made not only in monofrequency sound fields but also in more complex sound fields containing a fundamental component and additional harmonics. Theoretical and practical limitations of the techniques are investigated and used to assess the viability of applying them to studies in the bores and mouthpieces of brass instruments.

4aPAc12. Approach to characterize the sound field of pulse-excited ultrasonic sensors using a Laser-Doppler vibrometer. Bernd Henning, Stefan Prange, Mark Schuart, and Karsten Dierks (Institut für Automation und Kommunikation e.V. Magdeburg, Steinfeldstrasse 3, D-39179 Magdeburg-Barleben, Germany)

A new efficient approach to make visible sound fields in front of acoustic sensors in liquids will be presented. The goal is to investigate the sound field characteristic of real manufactured acoustic transmitters particularly by discontinuous excitation like pulse, burst, and other. Another point is the analysis of the influence of materials and sensor design on the final acoustic behavior. In comparison to the Schlieren technique it is possible to use moderate excitation of transmitters (only a few volts) and more variable signal shapes too. The experimental setup consists of a Laser Doppler vibrometer and a high-resolution three-dimensional positioning system. The control of the positioning system as well as the data acquisition and processing will be fully automatic realized by a personal computer. A software tool archives the data and calculates the complex sound field characteristic in dependence on time or place. The result is an interesting animation representing the real sound field characteristic of acoustic sensors. The advantages and limits of the approach and developed experimental setup will be described. Some examples will show the sound fields in front of cylindrical focusing transmitters. The influence of different excitation signal shapes on the final sound field will be discussed.

4aPAc13. Simulation of ultrasound propagation in a thermally turbulent fluid using Gaussian beam summation and Fourier modes superposition techniques. Christian Lhuillier, David Fiorina, Jean-Luc Berton (Commissariat a l'Energie Atomique, Cadarache, DER/SSAE/LSMR, F-13108 Saint-Paul-lez-Durance cedex, France, ssaelsmr@macadam.cea.fr), and Daniel Juve (UMR CNRS 5509, Ecole Centrale de Lyon, F-69131 Ecully, France)

The CEA (French atomic agency) develops ultrasonic devices for the control and the inspection of Fast Breeder Reactors. The ultrasonic pulses propagate in a liquid which may be turbulent. To predict or interpret the effects of a homogeneous isotropic thermal turbulence, a numerical model has been implemented. Independent realizations of random thermal field are generated by superposition of random Fourier modes. For each realization the complex pressure field is calculated by the Gaussian beam summation method. Statistics over the set of realizations provide an estimation of the variance of time of flight and intensity fluctuations. Simulations have been performed with a point source. When the level of fluctuations is small, the model yields results which are in a good agreement with the analytical solutions given by Chernov and Rytov. When the level of fluctuations grows, it predicts correct evolutions. Multiple eigenrays are well predicted and calculations in time domain with a pulsed source show multiple transmitted signals or modifications in the signal shape.

4aPAc14. Lattice Boltzmann methods in acoustics. James M. Buick, D. Murray Campbell, and Clive A. Greated (Dept. of Phys. and Astron., Univ. of Edinburgh, Rm. 4201 J.C.M.B., The Kings' Bldg., Mayfield Rd., Edinburgh EH9 3JZ, UK)

The lattice Boltzmann model is a recently developed technique for the numerical simulation of fluid motion. It has been successfully applied to a variety of incompressible fluid phenomena and has been shown to be an efficient simulation tool which is ideally suited for implementation on massively parallel computers. Here the application of the lattice Boltzmann model is considered to simulate acoustical processes in which the fluid can, to a first approximation, be considered to be incompressible. The validity and limitations of this approach are considered and the results of simulations performed on the Cray T3D at Edinburgh University are presented. These include sound propagation in tubes, modeling of air flow in the mouthpiece of brass instruments, and the simulation of acoustic streaming in tubes and around cylinders. In each case the fluid velocity, density, and pressure are found at each grid point and each time-step of the simulation, allowing flow patterns and density changes to be observed. The results of the simulations are compared to the predictions of theory and/or experimental observations and suggest that the lattice Boltzmann method can be used as a numerical tool to give insight into complex acoustical situations.

4aPAc15. Acoustical characteristics of the noise radiated from supersonic multijets. Yoshikuni Umeda and Ryuji Ishii (Div. of Aeronautical and Astronautics, Dept. of Eng. Sci., Kyoto Univ., Yoshida Hon-Machi, Sakyo-ku, Kyoto, 606 Japan)

The acoustical characteristics of the noise radiated from combined jets exhausted from one main nozzle and 1, 2, 4, or 8 subnozzles which were placed asymmetrically or symmetrically to the main nozzle were investigated experimentally. The diameter d of the main nozzle was 5 mm and that of subnozzle a/d was 0.2 or 0.6, where a is the diameter of the subnozzle. The center-to-center spacing b of the main and subnozzles was fixed to 5 mm (b/d=1.0). In this experiment, the pressure ratio R of the jets was varied from 2.00 to 6.33. From the frequency characteristics of the screech tone, it is found that in almost all cases, two oscillation modes appear, but when eight subjets with a/d=0.6 were used, the screech tone disappeared completely. Although the total cross-sectional area of the multiple nozzles was larger than that of a single jet, it is found that the sound-pressure level radiated from multijets becomes the same or lower than that from the single jet. [Work supported by the Ministry of Education and Culture of Japan, Grant No. C2-09650187.]

4aPAc16. Acoustic parameters of high functional plastics. I. Measurement of acoustic velocities (wave propagation velocities). Masahide Gakumazawa and Minoru Akiyama (Dept. of Systems Eng., Shibaura Inst. of Technol., 307 Fukasaku, Ohmiya-shi, Saitama, 330 Japan)

Recently, high functional plastics (so-called engineering plastics) which improved the weak point of conventional plastics have been used in the industry. The acoustic parameters of these engineering plastics in the high frequency have been sparsely reported in the literature using a variety of experimental techniques. In this study the acoustic velocities (wave propagation velocities) of engineering plastics are measured using the scanning acoustic microscopy in the range from 50 to 200 MH, and the velocities were compared with the velocities reported in the literature in the lower frequencies.

THURSDAY MORNING, 25 JUNE 1998 FIFTH AVENUE ROOM (W), 8:00 TO 10:45 A.M.

Session 4aPP

Psychological and Physiological Acoustics: Plenary Preview—Issues Related to Speech Perception and Hearing Impairment

Jacek Smurzynski, Chair
Department of Otorhinolaryngology, University of Basel, Kantonsspital, Petersgraben 4, CH-4031 Basel, Switzerland

Contributed Papers

8:00

4aPP1. Effects of stimulus level and background noise on vowel representations in the auditory brain stem of cats. Bradford J. May (Dept. of Otolaryngol.—HNS, Johns Hopkins Univ., Baltimore, MD 21205) and Murray B. Sachs (Johns Hopkins Univ., Baltimore, MD 21205)

Effects of stimulus level and background noise on vowel representations in the ventral cochlear nucleus (VCN) were investigated by recording single-unit discharge rates in anesthetized cats. The quality of vowel encoding by populations of neurons can be evaluated by plotting vowel-driven rates as a function of each unit's best frequency (BF, the most sensitive frequency); the resulting rate profiles are assumed to show good representations of formant structure if there are high driven rates at BFs near formant frequencies and low rates at BFs near spectral troughs. Among the VCN unit types recorded in the present study, chopper units exhibited the most detectable formant-to-trough rate differences across vowel levels and in background noise. By contrast, primarylike units produced vowel representations that reflected the limited dynamic range of auditory-nerve fibers [May *et al.*, Aud. Neurosci. **3**, 135–162 (1996)]. Primarylike units with low spontaneous rates (SR$<$18 sp/s) showed better representations than high SR primarylike units, but failed to respond at low vowel levels. These results suggest that peripheral representations of speech sounds are sharpened by patterns of auditory nerve convergence in the cochlear nucleus. [Work supported by NIDCD Grant No. DC00109.]

8:15

4aPP2. The role of the auditory periphery in the categorization of stop consonants. Robert I. Damper (Dept. ECS, Univ. of Southampton, Southampton SO17 1BJ, UK, rid@ecs.soton,.ac.uk)

A proper understanding of speech perception requires an understanding of the staged restructuring of information that occurs as an acoustic stimulus is transformed by the auditory system into a phonetic percept. It is now firmly established that a physiologically faithful computer simulation of peripheral processing feeding into a trainable artificial neural network (ANN) is capable of reproducing the important aspects of the categorization of initial stop consonants into voiced and unvoiced classes. Correct behavior is found to be insensitive to the ANN architecture and a precise training scheme used, suggesting that such categorization is very basic to the repertoire of auditory processing strategies. Unlike a human or animal listener, a computational model can be easily manipulated and interrogated to discover the underlying basis of category formation. In particular, the physiological-based simulation of the auditory periphery can be dramatically simplified and the effect on categorization observed. It is found that proper modeling of frequency scaling and neural adaptation are essential to correct simulation of the well-known shift of phoneme boundary with place of articulation for stop consonants. The role of the auditory periphery in this case seems to be to emphasize the region of first formant information around the time of voicing onset.

8:30

4aPP3. Integrating monaural and binaural spectral information. Michael A. Akeroyd, Quentin Summerfield, and John R. Foster (MRC Inst. of Hearing Res., University Park, Nottingham NG7 2RD, UK, michael@ihr.mrc.ac.uk)

When speech is presented at a different lateral position to an interfering noise, analyses of both intensity and interaural differences can recover parts of the speech spectrum. It would thus be advantageous if listeners could integrate these two types of information. Their ability to do so using two-formant vowel-like stimuli was tested. A diotic white noise was modified by either increasing the intensity of two 1-ERB-wide bands by 6 dB (''monaural'' condition), by interaurally decorrelating those two bands (''binaural'' condition), or by increasing one band's intensity but interaurally decorrelating the other (''mixed'' condition). Four center frequencies were used—250, 650, 950, and 1850 Hz—and combined in pairs to create four vowels: 250–950, 250–1850, 650–950, and 650–1850. Both bands must be detected to distinguish these stimuli. Identification performance was significantly above chance in all conditions but best in the monaural condition. It was as good in the mixed condition as in the binaural condition. Performance in the mixed condition was better than predicted form performance when only a single band was presented, and thus it could not be due to listening to the bands individually. These results indicate that listeners can integrate monaural and binaural evidence of speech spectra.

8:45

4aPP4. Gap detection threshold in ears with and without spontaneous otoacoustic emissions (SOAEs). Jacek Smurzynski and Rudolf Probst (Dept. of Otorhinolaryngology, Univ. of Basel, Kantonsspital, CH-4031 Basel, Switzerland, smurzynski@ubaclu.unibas.ch)

Gap detection was measured for two groups of normally hearing young adults using broadband (0.1–12 kHz) noise stimuli presented monaurally through an insert earphone. Group I consisted of subjects who exhibited both strong SOAEs and click-evoked otoacoustic emissions (CEOAEs). Group II included individuals with no SOAEs in either ear and CEOAE levels that were $<$50th percentile of a laboratory normative da-

tabase. An adaptive 2IFC procedure was used to determine the hearing threshold of the noise stimulus for each ear tested. Next, an adaptive 2IFC gap detection task was performed for the stimuli presented at either 10-, 20-, 30-, or 50-dB sensation level (SL). At the stimulus level of 10 dB SL, the ears in group I exhibited greater intersubject variability and higher mean gap detection thresholds than those in group II. The results at higher SLs did not show any differences between the two groups and were consistent with the data in the literature. A short gap in the broadband stimulus presented near the hearing threshold is masked by SOAEs and strong CEOAEs that are evoked by the stimulus. [Work supported by Swiss National Science Foundation.]

9:00

4aPP5. Modulation masking in a speech recognition task for normal-hearing and hearing-impaired subjects. Rene van der Horst (Dept. of Clinical and Experimental Audiol., Univ. of Amsterdam, Meibergdreef 9, Amsterdam, 1105 AZ, The Netherlands, r.vanderhorst@amc.uva.nl)

The effect of masking in the modulation domain was determined in a consonant recognition task. It is known from psychoacoustical experiments [T. Houtgast, J. Acoust. Soc. Am. **85**, 1676–1680 (1989)] that modulation detection can be masked by another modulation. An experiment was designed that tries to verify whether detection of the modulations in a speech signal is essential for speech recognition. The goal of this experiment was to determine whether stochastic modulations can mask the modulations in a speech signal and disturb speech recognition as effectively as filtering out speech modulations [R. Drullman *et al.*, J. Acoust. Soc. Am. **95**, 1053–1064 (1994)]. The masker used was a stochastic narrow-band signal with minimal inherent fluctuations. The masker signal was, in every critical band, multiplied with the speech modulations. Several masking conditions were created with different bandwidths and center frequencies. The center frequencies were all lying in the region of modulation frequencies which are important for speech (4–20 Hz). Normal-hearing subjects and subjects with sensorineural hearing loss participated in the experiment. The confusion matrices were analyzed by means of multidimensional scaling techniques. The results of the two groups of subjects will be compared with the results obtained with a psychoacoustic model.

9:15–9:30 Break

9:30

4aPP6. A mathematical model of consonant perception by cochlear implant users with the SPEAK strategy. Mario A. Svirsky and Ted A. Meyer (Dept. of Otolaryngol., Indiana Univ. School of Med., 702 Barnhill Dr., RR-044, Indianapolis, IN 46202)

The traditional approach to determine which psychophysical variables are important for speech perception by cochlear implant (CI) users has been to perform correlational analyses. However, the finding of a statistically significant correlation between psychophysical and speech perception variables does not explain the actual mechanisms CI users employ to identify speech sounds. A mathematical modeling approach is proposed that does provide a detailed description of how CI users identify speech sounds based on the acoustic information they receive. To use the proposed approach, it must be hypothesized what the relevant psychophysical dimensions are for a given stimulation strategy. In this study, it is hypothesized that users of the SPEAK stimulation strategy use three types of cues to identify consonants: first, the average cochlear locations stimulated in response to the first three formants; second, the silent gap duration (if present); and third, the ratio between total energy above and below 1000 Hz. Based on a subject's performance along these psychophysical dimensions, the model outputs a predicted consonant confusion matrix. Preliminary comparisons between model predictions and actual data indicate that the model was able to predict which pairs of consonant sounds would be most frequently confused by CI users. [Work supported by NIH.]

9:45

4aPP7. Categorical loudness perception in normal and hearing impaired subjects. Vishakha W. Rawool (Commun. Disord. and Special Education, Bloomsburg Univ., Bloomsburg, PA 17815)

This study compared categorical loudness perception in three subject groups: young (22–24 years) normal, older (43–60 years) normal, and older (60–80 years) hearing impaired (thresholds between 50 and 70 dB SPL at the test frequencies). Eleven subjects were included in each group. The subjects were asked to judge the loudness of warbled tones (2 and 4 kHz) presented at various sound-pressure levels in the following categories: not audible, very soft, soft, comfortable, loud, and very loud. The sound-pressure levels were calibrated for each individual ear. The multivariate analyses of variance was performed on the very soft, soft, comfortable, and loud variable. Average SPLs obtained in each category were used for the analyses. The results showed significant differences in the young normal and the hearing impaired groups in loudness perception across the four categories. No significant differences were apparent in the young normal and older normal groups. The older normal and the older hearing impaired groups differed across all the loudness categories except for the ''loud'' category where no significant differences were apparent. [Work supported by a Bloomsburg University grant for Research and Disciplinary Projects.]

10:00

4aPP8. Across-channel sensitivity to temporal asynchrony in cochlear implantees. Robert P. Carlyon (MRC Appl. Psych. Unit, 15 Chaucer Rd., Cambridge CB2 2EF, UK), Luc Geurts, and Jan Wouters (Univ. of Leuven, Leuven, Belgium)

Five post-lingually deafened users of the LAURAflex cochlear implant detected temporal asynchronies between two 400-ms, 100-Hz trains of 40-μs biphasic pulses, applied to two ''target'' electrode pairs. The pulses in the signal stimulus of each 2IFC trial were delayed by between 0.3 and 4.8 ms on either the basal or apical target, relative to those on the other target. The pulses in the standard stimuli were nearly synchronous, containing a 0.1 ms-delay to the electrode whose pulses were not delayed in the signal. All listeners performed substantially above chance over a wide range of delays. Four were more sensitive to basal delays than to apical delays; one showed the opposite trend. Presenting a ''masking'' 1000-Hz pulse train to electrode(s) between the two targets did not substantially affect the results. This finding, combined with the wide (6–10 mm) separation between the targets suggests that performance was not mediated by neurons responding to both target electrodes. Rather, it is argued that they represent a genuine sensitivity to differences in timing between discrete auditory channels.

10:15

4aPP9. The effect of fast-acting compression on psychoacoustic jnd's. Brent W. Edwards (ReSound Corp., 220 Saginaw Dr., Redwood City, CA 94063)

Fast-acting compression is used in hearing aids to compensate for loudness recruitment in patients with sensorineural hearing loss. Intensity difference limens are independent of the slope of the loudness function and it has been argued that compression normalizes loudness perception in hearing-impaired individuals at the expense of their intensity discrimination performance. Similar arguments have been made with regard to the effect of fast-acting compression on envelopes: Compression reduces the magnitude of envelope fluctuations and thus impairs the near-normal temporal processing capabilities of hearing-impaired individuals. The exact consequences of compression on these perceptual abilities are not clear, however, due to the level- and depth-dependent thresholds of intensity and modulation discrimination tasks, respectively. The effect of fast-acting compression will be quantified in terms of these psychophysical tasks and the results will be related to the speech recognition performance of hearing-impaired individuals.

10:30

4aPP10. Spectro-temporal modulation transfer and speech intelligibility for multiband amplitude compression. Joost M. Festen (Dept. of Otolaryngol., Univ. Hospital VU, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands, jm.festen@azvu.nl)

For ears with sensorineural hearing loss the growth of loudness with stimulus level is often larger than in normal ears. This phenomenon is called recruitment. To restore normal loudness perception and to fit speech in the limited dynamic range of an impaired ear many modern hearing aids apply some form of multiband amplitude compression. However, within the cochlea nonlinearity and frequency selectivity are closely linked and therefore recruitment often goes together with reduced frequency selectivity. While compression may correct the recruitment, it does not improve the frequency selectivity. The opposite is even more correct; multichannel compression reduces the spectro-temporal variations in speech. Therefore in terms of speech intelligibility, the net effect of compression is difficult to predict. In an experiment speech intelligibility after multiband compression in quiet and in noise for 16 hearing-impaired listeners was measured. For each individual listener the speech spectrum was shaped and presented monaurally halfway in the ear's dynamic range in all conditions. Results show a decrease in intelligibility with both increasing compression ratio and increasing number of channels. These results can be fitted with a single curve when intelligibility is plotted as a function of the transfer of phase-locked modulations in time and frequenc.

THURSDAY MORNING, 25 JUNE 1998 GRAND BALLROOM A (S), 9:15 TO 11:20 A.M.

Session 4aSA

Structural Acoustics and Vibration: Statistical Energy and Fuzzy Structure Analyses—Similarities and Differences

Joseph W. Dickey, Chair
CNDE, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218-2689

Chair's Introduction—9:15

Invited Papers

9:20

4aSA1. Resonators attached to a structure—a SEA canon. R. H. Lyon (RH Lyon Corp, 691 Concord Ave., Cambridge, MA 02138-0799, rhlyon@lyoncorp.com)

One of the earliest problems analyzed using SEA ideas was a plate structure connected to a single resonator. Then structure-to-structure problems replaced the single resonator with a multiplicity of resonators as a model for the second structure. Because of the desire to relate the problem to those involving real structural connections—to beams, plates, and sound fields—the resonators became models for the natural modes of the structure and the resonators lost their identity. Fuzzy structure models have emphasized the resonators themselves, with passing attention to actual structural arrangements that they apply to. Of particular interest to SEA, however, is the relationship between modal density and ''mass density,'' and the relative strengths of internal and coupling damping, particularly as they relate to the transient response of the system. This presentation will discuss these matters.

9:55

4aSA2. Rudimentary statistical energy analysis and structural fuzzies. G. Maidanik (David Taylor Res. Ctr., Bethesda, MD 20084) and J. Dickey (Johns Hopkins Univ., Baltimore, MD 21218)

Structural fuzzies (Sfs) is a dynamic complex that is attached to a primary structure in order to provide a higher degree of dissipation. The dissipation is with respect to the power that is injected into the driven primary structure. It is argued that care needs to be exercised in the definition of the loss factor that is associated with this dissipation. This definition is examined and discussed; then, the role of the Sfs is analyzed in terms of a rudimentary statistical energy analysis (SEA). It is shown that the effectiveness of the dissipation is enhanced if the coupling of the Sfs to the primary structure is strong enough to approximate equipartitioning of modal energies, if the modal density in the Sfs either approximates or exceeds that in the primary structure and, finally, if the inherent loss factor in the Sfs is sufficiently higher than that in the primary structure. The definitions of strong enough coupling, modal densities, and inherent loss factors will be given. Examples of Sfs will be cited and discussed.

Contributed Papers

10:30

4aSA3. Spatial distribution of energy in SEA acoustic volumes. Evan B. Davis (Boeing Commercial Airplane Group, P.O. Box 3707 M/S 67-ML, Seattle, WA 98124-2207, evan.b.davis@boeing.com)

Statistical energy analysis (SEA) methods predict the space average energy in acoustic volumes. The predicted SEA acoustic energy levels are converted to SPL levels using the assumption that the sound field is diffuse. The SEA prediction methods estimate a single, space averaged value for the acoustic energy in the enclosure, whereas the theory of diffuse sound fields in enclosures predicts a nonuniform energy distribution in the enclosure. The energy levels and SPLs increase near the boundaries of the enclosure. The theory of diffuse sound fields can therefore be applied to generate an estimate of the spatial distribution of sound energy from the predicted spatial average. The sound field distributions are based on the Waterhouse corrections for reverberation rooms [Waterhouse, J. Acoust. Soc. Am. **27**, 247–258 (1955)]. The use of the Waterhouse correction in

the interpretation of SEA predictions is most useful in acoustic volumes which are relatively small in relation to the acoustic wavelengths of interest. Enclosures which have dimensions on the same order as the acoustic wavelengths are commonly encountered in transportation systems such as automobiles and airplanes.

10:55

4aSA4. Fuzzy structure, initial problem. Poincare theorem. Samuil A. Rybak (N. N. Andreyev Acoust. Inst., Shvernik Str. 4, Moscow 117036, Russia)

A system consisting of a master structure—an oscillator, coupled with a number of oscillators, distributed in some frequency range—is considered. The master oscillator is initiated by a delta-function pulse, and damping is neglected. The energy flows from the master structure to the fuzzy system, the amplitude of its oscillations falls exponentially, and the amplitudes of fuzzy system oscillators are exponentially growing. This means that asymptotically the master structure becomes silent and the fuzzy oscillators have the maximum of their amplitudes. After the time T, depending on the full number of degrees of freedom, the picture almost fully returns back. But the time T becomes infinite when the number of oscillators is infinite (when their distribution is continual, for example). If at an arbitrary time one changes the signs of the velocities of the full system of oscillators, the phase trajectory turns back strongly the same way, because the system is conservative. The path tracing of the pole, which is originating in the integral describing the influence of the fuzzy system of oscillators on the master oscillator, is the same as usual, but the sign of the index of the exponential attenuation changes and the system goes to the initial state, with amplitude growing exponentially. [Work was partially supported by the Russian Foundation of Basic Research.]

THURSDAY MORNING, 25 JUNE 1998 GRAND BALLROOM I (W), 7:45 TO 10:45 A.M.

Session 4aSC

Speech Communication: Speech Perception Theories and Models (Poster Session)

Paul Iverson, Chair
House Ear Institute, 2100 West Third Street, Los Angeles, California 90057

Contributed Papers

All posters will be on display from 7:45 a.m. to 10:45 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 7:45 a.m. to 9:15 a.m. and contributors of even-numbered papers will be at their posters from 9:15 a.m. to 10:45 a.m. To allow for extended viewing time, posters will remain on display until 7:30 p.m. A cash bar will be set up in the poster area from 5:00 p.m. to 7:30 p.m. preceding the banquet.

4aSC1. Testing two information taxonomies in Spanish. Guillermo Andrés Toledo (Dépt. de langues et linguistique, Faculté des lettres, Univ. Laval, Cité Univ., Québec, QC G1K 7P4, Canada, ucineacad@sminter.cam.ar)

Two information taxonomies, i.e., in the first experiment, New/Given distinction and, in the second experiment, three cognitive categories of the mental representation of referents: activation, semiactivation, inactivation differences [W. Chafe, in Tomlin *Coherence and Grounding in Discourse* (John Benjamins, Amsterdam, 1987), pp. 21–51 and K. Lambrecht, *Information Structure and Sentence Form* (Cambridge U.P., Great Britain, 1994), pp. 1–1161 have been explored in Argentinian Spanish. In the first experiment, semispontaneous materials were recorded from communicative interactions between three male speakers, one female speaker, and the listener. Brand-New plus New (Inferrable) versus Given (Textually Evoked) items were selected for acoustic analysis. In the second experiment, activation, semiactivation, and inactivation states encoded in semispontaneous discourses were taken from interactions between three male speakers and the listener. In both experiments, tonal prominences were studied through intonational contours on pairs of contrasted items: natural values and data transformed through logarithmic z-score normalization. Results in the first experiment indicated a relevant tonal encoding in new items. Findings in the second experiment have shown higher prominences correlated with inactive states, lower prominences with active and semiactive states.

4aSC2. Perception of speech gestures. René Carré (ENST, Dept. Signal, 46 rue Barrault, 75634 Paris Cedex 13, France) and Pierre L. Divenyi (Exp. Audiol. Res., V. A. Med. Ctr., Martinez, CA 94553)

Perception of a synthetic French /aya/ sequence, generated by coproduction of two simultaneous gestures (tongue constriction and lip rounding), was tested in a series of experiments. The tokens consisted of eight different degrees of gesture reduction in which the target /y/ was not reached. The /ay/ trajectory in the plane $F1$–$F2$ was not rectilinear (it pointed first to /i/ before reaching /y/), having to do with the fact that the acoustic effect caused by lip rounding (lowering $F2$) becomes manifest later than the effect due to constriction (lowering $F1$ and raising $F2$). The /aya/ transition is, therefore, an appropriate stimulus to study the role of the linearity of trajectories in speech perception. Results show that unless /y/ is present as a steady-state vowel for at least 30 ms, the lip rounding gesture is ignored and the percept is /aia/. Thus the results strongly suggest

that formant trajectories are perceived as time averages through a mechanism that calculates the direction of transition toward an (unreached) target from a running average of the direction of the formant trajectory. These findings point to the importance of the trajectory direction in vowel reduction and of the time dimension in phonetic distinctions. [Work supported by NATO.]

4aSC3. Coarticulation in CV syllables: A locus equation and EPG perspective. Marija Tabain (Speech, Hearing and Lang. Res. Ctr., Dept. of Linguist., Macquarie Univ., Sydney, Australia 2109)

Stop, nasal, and fricative consonants of Australian English were paired with a wide variety of vowel monophthongs to form CV syllables which were placed in a carrier phrase. The EPG and acoustic data were obtained from four female speakers producing these phrases. Locus equations [Lindblom, Acoust. Soc. Am. **35**, 1773–1781 (1963)] were generated from the acoustic data, and coarticulation indices [Farnetani, *Speech Production and Speech Modelling*, edited by Marchal and Hardcastle (1990), pp. 93–130] were calculated from the EPG data. The coarticulation indices were compared with the slope values from the locus equations [Krull, PERILUS **10**, 87–108 (1989)] in order to determine the degree of coarticulatory resistance for a given consonant phoneme as measured in the articulatory and acoustic domains. Results suggest that the greatest degree of coarticulation is shown in the labial and velar consonants, and less in the apico-alveolar consonants. The least degree of coarticulation is shown by the lamino-postalveolar consonants /S Z/. These results agree with data from Australian Aboriginal languages and with results presented by Recasens for Catalan [J. Phon. **12**, 61–73 (1984); **19**, 177–192 (1991)].

4aSC4. A model for dependences in phonetic categorization. Roel Smits (Dept. of Phonet. and Linguist., Univ. College London, 4 Stephenson Way, London NW1 2HE, UK)

A quantitative model for human categorization behavior is proposed which can be applied to four-alternative forced-choice categorization data involving two binary classifications. A number of processing dependences between the two classifications are explicitly formulated, such as the dependence of the location, orientation, and steepness of the class boundary for one classification on the outcome of the other classification. The significance of various types of dependences can be tested statistically. Analyses of two data sets from the literature [Repp *et al.*, J. Exp. Psych: Human Percept. Perform. **4**, 621–637 (1978); Whalen, Percept. Psychophys. **46**, 284–292 (1989)] show that interesting dependences in human speech recognition can be uncovered using the model.

4aSC5. Familiarity and pronounceability of nouns and names. Aimée M. Surprenant (Dept. of Psychol. Sci., Purdue Univ., West Lafayette, IN 47907-1364, aimee@psych.purdue.edu), Susan L. Hura (Lucent Technologies, Holmdel, NJ 07733), Mary P. Harper, Leah H. Jamieson, Stephen A. Hockema, Tsung-Hsiang Hsueh, Michael T. Johnson, Pramila N. Srinivasan, Scott M. Thede, Christopher M. White, Glenis Long, and Ayasakanta Rout (Purdue Univ., West Lafayette, IN 47907-1353)

Proper names have several properties that create problems for speech recognition systems: the number of names is large and ever changing, names can be borrowed directly from other languages and may not conform to usual pronunciation rules, and the variety of pronunciations for names can be high. Because the set of proper names is so dynamic and machines are notoriously poor at phoneme recognition, a promising approach to designing a name recognition system is to incorporate statistical aspects of proper names (e.g., frequency, familiarity). Unfortunately, there exists relatively little data on the distribution of names. Ratings of familiarity and pronounceability were obtained for a randomly chosen sample of 199 surnames (from 80 987 entries in the Purdue phonebook) and 199 nouns (from Kucera–Francis). The ratings for nouns versus names are substantially different: nouns were rated as more familiar and easier to pronounce than surnames. Frequency and familiarity were more closely related in the proper name pool than the word pool, although the correlations were modest. Ratings of familiarity and pronounceability were highly related for both groups. The value of using frequency and the ratings of familiarity and pronounceability for predicting variations in actual pronunciations of words and names will be discussed.

4aSC6. Ranking the pitches of concurrent vowels. D. Dwayne Paschall (Texas Tech Univ. Health Sci. Ctr., Box 42073, Lubbock, TX 79409) and Peter F. Assmann (Univ. of Texas at Dallas, Richardson, TX 75083-0688)

When two vowels are presented simultaneously, listeners can label them more accurately if their fundamental frequencies differ, and under some conditions they can report which vowel has the higher pitch. In this study the interaction of fundamental frequency and binaural cues on judgments of the relative pitches of concurrent vowel pairs is examined. On each trial listeners were told which vowels were presented and were asked to indicate which vowel was on the higher pitch. The first vowel had an F_0 of 140 Hz and the second vowel was either 1, 2, 4, or 8 semitones lower or higher than the first. When the fundamental frequency difference was small (1 semitone), the ability to rank the pitches of the two vowels was near chance level, and dichotic presentation provided little advantage over monaural presentation. With monaural presentation performance improved gradually with increasing F_0 difference, reaching a maximum near 80% correct at $+8$ semitones. When the stimuli were presented dichotically there was a dramatic improvement between 1 and 2 semitones and performance was near 100% with larger F_0 differences of 4 and 8 semitones. The interaction of F_0 differences and binaural cues places important constraints on models of concurrent vowel segregation.

4aSC7. Selection of a tonotopic scale for vowels. Terrance M. Nearey (Dept. of Linguist., Univ. of Alberta, Edmonton, AB T6G 2E7, Canada)

Various distinct tonotopic scales (e.g., log frequency, ERB, Bark) have been used for the representation of fundamental and formant frequency data for vowels. In previous work in our laboratories [T. Nearey, Proc. ICSLP92, 583–586 (1992)], generalized linear modeling was used to assess the relative appropriateness of these scales on both production measurement and vowel categorization data on each scale using several alternative vowel normalization schemes. That work focused on F_0 and F_1 and used the classic data of Peterson and Barney [G. Peterson and H. Barney, J. Acoust. Soc. Am. **24**, 175–184 (1952)]. Analysis of that data suggested that the log scale as provided somewhat better fits to the data with less evidence of violations of modeling assumptions. However, results in modeling perceptual data (from Canadian English listeners) were less clear. The present study will extend the scope of the earlier work by examining more data and expanding the statistical modeling framework. The production analyses will be augmented using an additional large production data set for another dialect of American English [J. Hillenbrand *et al.*, J. Acoust. Soc. Am. **97**, 3099–3111 (1995)]. The perceptual analyses will be augmented using additional perceptual data sets collected in our laboratories.

4aSC8. An exemplar-based account of emergent phonetic categories. Francisco Lacerda (Stockholm Univ., Inst. of Linguist., 106 91 Stockholm, Sweden)

This paper describes how an exemplar-based model can be used to study the early stages of language acquisition and phonetic category formation. The paper discusses in particular how a (conceptually) very simple model, inspired by Edelmans notion of Neural Darwinism, may account for the emergence of phonetic categories arising from the interaction between ambient language input and limited memory-representation resources. It presents an explicit attempt to demonstrate how phonetic

categorization may emerge ''spontaneously'' from exposure to natural infant-directed speech data. The results suggest that phonetic categories may emerge if the auditory input is statistically correlated with one (or several) other sensory inputs. The perceptual-magnet effect is also discussed in the light of the present model. According to the model, the perceptual-magnet effect can be seen as a consequence of distributed memory representations. [Work supported by The Bank of Sweden Tercentenary Foundation, grant 94-0435.]

4aSC9. Effect of spectral distance on vowel height perception. Patrick C. M. Wong and Randy L. Diehl (Dept. of Psych., Univ. of Texas at Austin, Austin, TX 78712, pcmwong@ccwf.cc.utexas.edu)

Perceived vowel height has been reported to vary inversely with the distance (in Bark) between the first formant frequency ($F0$) and the fundamental frequency ($F0$) [H. Traunmueller, J. Acoust. Soc. Am. **69**, 1465–1475 (1981)]. However, in a study using back vowels, Fahey *et al.* [J. Acoust. Soc. Am. **99**, 2350–2357 (1996)] found that phonetic quality was not only determined by the $F1-F0$ distance, but also by the tonotopic distances between *any* adjacent spectral peaks (e.g., $F3-F2$, $F2-F1$, and $F1-F0$), with greater weight accorded to smaller distances. The present study further tested this possibility by using front vowels. Listeners identified three sets of synthetic vowels varying orthogonally in $F1$ and $F0$ and ranging from /i/–/ɪ/, /ɪ/–/ɛ/, or /ɛ/–/æ/. The results allowed this possibility to be rejected. Furthermore, they supported the study of Hoemeke and Diehl [J. Acoust. Soc. Am. **96**, 661–674 (1994)] that $F1-F0$ was the best predictor of perceived vowel height for the phonological distinction [+/−high] (i.e., the /ɪ/–/ɛ/ distinction), while $F1$ alone was the best predictor for the two other vowel set distinctions. [Work supported by NIDCD.]

4aSC10. Syllable as the segmentation unit in perceiving spoken Chinese. Chin-Hsing Tseng (Dept. of Special Education, Natl. Kaohsiung Normal Univ., Kaohsiung, Taiwan)

How spoken Chinese is segmented by native speakers has been addressed by the author in recent years. Evidence has accumulated to support the view that the syllable is the level at which speakers segment their spoken language. Earlier studies found that Mandarin–Chinese phonemes were detected faster in real syllables than in nonexistent syllables, that they were faster for level-tone and falling-tone syllables than for rising-tone or dipping-tone syllables, and that structural complexity had no effect on phoneme detection. These effects have preempted phoneme and word as the segmentation unit in real-time processing. This study was designed to investigate the role of stress foot in segmentation, to discover the locus of the tone effect, and to replicate the tone and complexity effects with Taiwanese speech materials. Stress assignment and phoneme detection tasks were employed as the experimental paradigms. The results indicated that stress assignment in Mandarin is not consistent enough to warrant a stress-foot segmentation procedure, that the tone effect originates from syllable duration variations, and that earlier effects were replicated with Taiwanese.

4aSC11. Phonological boundaries and the spectral center of gravity. Michelle R. Molis, Randy L. Diehl, and Adam Jacks (Dept. of Psychol., Univ. of Texas, Austin, TX 78712, molis@mail.utexas.edu)

A critical limit of 3–3.5 Bark has been reported for the ''spectral center of gravity'' effect between the second and third formants ($F2$ and $F3$) and is assumed to correspond to a perceptually natural boundary [A. K. Syrdal and H. S. Gopal, J. Acoust. Soc. Am. **79**, 1086–1100 (1986)]. This hypothesis was tested for both the [+/−back] and [+/−coronal] distinctions in English vowels. Subjects identified two sets of three-formant synthetic vowels which varied orthogonally in $F2$ and $F3$ and ranged between /ʊ/–/ɪ/ ([+/−back]) or /ɝ/–/ʊ/ ([+/−coronal]). For the /ʊ/–/ɪ/ distinction, a boundary shift was observed solely as a function of $F2$, ruling out an invariant $F3-F2$ boundary. For the /ɝ/–/ʊ/ series, both $F2$ and $F3$ influenced boundary location. There was a relatively stable $F3-F2$ boundary for the mean identification responses in this case, but it occurred at less than the predicted 3–3.5 Bark difference. Follow-up results will also be presented. [Work supported by NIDCD.]

4aSC12. An exemplar-based model of silent-center syllable perception. Stephen Winters and Keith Johnson (Dept. of Linguist., Ohio State Univ., 222 Oxley, 1712 Neil Ave., Columbus, OH 43202)

Rakerd and Verbrugge [J. Memory Lang. **26**, 558–563 (1987)] suggested that the perception of silent-center syllables—CVC syllables with the middle 60% of the vowel edited out and replaced with silence—had to involve some form of speaker normalization. Crucial to their argument was their subjects' ability to perceive cross-gender hybrids of such syllables. Rakerd and Verbrugge reasoned that even modest accuracy in perceiving these hybrids entailed some sort of normalization of the disparate acoustic qualities of male and female voices in order to derive one perceptual whole. Rakerd and Verbrugge's study has been replicated in order to test the ability of a computationally implemented exemplar model to perceive silent-center syllables. Exemplar-based modeling utilizes no ''speaker normalization'' as such, but, nonetheless, the model in this study approximated the performance of human subjects in perceiving silent-center syllables correctly. The model's performance on the important ''hybrid'' case was almost identical to subjects' responses. These results suggest that the human ability to perceive silent-center syllables does not necessarily entail some process of speaker normalization in speech recognition. [This material is based upon work supported under a National Science Foundation Graduate Fellowship and NIDCD Grant No. R29-DC01645-06.]

4aSC13. The time course of learning: Neurophysiologic changes during speech training. Kelly L. Tremblay, Nina Kraus, Therese McGee, and Steve Zecker (Dept. of Commun. Sci. and Disord., Northwestern Univ., 2299 North Campus Dr., Evanston, IL 80208)

The ability of normal-hearing adults to identify and discriminate novel speech sounds improves with training. Neurophysiological changes associated with this improvement include increased duration and area, as well as shortened onset latency, of the mismatch negativity (MMN) [N. Kraus *et al.*, J. Cog. Neurosci. **7**, 25–32 (1995); K. Tremblay *et al.*, J. Acoust. Soc. Am. **102**, 1–12 (1997)]. The MMN is a passively elicited, preattentive, auditory event-related potential that reflects auditory discrimination and echoic memory [R. Naatanen *et al.*, Acta Psychol. **42**, 313–329 (1978)]. To date, little is known about the time course of neurophysiological change and the manifestation of learning. The present study examined the time course of perceptual learning and its relation to changes in the MMN with the expectation that cortically evoked potentials could provide insight into the biological processes underlying perceptual learning and serve as a clinical tool for evaluating change during efforts of (re)habilitation in the communicatively impaired. Specifically, it was found that neurophysiological change is apparent in the MMN preceding behavioral improvement. This suggests that auditory training alters the neural activity that provides the necessary coding of the speech stimuli and those changes are later integrated into functional behavior.

4aSC14. Context effects in the auditory identification of Spanish fricatives /f/ and /θ/: Hyper and Hypospeech. Sergio Feijóo, Santiago Fernández, and Ramón Balsa (Dpto. Física Aplicada, Fac. de Física, 15706, Santiago, Spain, fasergio@uscmail.usc.es)

Twenty-eight subjects heard 40 Spanish words in which the initial fricative was /f/ or /θ/, combined with vowels /e/ and /u/. Ten words were used for each particular combination (2 fricatives$\times$2 vowels $\times$10 words). Two forms of speech (Hypo and Hyperspeech) and four conditions were considered: (1) Isolated fricative segment; (2) fricative segment + 51.2 ms of the following vowel; (3) fricative + whole following vowel; (4) whole word. The statistical analysis showed that, despite their differences in production and acoustic characteristics, isolated fricative segments were equally recognized in Hypo and Hyperspeech (cond. 1). Including the vowel (conditions 2 and 3) significantly improved recognition of both fricatives for both forms of speech, except for the combination /f/+/e/: While fricative identification improves slightly in Hyperspeech, in Hypospeech, recognition decreases with respect to cond. (1). For this particular combination, an acceptable recognition rate is only achieved in the whole word condition for both forms of speech. The results indicate that: (a) Clear speech is not necessarily more intelligible than hypospeech; and (b) The role of vocalic context in the recognition of /f/ and /θ/ is similar for both forms of speech.

4aSC15. Evidence of independent verbal processors for the same stimulus: Insights from dichotic verbal transformations. James A. Bashford, Jr. and Richard M. Warren (Dept. of Psych., Univ. of Wisconsin, Milwaukee, WI 53201, bashford@csd.uwm.edu)

A clearly enunciated word heard repeating without change undergoes illusory verbal transformations (VTs) to different words or syllables. Surprisingly, when a repeating word is delivered dichotically with an interaural asynchrony of half its duration, the separate lateralized images of the word are heard to undergo independent VTs (e.g., ''commence'' heard at one ear while ''tress'' is heard at the other) [see Warren and Ackroff, Nature **259**, 475–477 (1976)]. The present study eliminates a peripheral explanation attributing such independent changes to lateral asymmetries in repetition effects upon the lower auditory pathways, which might provide conflicting acoustic-phonetic information to a single verbal processor. A dichotic alternation paradigm, in which successive statements of a repeating word were presented to opposite ears (e.g., ''flame'' presented left, then ''flame'' presented right . . .) was also found to produce independent VTs. However, when higher amplitude noise was presented to the ear opposite the alternating speech signal, contralateral induction [Warren and Bashford, Percept. Psychophys. **20**, 380–386 (1976)] produced diffuse medial localization of the otherwise lateralized signals, and despite monaural verbal stimulation, VTs were linked across ears. This and other evidence to be discussed demonstrate independent preattentive verbal processors for spatially distinct sources. [Work supported by NIH.]

4aSC16. Inducing a ''perceptual magnet''-like effect in a nonspeech modality. Fatima Husain and Frank H. Guenther (Dept. of Cognit. and Neural Systems, Boston Univ., 677 Beacon St., Boston, MA 02215, fhusain@cns.bu.edu)

Kuhl and colleagues have described a ''perceptual magnet effect'' in which subjects show a decreased ability to discriminate between vowel-like stimuli when these stimuli fall near a prototypical vowel from their native language. Guenther and Gjaja [J. Acoust. Soc. Am. **100**, 1111–1121 (1996)] proposed that the effect results from experience-based map formation in auditory brain regions. Because this model attributes the effect to neural mechanisms that have been identified in cortical areas subserving different sensory modalities, it predicts that it should be possible to induce magnet effects in nonspeech modalities. To test this, adult subjects' discrimination sensitivities to narrow-band noise stimuli were tested before and after training sessions meant to approximate an infant's exposure to speech sounds. Training consisted of a categorization task that preferentially exposed subjects to stimuli from a small frequency range meant to correspond to a prototypical vowel. Sensitivity tests after training revealed that subjects had become worse at discriminating sounds within this ''prototype'' range as compared to a ''nonprototype'' range that was absent during training. This suggests that the magnet effect may not depend on phonetic processing and may reflect neural mechanisms common to different sensory modalities. [Supported by NIDCD and the Sloan Foundation.]

4aSC17. Interarticulator phasing and locus equations. Anders Lofqvist (Haskins Labs., 270 Crown St., New Haven, CT 06511, lofquist@haskins.yale.edu)

A locus equation plots the frequency of the second formant at vowel onset against the target frequency of the same formant for the vowel in a consonant–vowel sequence. The slope of the equation has been assumed to reflect the degree of coarticulation between the consonant and the vowel, with higher slopes associated with more coarticulation. This study examined the articulatory basis for this assumption, using VCV sequences where the consonant was a bilabial stop /p, b/ and the vowels one of /i, a, u/. Articulatory movements were recorded using a magnetometer system. Four subjects participated and produced ten repetitions of each sequence. One articulatory measure was the temporal phasing between the onset of the lip closing movement and the onset of the tongue body movement from the first to the second vowel. Another was the magnitude of the tongue movement from the onset of the second vowel to the tongue position for the vowel, averaged across four receivers placed on the tongue. When compared with the corresponding locus equation slopes, neither measure showed support for the assumption that the slope serves as an index of the degree of coarticulation between the consonant and the vowel. [Work supported by NIH.]

4aSC18. Latency of MEG M100 response indexes first formant frequency. Krishna K. Govindarajan (MIT, Cambridge, MA 02139), Colin Phillips (Univ. of Delaware, Newark, DE 19716), David Poeppel, Timothy P. L. Roberts (UCSF, San Francisco, CA 94143), and Alec Marantz (MIT, Cambridge, MA 02139)

Magnetoencephalography recordings from the auditory cortex of subjects listening to synthetic vowels show a close correlation between the timing of the evoked M100 response and the first formant frequency ($F1$). These results are consistent with evoked magnetic field latencies elicited by tone stimuli, which show 100 to 300-Hz tones associated with latencies up to 30 ms longer than 500 to 3000-Hz tones. In experiment 1, three-formant vowels /u,a,i/ were presented at two fundamental frequencies ($F0=100$ Hz, 200 Hz). The M100 latency was a function of the vowel identity and not $F0$: M100 was significantly shorter for /a/ than /u/, consistent with the spectral center of gravity being at a higher frequency for /a/. In experiment 2, single-formant vowels (/u/:$F1=330$ Hz, /a/:$F1=720$ Hz) were covaried with two $F0$ values. M100 latencies were shorter for /a/ (high $F1$) than for /u/ (low $F1$), at both fundamentals. In experiment 3, subjects listened to pure-tone complexes with frequencies and amplitudes matching the $F0$ and $F1$ energy peaks of the stimuli in experiment 2. M100 latencies showed the same pattern: latency covaried with the energy

peak corresponding to $F1$, suggesting that the sensitivity to the energy in the $F1$ range is not speech-specific. Finally, an experiment involving two-formant, amplitude-varying vowels will be presented.

4aSC19. Perceptual magnet effect in corner vowels. Robert Allen Fox and Lynn Carahaly (Dept. of Speech and Hearing Sci., Ohio State Univ., Columbus, OH 43210)

In previous studies of speech perception, Kuhl and colleagues have provided evidence that the category goodness of a vowel influences the perception of that particular vowel—the perceptual magnet effect. These data indicate that listeners demonstrate poor discrimination of vowel tokens when they are close to the prototype vowel but good discrimination when the vowel is close to a nonprototypic vowel. However, data will be provided from experiments conducted in this laboratory using an MDS paradigm on the /i/–/ɪ/ distinction that indicate this effect is eliminated or severely curtailed when all tokens are good exemplars of the /i/ category ($>75\%$) and the nonprototypic exemplars are not near a phoneme boundary. In addition, given the almost exclusive reliance in the literature on the /i/–/ɪ/ vowel pair, additional data will be provided on the magnet effect in the /u/–/ʊ/ and /ʌ/–/ɑ/. Similar to /i/, extreme tokens of both /ɑ/ and /u/ that are distant from the prototype, but not in the direction of another vowel category (and thus near a phoneme boundary). The results will be discussed in terms of whether the effect is linguistic (phonetic) or auditory in nature.

4aSC20. Surveying auditory space using vowel formant data. Matthew J. Makashay and Keith Johnson (Dept. of Linguist., Ohio State Univ., 222 Oxley Hall, 1712 Neil Ave., Columbus, OH 43210, makashay@ling.ohio-state.edu)

Recent research has suggested that vowel category learning and the perceptual magnet effect are the natural consequences of auditory neural map formation [F. H. Guenther and M. N. Gjaja, J. Acoust. Soc. Am. **100**, 1111–1121 (1996)]. As these effects are language specific, the organization of the perceptual map crucially is dependent on the input received during learning. Guenther and Gjaja implemented a working model of a self-organizing neural map to show that, given input from Gaussian distributions around vowel category centers, unsupervised training leads to a warping of the perceptual space toward the category centers. This work replicates Guenther and Gjaja's results and provides a more realistic test of the model by using real vowel formant data instead of idealized distributions. Single and multiple male talker vowel data were used to train auditory neural maps. Single-talker simulations resulted in map organization similar to Guenther and Gjaja's results. However, no clear vowel category clusters emerged in the multiple-talker simulations. [Work supported by NSF Graduate Fellowship and NIH FIRST Award R29-DC01645-06.]

4aSC21. Relation between discrimination and identification of English vowels. Diane Kewley-Port and Amy T. Neel (Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405)

Although resolution for formant frequency is considerably better than needed for vowel identification in English, adjacent vowels in the $F1$-by-$F2$ plane may not be equally discriminable. Kuhl has reported better discrimination for poor vowel exemplars than good exemplars (''perceptual magnet effect'') for a few vowels in high stimulus uncertainty tasks. The purpose of this experiment was to determine the relation between discrimination thresholds and the judgment of vowels as good, confusable, or non-English. A large $F1$-by-$F2$ vowel space (190 tokens) encompassing five English front vowels was synthesized in equal bark steps. Based on an identification task with goodness ratings, two vowels representing each goodness category were selected for the discrimination task. Eight listeners participated in a minimal uncertainty discrimination task to estimate thresholds for $F1$ and $F2$. Analysis of variance of the 12 thresholds (converted to Δ Barks) found that thresholds for non-English vowels were significantly poorer than the other two categories, whereas thresholds for good and confusable vowels were similar. These results are not consistent with the perceptual magnet effect. Vowels within areas of the $F1$-by-$F2$ space judged as good exemplars are not discriminated more poorly than those in confusable areas. [Work supported by NIH-NIDCD-02229.]

4aSC22. The effect of vowel prototypicality and extremity on discrimination sensitivity. Satsuki Nakai (Univ. of Edinburgh, Edinburgh EH8 9LL, UK, satsuki@ling.ed.ac.uk)

The present study examines the effect of vowel extremity and prototypicality on discrimination sensitivity. Kuhl [1991] has shown that listeners are not as capable of distinguishing other variations of /i/ from a prototypical /i/ (the stimulus with the highest goodness rating, henceforth P) as from a nonprototypical /i/ (a stimulus with a low goodness rating). However, considering evidence suggesting that discrimination sensitivity grows poorer towards the exterior of the vowel space [Schouten and van Hessen, 1992], it is unclear whether it is vowel prototypicality or extremity that correlates with low discrimination sensitivity, for P is the most extreme among the stimuli used in the discrimination task in Kuhl [1991]. In order to investigate the above issue, discrimination curves were obtained from native speakers of Japanese and Greek in the corner of the vowel space where Japanese /u/ and Greek /u/ are located. As Japanese /u/ is not cardinal, the effect of vowel prototypicality and extremity on discrimination sensitivity could be observed separately on Japanese subjects' discrimination curves. The results indicate that both vowel prototypicality and extremity contribute to low discrimination sensitivity, and that the extremity of P correlates with the magnitude of its effect on discrimination sensitivity.

4aSC23. Phonetic categories: Internal category structure and processing speed. Joanne L. Miller (Dept. of Psych., Northeastern Univ., Boston, MA 02115), Peter D. Elimas (Brown Univ., Providence, RI 02912), and Ethan Cox (Univ. of Arizona, Tucson, AZ 85721)

Phonetic categories have a graded internal structure, with some members perceived as better exemplars than others. In a project focusing on voiceless stop consonants, the relation between this structure and processing speed has been examined. Speech continua were constructed in which the voice-onset-time (VOT) value of a word-initial consonant systematically changed from very short to very long. Unspeeded identification tests confirmed that stimuli with very short VOT values are perceived as voiced, whereas those with longer VOT values are perceived as voiceless. Category-goodness rating tests further confirmed that only stimuli within a limit range along each continuum are perceived as good exemplars of the voiceless category; stimuli with shorter or longer VOT values are perceived as poorer exemplars. Initial results from an ongoing experiment involving speeded identification indicate that (1) compared to the best exemplars of the voiceless category, identification of poor exemplars in the voiced–voiceless boundary region is relatively slow, but (2) across a wide range of values, poor exemplars with VOT values longer than the best exemplars are identified as quickly as the best exemplars themselves. These initial findings indicate that internal category structure is not necessarily reflected in processing speed. [Work supported by NIH.]

4aSC24. The effect of speaking style alterations on locus equation stability. Harvey M. Sussman, Eileen Dalston, and Sam Gumpert (Dept. of Commun. Sci. and Disord., Univ. of Texas, Austin, TX 78712, sussman@mail.utexas.edu)

Locus equations, as a phonetic descriptor of stop place of production, have been systematically investigated across speakers, gender, languages, perturbation conditions, syllable position, and manner class. As locus equation slope is thought to directly reflect the extent of anticipatory coarticulation of the vowel on the prevocalic consonant, changes in coarticulation as brought about by altered speaking styles can effectively alter locus equation descriptors of stop place. This study attempted to assess the stability of locus equations within each speaker as citation-style hyperarticulation was compared relative to spontaneous, hypoarticulated output. Twenty speakers, ten male and ten female served as subjects. Traditional citation-style locus equations were generated from productions of CV /t/ words printed on lists with initial /bdg/ followed by ten medial vowel contexts. Locus equations were also derived from spontaneous speech output which was generated by rapid reading of discourse-like passages constructed to resemble normal conversational speech, but with numerous lexical items beginning with /bdg/ in many vowel contexts. Spectral analysis was performed with the Kay Elemetrics CSL system. Comparisons of slopes, y intercepts, R^2, and standard error of estimates across the two speaking conditions will be presented and discussed.

4aSC25. Modeling the perception of features in the identification of English consonants. Tobey L. Doeleman (Dept. of Linguist., Cornell Phonet. Lab., Cornell Univ., Ithaca, NY 14853)

The perception of voicing, place, and manner was investigated in a series of gating experiments in which subjects identified eight English consonants in six different duration conditions [T. L. Doeleman, J. Acoust. Soc. Am. **101**, 3111(A) (1997)]. Subjects heard either all eight stimuli or two mutually exclusive subsets such that certain feature information was given (e.g., manner given in sets [p, t, b, d] and [f, s, v, z]). Although all subjects listened to the same 48 stimuli, identification significantly improved for subjects given place information (i.e., sets of labials and alveolars). The results are modeled using a spreading-activation connectionist model, the Integration-Competition model [Spivey-Knowlton (1996)]. In this model, the phoneme and feature representations compete with one another such that featural distinctions correspond to different weights assigned to the connection between feature and phoneme layers. The target phoneme has high activation from all featural arrays, resulting in the highest activation in the phoneme array. The increase in identification scores across duration conditions is modeled by the increase in activation of the target phoneme with successive time steps. The model also predicts observed consonant confusion patterns in that distracter phonemes sharing more features with the target phoneme will have higher activation levels.

4aSC26. Additive effects of phonetic distinctions in word learning. Joseph V. Pater (Dept. of Linguist., Univ. of Alberta, Edmonton, AB T6G 2E7, Canada, joe.pater@ualberta.ca), Christine L. Stager, and Janet F. Werker (Univ. of British Columbia, Vancouver, BC V6T 1Z4, Canada)

While infants have been demonstrated to be sensitive to a wide variety of phonetic contrasts when tested in speech discrimination tasks [Eimas *et al.* (1971) et seq.], recent work [Stager and Werker (1997)] has shown that following habituation to a word–object pairing, infants of 14 months fail to notice when the place of articulation of the initial consonant is switched [b/d]. Using the same procedure, the present study has found that infants do not respond to a change in voicing [b/p]. They do, however, notice a switch between dissimilar words [lɪf/nim]. One interpretation of these findings is that 14-month-olds do not encode either place or voice distinctions in lexical representations, so that words differing in only these features are treated as identical. To test this hypothesis, the effect of combining featural contrasts is currently being investigated by examining whether infants do respond to a change in both place and voice [d/p]. If there is such an additive effect, the contrasts must be represented. This would entail that an explanation for the failure to distinguish words differing in only a single feature should invoke processing factors, rather than representational ones.

4aSC27. Vowel perception by children and adults: Based on steady states or transitions? Joan E. Sussman, Jason Steinberg, and Julie Fenwick (Dept. of Communicative Disord. and Sci., State Univ. of New York at Buffalo, 122 Park Hall, Buffalo, NY 14260, jsussman@acsu.buffalo.edu)

The current investigation studied whether children, aged 4–5 years, and adults identified vowels on the basis of steady-state or transitional formant frequencies. Four types of synthetic tokens, created with a female voice, served as stimuli: (1) steady-state centers—220-ms formant frequencies for the vowels [i] and [ae]. (2) vowelless tokens with transitions appropriate for [bib] and [baeb], (3) ''congruent'' tokens that combined the first two types of stimuli into [bib] and [baeb], and (4) ''conflicting'' tokens that combined the transitions from [bib] with the vowel in [baeb] and vice versa. Results showed that children identified vowels most accurately in steady-state centers and congruent stimuli (96%) and used steady-state cues in 87% of conflicting tokens for vowel identification. Adults were equally accurate for vowelless, steady-state, and congruent tokens (99%) and used steady-state and transition cues for vowel identification. Findings do not support Nittrouer's Developmental Weighting Shift Hypothesis, but findings for adults do lend some support to the Dynamic Specification theory of Strange and colleagues. [Work supported by a Research Development Fund award from the Research Foundation at the University at Buffalo and by the Center for Hearing and Deafness at UB.]

4aSC28. Lexical competition in spoken word recognition by younger and older adults: A comparison of the rime cognate, neighborhood, and cohort. Mitchell S. Sommers (Dept. of Psych., Washington Univ., Campus Box 1125, St. Louis, MO 63130) and Shigeaki Amano (NTT Basic Res. Labs., 3-1 Morinosato-Wakamiya, Atsugi, Kanagawa, 243-01 Japan)

A central issue in research on spoken word recognition has been to specify the lexical candidates that are initially activated and compete during lexical discrimination. Amano [J. Mem. Lang. (submitted)] has recently proposed a new candidate set, the rime cognate (RC), which suggests that the set of potential lexical candidates consists of all items containing the same phoneme sequence as the rime (vowel nucleus plus coda) of that word. The present study was designed to compare the ability of the RC and two other sets, the neighborhood and cohort, to account for spoken word recognition in both younger and older adults. Density measures derived from the neighborhood were significantly correlated with identification scores for younger and older subjects. In contrast, RC density metrics were significantly correlated with identification scores only for older adults. Both neighborhood and RC density measures accounted for a greater percentage of the variance in spoken word recognition for older than for younger listeners. Metrics derived from the cohort were not significantly correlated with identification scores for either younger or older adults. The implications of these findings for several current models of spoken word recognition are discussed. [Work supported by the Brookdale foundation.]

4aSC29. Integration of auditory and visual speech information. Michael D. Hall, Paula M. T. Smeele, and Patricia K. Kuhl (Dept. of Speech and Hearing Sci., Univ. of Washington, Seattle, WA 98195, mdhall@u.washington.edu)

The influence of visual information on auditory speech perception can be observed under conditions where the two sources of information are discrepant [McGurk and MacDonald (1976)]. Recent studies from our laboratory [Hall *et al.* (1996)] focused on one such condition, where a face producing /b/ is combined with a dubbed auditory /g/ token, resulting in the frequent auditory perception of /bg/. These studies used auditory tokens that varied in quality, and have suggested that the probability of auditory consonant identification as /g/ was sufficient to predict the probability of /bg/ responses. The current study extended this approach to the converse conditions, where an auditory /b/ was dubbed on a visual /g/ production, which typically results in the frequent auditory perception of /d/ or /th/. Additionally, similarity ratings were collected for sets of auditory /g/ and /b/ stimuli, as well as for both types of auditory-visual conditions (/g/+/b/ and /b/+/g/). Multidimensional scaling (MDS) of these ratings was used to determine whether auditory category structure predicts MDS solutions for combined auditory and visual information. Implications of the results for the nature of auditory-visual speech integration will be discussed. [Work supported by NICHD.]

4aSC30. Speech-sound perception in normal and learning-disabled children: Effect of lengthened CV transition duration. Ann R. Bradlow, Nina Kraus, Trent Nicol, Therese McGee, Jenna Cunningham (Auditory Neurosci. Lab., Northwestern Univ., 2299 N. Campus Dr., Evanston, IL 60208, abradlow@nwu.edu), and Thomas D. Carrell (Univ. of Nebraska–Lincoln, Lincoln, NE 68538)

Previous research has established that many children with learning problems exhibit perceptual deficits in response to certain auditory stimuli, such as stop consonants in CV context. In order to investigate the precise acoustic features of stop consonants that pose perceptual difficulties for these children, discrimination thresholds were compared for normal and learning-disabled children along two separate synthetic /da/–/ga/ continua that differed only in the duration of the formant transitions. Results showed that simply lengthening the formant transition duration from 40 to 80 ms did not result in improved discrimination thresholds for the learning-disabled children. Similarly, an electrophysiologic response that is known to reflect the brain's preconscious ability to detect a change from one auditory stimulus to another indicated diminished responses in these children to /da/ versus /ga/, regardless of transition duration. This finding suggests that the brevity of the CV formant transitions is not the sole acoustic basis for the observed perceptual deficits. Furthermore, signal processing techniques aimed at enhancing perception of stop consonants likely need to consider other acoustic features, such as relative amplitude and frequency characteristics, instead of (or in addition to) CV transition duration. [Work supported by NIH R01 DC 01510.]

4aSC31. Listener variability and multiple perception processes. Stephanie Lindemann (Program in Linguist., Univ. of Michigan, 213 W. Mosley St., Apt. 6, Ann Arbor, MI 48103, lindeman@umich.edu)

The relationship between the perception of individual phonemes and the perception of words is investigated to determine if these are similar tasks for listeners. Since listeners may be able to use top-down processes in full words, it is argued that, at least for some, the task of identifying isolated phonemes may be very different from the task of identifying a full word. Because greater variability in performance is expected from non-native speakers, 15 native speakers of Sudanese Arabic (as well as a control group of 12 native speakers of American English) were tested on identification of American English phonemes in word context and the same instance of these phonemes extracted from the words. The overall agreement between corresponding items on the two tests was significantly better than chance, but scores of individual subjects varied substantially, with phoneme-word test agreement falling below chance for many subjects. Additionally, not all non-native subjects performed better on the full word task, suggesting that they do not all use higher-level information. It is therefore argued that studies of phoneme-level perception are more relevant to some listeners' word perception than to others'.

4aSC32. Effects of affective tone of voice on spoken word recognition. Lynne C. Nygaard, Jennifer S. Queen, and S. Alexandra Burt (Dept. of Psych., Emory Univ., Atlanta, GA 30322)

Traditionally, the study of emotional tone of voice has been considered separately from the study of formal linguistic aspects of spoken language. Research has typically focused either on how listeners detect emotion in an individual's voice, or on how listeners extract abstract linguistic content. The present study was designed to investigate how these two sources of information interact. In two experiments, listeners were presented with homophones that had one affective (happy or sad) and one neutral meaning and with nonhomophones that had either a happy, sad, or neutral meaning. Within each experiment, words were spoken in a tone of voice that was either congruent, incongruent, or neutral with respect to affective meaning. In the first experiment, naming latencies were collected to determine if tone of voice affects the speed of lexical processing. In the second experiment, transcription accuracy was assessed to determine if tone of voice affects the selection of word meaning. The results suggest that emotional tone of voice was used to resolve lexical ambiguity, but did not appear to affect specifically the time course of lexical access. The implications of these results for current theories of spoken word recognition will be discussed. [Work supported by NIH.]

4aSC33. On the perception of qualitative and phonetic similarities among voices. Robert E. Remez, Jennifer L. Van Dyk (Dept. of Psych., Barnard College, 3009 Broadway, New York, NY 10027-6598), Jennifer M. Fellowes (Columbia Univ. College of Physicians and Surgeons, New York, NY 10032), and Philip E. Rubin (Haskins Labs., New Haven, CT 06511)

A perceiver who learns to recognize an individual talker becomes familiar with attributes of the talker's voice that are present in any utterance regardless of the linguistic message. Customary accounts of individual identification by ear presume that such durable personal aspects of an individual's speech are graded qualities (e.g., vocal pitch and pitch range; melodious, breathy, or raspy timbre; etc.) independent of the acoustic properties that evoke segmental phonetic contrasts. Alternatively, some classic and recent studies alike suggest that familiarity includes attention to attributes of dialect and idiolect conveyed in the articulation of consonants and vowels. These linguistic phonetic properties of speech are effective for recognizing a talker when voice quality is eliminated as a source of information. The present investigation sought direct evidence of attention to qualitative and phonetic attributes of speech. Natural samples and sine wave replicas of sentences spoken by five male and five female talkers were used in a similarity tournament with adult listeners. The results establish the differential perceptual resolution of qualitative and phonetic attributes in the perception and recognition of talkers. [Work supported by NIDCD and NICHD.]

4aSC34. Perception of American English /r/ and /l/ by Mandarin speakers: Influences of phonetic identification and category goodness. Feng-Ming Tsao, Michael Hall, Richard Eyraud, and Patricia K. Kuhl (Dept. of Speech and Hearing Sci., Univ. of Washington, Seattle, WA 98195)

Studies in speech perception show a perceptual magnet effect for listeners which is characterized by shrunken perceptual distances near excellent exemplars of phonetic categories and stretched distances near poor exemplars. This study extended this work by examining the influence of Mandarin speakers' phonetic identification and category goodness on their perception of American English /r/ and /l/. Eighteen /ra/ and /la/ tokens were used that varied *F*2 and *F*3 frequencies. Listeners identified stimuli as Mandarin /u/, /m/, or /l/, provided goodness ratings for them, and rated the similarity of all stimulus pairs. Multidimensional scaling analyses of similarity ratings exhibited that the perceptual ''map'' for Mandarin speakers differs substantially from that of American speakers [Iverson and Kuhl, J. Acoust. Soc. Am. (1996)] listening to the same stimuli. The results indicate that (1) Mandarin speakers perceived /u/ and /l/ in the space, (2) perceptual distance was shrunk near the best exemplars of phonetic categories and stretched near the category boundary, (3) variations of *F*2 were more important than *F*3, and (4) individual differences in /l/ identification corresponded to a degree of shrinking near the best exemplars of the /l/ category. Results indicate that listeners use their native-language representations in the perception of a non-native language.

4aSC35. Non-native listeners' representations of within-word structure. Takashi Otake (Dokkyo Univ. and Max-Planck Inst. for Psycholinguistics, Wundtlaan 1, 6525 XD, Nijmegen, The Netherlands, otake@dokkyo.ac.jp), Kiyoko Yoneyama (The Ohio State Univ., Columbus, OH 43210-1298), and Hideki Maki (Salem-Teikyo Univ., Salem, WV 25426-0500)

Human listeners may form conscious representations of potential within-word structure in which lexicon is represented by some phonological units. An earlier study examining monolingual speakers of Japanese and English with native inputs suggests that levels of representation by Japanese speakers may be involved with richer knowledge of word-internal structure, while English speakers are sensitive to syllables [Otake *et al.*, Proceedings of EUROSPEECH 95 **3**, 1703–1706 (1995)]. The present study investigated how monolingual speakers of English learning Japanese could form conscious representations of potential within-word structure in Japanese. Three groups of subjects (*N*: 36, 33, and 40 for three different levels) were presented with 150 Japanese spoken words and asked to mark on a written transcript of each word the second natural division point from the onset in the word. The statistical analysis showed that all groups exploited syllables to represent Japanese words irrespective of Japanese proficiency. These results suggest that the exploitation of phonological unit to represent within-word structure in foreign inputs may be the same as the one in the native input as proposed by our recent pilot work (Otake and Yamamoto, 134th Meeting of the ASA **101**). [Work supported by Fulbright senior research grant and TAF.]

4aSC36. The discovery of natural classes by non-native listeners. John Kingston and Elliott Moreton (Linguist. Dept., South College 226, Univ. of Massachusetts, Amherst, MA 01003)

Using natural disyllabic stimuli, English speakers were trained to sort one, two, or three pairs of German nonlow rounded vowels into two categories, then a new pair was added and their ability to sort it into the same categories was tested. Each pair contrasted in [high], [tense], and [back]. Category membership was determined by only one of these features. Three-pair training provided enough information to tell which one; two-pair narrowed it down to two possibilities; one-pair left it open. Listeners in three experiments using different pairs of test vowels sorted the test vowels significantly better as the number of training pairs increased. This suggests they were inducing feature-based natural classes. The rival exemplar model of Nosofsky [J. Exp. Psych.: Gen. **115**, 39–57 (1986)] would say that while more pairs bring more phonetic variety, they also exemplify more robustly the dimensions along which the categories differ phonetically, and thereby shift attentional weight to those dimensions. Similar experiments will pit these two hypotheses against one another with training pairs which contrast only in the feature that defines the sort categories. If listeners are using feature-based natural classes, their performance should not improve with more training pairs; if they are using exemplars, it should. [Work supported by NIH.]

Session 4aSP

Signal Processing in Acoustics, Engineering Acoustics, Architectural Acoustics and Education in Acoustics: Acoustics in Multimedia—Systems Issues I

Daniel R. Raichel, Chair
Department of Mechanical Engineering, Steinman Hall, The City College of The City University of New York, New York, New York 10031

Invited Papers

8:00

4aSP1. *The Journal of the Acoustical Society of America* and other technical journals of the future—a new multimedia format? Daniel R. Raichel (Dept. of Mech. Eng., Steinman Hall, The City College of The City Univ. of New York, New York, NY 10031)

The *Journal of the Acoustical Society of America (JASA)* is now available in two formats—the traditional printed issue and the CD-ROM. However, CD-ROMs are in the early process of being displaced by the newer, more powerful DVDs which feature a capacity of approximately 2.8 GB per disk side rather than the 650 MB standard for each CD-ROM. CD-ROMs, however, will not immediately be rendered obsolete by the use of a current DVD player since the latter is backward compatible for CD-ROM playback. However, even in the CD-ROM format, a technical journal can take on new aspects of multimedia, which were hitherto unavailable or impracticable in printed media. The greatly enhanced capacity of the DVD disk provides even greater opportunities for extended expositions. Animation scenes, photographs, extended databases which would not have otherwise been included in a printed article, dynamic plots, sound effects, virtual reality scenes which permit one to ''visit'' laboratory sites or ''witness'' the topical phenomena, and even computer programs and entire reproductions of cited references can be incorporated in an article slated for publication on the new multimedia disks or on future storage devices.

8:25

4aSP2. Mediacoustics, a software package on CD for teaching acoustics using multimedia techniques. Michael Wahlrab (01dB, Inc., P.O. Box 796, Skaneateles, NY 13152)

Mediacoustics uses multimedia techniques on CD for teaching acoustics. It is divided into four study areas: basic physics, noise and man, room acoustics, and noise control. Sound clips, text, pictures, photos, and video animation illustrate each module. The index and chapters may be edited. Individual text passages may be annotated and an included acoustical dictionary provides definitions of common terms and formulas. Any contemporary multimedia computer equipped with a sound card and amplified loudspeaker can be used to access Mediacoustics and its hundreds of video clips.

8:50

4aSP3. Animations created in *Mathematica* for acoustics education. Victor W. Sparrow (Grad. Prog. in Acoust., Penn State Univ., University Park, PA 16802) and Daniel A. Russell (Kettering Univ., Flint, MI 48504)

One of the unifying assumptions in acoustics and vibrations is that waves and structures move and vibrate. In teaching these topics it makes sense to show students exactly HOW things move. For the last several years the authors have been using the symbolic manipulation program *Mathematica* [Wolfram Research, Inc., Champaign, IL] to produce brief animations for teaching acoustics. World Wide Web sites have been created for easy access to the animations by all students. The present paper will explain how the animations are created and will demonstrate several including one and two degrees of freedom oscillators, a piston in a tube, reflected waves, doppler effect, elastic waves, circular membranes, and circular plates. The current animations are relatively simple constructions to keep the file sizes to a minimum. However, improved animations containing substantially enhanced graphics will be possible with Internet 2.

Contributed Papers

9:15

4aSP4. A software teacher for acoustical measurements. Ila Tokola, Matti Karjalainen, and Martti Rahkila (Helsinki Univ. of Technol., Lab. of Acoust. and Audio Signal Processing, P.O. Box 3000, FIN-02015, Helsinki, Finland)

The huge development in computer technology has made it tempting to use computers in acoustics education. QuickSig, an object-oriented Common Lisp based signal processing environment has been used for acoustical measurements. It provides an excellent means to present graphics and handle signals, and it is easy to program for educational purposes as well. A Simple Loudspeaker Measurement Program is a Computer Based Education (CBE) application for self-studying of acoustical measurements. The application deals with free-field measurements of loudspeakers and related theory. The measurements are actually made with the program and genuine results are achieved each time. The software teacher interactively helps the student to obtain results that make sense and monitors the progress with multiple-choice questions. The results can also be

saved for post-processing. The main difference between this application and on-line help systems in measurement packages is that the teacher emphasizes learning of acoustical measurements, not the measurement software itself! On the other hand, with this application, the students do real measurements, including the loudspeaker and microphone setup in an anechoic chamber, unlike most self-study CBE applications. Furthermore, the application can be used also by engineers, technicians, etc. for assisted measurements.

9:35

4aSP5. Design and development of PC-IMAT. John W. Schuler (Naval Personnel Res. and Development Ctr., Armstrong Lab., Brooks AFB, TX 78235) and Murray S. Korman (U.S. Naval Acad., Annapolis, MD 21402)

NPRDC has been tasked to provide empirical evidence of effective instructional strategies for the acquisition of conceptual knowledge under the project name PC-IMAT (Personalized Curriculum for Interactive Multisensor Analysis Training). The domain to be investigated and demonstrated includes those concepts required for the successful planning and execution of antisubmarine warfare (ASW), specifically, the conceptual knowledge underlying the prediction of sound transmission paths and detection ranges. Navy-standard models which are used in fleet SONAR prediction systems are available to support the learners' conceptual understanding of the elements of the oceanographic environment which affect acoustic propagation. To date, these models are employed in a microcomputer based, stand-alone delivery architecture in both a linear interactive courseware (ICW) format and in modules suitable for independent exploratory learning. Midshipmen taking SP411 (Underwater Acoustics and Sonar) are currently using PC-IMAT to help investigate what are the effective instructional strategies which convey understanding of a complex multivariate domain (like ray trace or propagation loss models). Research on ''scientific visualization'' (to enhance comprehension and retention) and student feedback will also be used to help develop and evaluate other training materials including beamforming, reverberation, target motion analysis, and scattering. [Work supported by ONR.]

9:55

4aSP6. Construction of a HATS and its HRTF measurement for 3-D sound. Kyeong Ok Kang, Dong-Gyu Kang, Min-Soo Hahn (Elec. and Telecom. Res. Inst., P.O. Box 106, Yusong Post Office, Taejon, 305-600, Korea, kokang@audio.etri.re.kr), Mun-Jae Jho (Korea Res. Inst. of Standards and Sci., Taejon, Korea), and Dae-Kwon Jeong (Korea Aviation Univ., Koyang, Korea)

Based on the anthropometric data on the head and torso dimensions in Korean male adults, a head and torso simulator (HATS) was constructed. Data, which have no counterparts in the Korean standards, were used with ANSI S3.36-1985. Measuring microphones can be positioned at the eardrum or at the ear canal entrance according to measuring purposes. Head related transfer functions (HRTFs) for the HATS, i.e., binaural impulse responses, were measured at 710 points on a spherical surface of radius 1.55 m using a burst maximum length sequence (MLS) signal of 65 535 samples in an anechoic chamber. The measurement system consists of one part to generate the MLS signal and to drive a Boss 101A loudspeaker and another part to record the output signal of a microphone in the HATS and consequently to measure the impulse response. Also measured were the impulse responses of the driving loudspeaker and some headphones for 3-D sound reproduction in order to get the exact HRTF of the HATS-alone. The impulse-version HRTFs at the sampling frequency of 44.1 kHz, which have filter lengths of 512 points with minimum phase characteristics and can be used for 3-D sound, were finally obtained through a post-processing procedure.

10:15

4aSP7. Measuring and modeling the effect of source distance in head-related transfer functions. Jyri Huopaniemi and Klaus A. J. Riederer (Helsinki Univ. of Technol., Lab. of Acoust. and Audio Signal Processing, P.O. Box 3000, FIN-02015, Helsinki, Finland)

Efficient modeling of human spatial hearing by digital filter approximations of head-related transfer functions (HRTFs) is the key technology in 3-D sound processing. It is well known that the HRTF bears the major static localization cues, the interaural time difference (ITD), and the interaural level difference (ILD) that are functions of frequency and the incident angle of arrival. The effect of source distance has, however, often been neglected in HRTF models. In this paper, a method for efficient distance-dependent HRTF modeling is presented, which is based on both theoretical and empirical data. HRTF measurements on eight human subjects and one dummy head were carried out at two source distances, 2 and 0.65 m. It has been argued in the literature that the distance changes mainly affect the ILD, whereas the ITD remains approximately constant. Based on this finding, which was also supported by the measurements performed in this study, a filter structure that models the ILD change as a function of distance was derived. The results of this study are applicable to many near-field listening applications.

Session 4aUW

Underwater Acoustics: 3-D Propagation Effects I: Where are We Today in Models and Measurements?

Michael B. Porter, Cochair
Scripps Institute of Oceanography, Marine Physical Laboratory, MC 0205, 8602 La Jolla Shores Drive, La Jolla, California 92037-0205

Alexandra I. Tolstoy, Cochair
4224 Waialae Avenue, Suite 5-260, Honolulu, Hawaii 96816

Chair's Introduction—9:15

Invited Papers

9:20

4aUW1. The penetrable wedge as a three-dimensional benchmark. Grant B. Deane (Scripps Inst. of Oceanogr., La Jolla, CA 92003)

In 1989, members of the Acoustical Society of America developed a suite of range-dependent benchmark geometries, allowing the underwater acoustics community to compare the performance of analytical and numerical solutions to two-dimensional propagation problems. These benchmarks were successful in providing the acoustics community with a set of standards, and continue to be useful as reference geometries. Over the past decade, there have been significant advances in analytical and numerical solutions to three-dimensional propagation problems, and establishing three-dimensional benchmarks is a timely development. A useful benchmark must eliminate the undesired complications of a real-word environment but retain the essential physics of the problem while being tractable to a variety of modelers. A geometry which satisfies these constraints is the ocean wedge with a penetrable basement. The three-dimensional Green's function for the penetrable wedge has already been studied by a number of investigators, and measured with scale-model tank experiments. The state of modeling and measurement for the wedge will be discussed in the context of its usefulness as a three-dimensional benchmark.

9:40

4aUW2. Three-dimensional propagation effects: Modeling, observations, and suggested benchmark cases. Kevin B. Smith (Dept. of Phys., Naval Postgrad. School, Monterey, CA 93943, kevin@physics.nps.navy.mil)

Over the past several years, many acoustic propagation models have been adapted to compute the influence of azimuthal coupling. These so-called three-dimensional (3-D) models should provide more accurate predictions of the acoustic field in 3-D variable environments than previous 2-D$\times$N models. However, such advanced 3-D propagation models are generally slower than their 2-D $\times$N predecessors. It is, therefore, important to understand the significance of the difference between 3-D and 2-D$\times$N and when such effects should be considered. Furthermore, as of this date, no formal set of benchmark cases and solutions has been defined to test the accuracy of the various 3-D models currently being used. This paper will provide a general overview of 3-D propagation models and how they can be used to assess the significance of azimuthal coupling in ocean acoustics. Specific examples of such influences will be provided from a 3-D parabolic equation model and compared to results from the corresponding 2-D$\times$N version. Of particular interest will be the pseudo-3-D effects predicted by 2-D$\times$N calculations and the ability to distinguish true 3-D effects in experimental data. Finally, several environmental scenarios will be suggested as possible benchmark cases for future studies. [Work sponsored by ONR Code 321OA.]

10:00

4aUW3. Experimental measurements of three-dimensional underwater sound propagation over a variable bathymetry. Stewart A. L. Glegg, Joseph M. Riley, and Antony LaVigne (Ctr. for Acoust. and Vib., Dept. of Ocean Eng., Florida Atlantic Univ., Boca Raton, FL 33431)

There appear to be very few experimental measurements of the three-dimensional features of sound fields in the ocean, especially in shallow water where bottom interactions can have a dominant effect. One of the reasons for this lack of experimental data can be attributed to the difficulty in designing an experiment to identify three-dimensional propagation effects. Bathymetric refraction occurs, for the most part, on a very large scale and so to differentiate between local variability of the propagation and refraction by a bathymetric feature, measurements have to be carried out simultaneously over a large area. This is easier to achieve in the laboratory environment where repeatability of the experimental results is not an issue. This paper will describe a series of laboratory scale experiments which were carried out to identify the major features of underwater sound propagation over a realistic three-dimensional bathymetry. The model used is based on the bathymetry of the Santa Lucia Escarpment, but is not unlike bathymetry found in other coastal regions in different water depths if the correct scaling is used. A number of three-dimensional sound maps will be shown of the field from a point source. [Work supported by ONR.]

10:20

4aUW4. Three-dimensional sound propagation modeling of the Santa Lucia escarpment data. Michael B. Porter (Marine Physical Lab., Code 0205, 8602 La Jolla Shores Dr., La Jolla, CA 92037, porter@mpl.ucsd.edu)

The KRAKEN normal-mode program includes options for both $N\times 2$-D and full 3-D modeling. The latter is implemented using Gaussian beams to treat the horizontal refraction of modes. Since the model allows for mixed acoustoelastic problems, 3-D sound propagation can be simulated over ocean bottoms supporting shear waves. To study its accuracy, the model is applied to several 3-D scenarios, including both a simple wedge and a topographically complicated case based on the Santa Lucia escarpment off the coast of California. The results are compared to the laboratory scale data presented by Glegg *et al.* in this session.

10:35–10:40 Break

10:40

4aUW5. Benchmarking two three-dimensional parabolic equation methods. Frederic B. Sturm (Lab. M.N.C., I.S.I.T.V, Université de Toulon, France), John A. Fawcett, and Finn B. Jensen (SACLANT Undersea Res. Ctr., 19138 La Spezia, Italy)

Results of acoustic propagation modeling using two different three-dimensional parabolic equation methods are presented. One model [J. Fawcett, J. Acoust. Soc. Am. **93**, 2627–2632 (1993)] uses standard splitting and operator approximations while the other method [F. Sturm, M. Pélissier, and D. Fattaccioli, *Proc. of the Third European Conference on Underwater Acoustics*, 243–248 (1996)] does not split the azimuthal operator and maps the bathymetry into an equivalent flat bottom problem. Examples of propagation over wedge and corrugated bathymetries are considered for different frequencies. Both point-source and modal initial fields are used. The three-dimensional solutions exhibit interesting three-dimensional refractive effects. The convergence of the solutions with respect to azimuthal discretization is also discussed.

10:55

4aUW6. Parabolic equation modeling with the split-step Fourier algorithm in four dimensions. F. D. Tappert (Appl. Marine Phys., Univ. of Miami, RSMAS, Miami, FL 33149, ftappert@rsmas.miami.edu)

Fully range-dependent 3-D (with coupled azimuths) single-frequency PE/SSF models have existed for many years, and 2-D (vertical plane at fixed bearing) wide-angle broadband PE/SSF models have existed for even longer. By combining these, a 4-D (three space dimensions plus time) PE/SSF model has been developed that runs on a desktop computer. This model is exactly reciprocal if ocean currents are omitted, and is exactly energy conserving if absorption is omitted. Examples of model predictions, projected onto two dimensions, are displayed for shallow water propagation. In a situation where ray chaos due to variable bathymetry is strong, plots of intensity as a function of travel time and bearing at fixed depth and range are especially interesting. [Work supported by ONR.]

11:10

4aUW7. Three-dimensional ocean acoustics: Techniques and examples. Michael D. Collins, Joseph F. Lingevitch, and Gregory J. Orris (Naval Res. Lab., Washington, DC 20375, collins@ram.nrl.navy.mil)

The method of matched asymptotics can be used to solve scattering problems involving a compact object in a waveguide [M. D. Collins and M. F. Werby, J. Acoust. Soc. Am. **85**, 1895–1902 (1989)]. This approach involves different asymptotic limits in different regions. In the inner region near the object, the Green's function can be approximated by either the free-space or half-space Green's function depending on the vertical location of the object. The inner problem is therefore the free-space or half-space scattering problem. The outer solution satisfies the three-dimensional parabolic equation. When horizontal variations in the waveguide are sufficiently gradual, the solution is asymptotic to a specular point source for $ka=O(1)$. This approximation can be improved by including a vertical dipole correction. The three-dimensional parabolic equation has been applied to free-space scattering problems by solving the exterior problem for the scattered field directly [M. F. Levy and A. A. Zaporozhets, J. Acoust. Soc. Am. **103**, 735–741 (1998)]. This approach is also useful for some waveguide scattering problems. [Work supported by ONR.]

11:25

4aUW8. Limitations on adiabatic normal modes in range-varying environments. Peter C. Mignerey (Acoust. Div. 7120, Naval Res. Lab., Washington, DC 20375-5350, mignerey@nrl.navy.mil)

Curvilinear coordinates are used to extend adiabatic normal modes to weak range-dependent environments including those with sloping bottoms. Reciprocal vertical wave numbers of a reference mode provide the coordinate scale factors. This approach is valid for negligible horizontal derivatives of intermodal vertical-wave number ratios. Large gradients in the bottom impedance make these ratios large, implying significant mode coupling in oceanic regions spanning both soft and hard terrains. Additionally, a large number of eigenvalues are needed to compute adiabatic normal-mode solutions for high-resolution acoustic fields. The direct solution of eigenvalue problems on a fine grid is often prohibitive. Alternatively, the eigenvalues may be interpolated to the fine grid. This approach usually requires extrapolation toward edges across which the horizontal wave number becomes imaginary at mode cutoff. One choice for an interpolation function is the integrated reciprocal vertical wave number. This function, proportional to water depth, is smooth and real throughout the cutoff region, while causing the horizontal wave number to vanish along the cutoff line. [Work supported by the Office of Naval Research and partially by DOD High Performance Computing facilities at NRL.]

11:40

4aUW9. Three-dimensional acoustic effects due to ocean currents. Oleg A. Godin (School of Earth and Ocean Sci., Univ. of Victoria, P.O. Box 1700, Victoria, BC V8W 2Y2, Canada, ogodin@uvic.ca)

In the presence of currents, vertically stratified fluids become acoustically anisotropic in the horizontal plane and the group and phase velocities of the acoustic normal mode are generally not parallel. In this paper, the resulting 3-D effects and their implications for acoustic monitoring of ocean currents are discussed in terms of the normal modes. The sensitivity of various acoustic quantities to the vertical and in-plane and out-of-plane horizontal components of the flow velocity is estimated. Despite anisotropy, there is no azimuth coupling at long-range sound propagation in a stratified moving medium. When a waveguide varies in the horizontal plane, the 3-D effects due to current-induced anisotropy and horizontal refraction are superimposed. The relative importance of these two phenomena in shallow- and deep-water scenarios is analyzed by considering horizontal (modal) rays in the presence of currents. For the purposes of numerical simulations, a parabolic approximation is considered for sound in a moving fluid. A new parabolic equation is proposed for efficient numerical modeling of the 3-D acoustic effects in the ocean with currents. [Work supported by NSERC.]

11:55

4aUW10. Three-dimensional solutions to range-dependent problems in shallow water by a pseudo-spectral method. Altan Turgut and Stephen N. Wolf (Naval Res. Lab., Acoust. Div., Washington, DC 20375)

A pseudo-spectral method is used to solve the acoustic and elastic wave equations in the time domain. The application of such a computation and memory-intensive method to shallow-water propagation and scattering problems is discussed. Three-dimensional solutions are given for two

different shallow-water related problems: pulse propagation through solitary internal waves and pulse scattering from bottom inhomogeneities. The size of a typical computational domain for both problems is about (30 wavelength)^3. However, for the propagation problem, distances up to several hundred wavelengths have been achieved by moving the computational domain with the wavefront. The accuracy of the pseudo-spectral method has been checked with analytical solutions to simpler problems. Two-dimensional results of moving-domain approach are compared to those of single-domain and good agreement is observed. The existing method is expected to confirm the accuracy of the approximate solutions such as PE in a complicated range-dependent environment. [Work supported by ONR.]

THURSDAY MORNING, 25 JUNE 1998 GRAND BALLROOM III (W), 11:00 A.M. TO 12:00 NOON

Session 4aPLb

Plenary Lecture

Donna L. Neff, Chair
Boys Town National Research Hospital, 55 North 30th Street, Omaha, Nebraska 68131

Chair's Introduction—11:00

Invited Paper

11:05

4aPLb1. Psychoacoustics of cochlear hearing impairment and the design of hearing aids. Brian C. J. Moore (Dept. of Experimental Psych., Univ. of Cambridge, Downing St., Cambridge CB2 3EB, England, bcjm@cus.cam.ac.uk)

Cochlear hearing impairment is usually associated with damage to the hair cells within the cochlea. When the damage is restricted to the outer hair cells (OHCs), the main consequence is disruption of the ''active'' mechanism which normally enhances the response of the basilar membrane to weak sounds and which sharpens the tuning (frequency selectivity) of the basilar membrane. Psychoacoustically, damage to OHCs results in loss of sensitivity (elevated absolute thresholds), loudness recruitment, and reduced frequency selectivity. Damage to the inner hair cells (IHCs) causes basilar membrane vibrations to be transduced less effectively, so absolute thresholds are elevated, but does not result in altered frequency selectivity or loudness recruitment. Sometimes, IHCs and/or neurons may be completely inoperative at certain places within the cochlea, giving rise to ''dead regions.'' Such regions can strongly influence the perception of pitch and loudness. Current hearing aids can partially compensate for the effects of loudness recruitment by using compression amplification, but there is much controversy about the ''best'' form of compression. The deleterious effects of reduced frequency selectivity on speech intelligibility in noise can be alleviated by various methods for improving the speech-to-noise ratio, although so far only directional microphones have given clear benefits.

Meeting of the Standards Committee Plenary Group

ORGANIZATION OF STANDARDS COMMITTEE PLENARY GROUP MEETING

S1 Acoustics—U.S. Technical Advisory Group (TAG) for IEC/TC 29 Electroacoustics and ISO/TC 43 Acoustics

S3 Bioacoustics—U.S. Technical Advisory Group (TAG) for ISO/TC 43 Acoustics, IEC/TC 29 Electroacoustics, and ISO/TC 108/SC4 Human Exposure to Mechanical Vibration and Shock

S12 Noise—U.S. Technical Advisory Group (TAG) for ISO/TC 43/SC1 Noise and ISO/TC 94/SC12 Hearing Protection

The meeting of the Standards Committee Plenary Group will precede the meetings of the Accredited Standards Committees S1, S3, and S12, to take place in the following sequence on the same day: S12—9:30 a.m. to 11:30 a.m., St. Helens Room; S3—2:00 p.m. to 3:15 p.m., St. Helens Room; S1—3:30 p.m. to 4:45 p.m., St. Helens Room.

Discussion at the Standards Committee Plenary Group meeting will consist of national items relevant to all S Committees, plus a review of the international standardization (U.S. TAG) activities including reports on recent meetings and planning for forthcoming meetings.

Members of S2 on Mechanical Vibration and Shock (and U.S. TAG for ISO/TC 108 and five of its Subcommittees, SC1, SC2, SC3, SC5, and SC6) are also encouraged to attend the Standards Committee Plenary Group meeting, even though the S2 meeting will take place one day earlier, on Wednesday, 24 June 1998 at 9:00 a.m.

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.S. TAG Chair/Vice Chair	TC or SC	U.S. TAG
ISO		
P. D. Schomer, Chair H. E. von Gierke, Vice Chair	ISO/TC 43 Acoustics	S1 and S3
P. D. Schomer, Chair H. E. von Gierke, Vice Chair	ISO/TC 43/SC1 Noise	S12
E. H. Berger, Chair	ISO/TC 94/SC12 Hearing Protection	S12
J. Erdreich, Chair H. E. von Gierke, Vice Chair	ISO/TC 108/SC4 Human Exposure to Mechanical Vibration and Shock	S3
D. J. Evans, Chair	ISO/TC 108 Mechanical Vibration and Shock	S2
R. H. Mehta, Chair K. Won, Vice Chair	ISO/TC 108/SC1 Balancing, including Balancing Machines	S2
A. F. Kilcullen, Chair	ISO/TC 108/SC2 Measurement and Evaluation of Mechanical Vibration and Shock as Applied to Machines, Vehicles and Structures	S2
B. E. Douglas, Chair	ISO/TC 108/SC3 Use and Calibration of Vibration and Shock Measuring Instruments	S2
D. J. Vendittis, Chair	ISO/TC 108/SC5 Condition Monitoring and Diagnostics of Machines	S2
G. Booth, Chair	ISO/TC 108/SC6 Vibration and Shock Generating Equipment	S2
IEC		
V. Nedzelnitsky, U. S. TA	IEC/TC 29 Electroacoustics	S1 and S3

Meeting of Accredited Standards Committee S12 on Noise

P. D. Schomer, Chair S12, and Chair U. S. Technical Advisory Group (TAG) for ISO/TC 43/SC1, Noise
U. S. CERL, P.O. Box 9005, Champaign, Illinois 61826-9005

B. M. Brooks, Vice Chair, S12
Brooks Acoustics Corporation, P.O. Box 3322, Vernon, Connecticut 06066

H. E. von Gierke, Vice Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 43/SC1, Noise
1325 Meadow Lane, Yellow Springs, Ohio 45387

E. H. Berger, Chair, U. S. Technical Advisory Group (TAG) for ISO/TC 94/SC12, Hearing Protection
E-A-R/Aearo Company, 7911 Zionsville Road, Indianapolis, Indiana 46268-1657

Accredited Standards Committee S12 on Noise. Working group chairs will report on their progress for the production of noise standards.

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 4pAAa

Architectural Acoustics: Case Study on the New Tokyo International Forum

J. Christopher Jaffe, Chair
Jaffe Holden Scarbrough Acoustics, Inc., 114A Washington Street, Norwalk, Connecticut 06854

Chair's Introduction—1:00

Invited Papers

1:15

4pAAa1. Role of the acoustic advisors in the design of the Tokyo International Forum. Teruji Yamamoto (Toho Univ., Funabashi, Japan), Hideki Tachibana (Univ. of Tokyo, Japan), and Christopher Jaffe (Jaffe Holden Scarbrough Acoust., Inc., 114A Washington St., Norwalk, CT 06854)

The Tokyo Municipal Government appointed two acoustical advisors to review the acoustic design criteria developed for the Tokyo International Forum, Dr. Teruji Yamamoto and Dr. Hideki Tachibana. This paper describes the role of acoustic advisors on such a large project and show how they were able to assist the design team throughout the design and construction of the complex. Items to be covered include: process of design competition, acoustic requirements, process of acoustical design, outline of construction; and acoustical check during and after construction.

1:45

4pAAa2. Room acoustic designs for the Tokyo International Forum. Christopher Jaffe, Mark Holden, Russell Cooper (Jaffe Holden Scarbrough Acoust., Inc., 114 A Washington St., Norwalk, CT 06854, dcoppola@jhsacoustics.com), T. Kobayashi, and F. Kawakami (Yamaha Acoust. Res. Labs., Hamamatsu, Japan)

The room acoustic design of the Tokyo International Forum centered around the four performance halls: the convention center, the atrium, and the large multi-use meeting/dining halls. This paper will describe the program use of each space and the acoustic criteria developed to meet these requirements. Major emphasis will be placed on Hall A, a 5000-seat congress hall that was to be used to unamplified symphonic concerts and operas as well as highly amplified Broadway musicals and rock concerts, and Hall B, a 1500-seat classical music hall and piccolo opera house that also would be used for conferences and popular entertainment. A unique architectural

feature of Hall B is the first design and installation of a Concert Hall Shaper, a device that enables a proscenium theatre with a stagehouse to be converted to a single acoustic volume for events such as symphony concerts requiring a longer reverberation time. Hall C, an extremely flexible space for fashion shows, summit meetings, and large receptions, and Hall D, another versatile space for theater, recitals, and live television productions, will also be described.

2:15–2:25 Break

2:25

4pAAa3. Sound and vibration control for the Tokyo International Forum. Fukushi Kawakami, Tetsu Kobayashi (Yamaha Acoust. Res. Labs., Hamamatsu, Japan), Christopher Jaffe, Mark Holden, and Russell Cooper (Jaffe Holden Scarbrough Acoust., Inc., Norwalk, CT 06854)

The acousticians for the Tokyo International Forum were faced with a number of very severe noise and vibration problems related to the siting of the complex in the center of downtown Tokyo, the placement of the performance halls next to one another, and the location of a portion of the mechanical equipment on the roof. In addition, subway lines ran parallel to three of the boundaries of the structure, another line ran under the building, and the main tracks of Japan Rail ran above ground alongside the major atrium. This paper will discuss the noise criteria set for the performance rooms and other noise sensitive spaces, the noise and vibration measurements taken to determine the severity of the train-related conditions, predictions of other noise and vibration sources, and the solutions developed by the design team to meet the criteria required by the program. Among the solutions are a secant wall surrounding the foundation, box in box construction for the four performance halls, blocking mass installations at key locations throughout the steel skeleton of the complex, and isolation joints and sound locks between the performance rooms.

2:55

4pAAa4. Design of sound reinforcement and acoustic field control systems for the Tokyo International Forum. Christopher Jaffe, Mark Holden, Paul Scarbrough, David Robb (Jaffe Holden Scarbrough Acoust., Inc., 114A Washington St., Norwalk, CT 06854, dcoppola@jhsacoustics.com), Yasushi Shimizu, Shinjirou Yamashita, and Fukushi Kawakami (Yamaha Acoust. Res. Labs., Hamamatsu, Japan)

The extremely varied program use of the four performance halls at the Tokyo International Forum provided the electroacoustic designers with a number of extremely challenging problems. This paper describes the events that were to take place in all of the rooms, the designs selected to accommodate these programs, and the coordination of the design of the electroacoustic systems with the room acousticians, the architects, and the theater consultants. Of special interest is the design of the sound systems for Hall A, the 5000 seat symphony/congress hall which was to be used for opera and symphony performances without employing close-in microphone amplification techniques, and Hall D, which was to be used for live theater and recitals as well as television productions. The final portion of the paper will discuss the sound contractor's role in completing such an immense project and in particular the relationship that contractors have with designers and owners in Japan.

3:25–3:55 Discussion

Session 4pAAb

Architectural Acoustics: Performance Spaces (Poster Session)

Michael R. Yantis, Cochair
Michael R. Yantis Associates, Inc., 1809 7th Avenue, Suite 1609, Seattle, Washington 98101

Basel Jurdy, Cochair
Michael R. Yantis Associates, Inc., 1809 7th Avenue, Suite 1609, Seattle, Washington 98101

Elizabeth Bozomolov, Cochair
Michael R. Yantis Associates, Inc., 1809 7th Avenue, Suite 1609, Seattle, Washington 98101

Contributed Papers

All posters will be on display from 1:00 p.m. to 5:30 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 3:15 p.m. and contributors of even-numbered papers will be at their posters from 3:15 p.m. to 5:30 p.m.

4pAAb1. Early sound distribution in auditorium. Jiqing Wang and Guorong Jiang (Inst. Acoust., Tongji Univ., 1239 Sipin Rd., Shanghai 200092, PROC)

It is known that the strength of the audience area of a hall is largely dependent on the early sound, i.e., the direct sound and the first reflections. They are varied not only with reverberation time, but also with the geometry of the room, the absorption arrangement of the room, and the location of the source and(or) the receiver. Results from digital simulation and field measurements in halls show that the attenuation rate of early sound with the distance between the source and the receiver is much larger than ordinary predictions. It is more rational in complying with the subjective judgment. Therefore it should not be ignored in the acoustical design of a hall.

4pAAb2. The late reverberation time as a new criteria for the evaluation of hall acoustics. V. J. Stauskis (Vilnius Gediminas Tech. Univ., Saulėtekio al. 11, 2040 LT Vilnius, Lithuania)

Investigations show that the field decay after-35 dB has its reverberation values, which differ considerably from the values of the standard reverberation time. The experiment was carried out in a rectangular hall measuring $13.6\times10.7\times7.0$ m. The sound absorption coefficients of all planes in the hall were small (0.02–0.05). One hundred seventy semi-upholstered chairs were used for the experiment. When the chairs were placed in the hall, the decrease in the field of a nonfiltered signal was sudden and pronounced, while in the interval from 1500–4000 ms it was slow. The experiments with the filtered signal showed that, at any frequencies, there existed both the initial sudden and the late slow decrease in the sound field energy. The higher the frequency, the longer the interval of the slow decrease in the sound field. When approximating the decrease from -35 to -40 dB, the reverberation time was 15 s at 250 Hz and 4000 Hz. When approximating the decrease from -45 to -50 dB, the reverberation time was as long as 20–21 s within the frequency range of 500 to -2000 Hz.

4pAAb3. Study on acoustic index variations due to small changes in an observation point. Katsuaki Sekiguchi (College of Sci. and Technol., Nihon Univ., Suragadai, Kanda, Chiyoda-ku, Tokyo, 101 Japan, sekiguch@arch.cst.nihon-u.ac.jp) and Toshiki Hanyu (Junior College of Nihon Univ., Narashinodai, Funabashi, Chiba, 274 Japan)

Acoustic index variations due to small changes in an observation point in a concert hall were studied. Three kinds of methods were used in order to measure the acoustic indexes: (1) one point measurement using a single channel microphone; (2) binaural measurement using a dummy head microphone; and (3) the regular tetrahedron peak point method using a 4-channel microphone system. The acoustic index variations in each of the methods were analyzed. In addition, subjective experiments were performed in order to evaluate the effects of the variations. Compared to these measurements, problems with the acoustic indexes which were analyzed by each of the methods were found. One point measurement may incorrectly evaluate a concert hall.

4pAAb4. Effects of the spatial information of a sound field on listener envelopment. Toshiki Hanyu (Junior College of Nihon Univ., Narashinodai, Funabashi, Chiba, 274 Japan, hanyu@arch.jcn.nihon-u.ac.jp) and Sho Kimura (Nihon Univ., Surugadai, Kanda, Chiyoda-ku, Tokyo, 101 Japan)

The spatial information of sound fields in various kinds of concert halls as virtual sound-source distributions and short-term directional diffusion factors were measured. It is clear that the spatial information in a concert hall is characterized by how the reflections surround a listener. Accordingly, the relationship between the spatial information and listener envelopment was studied. First, subjective experiments were performed using a simulated sound field in an anechoic room in order to study the effects of lateral energy ratios on listener envelopment. As a result, the lateral energy ratios were not sufficient to evaluate listener envelopment. The research demonstrates that listener envelopment is not only related to

late arriving reflections but also to early reflections. In addition, listener envelopment is related to the spatial distributions of the reflections that surround a listener. Furthermore, the implications for concert hall design were discussed.

4pAAb5. An evaluation method of concert hall acoustics. Y. Hirasawa (Faith, Inc., Form-Gokomachi 583, Gokomachi-Ebisugawa, Kyoto, 604 Japan) and Z. Maekawa (Environ. Acoust. Lab., Oyodonaka, 4-13-11, Osaka, 531 Japan, nd7j-mekw@asahi-net.or.jp)

A lot of objective physical parameters for acoustic evaluation of a concert hall have been proposed. However, no one can believe that they provide a perfect method for designing a concert hall. There is a problem with the measuring method of physical parameters, i.e., almost all acoustical measurements in a hall use an omni-directional loudspeaker, which is not similar to most musical instruments. Also, no parameter specifies tone-color or sonority of music performance. These are important factors in evaluating concert hall acoustics. The remaining problems must be considered as follows: (1) The position of sources on a stage changes tone-color and generating sound power. (2) The directivity of sources gives different effects on room acoustics at the same position on the same stage. (3) Each musical instrument has its own important frequency range. What instrument shall be selected properly and what execution of the instrument shall be adopted for evaluation? (4) What is a suitable program of music for evaluation of hall acoustics? Some examples of the problems upon sources and its performance with using real musical instruments, and the remaining problems for evaluations of the concert hall will be presented.

4pAAb6. Acoustical behavior of churches: Gothic-Mudejar churches. Juan J. Sendra, Teófilo Zamarreño, and J. Navarro (School of Architecture of Seville, Avda Reina Mercedes s/n, 41012 Sevilla, Spain)

Traditionally, a particularly good acoustic behavior has been assigned to Christian churches. But this statement, which is not true at all, has led to important errors in church rehabilitation works, more so in those that implicate changes of use, such as auditoriums or theatrical halls, which is the case of many churches in Spain. This research group has worked in acoustical analysis for more than ten years, and has participated in many cases of church rehabilitation, among which is acoustical behavior. We are talking about a type of church very common in the south of Spain, Gothic-Mudejar churches. They are churches with these common characteristics: relatively little volume, plants with three naves, and wooden ceilings. With this work, significant conclusions drawn from the acoustical analysis of ten churches of Seville that respond to this type are presented. Their characteristics are quite homogeneous, so it is possible to get conclusions about the acoustical conditions of this ecclesiastical type.

4pAAb7. Acoustical design of Sumida Triphony Hall. Toshiko Fukuchi and Hideo Nakamura (Nagata Acoust., Inc., Minami-Shinjuku-Hoshino Bldg., 8F, 5-23-13 Sendagaya Shibuya-ku, Tokyo, 151-0051 Japan, fukuchi@nagata.co.jp)

The Sumida Triphony Hall was opened near Kinshityo JR station, located in the eastern part of Tokyo, in October 1997. The building contains two concert halls, the Main Hall, with 1801 seats, as the home of the New Japan Philharmonic, and the Recital Hall, with 252 seats. The Main Hall has a rectangular shape with the sloped ceiling rising at 12.5 degrees toward the back of the hall in parallel to the main floor and with two balconies around the main audience area. The reflecting panels were hung over the stage to help the ensemble for players. The shape was basically checked with the technique of computer simulation and the detail of the ceiling and the wall was studied by use of a 1/10th-scale model. An organ with 66 stops was installed at the front of the hall. The Recital Hall is a typical ''shoebox'' hall. In order to attenuate the vibration caused by railroad tracks, the vibration isolating barriers were embedded around a part of the outer walls of the building, and also the vibration insulating structure was adopted at the Main Hall partly and the Recital Hall entirely. The outline of acoustical design will be reported.

4pAAb8. Subjective assessment of the uneven distribution of reflections within a concert hall. Eugenio Collados (Dept. of Phys., Univ. of Santiago, Santiago, Chile, ecollado@lauca.usach.cl)

Most physical variables other than RT are known to vary significantly within a concert hall. Namely, early to late energy ratio, lateral energy fraction, and sound strength. Less well known are the subjective correlates to the above variations. A series of experiments were carried out in order to assess the number and meaning of subjective dimensions throughout a single hall. Testing environments were both an anechoic chamber array and an actual hall with artificial sources, where subjects were asked to judge classical music samples. In both cases the standard spatial distribution was modified by the addition of discrete reflections while keeping level and RT constant. At least four subjective dimensions were found, as many as found in hall-to-hall comparisons. Up to 59% of variance was related to the spatial image of the source. After comparing results between the anechoic chamber and the actual hall, the visual impression seems to partially compensate for the variations due to distance. These results raise some doubts on the validity of listening tests performed without a visual stimulus. It is also suggested that these effects should be considered in designing halls with active systems.

4pAAb9. A neural network analysis of concert hall acoustics. Fergus R. Fricke and Young G. Han (Architectural and Design Sci. Dept., Univ. of Sydney, Sydney, NSW 2006, Australia)

A neural network analysis using data from 51 concert halls was undertaken. The analysis related the acoustic quality of halls, as judged by musicians, to ten hall parameters: volume, surface area, number of seats, length, width, height, mean rake angle of seats, a surface diffusion index, stage height, and extent of stage shell/enclosure. The surface diffusion index and the extent of the stage shell were determined by a group of architects who made judgments on the basis of photographs and drawings of the halls. The results of the analysis are tentative and difficult to generalize, as there are so many inputs and so many possible combinations of parameters, and the effect of a particular factor is, in some cases, neither linear nor monotonic. The results are applied to a particular hall, the concert hall of the Sydney Opera House, where changes are being contemplated. There are some unexpected results which, at this stage, give food for thought rather than the basis for action, as there are obvious limitations to this work, such as extraneous factors influencing judgments of acoustic quality, surface diffusion estimation, and extent of the stage shell and the limited number inputs used to describe the halls.

4pAAb10. Directional dependence of the change of auditory source width by very short time-delay reflections. Masayuki Morimoto and Mariko Watanabe (Environ. Acoust. Lab., Faculty of Eng., Kobe Univ., Nada, Kobe, 657 Japan, mrmt@kobe-u.ac.jp)

Barron [J. Sound Vib. **15**, 475–494 (1971)] showed that spatial impression (SI) corresponding to auditory source width (ASW) decreases when the time delay of reflections relative to the direct sound is shorter than 10 ms. This paper investigates the effect of the direction of lateral reflections on the decrease of ASW. The music motif used in this investigation was a 6.5-s section of the 4th movement of Mozart's Jupiter Symphony (No. 41), which was the same motif as Barron used. The sound fields used as stimuli consisted of a direct sound and two discrete reflections with a 1-ms interval between them. The time-delay of the first reflection relative to the direct sound was 2, 4, or 8 ms. The directions of the

reflections were ±18°, ±45°, or ±72°. The experimental results demonstrate that the decrease of ASW depends on the direction of the reflections. As the reflections arrive from a more frontal direction, ASW decreases more sharply relative to the delay time.

4pAAb11. Reverberation time in Serbian Orthodox worship spaces. Miomir Mijic (Faculty of Elec. Eng., Bul. Revolvcije 73, Belgrade, Serbia)

Seven years ago, at the ASA Meeting in November 1990, this author presented some preliminary results of acoustical research realized in the Serbian Orthodox worship spaces. At that time the results for only seven churches were included in the paper. This research has been continued until now and nearly 60 churches were analyzed, different in volume, age, and the architectural style. Such a large number of measurements allow some more general conclusions about acoustical characteristics of orthodox worship spaces. Reverberation times obtained in realized measurements show the important influence of several specific factors such as strong orthodox rules in church interior design and volume shapes. These factors are influenced by traditional architectural styles accepted as unalterable even in the contemporary architectural design. Some differences in the frequency characteristic of the reverberation time according to different historical periods were also discovered.

4pAAb12. An experiment to identify preference groups among concert hall listeners. M. Miklin Halstead, Johan L. Nielsen, and A. Harold Marshall (Acoust. Res. Ctr., The Univ. of Auckland, Private Bag 92019, Auckland, New Zealand, mm.halstead@auckland.ac.nz)

The task of specifying the acoustical behavior of a concert hall depends upon knowing the preferred subjective acoustical characteristics for the intended program material. Several studies have suggested that the listening population might be divided into groups with respect to their preferred acoustic, with different weightings applied to different subjective factors in each. Our project endeavors to identify and describe these groups, and to make their composition an integral part of hall design. This paper describes a subjective experiment designed to identify and measure individual listening preferences in a simulated concert hall setting, and presents the experiments first results. Following from these results, an attempt is made to place tested individuals into preference groups. The applicability of the experiment is considered in terms of the quality of hall simulation and of the impact of the simulator and experimental design on the listeners discrimination.

4pAAb13. On spatial variability of room acoustics measures. Johan L. Nielsen, M. Miklin Halstead, and A. Harold Marshall (Acoust. Res. Ctr., The Univ. of Auckland, New Zealand, j.nielsen@auckland.ac.nz)

The use of high-energy measurement signals combined with digital multichannel recording and postprocessing leads to a highly efficient collection of impulse responses in concert halls. During measurements in two similar concert halls, the Christchurch Town Hall, New Zealand, and the Hong Kong Cultural Centre (both described in companion papers), responses were collected in all seats in some areas and in different positions in some single seats, in addition to the usual sampling of all areas. An omnidirectional loudspeaker source was used, and the measurements were repeated in several source positions. The distribution of various parameters in these areas are compared to the results for the whole hall. Based on this comparison the use of mean and standard deviation as measures of location and spread for the parameters is discussed, and the variation of parameters with varying source locations is demonstrated. Approaches toward the reduction of parameter sensitivity to insignificant details in the early sound field are briefly discussed and illustrated using the measured data.

4pAAb14. Acoustical design of the Queensland Conservatorium of Music. Keiji Oguchi, Yasuhisa Toyota, and Minoru Nagata (Nagata Acoust., Inc., Minami-Shinjuku-Hoshino Bldg., 8F, 5-23-13 Sendagaya Shibuya-ku, Tokyo, 151-0051 Japan, oguchi@nagata.co.jp)

The auditorium of the Queensland Conservatorium of Music was planned for both classical music concerts and lyric performances in the newly constructed conservatorium in the Southbank of Brisbane, Australia. The retractable orchestra shell is moved forward to the proscenium on the stage, and forms a shoebox style concert space with 643 seats. Small doors are installed on the orchestra shell for acoustical fine tuning in the stage space. Absorptive materials appear when the doors are opened. In lyric performances, the orchestra shell is retracted backward. The frontal area of the stage floor is sunk as an orchestra pit. The reverberation time becomes short, from 2.0 to 1.6 s, with retracting the orchestra shell, and varies by 0.3 s with acoustic curtains on the audience side walls for both concert and lyric use. The acoustical design and characteristics of the auditorium and players' preference on the acoustical tuning elements are reported in this paper.

4pAAb15. Estimation and analysis of acoustic parameters of ancient churches for concert performances. Tiziana Ottonello, Enrico Dassori (IATU—Univ. of Genoa, Via Opera Pia, 16145 Genova, Italy), and Andrea Trucco (DIBE—Univ. of Genoa, Via Opera Pia 11A, 16145 Genova, Italy)

This paper presents a study on the propagation of musical sound, where occasionally concert performances are held. The five selected churches were built in Genoa between the eleventh and the seventeenth century. The analysis was aimed at the computing of acoustic parameters (e.g., reverberation time, early decay time, initial decay gap, speech trasmission index, definition index), based on structural modeling and computer simulation (RAMSETE software) as well as experimental texts. Analysis of results showed that romanic churches generally provide better acoustic characteristics. Furthermore, acoustic not standing modifications (e.g., absorbing panels, curtains) were designed in order to make the observed acoustic parameters comparable with those observed in a typical concert hall. This work was developed in the framework of a project aimed at creating new music halls using the existing old buildings, generally dedicated to other aims.

4pAAb16. Acoustic behavior of a tense-structure. Alessandro Cocchi and Lamberto Tronchin (DIENCA–CIARM Viale Risorgimento, 2, 40136 Bologna, Italy, tronchin@ciarm.ing.unibo.it)

During the design of a concert hall, the acoustical behavior of the sound field is often investigated by using numerical models. In the case of a restoration of the hall, or during the design of the orchestra chamber, acoustic measurements of impulse responses in the hall give precious information on the sound fields, e.g., RT, clarity, and so on. By using numerical simulation, it is then possible to verify the efficiency of the acoustic solution. Much commercial software is available, all of them following the geometric assumptions, as the Eikonal equation, while the diffraction and diffusion of sound are not usually taken into account, and all of them require the absorption coefficient of the walls. The tense-structure, on the other hand, is not so easy to model, owing to the peculiar characteristics, i.e., modal behavior and impossibility to measure the absorption of the membrane. In this paper, the design of an acoustic intervention in a tense-structure is shown. A procedure to find out the ''equivalent absorption'' of the membrane is shown. The results of such a model are compared with the experimental measures.

4pAAb17. Variations in the surface textures of choral reflectors. Elizabeth J. Lee and Campbell J. Yule (Acoust. Res. Ctr., Univ. of Auckland, New Zealand, elee@ccul.auckland.ac.nz)

The commission to design a portable choral reflector seemed a simple process of finding a balance between acoustical and structural requirements. Initial inquiries, however, found a lack of information on the acoustical effect of different surface textures normally used. An experiment was developed to assess both the subjective and objective performance of varying degrees of surface roughness. Eight possible surfaces were chosen and applied to standardized panels. The impulse response of each panel was measured in an anechoic room using a maximum length sequence signal. Early results indicate significant differences in the reflected energy and frequency response of the respective panels. An attempt to obtain a subjective evaluation of the reflections by convolving choral and solo voice recordings with the impulse response of the panels will be reported. A prototype choral reflector based on this work will be described. [The support of the Acoustics Research Centre for this project is acknowledged.]

4pAAb18. Acoustical design of Harmony Hall Fukui. Akira Ono, Suzuyo Yokose, and Katsuji Naniwa (Nagata Acoust., Inc., Minami-Shinjuku-Hoshino Bldg., 8F, 5-23-13 Sendagaya Shibuya-ku, Tokyo, 151-0051 Japan, ono@nagata.co.jp)

Construction of Harmony Hall Fukui, which was sponsored by the Fukui prefecture government, was completed on September 3, 1997. The facility contained a large hall, a small hall, and administrative offices. In the central area of the facility is an administrative room, rehearsal rooms, and machinery space. The two halls are situated at the east and west ends of the facility. This architectural zoning brings advantages and savings for effective sound insulation and noise control of the ventilation system. The large hall seats 1456 people and is a concert hall in the shoebox style. The room shape was studied by computer simulation techniques during the design stage, and the details of the wall and ceiling were studied by use of 1/10th scale model experiments before the construction work started. The small hall seats 610 people in an extremely intimate setting. The seating surrounds the stage, and even the height of the balcony seating is staggered to create what may be termed a ''mini-arena'' style. The unique layout heightens the audience's feeling of proximity to the stage. Acoustical design and acoustical characteristics will be presented at the meeting.

4pAAb19. Effects of the modulated delay-time interval of the single reflection on subjective preference. Junko Atagi, Yoichi Ando (Grad. School of Sci. and Technol., Kobe Univ., Rokkodai, Nada, Kobe, 657 Japan, 962t001n@kobe-u.ac.jp), and Yasutaka Ueda (Hazama Corp., Karima, Tsukuba, 305 Japan)

In the past, experimental results of the sound-pressure level (SPL) were obtained which showed that the indoor sound field is regarded as a time-variant system due to an air current. Using the time-variant model represented by the direct sound and the varying delay time of single reflection, an experiment was performed on an effect on subjective preference. This experiment was focused on the effect of the modulation interval (Δ) of the delay time of the single reflection. The delay time of reflection (Δt_1) was adjusted to the most preferred delay time of reflection which was obtained by the experiments for the time-invariant sound field. In the case of fast tempo music, where the minimum value of the effective duration of the autocorrelation function is $\tau_{e_{\min}} = 43$ ms, the scale value of subjective preference is higher in a time-variant sound field than a time-invariant one. On the other hand, in the case of slow tempo music, $\tau_{e_{\min}} = 127$ ms, such an effect could not be found. It is suggested that a time-variant sound field might have a positive effect on subjective preference in the performance of a fast tempo music.

4pAAb20. Individual differences of subjective preference for sound fields with a different preferred delay time of reflection. Souichiro Kuroki, Ippei Yamamoto (Faculty of Eng., Kagoshima Univ., 1-21-40, Korimoto, Kagoshima, 890-0065 Japan), Hiroyuki Sakai, and Yoichi Ando (Kobe Univ., Rokkodai, Nada, Kobe, 657-8501 Japan)

The purpose of this study is to clarify the relationship between the individual preferred delay time of reflection and the effective duration of the autocorrelation function (ACF) of sound source. Subjective preference tests were conducted by varying the delay time of reflection (Δt_1) which is one of the four objective parameters related to subjective preference of sound fields. In this paper, five dry sources of instrumental music with different effective duration are used. In addition to the global preference with a number of subjects, individual-preference-scale values also were obtained by the method of paired-comparison tests. An attempt is made here whether or not individual differences in subjective preference may be explained by the categorical classification of musical experience, sex, and age of subjects.

Session 4pAAc

Architectural Acoustics: Speech Intelligibility

Dennis Noson, Chair
107 North 77th Street, Seattle, Washington 98103

Chair's Introduction—4:00

Contributed Papers

4:05

4pAAc1. Modeling of acoustic parameters and speech intelligibility in long enclosures. Lening Yang and Bridget M. Shield (School of Eng. Systems and Design, South Bank Univ., London SE1 0AA, UK, shieldbm@sbu.ac.uk)

Ray tracing computer models have been developed for the prediction of the sound field and speech intelligibility in long enclosures of any cross-sectional shape. The models for enclosures of circular and rectangular cross section are particularly suitable for the prediction of speech intelligibility in underground railway stations. The model for use in long enclosures of rectangular cross section has been validated using data measured in a real underground station in Hong Kong. Measurements of acoustics parameters including sound propagation, early decay time, clarity index C50, and center time D50 were made along the length of the platform, with a single source towards one end of the platform. The model correctly predicts the acoustic parameters, the predictions being particularly accurate at midfrequencies. In addition the speech transmission index is accurately predicted at all frequencies, especially in the far field of the source.

4:25

4pAAc2. Relationship between speech transmission index and easiness of speech perception in reverberatory fields. Hiroshi Sato, Hiroshi Yoshino (Dept. of Architecture, Tohoku Univ., Aramaki Aza Aoba, Aoba-ku, Sendai, 980-77 Japan), and Muneshige Nagatomo (Kajima Tech. Res. Inst., 19-1 Tobitakyu 2-Chome, Chofu-shi, Tokyo, 182 Japan)

When evaluating speech intelligibility in rooms, the Tri-syllable articulation test is used as one of the standard indicators in Japan. Sometimes there are differences between auditory impression and the score of the articulation test. So the psychological scale values of easiness of speech perception (ESP) were measured by experiments, which indicated differences between ESP and articulation scores and a good correlation ($R=0.9$) between ESP and speech transmission index (STI) in reverberation. The purpose of this paper is to point out relationships between the condition of sound field and ESP. The first of two experiments in this study was done in reverberatory fields. Sound-pressure level (SPL) of direct sound, direct sound to first reflection ratio (D/R), initial delay time, frequency characteristics of reverberation, and reverberation time (RT) were varied. The other experiment was done in simulated fields of actual auditoriums. The result of the first experiment shows D/R and RT has alarge effect on ESP and while varying D/R, the correlation between ESP and STI is not good ($R=0.6$). The second experiment indicates that when the sound source is on the center of the stage, the correlation between STI and ESP is 0.8; and a different tendency is found while using the sound reinforcement system.

4:45

4pAAc3. Investigation of the speech intelligibility in the dome space. Satoshi Inoue, Masahiro Katoh, Kiyoshi Sugino (Inst. of Technol., Tokyu Construction Co., Ltd., 3062-1 Tana, Sagamihara, Kanagawa, 229-11 Japan, ino@etd.tokyu-cnst.co.jp), and Hiroyuki Imaizumi (Natl. Inst. for Resources and Environment, 16-3 Onogawa, Tsukuba, Ibaraki, 305 Japan)

In order to investigate the acoustical characteristics of an underground dome (2 320 m^3), objective and subjective properties were tested and the mutual relationship of them was examined. Then to improve the speech transmission quality there, some simple absorption treatments were given. A series of experiments, carried out under different acoustical conditions, has confirmed the effects of the treatment. The main results are as follows: (1) In the dome space without absorption treatment, the differences of some objective measures and subjective responses (speech intelligibility) were great in places where observation was made, e.g., RASTI value was distributed over a range of 0.17–0.59. (2) It is indicated that the speech transmission quality was high at the observation points where the energy, arriving in the first 0.06 s after the arrival of the direct sound, is predominant, and others are almost low. (3) Among the physical measures, EDT, early/late-arriving sound energy ratio, *ts* (at 1 kHz octave-band), and RASTI were correlated well with speech intelligibility. (4) Rubber tiles on the floor, 12-m cotton canvas hung from the ceiling, and glass wool board were found to be effective in decreasing reverberation time and diffusing the sound. Also, this treatment was useful in improving the speech intelligibility remarkably.

5:05

4pAAc4. Calculation of speech intelligibility using four orthogonal factors involved in autocorrelation functions. Tetsuichi Shoda and Yoichi Ando (Grad. School of Sci. and Technol., Kobe Univ., Rokkodai, Nada, Kobe, 657 Japan)

In this paper, a method of calculating speech intelligibility for the sound field with the direct sound and the first reflection is proposed. This method is based on the model of an auditory-brain system consisting of the autocorrelation mechanisms which is equivalent to the power spectrum. The autocorrelation function (ACF) may play an important role on the speech information processing. Four factors may be extracted from ACF, such that (1) τ_e : the effective duration of ACF (the ten percentile delay) meaning the repetitive feature of signals; (2) τ_1 : the delay time of the first peak of ACF signifying the pitch of signals; (3) ϕ_1 : the amplitude of the first peak indicating the strength of the pitch; and (4) $\Phi(0)$: the power of the signal frame. The distance between the syllable in sound fields and the syllable of the direct sound were calculated by the use of these four factors. Using these distances and its linear combination, speech intelligibility of the syllables in the sound fields may be calculated. Results of speech intelligibility tests as a function of the delay time of the reflection are in good agreement with the calculated values.

Session 4pAO

Acoustical Oceanography and Animal Bioacoustics: Acoustics of Fisheries and Plankton III

Orest I. Diachok, Chair
Naval Research Laboratory, Washington, D.C. 20375

Contributed Papers

2:00

4pAO1. Variation in acoustically measured abundance from repeated surveys of an isolated herring population. I. Huse and M. Ostrowski (Inst. of Marine Res., P.O. Box 1870, N-5024 Bergen, Norway)

An isolated population of herring was surveyed acoustically a total of 18 times while wintering in a fjord in northern Norway. Total abundance and variograms were computed for each survey. Three-dimensional distribution graphs are presented. The results are discussed in terms of total variation of acoustic survey estimates.

2:15

4pAO2. Acoustic estimates of zooplankton and micronekton biomass using an ADCP. Patrick H. Ressler, Douglas C. Biggs, and John H. Wormuth (Dept. of Oceanogr., Texas A&M Univ., College Station, TX 77843-3146, pressler@ocean.tamu.edu)

A calibrated 153-kHz narrow-band ADCP (Acoustic Doppler Current Profiler) has been used to collect acoustic backscatter intensity (Sv) data during several cruises in the Gulf of Mexico. Data have been gathered both while on station and while underway along transects through different hydrographic regimes, enabling an examination of both temporal and spatial trends in backscatter. In addition, zooplankton and micronekton stocks were intensively sampled with a 1 $m\hat{2}$ MOCNESS (Multiple Opening-Closing Net Environmental Sensing System). Empirical correlations between spatial and temporal variations in Sv and in standing stocks of zooplankton/micronekton will be presented. Sv measured with an ADCP may be useful as an index of zooplankton and micronekton biomass. The location and abundance of these food stocks are hypothesized to be important in determining the distribution and abundance of marine cetaceans in the Gulf of Mexico. In addition, Sv-derived estimates of such stocks may allow inferences about secondary biological productivity in the Gulf, and the role that mesoscale circulation features might play in driving large-scale patterns. [This ongoing research is supported by the USGS Biological Resources Division and the US Minerals Management Service under USGS BRD Contract No. 1445-C109-96-004. For more information, see http://www.tamug.tamu.edu/gulfcet.]

2:30

4pAO3. Monitoring the growth and mortality rates of pelagic fish with absorption spectroscopy measurements. Orest Diachok (Naval Res. Lab., Washington, DC 20375) and Paul Smith (Southwest Fisheries Sci. Ctr., La Jolla, California, 02038)

The experimental geometry for absorption spectroscopy measurements, which was employed during Modal Lion, can be adapted for monitoring the growth and mortality rate of pelagic fish. During the juvenile stage, in which sardines grow from 3.5 to 13.5 cm over a period of about 1 year and mortality rates are relatively high, absorption spectroscopy measurements (together with complimentary biological and oceanographic measurements) conducted for a few days approximately once per month would permit investigation of the environmental determinants of these parameters. Measurement sites would have to follow the migration route of juvenile sardines over a period of about 1 year. To minimize cost and simplify logistics, a broadband omnidirectional source (instead of a parametric source) would be suspended from a moored ship or buoy. A relatively low source level ($\sim$170 dB) would be adequate. Simulations will be presented of the temporal evolution of the resonance frequency, absorption coefficient, and transmission loss due to sardines between 0.6 and 6.0 kHz. Simulations will illustrate the sensitivity of transmission loss to source and receiver depth, and the average depth and thickness of absorbing layers. In addition, calculations that illustrate the potential of estimating the depth and thickness of absorbing layers by varying the depth of an omnidirectional source will be presented. [This work was supported in part by the Office of Naval Research.]

2:45

4pAO4. Measurements of snapping shrimp colonies using a wideband mobile passive sonar. Marc Olivieri, Stewart A. L. Glegg, and Robert K. Coulson (Ctr. for Acoust. and Vib., Dept. of Ocean Eng., Florida Atlantic Univ., 777 Glades Rd., Boca Raton, FL 33431)

A wideband passive sonar has been developed to study ambient noise levels in coastal regions. During the past two years a large number of measurements have been made with this system in shallow water off Boca Raton, Florida. The results have demonstrated that the noise from snapping shrimp dominates in this area during low sea states. Localization of the colonies has been possible by moving the sonar to different locations over the period of a few hours. Typically the shrimp are clustered close to man-made structures at or near the shore line, with additional smaller colonies on nearby coral reefs. To speed up the data collection the system has been mounted to an autonomous underwater vehicle and on-board intelligence is being developed to optimize the survey locations in real time. Results from a demonstration experiment will be presented. [Work supported by ONR.]

3:00

4pAO5. Using sound to map fish spawning: Determining the seasonality and location of spawning for weakfish and red drum (Family Sciaenidae) within Pamlico Sound, NC. Joseph J. Luczkovich (Dept. of Biol. and Inst. for Coastal and Marine Resources, East Carolina Univ., Greenville, NC 27858 luczkovichj@mail.ecu.edu), Stephen E. Johnson, and Mark W. Sprague (East Carolina Univ., Greenville, NC 27858)

Fish produce species-specific sounds during courtship, during aggressive encounters with other fish of the same species, and as a response to threats from predators. Particular sounds are species-specific and have been recorded from fish during spawning in captivity. Such acoustical data may be used to identify and locate fish that are reproductive in the sea. In this paper, nocturnal sound production and planktonic egg production are correlated for weakfish (*Cynoscion regalis*) and red drum (*Sciaenops ocellatus*) (Family Sciaenidae), two species for which there are concerns over declining stocks. Recordings of underwater sounds and ichthyoplankton net tows were made after sunset at ten sites in Pamlico Sound, NC. Sites were relocated with a differentially corrected Global Positioning System receiver on a biweekly basis from May through October 1997. Based on laboratory studies, weakfish ''purring'' sounds and red drum ''knocking'' sounds only occur during spawning. ''Purring'' was recorded along with sciaenid eggs at stations near Ocracoke and Hatteras Inlets during May through July. ''Knocking'' was recorded in September and October at sites away from the inlets, also occurring with sciaenid eggs in one case. Mapping of sciaenid spawning areas may be possible using acoustical data alone, but may overestimate egg production.

3:15

4pAO6. Using fish sounds to identify spawning activity of weakfish (*Cynoscion regalis*) and red drum (*Sciaenops ocellatus*) in nature. Mark W. Sprague (Dept. of Phys., East Carolina Univ., Greenville, NC 27858, spraguem@mail.ecu.edu), Joseph J. Luczkovich (Inst. for Coastal and Marine Resources, Greenville, NC 27858), and Stephen E. Johnson (East Carolina Univ., Greenville, NC 27858)

In the coastal systems of the southeastern United States, populations of weakfish (*Cynoscion regalis*) and red drum (*Sciaenops ocellatus*) are in decline, and knowledge of their spawning areas is important for the formulation of fisheries management plans. Weakfish and red drum, both members of Family *Sciaenidae*, use their swim bladders to produce species-specific sounds associated with spawning activity. Large spawning aggregations of these fish can produce sound levels as high as 145 dB (re:1 μPa), and these sounds can be used to identify spawning areas. Recordings and spectrograms of spawning sounds will be presented for each species. Water depth, bottom type and contour, sound-speed gradient, and water current gradient all affect the propagation of the fish sounds through the water. Measurements of these factors and the sound level of the fish calls are used to obtain an approximate range to the spawning aggregation. [Work supported by North Carolina Division of Marine Fisheries and U.S. Fish and Wildlife Service.]

Session 4pBV

Bioresponse to Vibration/Biomedical Ultrasound: Applications of Microbubble Based Echo Contrast Agents II

Nico De Jong, Cochair
Erasmus University, Room Ee2302, P.O. 1728, 3000DR, Rotterdam, The Netherlands

Inder Raj S. Makin, Cochair
Ethicon Endo-Surgery, 4545 Creek Road, Cincinnati, Ohio 45242

Invited Papers

12:50

4pBV1. Quantification of myocardial blood flow using ultrasound-induced destruction of microbubbles administered as a constant venous infusion. Sanjiv Kaul (Univ. of Virginia, Charlottesville, Virginia)

It was hypothesized that if microbubbles are administered as a continuous infusion, then destroying them in the myocardium and measuring their myocardial reappearance rate at steady state will provide a measure of mean myocardial microbubble velocity. Conversely, measuring their myocardial concentration will provide an assessment of microvascular cross-sectional area (CSA). Myocardial blood flow (MBF) can then be calculated from the product of the two. Accordingly, experiments were performed in 21 dogs where MBF was altered mechanically or pharmacologically in the absence or presence of coronary stenosis. Microbubbles were delivered as a constant infusion, and two-dimensional ultrasound was performed using different PI. The myocardial video intensity (VI) versus PI plots were fitted to an exponential function: $y=A(1-e^{-\beta\tau})$, where A is the plateau VI reflecting the microvascular CSA, and β reflects the rate of rise of VI and hence, microbubble velocity. Excellent correlations were found between flow and β, as well as the product of A and β. It is concluded that MBF can be quantified by destroying microbubbles with ultrasound during a venous infusion.

1:10

4pBV2. Gas-filled liposomes as ultrasound contrast agents for blood pool, thrombus-specific and therapeutic applications. Evan C. Unger (Dept. of Radiology, The Univ. of Arizona, Tucson, AZ 85724-5067, eunger@vms.ccit.arizona.edu), Thomas McCreery, Dekang Shen, GuanLi Wu, Robert Sweitzer, and Qui Wu (ImaRx Pharmaceutical Corp., Tucson, AZ)

Our group has developed technology for stabilizing microbubbles with phospholipid coatings (Aerosomes®-ImaRx Pharmaceutical Corp.). The first agent (MRX-115) is based upon lipid-coated microspheres filled with perfluoropropane gas and is in phase III clinical trials (October, 1997) for radiology and cardiology applications. Myocardial perfusion studies show the potential for the agent to detect ischemia in patients with myocardial infarcts. Targeting ligands have been covalently bound to lipids and incorporated into the stabilizing shells on the microbubbles. The first targeted agent planned for clinical trials is MRX-408 (perfluorobutane gas, linear hexapeptide-RGD analog). *In vitro* studies show enhanced visualization of thrombi, and *in vivo* studies in dogs with arterial and venous thrombi show selective enhancement, even in animals injected with heparin. *In vitro* studies have been performed to test sonothrombolysis comparing MRX-115 to MRX-408 with and without urokinase. The targeted contrast agent MRX-408 increases the rate of clot dissolution with ultrasound. Targeted, tissue-specific agents are feasible for ultrasound. A thrombus-specific ultrasound contrast agent has the potential to improve clot characterization and detection as well as clot lysis.

1:30

4pBV3. Targeted acoustic contrast agents: New opportunities for ultrasound in medical diagnosis and therapy. Gregory M. Lanza (Div. of Cardiology, Campus Box 8220, Barnes-Jewish Hospital, Washington Univ., 216 S. Kingshighway Blvd., St. Louis, MO 63110, greg@soundlab.wustl.edu), Kirk D. Wallace, Rebecca L. Trousil, James G. Miller (Washington Univ., St. Louis, MO), James H. Rose (Iowa State Univ., Ames, IA), Patrick J. Gaffney (NIBSC, Herts, UK), Dana R. Abendschein, Christopher S. Hall, Michael J. Scott, Christine A. Lorenz, Ralph Fuhrhop, and Samuel A. Wickline (Barnes-Jewish Hospital, Washington Univ., St. Louis, MO 63110)

For 30 years, medical professionals have sought the ''magic bullet'' which would allow sensitive detection of disease and facilitate localized administration of potent chemotherapeutic agents without systemic toxicities. Ligand targeted ultrasonic contrast agents have the potential to improve diagnostic sensitivity, port drugs to specific pathologic tissues, and provide dosimetry of the delivered therapeutics. The authors have recently developed a ligand directed, lipid encapsulated perfluorocarbon emulsion particle (250-nm diameter). These particles, when bound to biological surfaces, strikingly increase the acoustic reflectivity, although the particles have low inherent echogenicity. Contrast enhancement of the reflectivity of blood thrombi, fibrin clots, and nitrocellulose membranes has been tentatively explained by a transmission line model. Intravenous injection of the perfluorocarbon emulsion particles has localized and acoustically enhanced remote thrombi which were otherwise difficult to detect. The contrast agent has a 1-h circulatory half-life which greatly facilitates successful systemic targeting. The small particle size of these emulsions affords access to tiny capillary beds and vascular endothelial disruptions such as those following angioplasty. Collectively, these studies suggest perfluorocarbon contrast agents can target molecular epitopes within and proximate to the systemic circulation. These agents are creating new and exciting opportunities for ultrasound in medical diagnosis and therapy.

Contributed Papers

1:50

4pBV4. Disruption of contrast agents for monitoring blood flow. J. Brian Fowlkes (Dept. of Radiology, Univ. of Michigan, Medical Ctr., Kresge Res. Bldg. III R3315, Ann Arbor, MI 48109-0553, fowlkes@umich.edu), Richard T. Rhee, David W. Sirkin, and Paul L. Carson

An experimental system is being developed for controlling the flow of ultrasound contrast agent to create abrupt changes in contrast levels in tissues. Such changes can be used to monitor blood flow through bolus passage and to identify specific feeder vessels. Focused ultrasound (2.25 MHz) has been used to interrupt the flow of the contrast agent MRX-115 (ImaRx Pharmaceuticals, Tucson, AZ) in a tube (6 mm diameter, average velocity of 9.45 cm/s) using short bursts 20 cycles in duration with a pulse repetition frequency (PRF) of 0.75–6.0 kHz. Peak rarefactional pressures (PRP) for effective disruption of contrast agent were only 0.6–1.2 MPa. This corresponds to mechanical indices (MI) of 0.4–0.8 and intensity values which are well within the recommended limits for diagnostic ultrasound, up to 90% of contrast agent signal was eliminated and positive boluses were produced by turning the field off and on. Similar results have been achieved with a 1.8-MHz *in vivo* system in a rabbit model. Therefore the potential exists to use common diagnostic ultrasound to produce contrast disruption suitable for short bolus production. [Work was supported in part by the U.S. Army under Grant No. DAMD17-94-J4144.]

2:05

4pBV5. Effects of pulse width, phase, and intensity on echoes from ultrasound contrast agents. Karen E. Morgan, Paul A. Dayton, Katherine W. Ferrara (Dept. of Biomed. Eng., Univ. of Virginia, Box 377, Health Sci., Charlottesville, VA 22908, kej6u@virginia.edu), Alexander L. Klibanov, and Gary H. Brandenburger (Mallinckrodt, Inc., Hazelwood, MO 63042)

Contrast-assisted imaging shows great promise for the imaging of microvascular perfusion, although current techniques cannot differentiate the echoes from tissue and ultrasound contrast agents. Harmonic schemes for imaging with these agents have limited spatial resolution due to the transmission of a narrow-band pulse. The potential for wideband contrast-assisted imaging through a systematic evaluation of the effects of imaging parameters on the received echoes is evaluated. Our results indicate that the echo intensity received from contrast agents following wideband transmission is at least as large as or larger than that for narrow-band transmission, with both exhibiting nonlinear increases with increasing transmitted signal intensity. While the echo intensity at harmonic multiples of the bubble resonant frequency can be larger for narrow-band insonation, echoes received after wideband insonation demonstrate a broadband spectrum with significant amplitude over a very wide range of frequencies. The authors also demonstrate that the transmission of wideband high intensity pulses produces consistent time domain echoes from microbubbles. These echoes are unchanged whether the microbubbles break or remain intact over several pulses. Additionally, if the phase (order of compressional and rarefactional half-cycles) of the transmitted pulse is changed, the polarity of the received signal changes accordingly for a stationary reflector but remains unchanged for a microbubble.

2:20–2:30 Break

2:30

4pBV6. Characterization of ultrasound-contrast agents by short-burst excitation. Oliver D. Kripfgans (Dept. of Radiology, Univ. of Michigan, Medical Ctr., Kresge Res. Bldg. III R3315, Ann Arbor, MI 48109-0553, oliver.kripfgans@umich.edu), J. Brian Fowlkes, and Paul L. Carson

A two-frequency method has been adapted to the use of short ultrasonic bursts for the determination of bubble size. The advantage of a burst method is threefold. The overall energy transfer into the volume of interest is smaller as compared to cw. Second, it allows simple methods for improving spatial resolution. Finally, only one transducer will eventually be required in a pulse-echo mode. The experimental setup is designed to fully compensate for transducer characteristics and sample artifacts such as attenuation through window material and contrast agent solution. This leads to a constant sound pressure exposure for contrast agent in the focal volume. The data acquisition employs the frequency tracking procedure where spectra for narrow-band excition are analyzed for fundamental, second harmonic, and sidebands. A technique of nonconstant imaging frequency is used to avoid ambiguities due to coincidences between higher harmonics of the fundamental and multiples of the sideband frequencies. Measurements with shell-encapsulated microbubbles showed the behavior demonstrated by previous cw methods. Higher harmonics as well as sidebands were present due to the nonlinear dynamics of the contrast agents. [This work was supported in part by USPHS Grant Nos. RO1 DK42290 and 1RO1 HL54201.]

2:45

4pBV7. Acoustic detection of microbubble destruction in gaseous contrast agents. William T. Shi and Flemming Forsberg (Dept. of Radiol., Thomas Jefferson Univ., Philadelphia, PA 19107)

The destruction of microbubbles from ultrasound contrast agents has recently received much attention because the disintegration can be used to create new imaging modalities and may produce adverse bioeffects as

well. Both passive and active acoustic detection techniques were employed in our experiments to investigate the collective and individual behaviors of contrast microbubbles at different insonification levels. With the passive technique, the oscillation and destruction of microbubbles were studied by characterizing the waveforms of scattered acoustic signals and analyzing the harmonic and noise contents in the spectra of the signals. It was observed that discrete subharmonics and superharmonics (up to the 20th order) were generated at low acoustic pressures. When the bubble destruction occurred at high acoustic pressures, the noise level in the spectra was raised and the discrete harmonic structure in the higher frequency domain reduced to broadband noise. The noise spectrum in the higher frequency range was found to broaden as the acoustic pressure increased. A modified ACD (acoustic cavitation detector) with improved spatial resolution was utilized to study the disintegration process of an individual microbubble. The comparison of results from passive and active detections will be presented.

3:00

4pBV8. Acoustic and system parameters affecting destruction of ultrasound contrast agents. Peter P. Chang, Inder Raj S. Makin, and Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, pchang@apl.washington.edu)

In order to effectively use ultrasound contrast agents (UCAs), it is necessary to tune the ultrasound scanner parameters to a particular agent or change the characteristics of the UCA to target-specific applications. Both efforts are being sought by ultrasound manufacturers and pharmaceutical companies. The destruction of UCAs is usually undesirable, although it can sometimes be used advantageously such as in harmonic imaging mode. In the present work, changes in acoustic and system parameters such as frequency, transmitted acoustic pressure, pulse repetition frequency, flow rate, diluting media, concentration, and type of contrast agents were measured at both first and second harmonic. A prototype Doppler system previously described [Chang *et al.*, Ultrasound Med. Biol. **22**, 1205–1214 (1996)] has been adapted to carry out these measurements. Results with Albunex show that bubble destruction increases with an increase in flow rate, acoustic pressure, PRF, and decreases with an increase in the concentration of the contrast agent and oxygen content in the diluting medium. In conclusion, the rupture of UCAs are influenced by several imaging parameters. The threshold values when destruction of the agents occurs should be known in order to optimize the use of UCAs in clinical settings. [Work supported by DARPA-TRP.]

3:15

4pBV9. Observations of insonified contrast agents *in vitro* and *in vivo*. Paul Dayton, Karen Morgan, Margaretta Allietta, Kathy Ferrara (Univ. of Virginia, Box 377 HSC, Charlottesville, VA 22908), Alexander Klibanov, and Gary Brandenburger (Mallinckrodt, Inc., Hazelwood, MO 63042)

Ultrasound contrast agents, which may include a gas core and lipid or albumin shell, play an increasingly important role in medical imaging. However, a complex set of phenomena have limited quantitative evaluation of tissue perfusion based on echoes from these microbubbles. With the use of optical microscopy, mechanisms of microbubble destruction during insonation are evaluated, as well as the effects of ultrasonic radiation force. Depending on conditions such as acoustic intensity and dissolved gas concentration in the media, several mechanisms can change the microbubble and promote its destruction during insonation. Slow gas diffusion from the bubble is observed with low-intensity transmission, but with several pulses at a higher intensity, the albumin shell may weaken sufficiently to release the gas core. It is demonstrated that the primary acoustic radiation force has a significant effect on microbubbles when the ultrasound is transmitted with a low acoustic pressure and high pulse repetition frequency. Deflection of the contrast agent streamline to the wall of a small vessel is shown both *in vitro* and *in vivo*. Such manipulation of the contrast agents may be desired in drug delivery applications, as it provides a method of concentrating bubbles in a specific area.

3:30

4pBV10. Ultrasound contrast agent enhances vascular damage in mouse intestines. Richard A. Gies and Douglas L. Miller (P7-53, Battelle Pacific Northwest Natl. Lab., P.O. Box 999, Richland, WA, doug.miller@pnl.gov)

Either Albunex®, a gas-body based contrast agent, or 5% albumin in saline without gas bodies was administered to anesthetized hairless mice at 10 ml/kg by retro-orbital injection. The mice then were exposed at the midsection to 1.09 MHz of ultrasound for 100 s, either continuous (cw) or pulsed (pw, 10 μs pulses repeated at 1 kHz), in a 37 °C water bath. After exposure, the intestines were examined to count petechiae and hemorrhages into the intestinal lumen. cw thresholds for statistically significant petechiae production were 0.7 MPa spatial peak pressure amplitude without added gas bodies, and 0.5 MPa with added gas bodies. At 0.7 MPa the mean counts (five mice) were 22 without and 120 with added gas bodies. A statistically significant number of hemorrhages occurred at 0.7 MPa (three per mouse) with the contrast agent. The pw thresholds with added agent were 1.4 MPa for petechiae and 2.8 MPa for hemorrhage. At 2.8 MPa, average counts were 227 petechiae and 2.4 hemorrhages. No statistically significant vascular damage occurred for pw exposure up to 2.8 MPa without contrast agent. In terms of mechanisms, the effects appeared to be partly thermal for cw, but primarily associated with ultrasonic cavitation for pw exposure. [Work supported by NIH CA42947.]

Session 4pEA

Engineering Acoustics: Applications of Acoustic Technology

Robert D. Finch, Chair
Cullen College of Engineering, University of Houston, Mechanical Engineering, Houston, Texas 77204-4792

Contributed Papers

1:00

4pEA1. Determination of the noise attenuation of hearing protectors by numerical modeling of the outer ear. Samir N. V. Gerges and Elizabete Y. Bavastri (Federal Univ. of Santa Catarina Mech. Eng. Dep.-Acoust. and Vibration Lab., Cx. P. 476, Florianópolis, SC, Brazil, gerges@mbox1.ufsc.br)

This paper will address the important practical issues of hearing protectors, used in industry, to protect the workers from high noise levels. Comments on the difficulties for the measurements of hearing protector's noise attenuation is discussed. A new work on the numerical modeling of the outer ear canal, considering the eardrum acoustic impedance, using the finite elements (FEM), infinite FEM for the quantification of the noise attenuation of the protectors is presented. This numerical model considers the geometry of the outer ear, outer ear canal, and the eardrum acoustic characteristics. This numerical model can serve as a quick and low cost tool for the optimization of the protector project and the investigation of the effect of different parameters such as protector insertion, effect of leakage, materials, and others on the protector noise attenuation.

1:15

4pEA2. Improved TV/radio listening for hearing impaired. J. G. H. van Zutphen (Tech. Univ. Twente and Audiol. Dept. of the Free Univ. of Amsterdam, The Netherlands), W. F. Druyvesteyn (Philips Res. Labs., Eindhoven, The Netherlands), and C. H. Slump (Tech. Univ. Twente, The Netherlands)

In case different persons, having different preferences for the loudness level, want to listen to one sound program in one room, problems arise concerning the loudness of the reproduced program. An example: person A has a normal hearing threshold, while person B is hearing impaired. A sound reproduction system, consisting of a small array of loudspeakers at the TV/radio can solve this problem with digital signal processing; the directivity pattern of the array is manipulated such that the main lobe is directed to person B. The investigation and the presentation at the conference contains two parts: (1) physical acoustics: design and testing of the loudspeaker array including digital processing, and (2) subjective evaluation with normal-hearing subjects and hearing impaired subjects. The subjects listen (and view) a TV program consisting of a selection from 120 short spoken sentences. The normal-hearing subjects adjust the loudness level at their listening position. The hearing-impaired subjects test the effect of the array. These listening tests show a clear improvement in the speech intelligibility for the hearing-impaired subjects as expressed in the number of correct understood sentences and in subjective terms as ''easier to understand'' or ''less tiring to listen to.''

1:30

4pEA3. Piezoelectric transducer for a hearing aid using PZT thin film. Hidehiko Yasui, Minoru Kurosawa, Takeshi Morita, Takefumi Kanda, and Toshiro Higuchi (Dept. of Precision Machinery Eng., Grad. School of Eng., Univ. of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113 Japan, hide@intellect.pe.u-tokyo.ac.jp)

A piezoelectric transducer using a PZT thin film deposited by a hydrothermal method was fabricated for the investigation of ability for an earphone. The PZT material can produce sounds efficiently from an electrical signal. The thin film technology is suitable for the fabrication of a small transducer which is worn in the ear. Especially, the hydrothermal method has merits such as thick film fabrication and possibility of deposition on a round surface. Therefore we can design and fabricate a diaphragm for a small earphone which produces a high enough sound-pressure level for people who have an obstacle for hearing ability. The fabricated transducer had a bimorph construction. The base material was titanium foil 20-μm thick. Both sides of the titanium, PZT thin films 10-μm thick were deposited. The diameter of the transducer was 6 mm. We investigated the ability of the fabricated transducer. Displacement amplitude of the transducer was about 0.9 μm per 5 V driving voltage below the resonance frequency.

1:45

4pEA4. Low-frequency circuit noise in hearing aids. Vishakha W. Rawool (Commun. Disord. and Special Education, Bloomsburg Univ., Bloomsburg, PA 17815)

Some hearing impaired listeners complain about the circuit noise in hearing aids. This is specially true for individuals with high-frequency hearing impairments who have normal hearing in the lower frequencies. In this investigation, the equivalent input noise levels were measured in several different hearing aids: linear, wide dynamic range compression, high-level compression, programmable, and digital aids. For the linear, programmable, and digital aids the measurements were made with two types of frequency responses. One was suitable for a flat moderate hearing impairment and another for a high-frequency hearing impairment. For the linear and compression aids, the noise levels were measured by using two different volume control settings. All the measurements were made by using 11 different special purpose average frequencies ranging from 200 to 5000 Hz. The results suggest generally higher circuit noise in the lower frequencies in all the hearing aids. Changes in volume control settings showed no differences in noise levels. For the programmable aids, the flat versus high-frequency configurations yielded similar noise levels. For the linear and digital aids, the high-frequency configuration yielded higher noise levels in the lower frequencies when compared to the flat configurations.

2:00

4pEA5. High degree of freedom muffler optimization using genetic algorithms: Experimental verification. Siska Pottie and Dick Botteldooren (Information Technol. (INTEC) Univ. Gent st-Pietersnieuwstraat 41 B-9000 Gent, Belgium, pottie@intec.rug.ac.be)

This paper reports on an experimental verification of a nonconventional muffler design. The optimal design is searched using a genetic algorithm (GA). GAs are a powerful optimization technique that are well suited to handle multiparameter optimization problems. The degrees of freedom used to describe the structure is far beyond the usual set. The muffler has a fixed length and width and is described by a set of design rules defining the use of the available space. Different coding alternatives for the GA were analyzed. Best results are obtained with a coding that describes the position and size of the design rules. The performance of the mufflers is characterized by the theoretically achieved sound reduction and by the ease of construction of the optimized muffler. The theoretically achieved sound reduction is calculated using a time domain simulation. In

order to verify the promising theoretical performance of the optimized mufflers, a good performing design is built and tested in the laboratory. General agreement between the theoretical and experimental results is found. Since no conventional muffler design rules are imposed, the optimization is completely open for new approaches which can led to new insights for conventional muffler design.

2:15

4pEA6. Theoretical analysis of reactive silencers with two propagating modes. J. Kergomard and M. Pachebat (Lab. Acoustique, Universite du Maine, Av. O. Messiaen, 72017 Le Mans Cedex 9, France, kergo@laum.univ-lemans.fr)

A general modeling of sound ducts with reactive (or dissipative) treatments considers one main guide and a secondary one (the treatment), with another propagation constant, including discontinuities, coupled to the main one by perforations, i.e., a shunt impedance and a series impedance [see Kergomard *et al.*, Acta Acustica **2**, 1–16 (1994)]. For a periodical system the calculation can be made analytically for a succession of N cells by diagonalization: the impedance matrix of the portion of the duct with treatment can be deduced. At low frequencies only one mode propagates in each duct and the matrices are of fourth order. From the boundary conditions for the treatment the insertion loss (IL) for the main duct is deduced. Examples are an array of coupled resonators, a lining with perforated facing, and a perforated tube muffler with partitions. A conclusion is the proof of the interest of the interferences between two propagating modes of the periodical system: resonances occur and IL is proportional to $\log N$. It is more efficient than a single propagating mode (e.g., in expansion chambers). Compared to the evanescent case (IL proportional to N), it is less efficient but the frequency band is wider. Many facts of practical interest can be deduced.

2:30

4pEA7. Observations of characteristic transmission and reflection coefficients of a splitter duct attenuator. M. Terao and H. Sekine (Dept. of Architecture, Kanagawa Univ., Kanagawaku, Yokohama, 221 Japan)

The characteristic transmission coefficients (including the reflection coefficients) of a splitter duct attenuator with two air channels containing higher-order wave modes were investigated. These were evaluated at interfaces taken in the straight duct sections in both sides of the attenuator and were determined experimentally and numerically (by BEM). The determination method introduced here requires the solution of two sets of the simultaneous equations given by the superposition principle of traveling waves. One is that the sound pressure at a point is the sum of the pressures of all the traveling waves that could possibly exist there, and it was applied to decompose the incoming and outgoing waves of all modes at each interface. The other is that each outgoing wave pressure of each mode at each interface equals the sum of the products of the incoming wave pressures of all modes at all interfaces and the corresponding transmission factors. It was applied to determine, without using an echoic termination, each characteristic transmission coefficient by which each incoming wave contributes to each outgoing wave pressure. The transmission coefficients of the attenuator were obtained for the fundamental and the first-order modes. Agreement of those by experiment and numerical results is substantially good.

2:45

4pEA8. Applications of sonic soot cleaning techniques. Jing Tian (Inst. of Acoust., Academia Sinica, Beijing 100080, PROC)

Sonic soot cleaning is a method that is applied to clean the ash or fossil deposit on surfaces of heat exchangers in boilers. The technique includes a set of hardware and software for the designing, production, installation, debugging, and operation of sound wave generators and the controller. Although it has been applied in boilers in power plants, petrochemical works, and general industries worldwide, still most of the correlated basic problems, such as the relationships between sound field and the mass delivery, the heat conduction as well as the combustion process in boilers, have not been well solved. It is almost a purely empirical method that its application is far in the lead of fundamental research. Now in China more than ten institutions and companies are engaged in the research, development, and marketing of sonic soot cleaning techniques. These techniques cover sound generators ranging a frequency domain from infrasound up to ultrasound, under the protection of more than 20 patents. In this paper, several applications are analyzed as examples to demonstrate the effectiveness. Some basic phenomena and unsolved problems are revealed. The course of sonic soot cleaning development is discussed. [Project supported by the Natural Science Fund of China.]

THURSDAY AFTERNOON, 25 JUNE 1998 GRAND CRESCENT (W), 1:00 TO 3:00 P.M.

Session 4pMU

Musical Acoustics: Timbre of Musical Sound II

James H. Irwin, Chair
Electrical and Computer Engineering Technology, ECET, Bradley University, Peoria, Illinois 61625

Contributed Papers

1:00

4pMU1. Validation of a multidimensional distance model for perceptual dissimilarities among musical timbres. Nicolas Misdariis, Bennett Smith, Daniel Presssnitzer, Patrick Susini (Inst. de Recherche et de Coordination Acoustique/Musique (IRCAM), 1 place Igor Stravinsky, F-75004 Paris, France), and Stephen McAdams (IRCAM, F-75004 Paris, France and (CNRS), Univ. René Descartes, EPHE, F-75006 Paris, France)

Several studies dealing with the perception of musical timbre have found significant correlations between acoustical parameters of sounds and their subjective dimensions. Using the conclusions of some of these studies, a calculation method of the perceptual distance between two sounds has been developed. Initially, four parameters are considered: spectral centroid, irregularity of the spectral envelope, attack time, and degree of variation of the spectral envelope over time. For each of these, a transformation factor between the physical axis and the corresponding subjective dimension is obtained by linear regression. After a normalization of the data, the four coefficients then found are those of a linear combination that gives the final distance values. Since this model is based on numerical results derived from experiments that mostly used synthesized sounds, the application to a database of recorded musical instrument sounds needs a strong validation procedure. This procedure involves the adjustment of the

coefficients of the first four parameters as well as the eventual introduction of new ones to attain a perceptually relevant distance between two musical sounds. The progress of this research and the results of the database search engine built on this similarity model will be presented and discussed.

1:15

4pMU2. Timbre effects caused by drumstick tip shapes/sizes. James H. Irwin, Jr. (Elec. and Comput. Eng. and Technol., Bradley Univ., Peoria, IL 61625, jhirwin@bradley.edu)

The relationship between the shape/size of a drumstick tip and the resulting initial spectrum produced by hitting a drumhead was investigated. Contrary to tensioned plate theory, and popular opinion, the amplitudes of the various induced partials were not significantly different when the shape/size of the tip was changed dramatically. The work was extended by using various diameter steel balls to strike the drumhead and it was found that the initial spectra did change for extreme differences in ball diameter. Further investigation showed two, or possibly three, timbre groupings. Present results indicate that small diameter balls (less than 15 mm diameter) induce significant energy above the first few partials of the drumhead, medium diameter balls (15 to 25 mm diameter) induce most of their energy in the first few partials, and perhaps large diameter balls (greater than 25 mm diameter) induce little energy above the lowest partial. These effects are linked to the membrane and plate characteristics of the drumhead. Since most drumstick tips have diameters in the medium range little distinction occurs among their spectra. In all cases the spectra were normalized for energy input, and the drum tension and position of impact were fixed.

1:30

4pMU3. On the intentional use of instrument characteristics in contemporary classical compositions. Alexandra Hettergott (Inst. of Musicology, Univ. of Vienna, Spitalgasse 2, 9. Hof, A-1090 Vienna, Austria)

Compositional intention is likewise subjected to the (mechanical) instruments' sound characteristics and technical possibilities, for instance, the possibility to produce impulsive or stationary sounds, to form (pedaled) chord clusters, or to microtonally retune pitch during performance. These differing techniques may additionally allow several (intended) psychoacoustical effects, as a changing fluctuation (beatings, roughness) or pitch strength, which also have a strong impact on timbre: so in the latter case the pitches' varying distinctiveness may be combined with a ''metallic,'' ''plastic,'' or ''wooden'' sound coloring. Consequently, in composing the decision for strings, either *struck* (piano, also prepared), *bowed* (violin, microtonally tuned), or *plucked* (harpsichord), for *flue, reed* (organ), or *free-reed* (mouth organ) pipes, etc., is also subjected to (*a priori*) intentional reflections aimed on the very (new) sound experience imagined. By means of music and score examples [W. Rihm, J. Cage, G. Scelsi, G. Ligeti, O. Messiaen, T. Hosokawa] as well as spectra and spectrograms the paper sheds some light on the composers' sounding and purposeful use of the mechanical instruments' possibilities in sound production in order to get (close to) the corresponding aesthetics of expression intended.

1:45

4pMU4. Musical instrument perception in cochlear implant listeners. John J. Galvin III and Fan-Gang Zeng (House Ear Inst., 2100 W. 3rd St., 5th Fl., Los Angeles, CA 90057-1922)

To examine the degree to which the cochlear implant preserves acoustic cues for timbre perception, eight cochlear implant and three normal-hearing listeners participated in a closed-set musical instrument identification experiment. Eleven nonpercussive, single-note instrument samples including woodwinds, brass, and strings were selected from the MUMS Classical Sounds CDs and formed a six-token and a nine-token set condition. The stimuli were normalized to have the same duration and RMS amplitude, and edited to include either the natural attack or the steady portion only. Subjects were trained with trial-by-trial feedback until asymptotic performance was achieved on one pitch (a4) and were tested without feedback on another pitch (c4). In the six-instrument task, both normal and implant listeners scored an averaged 70%–80% correct identification performance for both the natural attack and steady-state conditions. In the nine-instrument task, the implant listeners scored 20%–30% lower than the normal control. These results indicate that in a relatively simple six-instrument task, the cochlear implant can provide sufficient cues to support normal-level timbre perception, whereas the distorted spectral transformation via the cochlear implant fails to support a normal-level timbre perception, as the task complexity is increased to require more spectral resolution in the nine-instrument task.

2:00

4pMU5. Signal processing techniques to analyze the effects of guitar geometry on musical timbre. Dominic DiDomenico, Kevin Morris, and Suzanne Keilson (Loyola College, Dept. of Elec. Eng. and Eng. Sci., 4501 N. Charles St., Baltimore, MD 21210, keilson@loyola.edu)

Musical acousticians seek to answer many questions. What makes a particular instrument sound like it does? How do listeners perceive musical sounds? As an undergraduate design project, three guitars of varying shapes (''box,'' ''triangle,'' ''octagon'') were constructed and then compared to the traditional design. Materials were not varied, only the geometry of the sound boxes. The overall volume of each sound box was approximately equal. Digitized samples of the plucked strings of the constructed and traditional guitar were obtained. Conventional Fourier signal analysis of the notes from the three constructed guitars showed similarities to those of the traditional guitar. Other signal processing methods, such as short-time Fourier transform spectrograms, appropriate for the transient signal of the plucked string, were also investigated. Although the signals were broadly similar, analysis revealed differences, which might be attributed to the different geometries. A survey was conducted to assess listener's preferences for the various instruments. The survey was conducted from taped samples of the plucked strings to correlate the importance of psychoacoustics and the physical signal in assessing instrument quality. Listener preference rated the octagon-shaped guitar as highly as the traditional one. [Work supported by the Loyola College Department of Electrical Engineering and Engineering Science, and the Loyola College Hauber Fellowship Fund.]

2:15

4pMU6. The subjective evaluation of resonating sound-art samples and its relation to psychoacoustic measures. Densil A. Cabrera (Dept. of Architectural and Design Sci., Wilkinson Bldg., G04, Univ. of Sydney, NSW 2006, Australia, densil@arch.usyd.edu.au)

A genre of sound-based art is identified, wherein the artist simply records the sound of a continuously resonating acoustic system, with a minimum of intervention in the process. The sound of such works is often close to steady state, and tends to evolve slowly. The sounds include nearly pure tones, tone combinations (sometimes with prominent beats), and sustained noiselike sounds. Fluctuations due to variations in the energy source (often the wind) also occur. It was thought that the abstract nature and simplicity of these sounds would lead to subjective evaluations based largely on fundamental psychoacoustic criteria. Fifty-three samples, of 20 s duration, were arbitrarily selected from the genre, including works by Alvin Lucier, Gordon Monahan, and Roger Winfield. Subjects rated these stimuli on five continuous bipolar scales: (dis)like, (un)interesting, (un)pleasant, emotional arousal, and emotional valence. The stimuli were measured for loudness, pitch characteristics, roughness, fluctuation, and lateralization. Analysis showed emotional arousal to be the strongest response, with significant correlations to loudness, roughness, and pitch. The results point towards the possibility of constructing a perceptual model of this type of minimalist art form. [Work supported by an Australian Postgraduate Award.]

2:30

4pMU7. The dependence of timbre perception on the acoustics of the listening environment. Dae-Up Jeong (Dept. of Architectural and Design Sci., Wilkinson Bldg., G04, Univ. of Sydney, NSW 2006, Australia, jeong_d@arch.su.edu.au) and Fergus R. Fricke (Univ. of Sydney, NSW 2006, Australia)

The acoustic quality of a room for music cannot easily be determined. As timbre is important in music it is possible that the perceived change in timbral quality of a sound in two spaces may give the basis for a measure of acoustic quality. The smallest perceivable change in the timbre of steady-state stimuli was measured in different listening environments through two experiments featuring a spectral centroid and rise time, respectively. Five component harmonic stimuli at 440 Hz were employed. The centroid of the standard stimuli was fixed at 2.95, in which the median amplitude of each component of the standard was set at 0.5, while those of comparison stimuli were varied from 2.93 to 2.41 in steps of 0.02 with a negative spectral tilt. The smallest perceivable change in the rise time was also measured using complex tones containing ten harmonics of 200 Hz (F_0) with linear spectral slopes of -6 dB/oct. Rise times of comparison stimuli were varied from 5 to 25 ms in 2-ms steps, while the fall time was fixed at 5 ms across the stimuli. Results are presented and comments are made on how the acoustics of the listening environment influences the perception of timbre.

2:45

4pMU8. Perceptual analysis of synthesized struck bars. Vincent Roussarie (Inst. de Recherche et de Coordination Acoustique/Musique (IRCAM), 1 place Igor Stravinsky, F-75004 Paris, France), Stephen McAdams (IRCAM, F-75004 Paris, France), and Antoine Chaigne (Ecole Nationale Supérieure des Télécommunications (ENST), F-75013 Paris, France)

This study is part of a project seeking to understand musical timbre and perceptual phenomena of the sound environment. It concerns the sensitivity of listeners to sounds emitted by simple vibrating structures, such as xylophone bars. Previous studies of timbre have characterized this attribute in terms of a small number of perceptual dimensions of an analytic nature (attack time, spectral centroid, spectral flux). Other studies have suggested a return to the physical characteristics of the sound sources themselves in order to understand their influence on timbre perception. Sounds were synthesized with a computer model of the vibrating bar developed at ENST that allows precise control of physical parameters. Listeners made dissimilarity judgments on all pairs of different sounds. Multidimensional scaling analyses of these data yield a perceptual space with dimensions that correspond strongly to the varied physical parameters (material density and the first term of internal damping) as well as to salient auditory dimensions (pitch for the former and a combination of spectral centroid and decay rate for the latter). These results argue in favor of a privileged processing of source characteristics by the auditory system.

THURSDAY AFTERNOON, 25 JUNE 1998

CASCADE BALLROOM I, SECTION A (W), 1:00 TO 5:00 P.M.

Session 4pNSa

Noise and Architectural Acoustics: Surface Transportation Noise: Trains and Automobiles

Carl E. Hanson, Chair
Harris Miller Miller and Hanson, 15 New England Executive Park, Burlington, Massachusetts 01803-5221

Chair's Introduction—1:00

Invited Papers

1:05

4pNSa1. Diagnosis of noise sources from high-speed trains using the microphone-array technique. B. Barsikow (akustik-data Eng. Office, Kirchblick 9, D-14129 Berlin, Germany)

To reduce noise radiated by high-speed trains, abatement at the source is becoming increasingly important. Techniques for accomplishing this task require an exact knowledge of the individual sources involved and their reduction potential. During the past several years, a number of sound-generating components on ICE trains operated by the DB AG (German Railway) have been investigated by the akustikk-data Engineering Office. Microphone arrays provide a useful technology for investigating noise radiated by individual sources. Among the arrays used for recent measurements are the horizontal and vertical interspersed line arrays as well as X-shaped arrays. Examples measured by each of these instruments are shown. Results of these measurements with arrays yield the strength of each source and their speed exponents. In addition, narrow-band analysis of the sound is possible because the effects of the Doppler shift are not present. With this information, the importance of each source to the total radiated noise can be estimated. Examples of such results are given for rolling noise as well as for components emitting aerodynamic noise. For both types of sources, the effectiveness of various abatement measures is discussed.

1:25

4pNSa2. Noise prediction and control for new Norwegian high-speed trains. Matias Ringheim (Kilde Akustikkas, P.O. Box 229, N-5701, Voss, Norway)

The prediction of future noise levels is an important part of the planning process for high-speed trains and tracks. The first Norwegian high-speed tracks will be opened in 1998, using new 200-km/h electric trains based on the Swedish X2000. Strict noise criteria have been defined and extensive noise control measures are necessary both on the train itself and along the track. The noise prediction was performed several years ago, with the early and fairly simple Nordic prediction method. In the final planning stages some of the more critical results were checked against the revised Nordic method. The method enables the calculation of outdoor

energy equivalent levels and maximum levels in octave bands, in simple and complex terrain. The first version of a computer implementation is available. The outdoor results have a form which is suitable for the further calculation of indoor levels, when the building properties are known. The basic aspects of the method, including a summary of available high-speed data, will be presented. Problems related to the prediction and reduction of future noise from new high-speed trains will be summarized.

Contributed Papers

1:45

4pNSa3. Ground-borne vibration measurements of high-speed trains. David A. Towers (Harris Miller Miller & Hanson, Inc., 15 New England Executive Park, Burlington, MA 01803, dtowers@hmmh.com) and Hugh J. Saurenman (Harris Miller Miller & Hanson, Inc., Burlington, MA 01803)

Ground-borne vibration measurements of high-speed train operations were carried out in France for the TGV and Eurostar trains, in Italy for the Pendolino trains, and in Sweden for the X2000 trains. The measurements were made to develop vibration-prediction models for a new guidance manual published by the U.S. Federal Railroad Administration on ''High Speed Ground Transportation Noise and Vibration Impact Assessment.'' The tests included measurements of ground-borne vibration at various distances from the track as well as a procedure to characterize the ground vibration propagation characteristics at each measurement site. The results indicated a wide spread in the vibration data, partly due to differences in the equipment and track condition and partly due to differences in site geology. Further analysis suggested that much of the difference between the trainsets is due to variations in the geology rather than differences in suspension, axle load, or wheel conditions of the trainsets. Applying the characteristics of one type of trainset to different test sites resulted in widely varying ground vibration levels and propagation rates, whereas normalizing the ground vibration from the trainsets to one site substantially reduced the differences in overall vibration level.

2:00

4pNSa4. New noise impact criteria for high-speed ground transportation systems in the United States. Carl E. Hanson (Harris Miller Miller & Hanson, Inc., 15 New England Executive Park, Burlington, MA 01803, chanson@hmmh.com)

New noise impact criteria have been developed by the U. S. Department of Transportation for the assessment of environmental impacts of high-speed rail projects. The environmental impacts of new high-speed rail projects in the United States are the responsibility of the Federal Railroad Administration (FRA). Adoption of these criteria will assure a uniform treatment of noise impacts of new high-speed rail systems throughout the country. The noise impact levels were developed based on well-documented criteria, and on comprehensive social survey data concerning annoyance due to transportation noise. Impact levels are based on the Schultz response curve and include consideration of the combination of project and ambient noise levels. Correction terms are included in the case of high-speed rail (and maglev) to account for startle effects from fast rise times. This presentation discusses the basis for the curves used as thresholds of impact—where noise mitigation is to be considered—and severe impact—where noise mitigation must be implemented. These criteria are included in a new guidance manual published by the FRA on ''High Speed Ground Transportation Noise and Vibration Impact Assessment'' and will be used throughout the USA.

2:15

4pNSa5. Wayside noise measurements of high-speed trains. David A. Towers (Harris Miller Miller & Hanson Inc., 15 New England Executive Park, Burlington, MA 01803, dtowers@hmmh.com)

Wayside noise measurements of high-speed train operations were carried out in the USA for the German ICE, Swedish X2000, and U.S./French RTL-2 Turboliner trains, in France for the TGV and Eurostar trains, in Italy for the Pendolino trains, and in Sweden for the X2000 trains. The measurements in the USA were made as part of the Northeast Corridor Project and the measurements in Europe were made to develop noise-prediction models for a new guidance manual published by the U.S. Federal Railroad Administration on ''High Speed Ground Transportation Noise and Vibration Impact Assessment.'' The results indicate that the TGV, ICE, and RTL-2 trains had similar noise emissions and were typically the quietest of the trains tested. A-weighted noise levels for the X2000 and Pendolino trains averaged about 5 dB higher, and data for the Eurostar trains were scattered over the range for the other trains. Although the higher-speed French trains generated some additional low-frequency aerodynamic noise, this noise did not appear to significantly affect the A-weighted noise levels for train speeds up to nearly 300 kph (186 mph). The results also suggest that at the speeds tested, the observed trains were generally in compliance with the U.S. Federal Railroad Noise Emission Standard.

2:30

4pNSa6. Identification of moving acoustic sources from the pass-by noise. Byoung-Duk Lim and Deok-Ki Kim (Dept. of Mech. Eng., Yeungnam Univ., Gyongsan, Kyungbuk, 712-749, Korea)

The identification of a moving noise source such as a transportation system is usually difficult at high speed since microphones attached to the moving noise source may induce a high level of aerodynamic noise, while the measurement using a fixed microphone suffers from the distortion of the source characteristics due to its motion. In this study a time domain technique based on the finite difference method and regularization is developed for the recovery of the stationary source characteristics from the distorted signal measured at a fixed point. The sources are assumed to be point sources with stationary frequency characteristics, and move along a line at a constant subsonic speed. In the previous work of the present authors recovery of the stationary source signal from a point source requires the information about time origin. For multiple sources it is impossible to set a single time origin that a source signal recovery method using multiple time origins are developed. In the numerical simulations the effects of estimation errors in the parameters, such as the speed of source, time origin, and background noise level, are investigated and the results show the usefulness of this method.

2:45–3:00 Break

3:00

4pNSa7. The acoustical impact of local railway lines. Cristina Pronello (Ing. Cristina Pronello C.so Duca degli Abruzzi n. 24-10129 Torino, Italy, pronello@polito.it)

Local railway lines serve a regional traffic. They have few users since the service offered presents insufficient quality standards. These lines could have more users and they could become metropolitan connections by adopting new vehicles which are faster and more modern. This study has dealt with the current acoustical impact that they produce. Therefore, a number of acoustical measurements have been made along a local railway line. They have been carried out by varying the speed of the trains at a fixed measuring distance of 7.5 m and by varying the measuring distance with a single train speed. The obtained results have shown that current trains produce in the average a greater acoustical impact when compared to modern trains. This has been observed even at slow speeds and with reduced train lengths. Measurements made on a new train, that should soon start its service on the line, have shown that there would be no increase of the produced acoustical level, even with increased speeds. As a consequence, it may be concluded that the current service on the analyzed line could be enhanced by increasing the train speed and the number of runs, maintaining an equivalent daily analogous to the one currently guaranteed.

3:15

4pNSa8. Limits to limits? Tor Kihlman and Wolfgang Kropp (Dept. of Appl. Acoust., Chalmers Univ. of Technol., S-41296 Gothenburg, Sweden)

Today, traffic noise due to cars, trains, and airplanes is the main noise source in urban areas. In accordance with the European Commission's Green paper on Future Noise Policy, 22% of the European population are exposed to outdoor A-weighted noise levels higher than 65 dB and more than 45% are exposed to levels between 55 and 65 dB. This corresponds in total to about 250 million people. To improve the situation is a tremendous challenge and the question arises as to whether a good environment be provided (i.e., levels below 55 dB) for all people in today's cities. Can the demands of 55 dBs only be fulfilled by structuring traffic in a smart way and using noise barriers? Or is a clear change in traffic policy needed from individual traffic to public transport? Different cities with different structures and automobile dependencies are compared in a general study based on statistical data. Detailed information about traffic flow, population structure, and land use is analyzed to verify these results.

3:30

4pNSa9. Effects of traffic noise within the Madrid region. Manuel Recuero, Constantino Gil, and Jorge Grundman (Inst. de Investigacion del Automovil (INSIA), Univ. Politecnica de Madrid, Ctra. Valencia km. 7, 28031 Madrid, Spain)

In this work the results of a 20-question survey about the acoustic environment made on the population of 17 towns are presented. Up to 7141 questionnaires were distributed, where 3272 were conducted in population centers with more than 100 000 inhabitants, 2695 in towns below 100 000 and above 50 000 inhabitants, and 1174 in towns with less than 50 000 inhabitants. The aim of this statistical work is to estimate the citizens' opinions about noise sources in their municipalities, the annoyance that these sources produce, where and when the effect seems stronger, and how the noise affects residential areas. There were also questions concerning the opinions about protection against noise in dwellings and the presumed effects of environmental noise. Finally, it was intended to know how important the residents think the environmental noise problem is, how deep their knowledge is about their rights and the law, and what do they think about measures to improve the situation.

3:45

4pNSa10. Measuring the traffic noise in Valencia. E. Gaja, J. L. Manglano, A. Reig, and S. Sancho (Univ. Politécnica de Valencia, Camino de Vera S/N, 46071 Valencia, Spain)

The City of Valencia has approximately 700 000 inhabitants. Due to this large population, the city is subjected to heavy traffic, and, as a consequence, high acoustic pollution. The Townhall of Valencia is very worried about this acoustic pollution and has commissioned the Laboratory of Acoustics in the Universidad Politécnica de Valencia to work on various projects within the city. Due to the work the department has carried out on behalf of the Townhall of Valencia, it has gained experience measuring with the sonometer. An efficient way of measuring noise within the city that avoids wasting long periods of time has been developed. Continuous 24-hour measurements have been taken by sonometers located in different stations spread out around the city and controlled by the Townhall. With these measurements a method that indicates when to measure, in the morning, afternoon, and at night, as well as the exposure time, has been designed. Other parameters, such as characteristics of the traffic, types of streets, etc., have been considered in order to obtain the time and the exposure time of measuring that represents the level of the noise.

4:00

4pNSa11. Characterization of vehicle noise in Hong Kong. W. T. Ng, M. M. F. Yuen, and W. M. To (Dept. of Mech. Eng., Hong Kong Univ. Sci. Technol., Clear Water Bay, Kowloon, Hong Kong, mewmto@usthk.ust.hk)

There is a common perception among residents of high-rise buildings facing expressways that noise emitted from moving vehicles would be dominated by tire/road interaction and at high-frequency regions. Local government officials believe that the impact on residents of low-frequency noise radiated from car engines would be insignificant, especially after the corrections for the *A*-weighting scale are applied. This paper reports on an extensive survey and some *in situ* sound measurements. It was found that road noise in Hong Kong is dominated by low-frequency noise emitted from heavy vehicles in expressways and by low-frequency noise emitted from heavy and light vehicles driven at a speed below 50 km/h in streets. The measured data were characterized by using time-averaged 1/1 octave band analysis, time-averaged 1/3 octave band analysis, and time-frequency analysis. The Doppler effect was clearly observed at low-frequency regions in joint time-frequency distributions. It is suggested that the *A*-weighting scale would not reflect the true annoyance level of traffic noise. Noisiness should be used to quantify the annoyance caused by moving vehicles. A new noise model is proposed to give a realistic description of noise radiated from a moving vehicle.

4:15

4pNSa12. Multivariate analysis of road traffic noise in Gandia (Spain) during 24 hours and its evolution in the last decades. Jose Romero (Lab. Acoust., Dept. Fisica Aplicada U.V., Facultad de Ciencia Fisicas, Valencia, Spain, Romerof@arrakis.es), Alicia Jiminez, Antonio Sanchis, Albert Marin (Polytecnic Univ. of Valencia, Spain), Amando Garcia (Lab. Acoust., Valencia, Spain), and Grover Zurita (Lulea Univ. of Technol., Sweden)

Road traffic noise is one of the most widespread and growing problems in urban areas. While it has long been known that hearing can be damaged by exposure to noise, it is also believed that continual noise, even at low or moderate intensity, can cause psychological discomfort and sleep disorders. Traffic noise levels in urban areas are increasing, and the areas affected by noise are spreading. As a result of these problems, legislation has been introduced by the city council to control the noise produced by individual vehicles with the aim of eventually producing traffic noise levels which are acceptable to the public. In this paper, therefore, aspects which connect the evolution of the road traffic noise in the last decades and the development of the infrastructure of the city are discussed. The study was carried out in Gandia (Spain), which is situated on the East Coast of the country, 65 kms from Valencia and 100 kms from Benidorn. The population has increased from 50 000 to 60 000 in the last 15 years, and the traffic distribution of the city has changed during this time, with a new bridge and some highways around the urban area. The main objective of this paper is to characterize the effects of the road traffic noise during 24 h (between 1983 and 1997), and including such aspects as traffic density and psychological aspects. The analysis part was performed by using multivariate analysis methods (MVA). Multivariate analysis methods can be used to investigate relationships between all the variables by extracting information from data with many variables and treating them simultaneously.

4:30

4pNSa13. Criteria and control for environmental noise emitted from motor vehicles. Ren Wentang (Noise Vib. Control Key Lab., Beijing Municipal Inst. Labour Protection, No. 55 Tao Ran Ting Rd., Beijing, 100054, China)

The present work investigated the environmental noise in a city from energy distribution and annoyance, and showed that the noise from different kinds of in-use vehicles is the main cause of noise pollution, especially in a city in a developing country. The technology and measures of noise control for in-use vehicles are urgently needed. According to measurement results of noise emission from 1000 vehicles based on ISO 5130, the statistical distribution parameters of stationary noise emission for six types of new vehicles and in-use vehicles are obtained. The statistical parameters show that the technology and measures of noise control for exhaust systems of in-use vehicles are key points. Some technological advances and noise-control measures have been proposed and practiced in China. A measurement procedure and system for vehicles in stationary operation is developed to distinguish some noisy vehicles with lost efficacy of a muffler. Regulations concerning stationary noise limits have been published and followed in China subsequent to the investigation. Reasonable design parameters of exhaust systems for different kinds of vehicles are proposed.

4:45

4pNSa14. A mathematical model for the evaluation and prediction of the mean energy level of traffic noise in Caracas. Nila Montbrun, Victor Rastelli, Alexis Bouza (Dept. de Mecánica, Univ. Simón Bolivar, A.P. 89000 Caracas 1080-A, Venezuela), Jenny Montbrun Di Filippo, and Yamilet Sánchez (Univ. Simón Bolivar, A.P. 89000, Caracas 1080-A, Venezuela)

The increase of the population in capital cities has provoked a serious increase in the acoustical pollution, where the real cause is the traffic noise. Over ten years, sound levels have been measured in several points in the city of Caracas. The main objective was to establish a mathematical model for the mean energy level of traffic noise. This model is a modification and improvement of the model that was used in Germany and Italy. Parameters of the fundamental equations were obtained by the use of large experimental data. In the particular case of Venezuela, an innovation was introduced by the consideration of the motorcycle noise. The obtained model can be used to predict noise pollution in cities and contour maps can be made as well.

THURSDAY AFTERNOON, 25 JUNE 1998

CASCADE BALLROOM I, SECTION C (W), 1:10 TO 3:45 P.M.

Session 4pNSb

Noise and Signal Processing in Acoustics: Active Noise and Vibration Control

Clemans A. Powell, Chair
NASA Langley Research Center, MS 462, Hampton, Virginia 23681-0001

Chair's Introduction—1:10

Contributed Papers

1:15

4pNSb1. Active local noise control in open space. Jingnan Guo and Jie Pan (Dept. of Mech. and Mater. Eng., Univ. of Western Australia, Nedlands, WA 6907, Australia)

Local noise control, which creates quiet zones in desired areas, is a practical option of active noise control in open space when the control sources and the primary sources are not able to be very close. In most cases, the local control system attenuates the sound pressure in some areas at the cost of total sound power output increase. As a result, while the control system creates a quiet zone in some areas, it causes the sound pressure to increase in others. Two indicators of control efficiency, quiet zone size and total sound power output, are dependent upon the configuration of the control system. It has been found that for the multichannel control system, there exists a range of optimal configuration of control system, in which the control system can create the largest quiet zone with the lowest increase of total sound power output. Outside this range, the control system either becomes useless in creating large quiet zones, or causes great increase of sound pressure in most areas.

1:30

4pNSb2. Adaptive control of sound transmission with neural network algorithms. Jing Tian, Hai Lin, and Mingkun Cheng (Inst. of Acoust., Academia Sinica, Beijing 100080, PROC)

The Artificial Neural Network is widely used for its self-learning, self-organizing adaptability as well as nonlinearity in nature. Active control of structural sound radiation and sound transmission through panels is widely researched recently for its effectiveness and efficiency in the low-frequency range. In this paper, a neural network-based feedforward adaptive controller using the modified Error Back Propagation Learning Algorithm is presented. The controller is realized with a TMS320C25 DSP Board monitored by an IBM PC compatible, and applied to control the sound transmission through a thin panel between two rooms. Experiments showed that the algorithm was superior in robustness and broadband performance. In the adaptive controller, the structure of a traditional multilayered feedforward neural network (MFNN) is modified. Both filtered-X and filtered-U adaptive filters are discussed. Two auxiliary filters are designed to compensate the secondary signal feedback and the error delay. Random gradient estimation method is used to update the weights of MFNN. Several methods to speed the learning rate are also introduced. The causality of the feedforward controller is analyzed, by which the system behavior is greatly improved. Further strategy to improve the controller is also proposed. [Project supported by the Natural Science Fund of China.]

1:45

4pNSb3. Real-time wave separation in a cylindrical pipe with applications to reflectometry and echo-cancellation. Jean Guerard and Xavier Boutillon (Lab. d'Acoustique Musicale, CNRS-Univ. Paris 6, Case 161, 4 place Jussieu, 75252 Paris Cedex 05, France)

Precise control of the acoustical pressure all along a pipe requires that backward and forward propagating waves are separated in real time. This has been achieved by means of an array of up to five microphones, a high-precision analog calculation, and a real-time numerical post-treatment. The separation factor exceeds 40 dB over several kHz, when

special attention to intercalibration issues and to the realization of the analog part is given. The theoretical limit at high frequencies has been reached experimentally. The advantage of a partial analog treatment over a fully numerical one will be discussed. One application consists in reflectometry (and consequently, impedancemetry) without use of the long pipe normally required. Associated with a real-time DSP processor (TMS 320C30), wave separation also permits one to efficiently control the wave reflected by one end. This is done with only one loudspeaker, whose transfer function has been inverted and implemented in the DSP as an IIR filter. This approach of active-control differs from adaptative procedures. Finally, a musical application of the system to a hybrid wind instrument will be presented. [Work supported in part by the French Ministry of Culture.]

2:00

4pNSb4. Active noise control of a single-engine light aircraft cabin. Colin D. Kestell and Colin H. Hansen (Dept. of Mech. Eng., Univ. of Adelaide, Adelaide, SA 5005, Australia, cdkestel@watt.mecheng.adelaide.edu.au)

Active noise control (ANC) shows success and potential in a growing number of commercial applications, one of which is aircraft cabin noise reduction. With the exception of utilizing ANC headsets, light aircraft, which to date offer a high noise environment, have been somewhat overlooked. The importance of weight minimization also prevents installing copious quantities of dampening and insulation materials as a passive noise control measure. While headsets are a pilot's necessity and an obvious target for ''localized'' noise reduction, they are not conducive to either operator or passenger comfort. High noise levels not only render communication difficult but also contribute toward stress and fatigue. A more globalized region of reduced noise will be less restrictive and no doubt provide the occupants with more freedom of movement and overall comfort. Light aircraft operators boasting quieter cabins with a focus on customer comfort will no doubt have a distinct commercial advantage. Using flight trials and laboratory experiments as a basis, this paper will discuss the introduction of ANC into a four-seater Piper Archer. Existing noise levels, objectives, equipment used, methods of approach, and results to date shall be reviewed, as well as the remaining work required to achieve the final goal.

2:15

4pNSb5. Control of exhaust noise from diesel electric locomotives using a hybrid active/passive system. Paul Remington (BBN Technologies, 10 Moulton St., Cambridge, MA 02138), Douglas Hanna, and Scott Knight (BBN Technologies, New London, CT 06320-6147)

Diesel electric locomotives in the US represent a significant source of environmental noise under a variety of operating conditions. There are two primary sources of noise on these locomotives that will have to be reduced before significant reductions can be achieved: exhaust noise and cooling fan noise. Here a system is proposed that uses both active and passive components to control broadband exhaust noise below 3 kHz. The active system utilizes eight actuators and eight residual microphones in a feedforward configuration to control tonal noise below 250 Hz. The passive system is a compact exhaust silencer designed to control exhaust noise above that frequency while still fitting within the limited space available beneath the locomotive hood. Performance estimates for both the active and passive system will be presented along with the results of field tests of a breadboard active system on an F59PHI locomotive. [Work was supported under contract to the Federal Railroad Administration.]

2:30–2:45 Break

2:45

4pNSb6. Anticipated effectiveness of active noise control in propeller aircraft interiors as determined by sound quality tests. Clemans A. Powell and Brenda M. Sullivan (NASA Langley Res. Ctr., M.S. 462, Hampton, VA 23681-0001, c.a.powell@larc.nasa.gov)

Active noise control technology offers the potential for weight-efficient aircraft interior noise reduction, particularly for propeller airplanes. However, there is little information on how passengers respond to this type of interior noise treatment. This paper presents results of two experiments which use sound quality engineering practices to determine the subjective effectiveness of hypothetical active noise control systems in a range of propeller airplanes. Binaural recordings were made in five different propeller airplanes and a commercial jet using an acoustic mannequin and a digital recording system. The noise stimuli for the tests were prepared by digitally modifying the propeller tones in recordings of each airplane type to simulate active noise control systems with a range of effectiveness and complexity. The noise stimuli were presented to the test subjects through electrostatic headphones to preserve the realism, spatiality, and directionality provided by the binaural system. The two tests differed by the type of judgments made by the subjects: pair comparisons in the first test and numerical category scaling in the second. Changes in subjective response relative to changes in measured sound characteristics afforded by the hypothetical active noise control schemes are presented and are contrasted between the two test methods.

3:00

4pNSb7. Robust feedback active noise control algorithm for impulsive additive noise. Sang-Wook Lee and Koeng-Mo Sung (School of Elec. Eng., Seoul Natl. Univ., Seoul 151-742, Korea, lsw@acoustics.snu.ac.kr)

Active noise control (ANC) is a method to reduce the unwanted noise level by introducing secondary noise which has the same amplitude and antiphase with the primary noise. The feedback ANC algorithm is a kind of linear predictor and it uses only one sensor to get the information about the error signal and reference signal. The filtered-x LMS (least mean square) algorithm is widely used for implementing the ANC system because of its simplicity and good performance. But when the additive noise is impulsive, the performance of LMS-type algorithms are known to be lowered. In this paper, the robust feedback ANC algorithm is proposed when impulsive additive noise is present. Instead of the LMS algorithm, the proportion-sign algorithm (PSA) which is a mixture of LMS and dual sign algorithm (DSA) is used to estimate the primary noise field robustly when impulsive additive noise is present. Computer simulations were performed under impulsive additive noise circumstances and showed better performance with the proposed algorithm than that with the conventional LMS algorithm.

3:15

4pNSb8. Active control of structural sound using an active constrained layer damping treatment. Jun Yang, Jing Tian, and Xiaodong Li (Inst. of Acoust., Academia Sinica, Beijing, 100080, PROC)

In this paper active constrained layer damping (ACLD) treatments are applied to control the noise transmitted into a cavity through a panel. Five of the enclosure walls are rigid and the other wall, which is flexible, is a simply supported panel. The panel is treated with ACLD, which consists of a viscoelastic shear layer sandwiched between a piezoelectric constraining cover sheet and the panel. The piezoelectric layer is utilized to control the panel vibration and the shear deformation of the viscoelastic layer to enhance its energy dissipation characteristics. The coupled structural sound equation of the composite panel and the cavity is derived and analyzed by the Galerkin approximation. The results indicate that combining active and passive controls, as in the case of ACLD treatment, is more effective in controlling the noise transmission. Moreover, such a control strategy is found to require lower control inputs than when the active and passive control actions are used separately.

3:30

4pNSb9. Active noise barrier efficiency improvement using multirate signal processing. Jelena Certic, Slobodan Kovacevic, and Petar Pravica (Univ. of Belgrade, Dept. of Elec. Eng., Bulevar Revolucije 73, Belgrade, Yugoslavia, cetic02996P@buef31.etf.bg.ac.yu)

In this paper, implementation of multirate signal processing techniques in an active noise barrier system was considered. Such a modification of a conventional multiple-error filtered-x LMS algorithm is supposed to give improved overall system performance, by separately controlling two different frequency ranges. The control system consists of one detection microphone, three secondary sources, and three error sensors. The primary signal was a random noise signal with a spectrum that coincides a typical road traffic noise spectrum. Locations of components were the ones found to be optimal in our previous work. Computer simulation was developed to compare the attenuation achieved with a control system that uses a multiple-error filtered-x LMS algorithm and the one involving multirate signal processing. Results of the computer simulation confirmed that attenuation is increased using multirate signal processing.

THURSDAY AFTERNOON, 25 JUNE 1998 ASPEN ROOM (S), 1:00 TO 3:45 P.M.

Session 4pPAa

Physical Acoustics and Bioresponse to Vibration/Biomedical Ultrasound: Cavitation Dynamics: In Memoriam Hugh Flynn II

Charles C. Church, Cochair
Acusphere, Inc., 38 Sidney Street, Cambridge, Massachusetts 02139

Werner Lauterborn, Cochair
University of Goettingen, Drittes Physikalisches Institute, Buergerstr. 42-44, D-37073 Gottingen, Germany

Invited Papers

1:00

4pPAa1. Acoustically induced cavitation fusion. Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

In 1982, Hugh Flynn was issued a patent (No. 4,333,796) for a Method of Generating Energy by Acoustically Induced Cavitation Fusion and Reactor Therefor. Although it was largely ignored at the time, there have been several recent papers that treat the subject of acoustically induced fusion as quite plausible. Such prescience in the area of cavitation research was typical of Hugh's work. This patent shows such an enormous grasp of detail, and suggests Hugh must have given a great deal of thought and energy to its composition. The author will review some of the interesting aspects of this patent as well as describe some of the most recent activity on this topic. He will also reflect on his many rewarding and stimulating experiences with Hugh over the past 30 years.

1:20

4pPAa2. The search for cavitation *in vivo*. Edwin L. Carstensen, Sheryl Gracewski, and Diane Dalecki (Depts. of Elec. Eng. and Mech. Eng. and the Rochester Ctr. for Biomed. Ultrasound, Univ. of Rochester, Rochester, NY 14627)

Until the mid-1970s, it was generally assumed that with the short pulses of ultrasound used in medical diagnosis, there was little need for concern about the possibility of inertial cavitation *in vivo*. Conversations with Hugh Flynn at about that time made it clear that complacency in this respect was not warranted. Flynn carried out a comprehensive theoretical study of the response of microbubbles to diagnostically representative pulse exposures culminating in a publication, which is now the classic basis for an understanding of cavitation in diagnostic medicine. Robert Apfel independently published similar conclusions at about the same time. Flynn's contributions continued in a collaboration with Charles Church. Experimental evidence that cavitation *in vivo* might be relevant to the use of diagnostic ultrasound began with observations of the killing of fruit fly larvae with microsecond length pulses. Later, the mammalian lung was shown to be particularly vulnerable to these exposures. Noting that purely positive pulses were as damaging as purely negative pulses, Ding and Gracewski showed that inertial cavitation is greatly inhibited by bubble constraints that might occur in tissue. Most recently, hemolysis and hemorrhage associated with the use of contrast agents have provided nearly incontrovertible evidence of the occurrence of cavitation *in vivo*.

1:40

4pPAa3. Nonlinear dynamics of bubbles with surfactants. Robert Apfel, Xiaohui Chen, and Jeffrey Ketterling (Yale Univ., New Haven, CT 06520-8286)

Hugh Flynn pioneered many aspects of cavitational phenomena. He even got a patent on the idea that large bubble collapse could be used for inertially induced thermonuclear fusion. Some researchers currently working in the area of sonoluminescence are asking the same questions. One of the key requirements for confining the energy in the collapse of a bubble is maintaining spherical symmetry. Instabilities, such as Rayleigh–Taylor and Kelvin–Helmholtz, can compromise this requirement. Therefore it is useful to examine the role of surfactants in conferring stability on the collapse of bubbles. Codes that have recently permitted the description of superoscillations of liquid drops with surfactants in air [e.g., Apfel *et al.*, Phys. Rev. Lett. **78**, 1912–1915 (1997)] can be applied to the case of a bubble in a liquid. Of particular interest are insoluble surfactants, such as bovine serum albumin, which do not reduce

the surface tension that much, but which greatly increase the surface damping. The role of damping in discouraging or delaying instability of collapsing bubbles in an acoustic field will be presented. [Work supported in part by NASA through Contract No. 958722 managed by the Jet Propulsion Laboratory.]

2:00

4pPAa4. Response curves of bubbles. Werner H. Lauterborn and Robert Mettin (Drittes Physikalisches Institut, Universität Göttingen, D-37073 Göttingen, Germany)

A spherical bubble in a liquid is a nonlinear oscillator that can be set into radial (and surface) oscillations by a sound field. The nonlinear response depends on a set of parameters, mainly the bubble radius, the driving amplitude, and the driving frequency. The response is involved. It consists of an intertwined set of nonlinear resonances, that besides turning over, comprises period doubling sequences to chaos when going up and down in bubble radius, driving frequency, or driving amplitude. A specific feature is the giant resonance appearing for small bubbles and being the host for single bubble sonoluminescence. A set of response curves will be given to acquaint the listener with some understanding of how bubbles behave in a sound field.

2:20–2:30 Break

Contributed Papers

2:30

4pPAa5. Bubble dynamics in non-Newtonian fluids. John Allen (Dept. of Mech. Eng., Univ. of Washington, Seattle, WA 98195) and Ronald A. Roy (Boston Univ., Boston, MA 02215)

Nonlinear oscillations of gas bubbles in viscoelastic non-Newtonian fluids remains a relatively unexplored area, but one of increasing importance with the growing use of high intensity ultrasound and contrast agents for imaging enhancement. Previous work by the authors focused on an analytical and numerical study of bubbles governed by linear viscoelastic constitutive equations. This work was constrained by the small deformation assumptions inherent in the equations. New numerical studies of objective viscoelastic constitutive equations reveal the limitations of the linear viscoelastic models and show many novel results. In particular, the risk of bioeffects from medical ultrasound and harmonic imaging applications are stressed. Questions about the trace of the stress tensor are also discussed. [Work supported by NIH and DARPA.]

2:45

4pPAa6. Giant resonance in the dynamics of small bubbles. Iskander Sh. Akhatov (Ufa (Bashkortostan) Branch of Russian Acad. of Sci., K. Marx St. 6, Ufa, 450000, Russia, iskander@ncan.ufanet.ru), Claus-Dieter Ohl, Robert Mettin, Ulrich Parlitz, and Werner Lauterborn (Universität Göttingen, Göttingen, Germany)

For bubble oscillations under medium and large pressure amplitudes a complicated scenario of bifurcations and coexisting (chaotic) attractors exists. However, for very small bubbles in a very strong sound field the dynamics becomes regular and a new type of strong resonance with a thresholdlike increase in oscillation amplitude occurs [E. A. Neppiras and B. E. Noltingk, Proc. Phys. Soc. London Ser. B **64**, 1032 (1951); H. G. Flynn, in *Physical Acoustics Vol. 1*, edited by W. P. Mason (1964), p. 57; W. Lauterborn, Acustica **20**, 105 (1968)]. This phenomenon has a strong influence on many properties of cavitation bubbles. The following aspects are considered: rectified diffusion, stability of the bubble under SBSL conditions, primary and secondary Bjerknes forces, and interpretation of bubble size measurements.

3:00

4pPAa7. Cavitation bubble dynamics induced by ultrasound waves. J-L. Laborde (EDF-DER ADEI, Rte. de Sens-Ecuelles, 77818 Moret-sur-Loing Cedex, France, jean-luc.laborde@edfgdf.fr), C. Bouyer (EDF-DER ADEI, 77818 Moret-sur-Loing Cedex, France), J-P. Caltagirone (ENSCPB-MASTER, 33402 Talence Cedex, France), and A. Gérard (LMP, Université de Boreaux I, 33405 Talence Cedex, France)

Propagation of power ultrasound (from 20 to 800 kHz) through a liquid initiates a not-so-well-known phenomenon called acoustic cavitation. Inceptions and germs grow into bubbles which collapse, possibly giving rise to extreme conditions of temperature and pressure (assessed to be up to 10 000 K and 500 atm). For instance, these conditions initiate and greatly enhance chemical reactions. A high-speed film shot at 500 fps clearly identifies stable and transient cavitation and shows bubble population phenomena. Clouds of bubbles grow up to ten times their emergence size during 15 ms, and move at velocity around 50 cm/s. Mathematical modeling is performed as a new approach to predict where active bubbles are and how intense cavitation is. Then, a first computation, based on Euler's linear acoustic equations, is used to calculate the pressure field in the case of a cylindrical sonoreactor. A second calculation based on fluid dynamics equations is undertaken as CFD codes are very interesting because they also provide velocity and temperature fields, and two phase flows (liquid and bubbles) could be modeled. The comparison with experimental observations (photographs, high-speed film) and measurements (Particule Image Velocity, temperature) shows good agreement with both calculations.

3:15

4pPAa8. Cavitation and capillary wave from a parametric decay scheme. Masanori Sato (Honda Electron. Co., Ltd., 20 Oyamazuka, Otwa-Cho, Toyohashi, Aichi, 441-31 Japan) and Toshitaka Fujii (Toyohashi Univ. of Technol., Hibari-ga-oka, Tempaku-Cho, Toyohashi, Aichi, 441 Japan)

Acoustic cavitation in liquids and capillary waves which cause ultrasonic atomization are considered to be related to the 1/2 subharmonic of ultrasonic surface waves via parametric decay instability. The cavitation bubble oscillation mode is an asymmetric surface bubble oscillation that does not easily emit ultrasonic waves into water and accumulates acoustic energy from longitudinal waves. The capillary waves accumulate acoustic energy in a surface of liquid, enough to be used not only for atomization but also the separation of water-ethanol solution. These were confirmed by experiments using ultrasonic longitudinal waves in water.

3:30

4pPAa9. An optical transducer for the study of the pressure field around a collapsing cavitation bubble. David C. Emmony and Robin D. Alcock (Dept. of Phys., Loughborough Univ., Loughborough, Leicestershire LE11 3TU, UK, DCEmmony@lboro.ac.uk)

An optical reflection transducer using a critical angle technique has been built to study the pressures developed around a single cavitation bubble in water. The bubble was generated at different distances from a planar glass surface, and total internal reflection at the glass water interface gives the pressure through the change in refractive index in both the water and the glass. A simple hemispherical pmma block system has been built to prove the feasibility of the system. In this case the change of the critical angle is not linear, but the use of a modified equilateral glass prism gives the pressure as a function of time around cavitation bubbles with maximum radii of approximately 1.0 mm and gamma values down to 0.3 with a temporal resolution only limited by the photodiode detection electronics (~10 ns). The transducer shows the development of positive pressure as well as the migration of the bubble to the surface and is capable of recording the rarefaction waves due to free surface reflection.

THURSDAY AFTERNOON, 25 JUNE 1998 METROPOLITAN BALLROOM (S), 12:45 TO 3:45 P.M.

Session 4pPAb

Physical Acoustics: General Topics (Poster Session)

Michael R. Bailey, Chair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

Contributed Papers

All posters will be on display from 12:45 p.m. to 3:45 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 12:45 p.m. to 2:15 p.m. and contributors of even-numbered papers will be at their posters from 2:15 p.m. to 3:45 p.m. To allow for extended viewing time, posters will remain on display until 12:00 noon on Friday, 26 June.

4pPAb1. Enhancement of acoustic cavitation effects by simultaneous multifrequency excitation. Qiping Fang, Beixing He, and Zhongmao Lin (Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC, qpfang@public.fhnet.cn.net)

Acoustic cavitation can produce many effects. Promoting or inhibiting cavitation depends on different practical purposes. The production of as many cavitation activities as possible is always expected in sonochemistry. Here, the enhancement of acoustic cavitation effects by simultaneous multifrequency excitation, with frequencies in the power ultrasound ranges (20–100 kHz), is studied. One ultrasonic transducer is used to produce multiple ultrasonic frequencies at the same time, rather than several transducers which work at respective frequencies. A probe hydrophone is utilized to measure the subharmonic emission of cavitation and distribution of the acoustic field. The chemical effect of cavitation is determined by the method of iodine release. The results show that simultaneous multifrequency excitation can reduce the cavitation threshold and enhance the chemical effect of cavitation. On the basis of the cavitation bubble dynamics theory, the theoretical analysis proposed attempts to explain this phenomenon, and further discussions are presented. [Work supported by National Natural Science Foundation of China.]

4pPAb2. Nonlinear wave propagation in cylindrical air-filled tubes. Ludovic Menguy and Joël Gilbert (Inst. d'Acoustique et de Mécanique, LAUM UMR CNRS 6617, Ave. O. Messian, BP 535, 72017 Le Mans Cedex, France)

Three dimensional conservation laws are reformulated in a dimensionless form for a cylindrical air-filled tube. Three nondimensional numbers therefore appear, and are used to evaluate the order of magnitude of each term. Simplifying leads to two quasi-unidimensional generalized Burgers equations, which take into account both nonlinear phenomena and viscothermal losses. Special attention is paid to the validity of the different approximations we proposed. The method of multiple scales indicates that the interaction between the two opposite propagative waves is noncumulative and remains a *local* effect. The two Burgers equations are therefore *globally* independent. A numerical computation in the frequency domain solves the Burgers equations in the case of a stationary wave. The signal is supposed to be known at the two ends of the tube (or impedance instead of signal for one extremity). The computation uses the harmonic balance method to calculate the two opposite propagative waves which makes it possible to verify the right conditions in the two extremities. The signal everywhere in the tube is so pieced together. Some experiments are performed with a siren as source in one extremity, and measurements are compared with theoretical results.

4pPAb3. Numerical simulation with experimental validation for nonlinear standing wave phenomena. Zhichi Zhu, Anqi Zhou, Dongtao Huang (Dept. of Eng. Mech., Tsinghua Univ., Beijing, 100084, PROC, zzc-dem@mail.tsinghua.edu.cn), and Ke Liu (Inst. of Acoust., Academia Sinica, Beijing 100080, PROC)

The study of nonlinear standing waves is of great significance to nonlinear acoustics. The strong nonlinear standing wave phenomena are constantly encountered in the experimental studies of acoustics, such as multiplying growth of higher harmonics, saturation of harmonics, bifurcation, and chaos. The experimental study of strong nonlinear standing waves is highly restricted by the sound source, which is very difficult to build up when its sound power becomes very strong. Therefore, in this paper, a numerical approach to the problem was adopted. By the use of Euler equations and MacCormack fourth-order difference method, the multiplying growth and saturation of higher harmonics in a nonlinear standing wave have been numerically simulated. By increasing the sound-pressure level of the excited source from 125 to 190 dB, the entire development

process of the nonlinear standing wave was clearly illustrated. Some new interesting results were obtained from the simulation. For example, the sound-pressure level at zero frequency enlarges quickly with increasing the intensity of the excited source. A detailed comparison between numerical simulation and the relevant experimental results shows that the numerical investigation is successful. [Work supported by NSFC.]

4pPAb4. Acoustics wave propagation from an underground waveguide. Stanislav A. Kostarev (Tunneling Assoc., Sadovaja-Spasskaja 21, 107217, Moscow, Russia)

A simple enough model of an underground acoustic waveguide is proposed. A special structure of nonhomogeneity of the subsurface layer in the soil may lead to the capture of acoustic waves within a channel. In this case, a sound signal will propagate with a sufficiently small decay in comparison with homogeneous media (due to the lack of discrepancy of the wavefront). In this case vibration is observed at large distances from the source. A maximum value of the acoustic level is reached near the axes of the waveguide. In practice, this effect may be significant for the case of vibration generation for buildings on the piles. Under these conditions, enhanced values of vibration may take place even when the vibration of the ground surface is negligible.

4pPAb5. Diffraction of a sound wave on an open end of a half-infinite waveguide with impedance walls and impedance flanges. Evgeni L. Shenderov (Res. and Development Inst. ''Morfizpribor,'' 46, Chkalovski pr. 197376, St. Petersburg, Russia, shend@fs.spb.su)

Diffraction of a plane sound wave on an open end of a waveguide with impedance walls, connected with impedance flanges, is considered. A plane wave is incident upon the waveguide from open space. As a result of diffraction, part of the sound energy is scattered in the half-space and the other part of the energy penetrates into the waveguide. Results of the solution of this problem can be used for calculation of the amplitude of a sound wave, scattered by a hole, in particular, in the backward direction. Besides, one can calculate the sound energy penetrating into the waveguide. Absorption of the energy in the waveguide can be taken into account by introducing impedances with nonzero real parts. The distortion of the directivity pattern of a system of receivers located in the waveguide can also be obtained. The solution of the problem has been carried out by the integral equation method using Green's function of the half-space with an impedance boundary. The equation was solved by expanding an unknown function in the eigenfunctions of the waveguide. It reduces to an infinite set of linear algebraic equations. Results of computation of bistatic diagrams of scattering on the hole and diagrams of backward scattering are presented.

4pPAb6. Dependence on waveguide width of elastic convolver efficiency. Yasuhiko Nakagawa and Suguru Kitabayashi (Faculty of Eng., Yamanashi Univ., Kofu, 400 Japan)

An elastic convolver based on a nonlinear effect in surface acoustic wave (SAW) on a Y-Z·$LiNbO_3$ is widely used as a functional device in the area of communication system. Therefore, the efficiency, which is an important parameter for the SAW convolver, is quite low. There have been many reports on electrode structures and convolver materials for improving the efficiency. In this work, theoretical and experimental results are reported for the optimum width of a ($\Delta = v/v$) waveguide which has high convolver efficiency. The theoretical calculation was obtained by using the scalar potential approximation. The efficiency has a peak at a waveguide width of around 1λ. In order to increase the power density multistrip beam width compressors have been used. Performance of a device having a 4-μs waveguide length, a central frequency of 43 MHz and using 10:1 beam width compressors is described. The validity of the theory is proven experimentally, and, therefore, the optimum waveguide width for the convolver efficiency in a conventional elastic convolver is around one wavelength of the input SAW.

4pPAb7. A new time-domain approach for nonlinear wave propagation: Comparison with the KZK equation approach in the case of an unfocused cw beam. Jahangir Tavakkoli (Inst. of Biomed. Eng., Univ. of Toronto, Toronto, ON M5S 3G9, Canada), Oleg A. Sapozhnikov (Moscow State Univ., Moscow 119899, Russia), Remi Souchon, and Dominique Cathignol (INSERM, 69424 Lyon Cedex 03, France)

A time-domain numerical model for simulating acoustic pulse focusing in soft tissue has already been presented. In this model, the main effects responsible for finite-amplitude wave propagation, i.e., diffraction, frequency-dependent attenuation, and nonlinearity, were taken into account. Here, a comparison between the results of this model with those of the KZK model is presented. Using this model, the acoustic pressure field of a plane circular transducer was simulated in water for a large range of cw-mode input excitation levels. The on- and off-axis pressure-time waveforms and their corresponding harmonic components were calculated using this model and the KZK model, and very good agreement was obtained between the two models. Furthermore, these results were verified by previously published experimental measurements. At present, all implementations of the KZK model are limited to planar or weakly focused sources. This study, along with this group's previous work, reveals the validity of this finite-amplitude wave propagation model for a wide degree of source focusing (from a planar to a highly focused source), and in cw as well as PW modes. It is also shown that the computation time for this model is much less than for the KZK, using the same input parameters.

4pPAb8. Description of nonlinear planar traveling waves in a gas filled tube. Michal Bednarik (CTU-FEE, Technicka 2, 16627 Prague, Czech Republic)

This paper deals with problems for the description of nonlinear planar traveling waves in a gas-filled tube. These waves, which are influenced by the tube-wall boundary layer, can be described by both the Burgers equation and the Khokhlov–Zabolotskaya–Kuznecov's equation with the inclusion of a hereditary integral known as the fractional derivate of order $\frac{1}{2}$. Emphasis is placed on comparison of applicability of both the mentioned equations with respect to experimental data. The applicability of the Burgers model equation is limited by the cutoff frequency of the tube which is exceeded due to the growth of higher harmonic components that is caused by the nonlinear distortion of the primary harmonic shape of waves. The next part of this contribution is dedicated to the presentation of some new approximate solutions of the Burgers equation with a fractional derivate and a brief description of used numerical methods.

4pPAb9. Vibratory gyro-sensor using a trident tuning fork resonator with lateral width set in parallel with a rotationary axis. Yoshiro Tomikawa, Toshiharu Ogasawara, and Norikazu Ishida (Dept. of Elec. Eng. and Information Sci., Yamagata Univ., 4-3-16 Johnan, Yonezawa, Yamagata, 992 Japan)

This paper deals with a new type of piezoelectric vibratory gyro-sensor using a trident tuning fork resonator, which is supported laterally at its base portion, the width direction of which is in parallel with a rotationary axis. The vibratory gyro-sensor is an important and attractive device as an angular velocity sensor, especially, in the vehicle stability control (VSC) system schemed for a practical use in the not so far future. Therefore, many researchers have been engaged in development of such a key device in the VSC system. Some types of piezoelectric vibratory gyro-sensors have also been investigated; for example, two types of vibratory gyro-sensors have been studied using a trident tuning fork resonator: one of

them was vertically supported at its base portion and the other was flatly supported at the same portion. The gyro-sensor dealt with here is one of such types of gyro-sensors; however, its operation principle is different from theirs. That is, this new gyro-sensor is aimed at being constructed using two-mode coupling of flexural and torsional vibrations in tuning fork arms. In this paper, the structure and operation principle of such a new trident tuning fork resonator gyro-sensor are described, with simulated results of vibrational characteristics for designing the resonator, and moreover, experimental results to confirm its operation principle are shown.

4pPAb10. Effects of diffraction on the sensitivity of needle-type ultrasonic receivers. Gerald R. Harris, Paul M. Gammell, and Jeffrey M. Porter (FDA/CDRH, 9200 Corporate Blvd., HFZ-132, Rockville, MD 20850, grh@cdrh.fda.gov)

The low-frequency sensitivity of piezoelectric receivers usually is assumed to be represented accurately by the -3-dB cutoff frequency $f_{co}=(2\pi R_i C_i)^{-1}$, where R_i is the loading (e.g., amplifier) resistance and C_i is the total capacitance. For typical needle-type hydrophones such as used in medical ultrasound exposimetry, f_{co} is less than 50 kHz. However, theoretical studies have shown that diffraction effects at the needle tip can cause a low-frequency rolloff in sensitivity at frequencies much higher than that predicted by this simple electrical model. To examine this effect, broadband frequency response measurements of several needle-type hydrophones were made in the frequency range 0.2–2 MHz. The active sensor material was polyvinylidene fluoride, and needle diameters ranged from a few tenths of a millimeter to approximately 1 mm. In all cases the sensitivity decreased with decreasing frequency, with the -3 dB points all lying above 400 kHz. Such behavior calls the use of these devices into question when accurate knowledge of the pressure waveform is required, particularly with regard to measuring the peak rarefactional pressure in pulsed waveforms displaying significant finite amplitude distortion.

4pPAb11. Study of a volcano-profiled ultrasound field. Gonghuan Du, Enyu Wang, and Yu Zhang (Inst. of Acoust., Nanjing Univ., Nanjing 210093, PROC)

As is known, the volcano-profiled ultrasound field is of interest in the application of ultrasonic therapy for cancer. This paper indicates that when the radial distribution of vibration velocity on the surface of the source satisfies the expression $q(\xi)=(B\xi^2)^n\exp(-B\xi^2)$, the radiation field, which may be expressed analytically, will maintain approximately the volcano profile in the near field. By employing the theory of the electric field and controlling the shape of the back electrode, a volcano-profiled distribution of the normal component of the electric field and accordingly a same-shaped distribution of the vibration velocity are achieved on the surface of the transducer. In order to improve the radiation efficiency of the transducer, the same material (ceramics) as that of the piezoelectric component, however, unpolarized, is chosen for the back electrode. A once-shaping technique is utilized to make the special-shaped back electrode. The test of the ultrasound field is of satisfying result. The last part of the article gives a brief theoretical description of a nonlinear volcano field. [Work supported by NSF of China.]

4pPAb12. The second harmonic component in the focused sound field diffracted by a straight edge. Shigemi Saito and Jung-Soon Kim (Faculty of Marine Sci. and Technol., Tokai Univ., 3-20-1 Orido, Shimizu, Shizuoka, 424 Japan, ssaito@scc.u-tokai.ac.jp)

To analyze the characteristics of the acoustical imaging utilizing higher harmonic components in the focused sound field, the influence of a diffracting edge put into the focal region on the fundamental and second harmonic fields is theoretically and experimentally investigated. A straight rigid edge put normal to the acoustic axis is assumed to diffract a focused Gaussian beam. The second harmonic sound generated both in front of and beyond the edge is taken into account to calculate the diffracted field through the manner using the Green's function for the governing approximate equation. In the experiment, a knife edge is inserted normal to the sound beam formed by the 1.9-MHz focusing Gaussian source employing a concave piezoelectric transducer with an asteroid electrode. The axial field detected by a needle-type hydrophone as well as the signal obtained with another concave transducer is discussed regarding the amplitudes and phase difference of the fundamental and second harmonic components.

4pPAb13. Acoustic streaming and temperature elevation in a high viscous fluid by irradiation of ultrasound beams. Tomoo Kamakura, Hai-Ying Huang (Dept. of Electron., Univ. of Electro-Commun., 1-5-1, Chofugaoka, Chofu-shi, 182 Japan, kamakura@ee.uec.ac.jp), and Kazuhisa Matsuda (Koganei Tech. High School, Tokyo, 184 Japan)

Sound energy losses due to viscosity and heat conduction in a fluid cause the medium movement as acoustic streaming and cause temperature elevation in beams. In most fluids, the buoyancy force by heat expansion is generally weaker than the driving force of the streaming, so the temperature field is basically determined by the heat transfer equation with convective terms. As some physical parameters such as viscosity depend generally on temperature, the sound field is changed with the irradiation time of cw beams. In addition to the viscosity, two other parameters of sound absorption and sound speed are taken account of their temperature dependency in the present theory. The hydrodynamic flow, forced convection heat transfer, and wave equations are simultaneously solved for axisymmetric ultrasound beams by a finite difference scheme. Numerical examples of the streaming velocity and temperature rise are given. They demonstrate that sound pressure amplitude is intrinsically changed with time and sound self-action is generated. The present method of numerical calculation is easily extended to the problems of focusing and multilayered beam systems.

4pPAb14. Inversion of velocity statistical parameters from travel time measurements. Bertrand Iooss (Ecole des Mines de Paris, Ctr. de Géostatistique, 77305 Fontainebleau, France, iooss@cg.ensmp.fr)

Velocity estimation remains one of the main problems of seismic exploration works. Below a certain range of velocity heterogeneities, deterministic methods become impracticable and the small-scale fluctuations of the velocity field can be described as a random field defined by its first and second statistical moments. It reduces the problem to the determination of a few unknowns, like the mean value, the correlation lengths in the various directions of space, the standard deviation, and the type of correlation function. Theoretical formulas link the statistical moments of an anisotropic random velocity field and those of the travel times of an acoustic plane or spherical wave [A. Ishimaru, ''Wave propagation and scattering in random media'' (1978)]. They are based on the smooth perturbation method of Rytov and on the parabolic approximation which neglects the heterogeneities with scales smaller than the wavelength. Validity domains of these approximations are studied. Therefore, inversion formulas are derived and inversion procedures of the velocity statistical parameters are developed. Finally, comparisons are made between theoretical predictions and synthetic results performed via finite difference algorithms.

4pPAb15. Inverse scattering for porous media with rigid frame. Claude Depollier, Zine el Abidine Fellah, and Achour Aknine (Lab. Acoustique, Univ. Le Maine, B.P. 535, Le Mans Cedex 9, France)

The sound propagation in porous materials having a rigid frame filled by air is well described by the equivalent fluid model. In this framework, the interactions between fluid and structure are taken into account in two response factors, the dynamic tortuosity $\alpha(\omega)$ and the dynamic compressibility $\beta(\omega)$ defined by the equations

$$\rho\alpha(\omega)\frac{\partial\langle \mathbf{v}\rangle}{\partial t}=-\nabla\langle p\rangle, \quad \frac{\beta(\omega)}{K_a}\frac{\partial\langle p\rangle}{\partial t}=-\nabla\langle \mathbf{v}\rangle.$$

These functions are related to the thermal exchanges between fluid and frame and to the geometry of the pores and characterize the sound propagation in the porous media. For homogenous porous materials, these response factors can be measured by different methods, but when the material is inhomogeneous, these methods generally fail. A method is proposed for computing the sound field within a one-dimensional porous medium characterized by a spatially varying tortuosity and compressibility profiles. The spatial variation of the medium is assumed to be along the direction of propagation, i.e., it is stratified. Scattering operators and propagation operators are defined and their properties are studied. An inversion algorithm which leads to the reconstruction of the tortuosity and the compressibility is presented. Numerical results are compared to experimental data.

4pPAb16. *A-wave spatial resonances on thin cylindrical shells: Experimental study.** Loïc Martinez, Jean Duclos, and Alain Tinel (L.A.U.E., Pl. R. Schuman, BP 4006, 76610 Le Havre cedex, France)

An experimental study of a thin cylindrical shell in contact with two different fluids is proposed. Using surface wave analysis methods (SWAM), the *A*-wave and *A**-wave attenuation and phase velocity are first measured. The experiments show that when the internal fluid has a sound velocity slower than the external fluid, internal reflection of energy inside the shell generates an *A**-wave in multiple space and time positions. Using those results as a basis hypothesis, the 2-D $Ksi(k,\omega)$ wavenumber-frequency response of the fluid filled target is performed. On this *Ksi* representation, the multiple space *A**-wave echoes cause multiple space resonances on a *Ksi*-cut versus *k*. Those numerical results are investigated using SWAM. The so-identified multiple spatial resonances correspond very well to the earliest shell whispering gallery waves families found by Veksler *et al.* in terms of RST.

4pPAb17. Diffraction and conversion of the *A* wave on a *T* structure. Bruno Morvan, Alain Tinel, and Jean Duclos (L.A.U.E, Universite du Havre, Pl. R. Schuman, BP 4006, 76610 Le Havre, France)

The *A* wave propagates without attenuation in an immersed plane plate. So it can be potentially used in nondestructive inspection for great dimension structure. The *A* wave propagation is well understood in canonical structures (cylindrical shell, infinite plate). In this paper, a *T* structure is studied: a brazed perpendicular plate positioned on another one. This stucture is identified as an attached substructure like stiffeners, bulkheads, rib, etc. The experimental study has been performed with a sample made of different metals (stainless steel, brass) and with plates of different thicknesses (frequency-thickness product $\sim$0.5 to 1.5 MHz mm). The *A* wave is generated at the end of the plate by the conversion of a pulse bulk wave. A second immersed transducer collects the *A*-wave diffracted signal, at the plate's junction, at a given observation angle. It should be noted that the *A* wave is reflected and transmitted; it is also converted into a Lamb wave on the principal plate and the perpendicular plate. The spectral analysis of received signal, for different angles, allows the reconstitution of the Lamb wave dispersion curves.

4pPAb18. Formulation of a boundary value problem by a new principle of diffraction. M. Ueda (Dept. of Intl. Development Eng., Tokyo Inst. of Tech., 2-12-1 O-okayama, Meguro-ku, Tokyo, 152 Japan, ueda@ide.titech.ac.jp)

A new diffraction principle, that is formulated from a viewpoint centered at an observation point by considering wave propagation in a space seen by the observer virtually, has been proposed [M. Ueda, J. Acoust. Soc. Am. **95**, 2354–2361 (1994)]. This new approach allows us to grasp a simple understandable concept of diffraction and to formulate the first rigorous and unique representation of diffracted waves, that is, it satisfies both the wave equation and hard or soft boundary conditions for 3D objects. Thus the diffraction problem for scalar waves is solved by this approach. The presentation is expressed in a form of integral equation. Thus it is another problem to formulate the algorithm to find the solution. In this paper the numerical procedures to formulate the algorithm are described. The boundary value problem is formulated by a new principle of diffraction. The structure of the response function, that shows mutual dependence of potentials at the surface of the object, is made clear by this new formulation and it reveals a special role of the edge point in the diffraction process. The usefulness of the response function is made clear by the analysis of diffraction arisen from a semi-infinite plane.

4pPAb19. Modelization of Kirchhoff scattering by a sound ray algorithm. Jean J. Embrechts (Natl. Fund Sci. Res., Dept. of Acoust., Univ. Liege, Sart Tilman, B28, B4000 Liege 1, Belgium, jjembrechts@ulg.ac.be)

The problem of the modelization of sound scattering by rough surfaces arises in room acoustics computer simulations, when sound diffusion is taken into account, as well as in other acoustical fields of interest. It is shown here how a sound ray algorithm can be applied to simulate a scattering process obeying the Kirchhoff tangent plane approximation. The Kirchhoff formulation for a random rough surface is briefly recalled, with the shadowing function proposed by Bass and Fuks. The scattered intensity is directly proportional to the statistical distribution of the surface slopes, which here is kept as general as possible. Then, it is established that this formulation is strictly equivalent to the result of a suitable sound ray algorithm, when the wavelength approaches zero in the Kirchhoff formulation. This is demonstrated in theory, and then applied to several Gaussian rough surfaces. Moreover, hints are given to extend the mathematical analysis of the sound ray technique to the second-order reflections on the rough surface. This is viewed as a way to take into account multiple scattering, which is neglected by the Kirchhoff approximation. [Work supported by FNRS, Natl. Fund for Scientific Research.]

4pPAb20. Analytical method for radiation and scattering problems in noncanonical domains. Victor T. Grinchenko (Inst. of Hydromechan. NAS of Ukraine, 8/4 Zhelyabov St., 252057 Kiev, Ukraine, vgr@ihm.kiev.ua)

One of the tendencies in modern acoustics is to use analytical solutions of the Helmholtz equation to present the sound field around radiators and scatterers of noncanonical shapes. The *T* matrix and internal source density methods are examples. In the present paper a method called a partial domain method is discussed. An essential point of the method is the notion of general solution of the boundary-value problem for the Helmholtz equation. The boundaries of the objects under consideration are parts of coordinate surfaces in separable systems. However, surfaces may be either coordinate surfaces of different types (as in the case of a finite cylinder) or coordinate surfaces of different systems (as in the case of a cylinder with spherical lids). Theoretical problems of both the method and numerical implementation of one are discussed. Consideration of such aspects of the method as singularity of the field near corner points, nonuniqueness of the wave functions, and extension of boundary conditions gives a basis to make the computation process more effective. The numerical results for some cases of radiation by finite cylinder and cylinder with hemispherical lids are used to illustrate the method.

Session 4pPP

Psychological and Physiological Acoustics: Auditory Attention II (Poster Session)

Christine R. Mason, Chair
Department of Communication and Hearing Research, Boston University, Boston, Massachusetts 02215

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. To allow for extended viewing time, posters will remain on display until 5:00 p.m. on Friday, 26 June. A cash bar will be set up near the poster session from 5:00 p.m. to 7:30 p.m. preceding the banquet. In order to provide extended viewing of posters in session 5aPP, authors of posters in that session have been requested to display their posters beginning at 1:00 p.m. in Grand Ballroom II.

4pPP1. The auditory attentional blink in a congenitally blind population. Kim M. Goddard, Elzbieta B. Slawinski, and Matthew I. Isaak (Dept. of Psych., Univ. of Calgary, Calgary, AB T2N 1N4, Canada, kgoddard@acs.ucalgary.ca)

The attentional blink (AB), a temporary information processing deficit that follows an attended-to target, has been demonstrated most recently using pure-tone stimuli in a rapid auditory presentation (RAP) paradigm. The current study compared this auditory AB in a group of congenitally blind individuals to nonblind individuals. RAP streams were created using 25 equally loud, randomly generated 1000–2490-Hz tones (11 tones/s). Tones were 85 ms in duration and were separated by a 5-ms silent interval. Target and probe tones were 1500-, 2000-, or 2500-Hz tones increased in intensity by approximately 10 dB SPL relative to the stream tones. In the experimental condition, participants identified the target and probe according to pitch (low, medium, or high). In the control condition, participants performed only the probe task. Probe detection accuracy was measured. Results revealed that in the control condition, participants reported a single loud tone with near ceiling accuracy. In contrast, in the experimental condition, probe detection accuracy was impaired for approximately 270 ms following the target and this impairment was significantly attenuated in the blind population. Findings are discussed in terms of compensatory attentional and temporal processing mechanisms. [Work supported by NSERC.]

4pPP2. The use of visible speech cues (speechreading) for directing auditory attention: Reducing temporal and spectral uncertainty in auditory detection of spoken sentences. Ken W. Grant and Philip F. Seitz (Walter Reed Army Medical Ctr., Army Audiol. and Speech Ctr., Washington, DC 20307-5001)

It is well established that auditory-visual speech recognition is far superior to auditory-only speech recognition. Classic accounts of the benefits of speechreading to speech recognition treat auditory and visual channels as independent sources of information that are integrated early in the speech perception process, most likely at a precategorical stage. The question addressed in this study was whether visible movements of the speech articulators could be used to improve the detection of speech in noise, thus demonstrating an influence of speechreading on the processing of low-level auditory cues. Subjects were required to detect the presence of spoken sentences in noise under three conditions: auditory-only, auditory-visual with a visually matched sentence, and auditory-visual with a visually unmatched sentence. The potential benefits of congruent visual speech cues to auditory detection will be discussed in terms of a reduction of signal uncertainty, in both temporal and spectral domains. In effect, the visual channel may serve to inform the auditory system about when (in time) and where (in frequency) to expect changes in the acoustic signal. [Work supported by NIH Grant DC00792.]

4pPP3. Order effects in the measurement of auditory thresholds during bimodal divided attention. Vishakha W. Rawool (Commun. Disord. and Special Education, Bloomsburg Univ., Bloomsburg, PA 17815)

This study investigated the effect of a visual distraction task on the detection of warbled tones presented at 0.5, 1, 2, and 4 kHz in two orders of presentation. Fourteen subjects responded to warbled tones in two conditions. In the first (attention) condition, the subjects pressed a button everytime they heard the tone. In the second (distraction) condition, they were asked to respond to the tones by turning their heads towards the speaker while solving a cardboard jigsaw puzzle as quickly as possible. The order of attention and distraction conditions was random. Following each correct response, visual reinforcement was presented. For seven of the subjects, the order of presentation of the stimuli was 4, 2, 1, and 0.5 kHz (order 1). For the remaining seven subjects, the order of presentation was 0.5, 1, 2, and 4 kHz (order 2). The same order was maintained in the attention and distraction conditions. Results showed a significant order, attention–distraction, and frequency interaction. The results suggest that distraction can worsen the thresholds at the beginning of the task, but it may cause improvement in thresholds over time. These results will be discussed with reference to dual task attention limits and intersensory facilitation.

4pPP4. The role of memory in the dual-task: Evidence from a frequency/amplitude judgment. Erick Gallun, Anne-Marie Bonnel, and Ervin R. Hafter (Dept. of Psych., Univ. of California, Berkeley, CA 94720, erick@ear.berkeley.edu)

Subjects judged changes in the amplitude and/or frequency of a 100-ms signal either within a trial (experiment one) or between trials (experiment two). Within-trial changes featured roving levels and frequencies, whereas between trial changes were always from the same pool of four stimuli. The cost of making two judgments as opposed to only one was measured by comparing performance in the three conditions (judge amplitude, judge frequency, or judge both). The cost of roving levels and frequencies was measured by comparing performance between the two experiments. It was found that two judgments resulted in a decrease in performance only when the rove was not employed. This agrees with

findings in the bimodal dual-task suggesting that the cost of dividing attention is a result of the use of long-term memory, and that the sensory-trace does not incur a cost in the dual-task. [Research supported by the Nat. Inst. of Health (NIDCD 07787), USA, Univ. of California, USA, and CNRS, France.]

4pPP5. Monitoring the simultaneous presentation of multiple spatialized speech signals in the free field. W. Todd Nelson (AL/CFBA, 2610 Seventh St., Bldg. 441, Wright-Patterson AFB, OH 45433-7901), Robert S. Bolia (Systems Res. Labs., Dayton, OH 45440), Mark A. Ericson, and Richard L. McKinley (AL/CFBA, Wright-Patterson AFB, OH 45433-7901)

The effect of spatial auditory information on listeners' ability to detect, identify, and monitor the simultaneous presentation of multiple speech was evaluated in the free field. Factorial combinations of four variables, including the number of localized speech signals, the angular separation of the speech signals, the location of the speech signals along the horizontal plane, and the sex of the speaker were employed using a completely within-subjects design. Participants were required to detect the presentation of a critical speech signal against a background of nonsignal speech events. Speech stimuli were derived from a coordinated call sign test which consisted of a call sign (''Ringo''), a color (''red''), and a number (''five''). In addition to having high face validity for aviation communication tasks, this measure has been successfully employed in competing message experiments. The experiment was conducted at the USAF Armstrong Laboratory's Auditory Localization Facility—a 277-speaker geodesic sphere housed within an anechoic chamber. Performance efficiency was evaluated in terms of percentage of correct detections and false alarms, reaction time, and signal detection theory indices of perceptual sensitivity and response bias. Implications for the design of spatial auitory displays to enhance communication effectiveness and situation awareness are discussed.

4pPP6. Sequential interactions in discrimination of target frequency increments trailed by irrelevant frequency increments: Effects of target duration and target-irrelevant frequency separation. Blas Espinoza-Varas and Hyunsook Jang (Commun. Sci. and Disord., Univ. of Oklahoma Health Sci. Ctr., Oklahoma City, OK 73190, blas-espinoza-varas@UOKHSC.edu)

Discrimination thresholds were measured for target frequency increments trailed by irrelevant frequency increments. Pairs of pure tones (t_1,t_2) separated by 20 ms of silence were presented on each interval of a three-interval, two-alternative, forced-choice task. Duration was 40 or 80 ms for t_1, and 80 ms for t_2; frequency was 1500 Hz for t_1, and either 500, 1500, or 2500 Hz for t_2; level was 70 dB SPL for t_1 and t_2. On each trial, a ''standard'' pair of tones (interval 1) was followed by two ''comparison'' pairs (interval 2 and 3). One (randomly chosen) ''comparison'' pair contained an adaptively varied target increment in the frequency of t_1. In addition, both comparison stimuli contained identical irrelevant increments in t_2 frequency. Listeners had to determine which ''comparison'' interval contained the target increment. Irrelevant frequency increments elevated target-frequency discrimination thresholds. Threshold elevation was greater with 40-ms than with 80-ms targets. Equal-frequency conditions (1500 Hz for t_1 and t_2) yielded larger threshold elevation than unequal frequency conditions (1500 and 500 Hz or 1500 and 2500 Hz for t_1 and t_2). Irrelevant-increment effects decreased with training. [Funded by OCAST Project HR4-064.]

4pPP7. Interactions of pairs of target and context tones based on relative frequency relation and degree of frequency uncertainty. Donna L. Neff and Rebecca L. Wrage (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131, neff@boystown.org)

Four normal-hearing listeners completed 2IFC sample-discrimination tasks for frequency, in which they judged which of two tone pairs was drawn from the higher of two Gaussian frequency distributions. The center frequencies of the two target distributions were placed at low (400/500 Hz), middle (1128/1410 Hz), or high (3200/4000 Hz) frequency regions. All distributions were equally spaced and equivariant on a logarithmic frequency scale. Pairs of extraneous (to-be-ignored) context tones were added above, flanking, and below the low-, middle-, and high-frequency target regions, respectively. Across conditions, target and context tones used the same frequency regions, but never occupied the same frequency region within condition. Context tones were fixed in frequency, varied between two known frequencies, or were Gaussian distributed. A level jitter was added to derive perceptual weighting functions across stimuli. The results showed little detrimental effect of fixed-frequency context tones, but large individual differences in the effects of known pairs of tones or tones with Gaussian variation. Weighting functions and patterns of frequency interactions among stimuli will be discussed. [Work supported by NIDCD.]

4pPP8. Effects of tonal masker uncertainty on detection. Elzbieta B. Slawinski (Dept of Psych., Univ. of Calgary, Calgary, AB T2N 1N4, Canada, eslawins@acs.ucalgary.ca) and Bertram Scharf (Ctr. de Recherche en Neurosci. Cognit., CNRS, Marseille, France)

The present study was inspired by results of Allen and Wightman [J. Speech Hear. Res. **38**, 503–511 (1995)]. The goal of the present study was to explore the effects of uncertainty of a weak tonal masker (or distractor) on the detection of a tone in noise. The thresholds of three listeners were measured for 1-kHz tone burst (350-ms duration) in broadband noise (300–1800 Hz, at 60 dB SPL). A 2IFC tracking procedure (3 down, 1 up) with six interleaved tracks was used. On each trial, the distractor frequency was selected randomly from six frequencies (525, 800, 925, 1075, 1280, and 1600 Hz) outside of the critical band surrounding the signal. The distractor came on simultaneously with the signal in the signal-plus-noise interval and also at the corresponding moment in the noise-alone interval. The preliminary results indicated that thresholds for the signal increased in the presence of distractors by 3–10 dB, depending on the frequency of the distractors and type of noise. Thresholds were higher for distractors close to the signal frequency, more so in intermittent noise than in continuous noise.

4pPP9. Identification of brief auditory patterns. Gerald Kidd, Jr., Christine R. Mason, and Chung-Yiu P. Chiu (Dept. of Commun. Disord. and Hearing Res. Ctr., Boston Univ., Boston, MA 02215, gkidd@bu.edu)

The identification of brief nonspeech sounds poses a challenging problem for machine-based recognition systems, especially at low signal-to-noise ratios. In contrast, human observers are thought to be remarkably adept at such tasks. However, there are few systematic studies in the literature that have attempted to quantify human psychophysical performance, or the factors that influence performance, for the identification of brief nonspeech sounds in noise. This study examined human identification of brief sequences of tones arranged in six frequency patterns in masked conditions. The stimuli and methods have been described previously [Kidd, Jr. *et al.*, J. Acoust. Soc. Am. **98**, 1977–1986 (1995)]. Several acoustical factors were varied including pattern duration, signal-to-noise ratio, and range of frequencies forming the patterns. Identification performance declined as duration was shortened below 100 ms even when signal energy, or detectability, were equated. For brief patterns at low signal-to-noise ratios, increasing the range of frequencies comprising the patterns improved identification performance. These results provide in-

sight into the limits of human performance in the identification of brief sounds and may prove useful in the design of algorithms for machine-based recognition. [Supported by ONR and NIH.]

4pPP10. Selective attending to auditory streams in complex sequences: Frequency and/or temporal expectations? Carolyn Drake, Renaud Brochard, and Matthieu Adenier (Lab. de Psych. Exp., CNRS URA 316, Univ. Rene Descartes, 28 rue Serpente, 75006, Paris, France)

Both frequency and temporal information influence the perceptual organization of complex sequences: stream segregation is more likely with wider frequency separations and at faster rates. The influence of frequency and temporal information in directing attention to a stream within a complex sequence is investigated. Complex sequences of three subsequences (each with a specific frequency/tempo combination) were preceded by a cue sequence containing different types of information (FT=frequency + tempo, F=frequency, T=tempo, S=silence). The ability to selectively attend to one subsequence was measured by correct detections of a small temporal irregularity in the cued subsequence. When stream segregation was easy (large frequency separations), performance was similar for FT and F (93%), intermediate for T (79%) and random for S (64%): frequency information facilitated selective attending more than temporal information. When stream segregation was hard (small frequency separations), performance was poorer and similar for all conditions (65%): temporal information became predominant when frequency information was less useful. Thus listeners adapt their attending strategies to coincide with the most pertinent information. Results are discussed in terms of diverging predictions of the theories of Bregman [1990] and Jones [1993].

4pPP11. Application of a peripheral model to an attentional filter of missing-fundamental tone. Hiromitsu Miyazono (Dept. of Administration, Pref. Univ. of Kumamoto, 3-1-100 Tsukide Kumamoto, 862 Japan, miyazono@pu-kumamoto.ac.jp), Tsuyoshi Usagawa, and Masanao Ebata (Kumamoto Univ., 2-39-1 Kurokami Kumamoto, 860 Japan)

An attentional filter measured by the probe-signal method is compared with the peripheral auditory filter measured by the notch-noise method. The shape of attentional filter is estimated by the roex (rounded exponential) filter model. However, the attention can be given on a frequency without the signal on the actual frequency. In this study, the attentional filter was measured using a missing-fundamental tone which has no energy on an actual frequency. The missing-fundamental tone was produced by a harmonic signal with 13 components from 1000 Hz to 4000 Hz. The fundamental frequency was set to 250 Hz. The psychometric function is measured on various frequencies and the shift of the function from that of control condition is regarded as the effect of the attentional filter. The filter shape is estimated by the roex model and compared with that of the pure tone. In the frequency region apart from the fundamental, the dynamic range of the filter is smaller than that of pure tone. Near the fundamental, the correlation between the measured value by the probe-signal method and the estimated value by the roex filter for the missing-fundamental tone is smaller than that for the pure tone.

4pPP12. Short-term auditory memory interference: The effect of speech pitch salience. Kazuo Ueda and Naoko Seo (Dept. of Psych., Kyoto Pref. Univ., Hangi-cho, Shimogamo, Kyoto, 606 Japan, h50015@sakura.kudpc.kyoto-u.ac.jp)

The interference effect of speech pitch salience on tone pitch recognition and digit recall was observed. For the pitch recognition task, 11 subjects had to recognize pitches of two harmonic complex tones separated by a 5-s retention interval. During the interval, no sounds or six randomly chosen intervening sounds, either complex tones from a semitone scale or spoken digits, were presented. The speech included natural stimuli and synthesized stimuli with the original and with flattened pitch contours. For the digit recall task, the subjects had to recall serially the intervening digits. They had to perform both tasks in the dual task conditions. The results were: (1) the synthesized speech was equally intelligible as natural speech, (2) both pitch recognition errors and digit recall errors were significantly increased in the dual task conditions compared to the single task conditions, and (3) the synthesized speech with the flattened pitch contours as well as the complex tones had the most deteriorating effect on tone pitch recognition. Thus it became obvious that the weak interference effect of natural speech on tone pitch recognition was due to their pitch ambiguity. [Work supported by Grant-in Aid for Scientific Research and by Nagai Research Fund.]

4pPP13. Changing frequency after-effect of a linear frequency glide. Takuro Kayahara (Intelligent Modeling Lab., Univ. of Tokyo, 2-11-16, Yayoi, Bunkyo-ku, Tokyo, 113 Japan, kayahara@iml.u-tokyo.ac.jp)

After listening to a linear frequency glide in a single direction ten times, subjective frequency change of stationary sine wave in the opposite direction was observed (changing frequency after-effect). Adapting glides had a fixed frequency excursion (ΔF: 960–1040 Hz), five different durations (250, 500, 750, 1000, or 1250 ms), and glided either upward or downward in frequency. Test glides starting from 1000 Hz had three different durations (500, 750, or 1000 ms), and its ΔF increased or decreased according to subject's response. ΔF at subjective stationary frequency was measured using a three-category (''up,'' ''down,'' or ''no change'' in frequency) double-staircase procedure. In the ipsilateral condition (presenting both adapting and test stimuli to the same ear), ΔF in the same direction with the adapting glides was obtained and maximum ΔF (12 Hz) was observed when the duration of each test glide was the same as adapting glides, although, in the contralateral condition, ΔF was very small or none. These results suggest that there is a feature extractor which is selective in changing frequency and its direction at the level before integration of interaural information, and this extractor changes its own property after a relatively short time exposure of a single direction frequency glide.

4pPP14. Dynamic intensity change influences perceived pitch: Attentional differences between musicians and nonmusicians. John G. Neuhoff (Lafayette College, Easton, PA 18042)

The influence of dynamic intensity change on the terminal pitch of ascending frequency glides was tested as a function of terminal frequency and musical experience. Previous work has shown that dynamic intensity change influences perceived pitch [J. G. Neuhoff and M. K. McBeath, J. Exp. Psychol.: Human Percept. Perform. **22**, 970–985 (1996)]. In the current study, musicians and nonmusicians were presented with dynamic tones that changed in both frequency and intensity, and made real-time judgments about perceived pitch change. Results show that for both musicians and nonmusicians, the influence of intensity change on perceived pitch lessened as the terminal frequency approached an octave. Musicians were less susceptible to the effects of intensity change for falling intensity but not for rising intensity. The findings suggest that musicians may have better selective attention to dynamic auditory stimulus dimensions, and the influence of a tonal schema may reduce the influence of intensity change on perceived pitch. The results imply an interplay of musical and nonmusical listening strategies and demonstrate that the two are not mutually exclusive.

4pPP15. Temporal properties of loudness recalibration. Dan Mapes-Riordan and William A. Yost (Parmly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626)

Loudness recalibration occurs when a loud (recalibration) tone at frequency $f1$ precedes quieter test tones at frequencies $f1$ and $f2$. Recalibration is a reduction in the perceived loudness of the test tone at $f1$ due to the addition of the recalibration tone. Mapes-Riordan and Yost [J.

Acoust. Soc. Am. **101**, 3170(A) (1997)] showed that the temporal onset of loudness recalibration is relatively fast. The current set of experiments addressed the temporal decay properties of loudness recalibration. The first set of experiments examined the ''local'' temporal characteristics by measuring how the average amount of loudness recalibration varies with the duration of silent gap between the recalibration tone and the first comparison tone. A second set of experiments examined the ''global'' nature of the decay of loudness recalibration by inducing recalibration and then monitoring the decay using an adaptive tracking procedure. The results of these experiments will be discussed in terms of how assimilation processes influence loudness recalibration, what factors make the decay of loudness recalibration relatively slow, and the effect of attention on loudness recalibration. [Work supported by a Program Project Grant from NIDCD.]

THURSDAY AFTERNOON, 25 JUNE 1998 GRAND BALLROOM A (S), 1:00 TO 3:45 P.M.

Session 4pSA

Structural Acoustics and Vibration: Vibrations of Complex Structures I

Donald B. Bliss, Chair
Department of Mechanical Engineering and Material Science, Duke University, Hudson Hall, Science Drive, Box 90300, Durham, North Carolina 27708

Contributed Papers

1:00

4pSA1. Uncertainty in modeling the dynamics of structures in midfrequency range. Kai-Ulrich Machens and Eike Brechlin (Inst. of Tech. Acoust., Univ. of Technol., Berlin, Germany)

Based on measurements of the energy contents in single struts in a complex truss structure (109 struts, 35 joints) [Machens, Dyer. Acustica **81** (1995)], a numerical simulation of the truss dynamics using the direct dynamic stiffness method was performed. Great care was taken to account for flexural, longitudinal, and torsional waves in the struts, moments of inertia and masses of the joints, complex coupling stiffness (translational and rotational) including a loss factor between joints and struts, as well as material damping in the struts. The struts were modeled as Euler–Bernoulli–Beams and additionally as Rayleigh–Timoshenko–Beams. Frequency averaging proved to be important in the midfrequency range as a meaningful comparison with the measured data depends on a high accuracy in modeling and choice of material parameters. The broadband frequency averaged energy in single struts show a high sensitivity towards various factors such as small random variations of the geometry and of the complex coupling stiffnesses as well as the use of first- and second-order beam theory. The results show that a truss, which is built up from simple elements, becomes fuzzy merely by its high number of coupled elements.

1:15

4pSA2. Reduction of fluid-loaded ribbed shells to equivalent uniform structures for mid- to high-frequency structural acoustic analysis. Donald B. Bliss (Dept. of Mech. Eng. and Mater. Sci., Duke Univ., Durham, NC 27708) and Linda P. Franzoni (North Carolina State Univ., Raleigh, NC 27695)

The acoustic and vibratory behavior of fluid-loaded shells with ribs and other structural discontinuities is of great interest in naval applications. These structural discontinuities lead to complex phenomena, especially at mid- to high-frequencies where shell wavelengths are shorter than the discontinuity spacing. The methods of homogenization and local/global decomposition can be used to divide this problem into two parts. In the global problem, periodic discontinuities are replaced by an equivalent distributed suspension with slowly varying properties. This problem can be solved much more efficiently than the original problem since all rapidly varying scales have been removed. The local problem is solved separately and independently, except for amplitude information from the global problem. The local problem formulation provides transfer function information that defines the suspension in the global problem. Once the formulation has been developed for a specific structure, the global problem is solved first, and the local solution can be reconstructed afterward. The two constituent parts are then recombined to form a composite solution. As an illustration, this methodology is used to solve for Bloch waves on a structure with identical periodic discontinuities. The application of the theory to nonidentical discontinuities is also explained.

1:30

4pSA3. Toward a design methodology for equipment emulators in the shock testing of large structures. Pierre DuPont and J. Gregory McDaniel (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215)

In a variety of situations, an undesired shock excitation is applied to a master structure which supports shock-sensitive equipment. Often, one wishes to design and test a master structure which transmits the least amount of shock energy to the attached equipment. In scaled testing of new designs, a major task is to design and construct ''equipment emulators,'' inexpensive mechanical systems which approximately mimic the dynamic behavior of the actual full-scale equipment *as seen by the master structure*. This presentation will present new methodologies for designing equipment emulators, assessing their fidelity, and interpreting test data taken in the presence of imperfect emulators. Starting with frequency-domain impedance descriptions of the master structure and actual equipment at the attachment points, this work develops sensitivity metrics which directly relate the fidelity of the emulator to its structural complexity. These ideas may provide a path by which experimentalists can efficiently arrive at conceptual designs of emulators which promise a specified degree of fidelity in terms of system velocities and their associated shock spectra. The ideas are illustrated on simple examples and applied to the emulation of commercial-grade electronic cabinets for the testing of novel ship deck structures. [Work supported by ONR.]

1:45

4pSA4. A new finite-element analysis approach for the prediction of minimum and maximum vibroacoustic response of solid–fluid systems. Jean-Sébastien Genot, François Charron, and Noureddine Atalla (Univ. of Sherbrooke, Dept. of Mech. Eng., Sherbrooke, QC J1K 2R1, Canada)

Engineers are aware that identical products do not have necessarily the same dynamic response because of small differences in their dimensions or material properties. The classical finite-element method cannot take account of this fact. Statistical tools are needed. In order to consider the input data variability, new modules have been integrated in a finite-element code to estimate the maximum and minimum response levels (mean quadratic velocity and radiated power) of a structure. Up to now,

numerical modules were generally based on Monte Carlo or perturbation methods. Both have limiting and complementary drawbacks: The Monte Carlo method is accurate but far too slow, even for small finite-element models, while the perturbation scheme is fast, but unable to predict nonlinear behavior of the dynamic response properly with respect to the input variables. The developed approach takes advantage of the response modal decomposition to identify intermediate variables which are nearly linear. The intermediate variables (eigenfrequencies, dynamic factors) are easily interpolated, thus requiring little computation. Results are presented to prove the efficiency of the method.

2:00

4pSA5. FEM aided structural intensity measurement method for thick body. Yoshikazu Koike, Chie Aramaki, Kentaro Nakamura, and Sadayuki Ueha (Precision and Intelligence Lab., Tokyo Inst. of Technol., 4259 Nagatsuda, Midori-ku, Yokohama, 226 Japan)

Many studies about structural intensity(SI) measurement, which is useful to know transmission paths of the vibration energy in a body, have been reported. Most of the works, however, have been done only for thin bodies, such as beams or plates because of the difficulty in measuring interior vibration distribution in thick structures. In this paper, a new SI measurement method applicable to thick bodies is proposed. The entire vibration distribution in the structure is calculated by dynamic analysis of the finite element method (FEM) from the partial surface vibration data if the elastic parameters, the sound velocity, sizes of a body under study, and the locations of excitation and absorption are given *a priori*. The entire SI distribution can be estimated using the results. Based on this method, several SI measurements for thick vibrating bodies have been successfully carried. The merits of this method are as follows. (1) Only surface vibration data of the structure is needed to estimate the entire SI distribution. (2) It is applicable for arbitrary 3-D structures if the conditions mentioned above are satisfied *a priori*. The principle and the experimental results are described in the report.

2:15

4pSA6. Resilient mounting of engines. Anders Nilsson, Leif Kari, and Romain Haettel (MWL, Dept. of Vehicle Eng., KTH, 10044 Stockholm, Sweden)

In any type of vehicle a number of mechanical sources generate structure-borne noise. The energy is transmitted from excitation points to surrounding structures. These radiate noise externally and internally. The most commonly used method to reduce the radiated noise is to mount the sources resiliently to the supporting foundation. The insertion loss due to the resilient mounting can in many cases be extremely poor. This could be due to the break down of the stiffness of the foundation in the audible frequency range. At the same time the stiffness of the mounts could be increased drastically as compared to the static stiffness. A test rig and measurement procedures for determining the dynamic properties of mounts are described. Measured and predicted results are compared for rubber mounts with simple geometries. The insertion loss of a resilient mounting very much depends on the location of the engine feet on a foundation. Case studies relating to passenger vessels and catamarans built with aluminum and sandwiched are presented. Reduction of both noise and weight is discussed.

2:30–2:45 Break

2:45

4pSA7. Assessment of construction vibration impacts on historic structures. Chetlur G. Balachandran (Parsons Brinckerhoff, Inc., One Penn Plaza, New York, NY 10119)

This paper addresses construction-related ground vibration impacts to historic structures located in an area where three options are being considered for providing capacity improvement to an existing roadway. Historic structures within the project area are: a Fort, a Legislature building, an electrical substation, and an auxiliary building. The above fragile structures are more likely to be affected by construction-related ground vibration impacts than by operational traffic-induced vibration impacts. The buildings are in close proximity to the existing heavily traveled roadway and are already experiencing moderate levels of ground vibration from road traffic. Potential vibration effects from heavy construction equipment include annoyance to people inside these buildings and potential for minor damage to the buildings. The particular concern in the present study is to comply with ground-vibration criteria to safeguard identified structures within the project corridor against construction-induced ground vibration. The various national and international criteria that were used to assess vibration impacts together with the results of vibration impact assessment will be presented. Construction of the roadway alternative that would cause the least vibration impact was identified and mitigation was considered.

3:00

4pSA8. Hybrid control for vibration and acoustics. Robert L. Clark (Dept. of Mech. Eng., Duke Univ., Box 90300, Durham, NC 27708) and Dennis Bernstein (Univ. of Michigan, Ann Arbor, MI 48109-2118)

The ''standard problem,'' frequently discussed in the controls literature, is detailed for application to noise and vibration control. Specifically, the feedback, feedforward, and hybrid (combination of feedback and feedforward) control system architectures are developed in the framework of the standard problem. The hybrid measurement structure and problem was formulated to demonstrate that the LQG compensator design for control of stochastic inputs is separable with respect to inputs which can be measured directly. Thus, one can design a feedback controller in the absence of measurable input disturbances if the objective is to combine both feedback and feedforward or adaptive feedforward control for stochastic and measurable inputs, respectively. Once the feedback controller is designed for the stochastic inputs, the feedforward controller required for the measurable input disturbances can be designed in a subsequent formulation of the control problem. Realizing that the design of such compensators is separable serves to simplify the hybrid design process.

3:15

4pSA9. Active control of total structural intensity in a T beam: General case. Sabih I. Hayek and Jungyun Won (Active Vib. Control Lab., Dept. of Eng. Sci. and Mech., Penn State Univ., University Park, PA 16802, sihesm@engr.psu.edu)

The active control of power flow in a T beam was presented at the 134th Meeting of the Acoustical Society of America. In that paper, the mechanical noise source was a shear force in the plane of the T beam. In this paper, the active control of structural intensity in the T beam is achieved for a general shear mechanical noise source that has in-plane and out-of-plane components. The reduction of the total power flow at a control point located midway in the vertical leg of the T beam is achieved through control actuators that are located at one end of the straight part of the T beam. The vector control actuators have components in three-dimensional space in various combinations of normal forces, shear forces, torques, and moments. The efficiency of the structural intensity reduction is considered for local versus global control. Various control strategies are compared which are also aimed at minimizing the cost function that represents the total power requirement for the control process versus the noise input power.

3:30

4pSA10. Vibration isolation of two elastic structures using active compliance at the isolation mounts. Kenneth E. Jones and Y. F. Hwang (Carderock Div., Naval Surface Warfare Ctr., 9500 MacArthur Blvd., West Bethesda, MD 20817-5700)

A mathematical model of the vibration isolation of two elastically connected structures (via isolation mounts) employing active springs is presented. Two elastic structures are connected by an arbitrary number of springs, each made active by the inclusion of a piezoelectric pad at one of

the spring and structure interfaces. Single-input–single-output feedback controllers at each spring support seek to minimize the force transmitted from one structure to the other. Solutions are based upon the formalism of Lagrange's equations using the normal modes of the uncoupled structures as the generalized coordinates. The constraints by the springs and supports are enforced by Lagrange multipliers. A 2-D numerical example of two elastic beams is presented. One beam, driven by a harmonic disturbance force, is coupled to the second elastically supported beam through three active springs. A beam-relative displacement sensor colocated at each spring is the input to an arbitrary feedback filter which can incorporate practical amplifier characteristics and actuator dynamics. Results of the numerical simulation of the feedback control case are discussed.

THURSDAY AFTERNOON, 25 JUNE 1998 GRAND BALLROOM III (W), 1:00 TO 4:50 P.M.

Session 4pSC

Speech Communication: A Half-Century of Speech Research

Patricia K. Kuhl, Chair
University of Washington, Eagleson Hall 204, Box #354875, Seattle, Washington 98195

Chair's Introduction—1:00

Invited Papers

1:05

4pSC1. Fifty-four years in speech research. Gunnar Fant (Dept. of Speech, Music and Hearing, KTH, Stockholm, S-10044, Sweden, gunnar@speech.kth.se)

This is a brief presentation of my scientific career with outlooks on the development of speech research in the last 50 years. It started with an electrical engineering thesis at KTH in 1944–1945. The topic was intelligibility loss as a function of bandwidth reduction in telephony and related problems in assessing the effects of various types of hearing loss. In 1945–1948 I was employed by the Ericsson Telephone Company where I conducted basic studies of the spectral characteristics of Swedish speech sounds. This early period and my two years at MIT, 1949–1951, provided an entry to a pioneering era of speech research with new tools for speech analysis and synthesis and speech production modeling. It was during this time I started my work on the acoustic theory of speech production and established my cooperation with the linguists Roman Jakobson and Morris Halle and with Ken Stevens. I have followed the advent and the progress of speech research through the analog, the digital, and the computer age. Speech research now has the challenge to provide a deeper and more integrated understanding of all levels of the speech communication process. This is the ultimate requirement for the realization of advanced applications.

1:25

4pSC2. Toward models for human production and perception of speech. Kenneth N. Stevens (Res. Lab. of Elec. and Dept. EECS, MIT, 50 Vassar St., 36-517, Cambridge, MA 02139)

Two of the goals of research in speech communication are to develop models of normal speech production and normal speech perception. Related objectives are to uncover the process by which children acquire the knowledge implicit in these models and to determine how the models are modified for disordered speech. Even partial achievement of these goals can have significant practical consequences, including machine recognition and synthesis of speech, and improved methods for diagnosis and remediation of speech disorders. In this paper, a current view of a framework for models of speech production and perception will be described, and some of the steps that have led to refinement of these models over the past 50-odd years will be described. Advances have been made in quantifying acoustic mechanisms of speech production and in specifying the nature of the discrete linguistic representation of an utterance in memory. From studies of speech perception and speech motor control, some understanding has been gained of how properties of the sound are related to the linguistic representation. There are large deficiencies, however, in our understanding of variability in speech due to speaker differences, speaking style, and context. [Work supported in part by NIH Grants DC00075 and DC02525.]

1:45

4pSC3. Junctures in speech communication. J. L. Flanagan (CAIP Ctr., Rutgers Univ., 96 Frelinghuysen Rd., Piscataway, NJ 08854-8088)

Several junctures in research on speech communication seem salient over the past 50 years. Under the goad of expensive long-distance transmission, and the criticality of voice communication in a world in turmoil, bandwidth compression methods received strong emphasis in the 1940s. Because transoceanic communication was by HF radio, privacy and encryption were also central issues. These incentives resulted in the Channel Vocoder and a variety of offshoots. By the 1960s, digital computers were being deployed widely and sampled-data theory was well established, providing the systems researcher with a new environment for experimentation—one in which concepts could be rapidly simulated and evaluated. Bandwidth compression continued to be a focus, as computer simulation techniques progressed from classical filter emulation to new designs based solely on digital theory. By the 1980s, broadband connectivity (including transoceanic cable) was becoming pervasive and economical enough that emphasis on bandwidth conservation waned. Concomitantly, digital machines were expanding in capability—but were markedly limited by their inability to

communicate with human users in natural ways. The focus consequently shifted to human/machine communication—especially by conversational means. As we approach the year 2000, a strong research emphasis remains on human/machine communication—but with expanded aspirations for natural interaction. Computer interfaces with multiple modalities—utilizing sight, sound, and touch in combination—are beginning to serve multiple collaborating users. In the same era, explosive deployment of two technologies—cellular telephony and computer networking—has rejuvenated interest in bandwidth conservation and coding for privacy—with many of the early vocoder concepts evolving to sophisticated digital forms implemented on single-chip processors. As we move toward 2020, a special emphasis is likely to be on multilingual communication. And in this time, continued advances in computing, acoustic signal processing, and language modeling should carry translating telephony far beyond the ''phrase book'' stage.

2:05

4pSC4. Three questions for a theory of speech. Alvin M. Liberman (Haskins Labs., 270 Crown St., New Haven, CT 06511, liberman@haskins.yale.edu)

On reviewing my 50 years of research, I see that in trying to head theory in the right direction I have been led ever more compellingly by the need to find plausible answers to three questions: (1) What is the basis for the parity between primary motor and perceptual representations that explains how, in the evolution of speech, production and perception managed to proceed in lock step, thus allowing the parties to each two-way exchange to communicate, then as now, in a common code? (2) How does speech meet the requirement, unique to its combinatorial mode, that discrete, invariant, categorical, and distinctly linguistic segments be produced and perceived in strings at rates as high as 20 segments per second? (3) What biological and psychological properties account for the vast difference in the difficulty of acquisition and use between speech and the reading or writing of its alphabetic transcription? Two theories of speech—one quite conventional, the other much less so—provide answers that differ greatly in their plausibility.

2:25

4pSC5. Speech research at the I. P. Pavlov Institute in Leningrad/St. Petersburg. L. A. Chistovich (Early Intervention Inst., St. Petersburg, Russia), J. M. Pickett (Windy Hill Lab, Surry, ME), and R. J. Porter, Jr. (Univ. of New Orleans, New Orleans, LA)

The ''Leningrad Group,'' led by L. A. Christovich and V. A. Kozhevnikov, carried out an extended research program to describe relations between speech perception and speech production and build a model of perception. The work first explored auditory processing mechanisms that could afford closely interdependent connections between auditory information and the control signals for motor production. The resulting model for auditory analysis employed amplitude modulation detectors as well as spectral measurements. The model was evaluated in perception tests with listeners mimicking and shadowing speech sounds and was also employed in a spectrographic display and a scheme for automatic speech recognition.

2:45–3:00 Break

3:00

4pSC6. Speech research from acoustic transmission to gestural analysis. Katherine S. Harris (Dept. of Speech and Hearing Sci., City Univ. of New York, 33 West 42 St., New York, NY 10036, loumau@erols.com)

My training was in experimental psychology at Harvard, where the focus of the department's speech work was the characteristics of sounds affecting their error-free transmission, essentially an extension of the Bell Telephone Laboratories' approach to telephone communication. At Haskins, I began research on the fricative sounds, and discovered the important role of the dynamic characteristics of the sounds in distinguishing among them. In turn, this focus led to an interest in the role of formant movement in speech perception, and ultimately to a study of the characteristics of speech movements themselves. Over the last 50-year period, thanks to some of those on this panel, we have developed a good understanding of the static characteristics of the vocal tract as a producer of speech sounds and the acoustics of speech. However, we still have but a rudimentary account of how the moving articulators develop a dynamic output, and how segmental and suprasegmental transformations of the phonetic message affect this output. Recent instrumental advances should allow the development of an understanding of the dynamics of articulation that parallels our understanding of speech acoustics.

3:20

4pSC7. Chasing ideas in phonetics. Peter Ladefoged (Phonet. Lab., Linguist. Dept., UCLA, Los Angeles, CA 90095-1543, oldfogey@ucla.edu)

In my early career I never stayed long enough in a particular field to be contradicted. I started as a poet learning about the sounds of words with David Abercrombie. Then, remembering my background in physics, I moved to studying acoustic phonetics. From there I became a pseudo-psychologist testing perceptual theories, until a meeting with a physiologist led to work on the respiratory muscles used in speech. Eventually I landed in Africa teaching English phonetics and learning about African languages. So, by the time I was asked to set up a lab at UCLA, I was a specialist in nothing. However, I was able to use my background to describe the sounds of a wide range of languages, becoming a sort of linguist. Computers and bright students led to other ways of analyzing sounds. Building a research group who felt that they had a stake in the development of the lab taught me their varied ideas from statistics to engineering, and the philosophy of linguistics. Now, still looking for the growing edge of the field, I think it might be in the physically observable activity of the brain; but perhaps this is because I have never been a pseudo-neurologist.

3:40

4pSC8. Language and speech. Victoria Fromkin (Dept. of Linguist., Univ. of California, Los Angeles, CA 90024)

It is exciting to note that I have been involved for thirty-six of these important fifty years of phonetic research. In 1962 I entered my first class in Phonetics and Phonology taught by the newly arrived Peter Ladefoged from Edinburgh. That course, and Peter himself, set my goals which have lasted all these years—the relationship between the cognitive knowledge that lay behind our ability to speak and understand. The field of phonetics is varied and researchers have their own agenda and questions to answer. While one's goals may be practical or theoretical, the answers we all find help us understand the incredibly unique ability of the human animal to communicate in a special, linguistic way. Whether studying normal or abnormal speech (speech errors and aphasia), child language, or language decay, we gain insights into the complexities of the process. In recent years, due to new technological tools such as CAT, MRI, fMRI, PET, ERPs, etc., we have the chance to ask new questions. But still it takes an understanding of the linguistic systems underlying the ''speech chain'' to know what questions to ask. Linguists therefore are now, as they have always been, important in the research into speech communication.

4:00

4pSC9. Communication between minds: The ultimate goal of speech communication and the target of research for the next half-century. Hiroya Fujisaki (Dept. of Appl. Electron., Sci. Univ. of Tokyo, 2641 Yamazaki, Noda, 278 Japan)

Needless to say, speech is a manifestation of the information intended by the speaker and is perceived by the listener who tries to retrieve it. Elucidation of the entire process of speech communication thus requires a multidisciplinary approach, as was eminently illustrated by the work of Chiba and Kajiyama about a half-century ago [*The Vowel, Its Nature and Structure*, Tokyo-Kaiseikan (1941)]. Through the efforts of distinguished researchers, supported by the rapid progress in electronics and computers as well as in other means, the ensuing half-century has seen remarkable advances in our understanding of the linguistic, physiological, articulatory, acoustic, sensory, and cognitive processes in normal, developmental, and pathological cases, as well as in technologies for its utilization. As a new millenium of man–machine symbiosis is entered into, however, still deeper understanding becomes necessary in order to establish a friendly human–machine interface through spoken language, often in multimodal and multilingual environments. In the author's view, this can only be achieved by investigating and modeling the mind itself as the ultimate source and destination of information. To that end, speech science will need closer cooperation with related disciplines such as brain science, cognitive science, and artificial intelligence.

4:20–4:50 General Discussion

THURSDAY AFTERNOON, 25 JUNE 1998

CASCADE BALLROOM I, SECTION B (W), 1:30 TO 5:20 P.M.

Session 4pSP

Signal Processing in Acoustics, Engineering Acoustics, Architectural Acoustics and Education in Acoustics: Acoustics in Multimedia—Systems Issues II

David I. Havelock, Chair
National Research Council, M-36 Montreal Road, Ottawa, ON K1A 0R6, Canada

Invited Papers

1:30

4pSP1. Acoustics and psychoacoustics of workstation audio systems. Floyd E. Toole (Harman Intl. Industries, Inc., 8500 Balboa Blvd., Northridge, CA 91329, ftoole@harman.com)

Workstations are widely used as production environments for multimedia, film, and audio programs. For many consumers, they are also the delivery systems for these programs, as well as for others, like music, television, and films, intended for reproduction in various conventional systems and environments. Listeners in workstations therefore must be able to perceive sounds with the appropriate sense of timbre, direction, distance, and ambiance whether the program was created in two-channel stereo, or any of the multichannel surround or binaural 3-D formats. The space constraints, and the usual restriction to two loudspeakers, put special demands on the design of these audio systems. Binaural image-steering techniques will be the basis for most interactive games and, through crosstalk-cancellation, binaural technology will enable realistic replications of two, four, and five-channel audio for music, films, etc. This paper reviews the requirements for workstation audio systems, discusses the key enabling technologies, and summarizes what is known about the appropriate technical specifications, measurement techniques, and subjective evaluation criteria and methodologies for evaluating workstation audio systems.

2:00

4pSP2. Using a personal computer platform to develop an information-rich learning environment for instruction in acoustics. Robert Celmer (Acoust. Prog. and Lab., College of Eng., Univ. of Hartford, 200 Bloomfield Ave., West Hartford, CT 06117, celmer@uhavax.hartford.edu)

Teaching students the subject of acoustics involves large amounts of technical information, including equations, derivations, and example problems. It also requires the explanation of a multitude of concepts, many of which are often difficult for students to grasp and equally frustrating for the instructor to disseminate. Part of the difficulty can be traced to the use of *text* (a visually based medium) to teach about *sound* (an aurally based medium). This presentation will describe the development of multimedia techniques for in-class presentation of acoustics instruction. Some of the materials were developed using certain authoring applications, drawing and animation programs, sound manipulation software, as well as 3D CAD and spectral analysis applets. In addition to the classroom materials, self-paced exercises have been developed to review the material in a student-centered learning environment, along with scheduled interaction time with the instructor for questions, clarifications, and test administration. Demonstrations of the materials for the instruction of acoustical concepts, as well as case studies, will be presented.

Contributed Papers

2:30

4pSP3. The impact of system latency on dynamic performance in virtual acoustic environments. Elizabeth M. Wenzel (NASA–Ames Res. Ctr., M.S. 262-2, Moffett Field, CA 94035-1000, bwenzel@mail.arc.nasa.gov)

Engineering constraints that may be encountered when implementing interactive virtual acoustic displays are examined. In particular, system parameters such as the update rate and total system latency are defined and the impact they may have on perception is discussed. For example, examination of the head motions that listeners used to aid localization in a previous study suggests that some head motions may be as fast as about 175°/s for short time periods. Analysis of latencies in virtual acoustic environments (VAEs) suggests that: (1) commonly specified parameters such as the audio update rate determine only the best-case latency possible in a VAE, (2) total system latency and individual latencies of system components, including head-trackers, are frequently not measured by VAE developers, and (3) typical system latencies may result in undersampling of relative listener-source motion of 175°/s as well as positional instability in the simulated source. To clearly specify the dynamic performance of a particular VAE, users and developers need to make measurements to average system latency, update rate, and their variability using standardized rendering scenarios. Psychoacoustic parameters such as the minimum audible movement angle can then be used as target guidelines to assess whether a given system meets perceptual requirements.

2:50

4pSP4. Immersive audio for desktop systems. Chris Kyriakakis, Tomlinson Holman, Hartmut Neven, and Christoph von der Malsburg (USC, 3740 McClintock Ave., EEB432, Los Angeles, CA 90089-2564)

Numerous applications are envisioned for integrated media workstations to create, edit, and accurately monitor digital video and audio. There are two major classes of limitations that impede the performance of current desktop loudspeaker-based sound systems. The first encompasses problems that arise due to the local acoustical environment, such as early reflections from the CRT and nearby flat surfaces, as well as the design characteristics and performance of the loudspeakers. The second class involves limitations that arise from variations in human listening characteristics and listener movement relative to the loudspeakers. In this paper several findings are presented that are based on acoustical, psychoacoustical, and signal processing methods for delivering accurate sound field representations. Furthermore, a novel vision-based method for accurate listener tracking is presented that eliminates the need for headgear or tethered magnetic devices. Current findings show that this algorithm can be implemented without imposing a significant computational overhead to the host processor. Research directions include adaptive real-time HRTF synthesis based on the tracking information, as well as vision-based pinna shape classification for improved performance.

3:10–3:20 Break

3:20

4pSP5. Distributed-mode loudspeakers and their impact on intelligibility in multimedia and sound distribution. Peter Mapp (Peter Mapp Assoc., Colchester, UK) and Henry Azima (New Transducers Ltd., Stonehill, Huntingdon, PE18 6ED, UK)

A new class of acoustic radiator has been investigated and shown to exhibit unusual radiation characteristics that can produce enhanced intelligibility under given listening conditions. The improvement mechanisms are investigated and shown to be related to a number of unique features. These include transient response, synergetic decay, lower inherent distortion, directivity, and the diffusivity of acoustic radiation. In a sound system design, it is desirable to produce an even coverage in both spatial and spectral domains. Conventional radiators, the majority of which work under a pistonic regime, exhibit two important characteristics that limit their effectiveness in real-world applications. These are the well-known narrowing of the acoustic radiation (dispersion) angle with increasing frequency and the coherency of the radiation. In contrast, the distributed-mode radiators, an emerging new technology, tend to overcome these obstacles and offer a new approach to improved sound distribution and multimedia. They combine a flat power response with an essentially constant and wide directivity (independent of frequency), augmented by a spatially and temporally diffuse radiation characteristic. The fast transient response results in a signal broadcast with very high clarity. The diffuse nature of the radiation suggests a lower degree of boundary interaction and spectral interference.

3:40

4pSP6. The measurement of coloration in electroacoustic enhancement systems. Mark Poletti (Industrial Res. Ltd., P.O. Box 31-310, Lower Hutt, New Zealand)

The use of electroacoustic systems to enhance the acoustics of concert halls is becoming increasingly prevalent. Regenerative systems are often used for obtaining a source position independent enhancement of reverberation time. However, these systems can produce unnatural acoustics due to the unequal enhancement of the response at different frequencies, a phenomenon termed coloration. Coloration alters the natural Rayleigh statistics of the transfer function magnitudes, and the variation in statistics can be used as a method for quantifying it. This approach avoids the difficulties of examining only the peaks in the transfer function magnitude. A method is developed for measuring the deviation of room transfer functions from Rayleigh statistics. Variance measures are derived. The underlying probability density function is modeled by fitting a three-parameter density termed the generalized gamma function. The maximum deviation of the modeled cumulative distribution from the Rayleigh distribution provides an alternative measure of coloration. The sensitivity of the measures is improved by eliminating the dominant early part of the impulse response, which is largely unaffected by non-in-line systems. An assisted reverberation system simulator is used to simulate the VRA system, which includes a unitary multichannel reverberator in the feedback loop to pro-

vide room volume expansion. The simulations show that the coloration measures increase with loop gain as expected, and are independent of other parameters in the system such as the secondary room reverberation time.

4:00

4pSP7. Criteria for accurate acoustic emulation. David McGrath (Lake DSP pty. Ltd., Sydney, Australia)

The technique of auralization was pioneered by acoustics researchers in the field of concert hall simulation, where an accurate rendering of the acoustics implied that all characteristics of an acoustic space were reproduced, wherever possible. As the convolution technology is pushed into more cost-effective implementations, the broader potential for this technology is being realized. Now the criteria for acoustic emulation must be reevaluated to take into account a number of new factors. This paper explores many relevant criteria for repeatability testing and characterizing the performance of equipment in the presence of realistic three-dimensional noise fields and complex acoustic environments. Such an acoustic system can be accurately emulated using real-time simulation hardware, if a complete characterization of the system can be obtained. This paper discusses methods that are appropriate for the implementation of accurate acoustic emulation, with the emphasis on real world problems and solutions. It is shown how the entire acoustic signal path can be replaced by an appropriate low latency FIR filter, defined using new techniques for impulse response measurement.

4:20

4pSP8. Multichannel audio signal compression and quality assessment for AV communications. Jin-Woo Hong, Dae-Young Jang, and Seong-Han Kim (Electron. and Telecommunications Res. Inst., Yusong, P.O. Box 106, Taejon, 305-600, Korea, jwhong@etri.re.kr)

This paper describes an algorithm for multichannel audio signal compression using the psychoacoustic model, an implementation and the result on quality assessment of real-time operating multichannel audio codec, and a multiplexing technique of synchronization for AV communication. The algorithm which is based on the MPEG-2 audio standard, utilizes psychoacoustic modeling, sub-band analysis and synthesis, and bit allocation for information as key technology. The audio codec was implemented on real-time operation by using a flexible DSP system in which an encoder and a decoder are independently designed for processing a formatted digital audio bit stream, and can compress the multichannel audio data of about 3848 kbits/s into 384 kbits/s. A subjective test for quality assessment of multichannel audio codec was performed using double blind and triple stimuli with hidden reference method. In the test condition with 17 subjects and ten test materials, the test result with 95% confidence intervalshows that all test materials except one (pitch pipe sound) are awarded mean difference grades better than -0.6 and this audio codec is acceptable for AV communication or broadcasting. Finally, a multiplexing method to control time stamp for synchronization of AV communication is, in this paper, introduced from the implementation of HDTV system using this codec.

4:40

4pSP9. Reproduction of CD by providing ultrasonic atmosphere. Jouji Suzuki, Shin'ichiro Kaneko (Dept. of Information and Comput. Sci., Saitama Univ., 255 Shimo-okubo, Urawa, 338 Japan), and Kiyomitsu Ohno (Nihon Tech. Ctr., Tachikawa, Tokyo, 190 Japan)

It is said that the LP sound reproduced by a well-prepared system is superior to the sound of a CD and the difference stems from whether the signal is recorded by analog or digital form. Considering master tape of LP is recorded in digital form, the main difference exists in the bandwidth of the reproduced signal. The signal of a CD is restricted to 20 kHz. On the other hand, the spectrum of an LP signal is extended to 100 kHz. Though the upper limit of hearing acuity of a human is 20 kHz, superior components will contribute to offer more rich and realistic atmosphere. In order to produce a higher-frequency component from CD sound, the folding spectrum of the sampling frequency of 44.1 kHz is utilized. The output signal of a DA converter of a CD player is filtered with BPF of 24–42 kHz. This signal is added to the CD output and presented to listeners. This system is named PULSA (providing ultrasonic atmosphere). A preference test is conducted on conventional CD sound and the sound by PULSA through a wideband static headphone on three minuets of orchestra and opera. It is revealed that PULSA sound is better than the conventional by the ratio of 65 to 35.

5:00

4pSP10. A multi-processor system for the production of virtual sound fields. Felipe Orduña-Bustamante, Ricardo R. Boullosa, and Antonio P. López (Centro de Instrumentos UNAM, Circuito Exterior CU, CP 04510, México DF, Mexico, felipe@aleph.cinstrum.unam.mx)

A multi-processor system is described for the production of virtual sound fields using small arrays of up to four high-quality loudspeakers and up to four control microphones distributed around an artificial head and torso in a lightly damped listening room. The digital signal processing engine consists of four floating point processors interconnected for parallel computation, together with digital and analog converters. The performance of the system is analyzed in terms of the tradeoff between the maximum sampling rates that can be used and the complexity of the digital filter matrices in the different operating modes of the system such as system identification, single-input multiple-error adaptive filtering, and full multichannel filtering with fixed coefficients.

Session 4pUW

Underwater Acoustics: 3-D Propagation Effects II: Where are We Today in Models and Measurements?

Michael B. Porter, Cochair
Scripps Institute of Oceanography, Marine Physical Laboratory, MC 0205, 8602 La Jolla Shores Drive, La Jolla, California 92037-0205

Alexandra I. Tolstoy, Cochair
4224 Walalae Avenue, Suite 5-260, Honolulu, Hawaii 96816

Chair's Introduction—1:00

Invited Papers

1:05

4pUW1. The 3-D effect on underwater acoustic propagation in the offshore area of Taiwan's coast. Chi-Fang Chen, Jang-Jia Lin (Dept. of Naval Architecture and Ocean Eng., Natl. Taiwan Univ., 73 Chou-Shan Rd., Taipei, Taiwan, ROC), Chung-Wu Wang (Naval Oceanograph. and Hydrographic Office, ROC), and Ding Lee (Yale Univ./Naval Undersea Warfare Ctr., New Haven, CT)

Lately 3-D effects on underwater acoustic propagation have frequently been reported. The major causes for the 3-D effects are the variations in azimuth of bottom topography or water column properties. The offshore regions of Taiwan's coast are of the similar nature. There are submarine canyons in the offshore region of southwest Taiwan. The acoustic propagation is studied numerically using FOR3D. The 3-D effect is found to be more serious along the axis of the canyon's axis than cross the canyon. The field test is also conducted in the region; the measurements are well matched with the numerical results calculated by FOR3D, which indicates the actual 3-D effect exists in the real ocean environment. [Work supported by National Science Council of Republic of China.]

1:25

4pUW2. High-resolution, three-dimensional measurements of low-frequency sound propagation in shallow water. George V. Frisk, Kyle M. Becker, Laurence N. Connor, James A. Doutt, and Cynthia J. Sellers (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

An overview is presented of the modal mapping experiment (MOMAX), which was conducted in March 1997 in the vicinity of the East Coast *STRATAFORM* site. Both fixed and moving source configurations were used to transmit several pure tones in the frequency range 50–300 Hz. The magnitudes and phases of these signals were recorded on several freely drifting buoys, each containing a hydrophone, GPS and acoustic navigation, and radio telemetry. High-resolution, three-dimensional measurements of the sound field were made out to ranges of 10 km and illustrate the influence of the laterally varying seabed, measured with a chirp sonar system. The precision navigation also enabled the creation of a two-dimensional, synthetic aperture planar array, parallel to the ocean surface. The pressure field data measured on the array were transformed into the wave number domain, where the lateral variability of the waveguide manifests itself in the spatially evolving spectral content of the modal field. Finally, the phases of the measured signals show remarkable stability and regularity, even in the context of complex, multimodal fields. This behavior can be exploited to make accurate estimates of the relative source/receiver speed from measurements of the time rate-of-change of the phase. [Work supported by ONR.]

1:45

4pUW3. A fully global hydroacoustic monitoring system for the Comprehensive Nuclear-Test-Ban Treaty—Plans and progress. Martin W. Lawrence and Marta Galindo Arranz (Provisional Tech. Secretariat, Comprehensive Nuclear-Test-Ban Treaty Organization, Vienna Intl. Ctr., A-1400 Vienna, Austria, mlawrence@ctbto.un.or.at)

The first ever fully global hydroacoustic monitoring system is being planned and implemented for use in the verification of a new international treaty, the Comprehensive Nuclear-Test-Ban Treaty (CTBT). This system will provide hydroacoustic monitoring of all the world's oceans for 24 h a day, every day of the year, into the indefinite future. This unique resource will utilize two types of station. One type will be based on a hydrophone at the SOFAR axis depth, cabled back to shore. The other will be based on a seismometer on a small island using detection of the T-phase signal. This latter station relies on a signal which has propagated predominantly through the ocean, but has been converted to seismic energy at the margin of the island. Although this network is being installed for monitoring of nuclear explosions, it will also be a unique resource for scientific investigation of various phenomena. Data will be available to all State Signatories to the CTBT.

2:05

4pUW4. Validation of source region energy partition calculations with small-scale explosive experiments. Douglas B. Clarke, Philip E. Harben, Steven L. Hunter, and Donald W. Rock (Lawrence Livermore Natl. Lab., Livermore, CA 94551)

The decrease in signal energy as the location of a nuclear explosion varies from deep in the ocean to above the ocean surface is a concern for the planned ocean monitoring component of the Comprehensive Test Ban Treaty. Small-scale experiments were designed to validate predictions of energy coupling by nuclear explosions in the ''source region,'' the origin of the signals that propagate in the deep underwater sound (SOFAR) channel. The experiments were performed in a biologically dead lake at a scale length of 1/50 (1/125 000 in explosive energy) relative to one kiloton using 6.82-kg charges of Pentolite 50/50. The acoustic energy coupled into the water was monitored at a 60-m range by a hydrophone string with eight piezoelectric sensors spaced from near-surface to a 30-m depth. Useful data were obtained at five burst locations: 5, 2, 0, −2, and −15 m. Results from the experiments and new calculations support the predicted energy partitioning for above-surface explosions with model and experiment peak pressures agreeing within a factor of two over three orders of magnitude variation. [Work performed under the auspices of the U. S. Department of Energy by the Lawrence Livermore National Laboratory under Contract W-7405-ENG-48.]

2:20

4pUW5. Three-dimensional propagation modeling in shallow water. Gregory J. Orris and John S. Perkins (Naval Res. Lab., Washington, DC 20375)

Much interest exists in problems of shallow-water propagation, where three-dimensional spatially dependent waveguide effects are sometimes very large. In these cases the three-dimensional dependence of the acoustic field can be accounted for in a variety of ways other than solving the often intractable fully coupled problem. There exist several pseudo three-dimensional models that are based on the parabolic equation method (e.g., $N\times 2$D PE [J. S. Perkins and R. N. Bear, J. Acoust. Soc. Am. **72**, 515–522 (1982)], the adiabatic PE [M. D. Collins, J. Acoust. Soc. Am. **93**, 2269–2278 (1993)], the coupled-mode PE [A. T. Abawi *et al.*, J. Acoust. Soc. Am. **102**, 233–238 (1997)] and the spectral PE [G. J. Orris and M. D. Collins, J. Acoust. Soc. Am. **96**, 3499–3503 (1994)]). Yet, because these models are extensions of efficient two-dimensional models they have explicit and implicit limitations on their applicability. Some of the advantages and limitations of these pseudo three-dimensional models will be demonstrated through their application to data collected during the SWellEx-96 and RDS-1 matched-field experiments.

2:35–2:40 Break

2:40

4pUW6. 4-D modeling of sound propagation in shallow water with anisotropic sediment layers. X. Tang, M. Badiey (Marine Studies, Univ. of Delaware, Newark, DE 19716), and W. L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180)

A highly efficient 4-D PE model is developed to study broadband sound propagation at the Atlantic Generating Station (AGS) site. Extensive core data has shown that the sediment at this AGS site consists of multiple layers which vary in both range and azimuth. Consequently, sound propagation in the water column in this area is significantly affected by the complex sediment structure due to the fact that the layered sediment acts like a bandpass filter in which broadband acoustic energy is selectively broken into narrow-band components propagating in different layers, usually referred as sediment mode trapping. The details of the range, azimuth, frequency, and bandwidth dependences as well as their relationships to the sediment structure of such mode trapping effect is numerically explored by performing the 4-D PE model. The empirical orthogonal function representation method is employed to transfer the geoacoustic parameters of the sediment core data into PE inputs. The broadband travel time results at eight different bearings are compared with the existing experimental data. The importance of the 3-D effects is also examined by comparing the numerical results of azimuthal coupled and uncoupled models.

2:55

4pUW7. Wave number extraction techniques for a three-dimensionally varying shallow-water waveguide: A comparison. Kyle M. Becker and George V. Frisk (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, kbecker@whoi.edu)

Theory shows that the wave number spectrum of a propagating normal mode field is a function of position for a complex, shallow-water waveguide whose acoustic properties vary in three spatial dimensions. By describing the spatially varying content of the modal field, a direct measure of the propagation characteristics of the waveguide can be made. Consequently, to obtain greater sensitivity of the spectrum as a function of position, it is of interest to resolve the contributions from individual modes using as short a data record as possible. To this end, several wave number extraction methods have been applied to synthetic data, as well as real data measured using a synthetic aperture planar array created by several freely drifting buoys. Wave number content as a function of position was extracted using Hankel transforms, best linear estimator techniques, and autoregressive algorithms. Results obtained for frequencies in the range 50–300 Hz are compared for resolution and consistency using data records of various lengths. Based on these results, recommendations are made for the selection of an appropriate extraction algorithm. [Work supported by ONR.]

3:10

4pUW8. Range-dependent matched-field source localization and tracking in shallow water on a continental slope. Lieutenant(N) Martin L. Taillefer and N. Ross Chapman (SEOS, Univ. of Victoria, P.O. Box 3055, Victoria, BC V8W 3P6, Canada)

This paper describes matched-field source localization and tracking of a sound source towed over a continental slope off the western coast of Vancouver Island. The data were collected using a multielement vertical line array (MEVA) as part of the PACIFIC SHELF sea trial carried out during September 1993. The MEVA is a 16-element array with sensors evenly spaced at depths between 90 and 315 m. The target was towed along a linear radial track over a steep slope out to ranges of 5.5–6 km from the array, and then along a navigational arc in water depths from 250 m at the array to 750 m. The target emitted three continuous wave tones in the band 45–72 Hz. In order to model the environment as accurately as possible, the replica fields were calculated using an adiabatic normal-mode approximation which accounted for elastic wave propagation in the bottom. In addition, a bathymetric database was compiled from various echo-soundings of the area. Using conventional linear MFP based on two-dimensional replica fields to determine ambiguity surfaces for any target radial from the array, the target was tracked in range, depth, and bearing along the entire experimental track, with Bartlett processor values as large as 0.87.

3:25

4pUW9. Experimental investigation of matched-field processing in a wedgelike shallow-water environment. Paul A. Baxley (Space and Naval Warfare Systems Ctr., San Diego, Code D881, 49575 Gate Rd. Rm. 170, San Diego, CA 92152-6435, baxley@spawar.navy.mil)

Underwater acoustic propagation in shallow-water wedgelike environments will experience ''bending'' out of the vertical plane containing the source and receiver. Consequently, the time for the energy to travel from

the source to the receiver will be altered, or some paths may miss the receiver altogether. Localization estimates from array processors will be in error if this horizontal multipath propagation is ignored. This phenomenon is investigated experimentally via the analysis of vertical-line-array source-tow data recorded in a steep wedgelike environment near San Clemente Island during the fourth Shallow Water evaluation cell Experiment (SWellEX-4). The source-tows examined were approximately in the cross-slope direction. The objective of this study is (1) to observe the horizontal ''refraction'' effects via an examination of the matched-field localization errors resulting from a neglect of that phenomenon, and (2) to study the feasibility of including modeled horizontal refraction effects into the matched-field processor to enhance localization performance. Geoacoustic parameter inversion was performed at short range so that mismatch in this parameter would not be misinterpreted as a horizontal ''refraction'' effect. A three-dimensional Gaussian beam program was used to generate replica fields, and to perform investigative simulations for the experimental environment. [Work supported by ONR/SPAWARSYSCEN San Diego.]

THURSDAY AFTERNOON, 25 JUNE 1998 GRAND BALLROOM A & B (S), 4:00 TO 5:00 P.M.

Session 4pPL

Plenary Lecture

Sabih I. Hayek, Chair
Department of Engineering Mechanics, Pennsylvania State University, 227 Hammond Building, University Park, Pennsylvania 16802

Chair's Introduction—4:00

Invited Paper

4:05

4pPL1. Trends in modeling of structural-acoustics systems with structural complexity in low- and medium-frequency ranges.
Christian Soize (ONERA, BP 72, 92322 Chatillon Cedex, France, soize@onera.fr)

This paper gives a comprehensive survey of a theoretical approach for predicting the frequency response functions of general complex structural acoustics systems (CSAS) in the low- and medium-frequency ranges. For such a CSAS, the complex structure is made of a master structure coupled with internal substructures presenting structural complexity and corresponding to a large number of secondary subsystems, such as equipment units or internal secondary structures, attached to the master structure. The master structure is a general 3-D dissipative structure with an arbitrary bounded geometry and is made of an anisotropic, inhomogeneous, viscoelastic medium. This complex structure is coupled with an external acoustic fluid via the master structure. The approach is based on the use of fuzzy structure theory, introduced by the author in 1985, including recent developments concerning identification of the fuzzy substructure model parameters. This theory allows the effects of the internal structural complexity on the master structure to be modeled. In addition, the construction of an intrinsic reduced model of the frequency response functions of this CSAS adapted to the medium-frequency range (recently proposed by the author) is presented, which allows simplifications in the calculation of the responses to any deterministic or random excitations.

Meeting of Accredited Standards Committee S3 on Bioacoustics

L. S. Finegold, Chair S3
USAF Armstrong Laboratory, Noise Effects Branch AL/OEBN, 2610 Seventh Street, Wright Patterson Air Force Base, Ohio 43433-7901

R. F. Burkard, Vice Chair S3
Hearing Research Laboratory, State University of New York at Buffalo, 215 Parker Hall, Buffalo, New York 14214

P. D. Schomer, Chair, U. S. Technical Advisory Group (TAG) for ISO/TC 43, Acoustics
U.S. CERL, P.O. Box 9005, Champaign, Illinois 61826-9005

J. Erdreich, Chair, U. S. Technical Advisory Group (TAG) for ISO/TC 108/SC4, Human Exposure to Mechanical Vibration and Shock
Ostergaard Acoustical Associates, 100 Executive Drive, West Orange, New Jersey 07052

H. E. von Gierke, Vice Chair, U. S. Technical Advisory Group (TAG) for ISO/TC 43, Acoustics and ISO/TC 108/SC4, Human Exposure to Mechanical Vibration and Shock
1325 Meadow Lane, Yellow Springs, Ohio 45387

V. Nedzelnitsky, U.S. Technical Advisor (TA) for IEC/TC 29, Electroacoustics
National Institute of Standards and Technology (NIST), Building 233, Room A149, Gaithersburg, Maryland 20899

Accredited Standards Committee S3 on Bioacoustics. The current status of standards under preparation will be discussed. In addition to those topics of interest, including hearing conservation, noise, dosimeters, hearing aids, etc., consideration will be given to new standards which might be needed over the next few years. Open discussion of committee reports is encouraged.

SCOPE OF S3: Standards, specifications, methods of measurement and test, and terminology in the fields of mechanical shock and physiological acoustics, including aspects of general acoustics, shock, and vibration which pertain to biological safety, tolerance, and comfort.

Meeting of Accredited Standards Committee S1 on Acoustics

J. P. Seiler, Chair S1
U. S. Department of Labor, Cochran Mill Road, P.O. Box 18233, Building 038, Pittsburgh, Pennsylvania 15236

G. S. K. Wong, Vice Chair S1
Institute for National Measurement Standards, (INMS), National Research Council, Ottawa, ON K1A 0R6, Canada

P. D. Schomer, Chair, U. S. Technical Advisory Group (TAG) for ISO/TC 43, Acoustics
U. S. CERL, P.O. Box 9005, Champaign, Illinois 61826-9005

H. E. von Gierke, Vice Chair, U. S. Technical Advisory Group (TAG) for ISO/TC 43, Acoustics
1325 Meadow Lane, Yellow Springs, Ohio 45387

V. Nedzelnitsky, U. S. Technical Advisor (TA) for IEC/TC 29, Electroacoustics
National Institute of Standards and Technology (NIST), Building 233, Room A149, Gaithersburg, Maryland 20899

Accredited Standards Committee S1 on Acoustics. Working group chairs will report on their preparation of standards on methods of measurement and testing, and terminology, in physical acoustics, electroacoustics, sonics, ultrasonics, and underwater sound. Work in progress includes measurement of noise sources, noise dosimeters, integrating sound-level meters, and revision and extension of sound level meter specifications. Open discussion of committee reports is encouraged.

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.

Session 5aPLa

Plenary Lecture

Andrea Prosperetti, Chair
Mechanical Engineering, Johns Hopkins University, 122 Latrobe Hall, 3400 North Charles Street, Baltimore, Maryland 21218

Chair's Introduction—8:00

Invited Paper

8:05

5aPLa1. Acoustics of two-phase fluids and sonoluminescence. Robert I. Nigmatulin (Ufa (Bashkortostan) Branch of Russian Acad. of Sci., K. Marx Str. 6, Ufa 450000, Russia, iskander@ncan.ufanet.ru)

The basic equations for dynamics of two-phase mixtures like gas–particle suspensions and bubbly liquids are presented. Heat and mass exchange phenomena near a drop, particle, and gas or vapor bubble are discussed. It is shown that these phenomena play an important role in propagation, attenuation, and amplification of sound waves and shock waves in gas–liquid systems, sometimes leading to paradoxical effects. The dynamics of a sonoluminescing bubble is discussed. There are two stages of the bubble oscillation process: The low Mach number stage when the velocity of the bubble interface is small compared with sound speed in the liquid, and the stage corresponding to the collapsing bubble's compression, when the velocity of the interface may be larger than the sound speed. The analytic solution for the low Mach stage is presented. This solution provides the economical boundary condition around the bubble for an effective numerical code.

Session 5aAA

Architectural Acoustics: Case Studies of Performance Spaces

Jerald R. Hyde, Chair
Consultant in Acoustics, Box 55, St. Helena, California 94574

Chair's Introduction—7:20

Invited Papers

7:25

5aAA1. The Hong Kong Cultural Centre halls: Acoustical design and measurements. A. Harold Marshall, Johan L. Nielsen, and M. Miklin Halstead (Acoust. Res. Ctr., Univ. of Auckland, New Zealand)

The main performance spaces in the Hong Kong Cultural Centre comprise the Concert Hall, seating 2000 and the Grand Theatre seating 1700. The design of the former is derived from the Christchurch Town Hall which it develops in several important ways, while the latter is a lyric theatre/opera hall with several unique features which are described briefly in this paper. The Centre opened in 1989 following extensive commissioning measurements in the halls. Recently the opportunity has arisen to repeat the measurements using an improved technique in an attempt to document changes following a stage extension. Results from these measurements are compared with the earlier series in each hall. Marshall Day Associates, Auckland, New Zealand was responsible for the room acoustical design of these spaces. [The recent visit to Hong Kong was supported by the Acoustics Research Centre, the Hong Kong Cultural Centre, and the Marsden Fund administered by the Royal Society of New Zealand.]

7:45

5aAA2. Modern measurements, optimized diffusion, and electronic enhancement in a large fan-shaped auditorium. John P. O'Keefe (Aercoustics Eng. Ltd., 50 Ronson Dr., Ste. 127, Toronto, ON M9W 1B3, Canada), Trevor Cox (Univ. of Salford, UK), Neil Muncy (Neil Muncy Assoc., Toronto, Canada), and Steve Barbar (LARES Assoc., Cambridge, MA)

The Hummingbird Centre in Toronto Canada, formerly known as the O'Keefe Centre, is a 3000 seat fan-shaped auditorium, typical of the postwar era. Since it opened in 1960 it has been plagued by complaints about its acoustics. A LARES electronic enhancement system was proposed for the building and a feasibility study was carried out. Modern acoustical measurements lend

credence to the years of complaints. Among other interesting findings, late energy (Glate) is 10–20 dB lower than a traditional shoe-box-shaped concert hall. Such low levels of late energy explain, in part, the image shift and echoes that are found throughout the hall. A side wall echo threatened to compromise the proposed enhancement system. In an effort to reduce the echo, a crescent-shaped diffuser was developed using a BEM optimization routine. Four LARES mainframes are employed in the enhancement system with more than 280 front ported two-way loudspeakers, each of which can be addressed individually. The renovated acoustics have been well received by the owners, users, and patrons.

Contributed Papers

8:05

5aAA3. Recent acoustical measurements in the Christchurch Town Hall Auditorium. A. Harold Marshall, Johan L. Nielsen, and M. Miklin Halstead (Acoust. Res. Ctr., Univ. of Auckland, Private Bag 92019, Auckland, New Zealand, h.marshall@auckland.ac.nz)

Two significant changes have occurred in the Christchurch Town Hall during 1996-97: the original seats have been replaced with a new type and a new Rieger pipe organ has been installed behind the choir seats. A new comprehensive set of acoustic measurements has been made in the unoccupied hall. These are reported and compared in this paper with earlier measurements. A related study on the usefulness of C80 as a clarity measure in concert halls and which makes use of these results is reported in a companion paper. [This work is supported by the Marsden Fund administered by the Royal Society of New Zealand.]

8:20

5aAA4. Calculation and measurement of acoustic factors for the Kirishima International Concert Hall. Tatsumi Nakajima (Takenaka Res. and Development Inst., 5-1, 1-chome Ohtsuka, Inzai, Chiba, 270-13 Japan) and Yoichi Ando (Kobe Univ., Rokkodai, Nada, Kobe, 657 Japan)

The Kirishima International Concert Hall was designed by architect F. Maki with Ando's acoustic design theory. It opened on July 22, 1994. The 770 seat main hall which has a volume of 8475 m^3 was primarily designed for chamber music. Computer simulations made it clear that the acoustical design of the main hall would be an optimum solution with a plan like a leaf shape and a cross section like a bottom of a ship. These were solutions which satisfied requirements for both acoustical and architectural design concepts of the main hall. Getting a low IACC sound field at the front seat was most important. In order to get a low IACC, sound reflections from the stage to the audience floor were examined and studied in the early design processes. Calculation and measurement of acoustic factors of the sound fields in the main hall show: (1) the calculated and the measured IACCs are good in agreement, (2) IACCs tended lower values of about 0.1 than the calculated ones throughout the seats, and (3) the number of reflections needed for predicting IACC is recommended to be about 40 for the practical method of IACC calculation at the design stage of a hall.

8:35

5aAA5. Interrelationship of musical excellence and acoustical excellence: A case study of the Gewandhaus, Leipzig. Pamela Clements (Acoust. Consultant, Jaffe Holden Scarbrough Acoust., Inc., 114A Washington St., Norwalk, CT 06854, pclements@jhsacoustics.com)

A study of changes in orchestral performance practice and repertoire in the Altes Gewandhaus and the Neues Gewandhaus (opened 1884) under the conductor Carl Reinecke was performed. A parallel study of changes in orchestral performance practice and repertoire in the Neues Gewandhaus was also performed when Carl Reinecke was succeeded by Arthur Nikisch in 1895. In the first instance, the orchestra and conductor are the same but the hall changes; in the second instance, the orchestra and hall are the same but the conductor changes. Both old and new halls were regarded as outstanding acoustically, but under Reinecke the orchestra's reputation waned because of conservatism in programming and interpretation. Historical evidence shows that the orchestra went through a difficult period accommodating the new hall's acoustics and emerged with a new sound. Yet, despite the hall's excellent acoustics, the orchestra performed better in another, poorer, acoustic environment under the more dynamic conductor Arthur Nikisch. The orchestra only regained its international reputation later when Nikisch became conductor at the Gewandhaus. The study explores Nikish's response to the hall's acoustic possibilities, and suggests reasons why a hall with a relatively dry acoustic could be pivotal in the orchestra's then becoming a foremost interpreter of the late Romantic repertoire.

8:50

5aAA6. Acoustical design and the characteristics of Sapporo Concert Hall. Yasuhisa Toyota and Katsuji Naniwa (Nagata Acoust., Inc., Minami-Shinjuku-Hoshino Bldg. 8F, 5-23-13 Sendagaya, Shibuya-ku, Tokyo, 151-0051 Japan, toyota@nagata.co.jp)

Sapporo Concert Hall, with 2008 seats in the large hall and 453 seats in the small hall, was opened on July 4, 1997. Both halls were designed for classical music, the large one mainly for orchestral music and the small one for chamber music and recitals. The large hall was designed with the concept of the surrounded stage and ''vineyard steps'' seating arrangement like Philharmonic Hall in Berlin and Suntory Hall in Tokyo. The emphasized points in acoustical design were as follows: (1) Effective use of the ''vineyard steps'' walls which were introduced into the audience area to provide the effective early reflections; (2) installation of the big sound reflectors suspended above the stage area to provide the effective early reflections to musicians and audience around the stage; (3) introduction of heavy material, 150-millimeters-thick concrete, into the main ceiling to get enough response in the low frequencies. The small hall was designed as a typical shoebox hall with narrow width and high ceiling. Acoustically tunable curtains were introduced to make the wide range of acoustical possibilities. The results of the acoustical design and its characteristics will be reported.

9:05

5aAA7. Acoustical renovation of The Opheum Theatre, Vancouver Canada. John P. O'Keefe (Aercoustics Eng. Ltd., 50 Ronson Dr., Ste. 127, Toronto, ON M9W 1B3, Canada), Gilbert A. Soulodre (Commun. Res. Ctr., Ottawa, ON K2H 8S2, Canada), and John S. Bradley (Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

The Orpheum Theatre is a 2800 seat vaudeville house that was renovated for the Vancouver Symphony in the 1970s. Funds ran out prior to completion and some problematic conditions remained for the following fifteen years. Perhaps the most significant was an image shift heard on the balcony. Two focused reflections arriving at approximately 50 ms made some source locations on the stage appear to be perched in the ceiling. The efficacy of the plastic reflectors above the stage was also questioned. These and other issues were addressed with a complete set of stage and audience measurements in the full-scale room, experiments with a 1:48 small scale model, and listening tests using anechoic music convolved with the full-scale binaural impulse responses. The latter was particularly important because it allowed for reliable quantification of image shift thresholds that were unique to this room. Small scale modeling proved useful in matching the new convex reflectors' radii of curvature to the image shift threshold. The renovation work is complete and the image shift has been completely eradicated.

9:20–9:30 Break

9:30

5aAA8. The best remaining seat. Bodil Vaupel (Arkitekturt Akuslik, Paludan Müllers Plads 1, 5300 Kerteminde, Denmark)

It is important that an auditorium be designed to have as many good seats as possible. Not all seats in an auditorium are equally good. This is manifested in that the audience does not choose its seats randomly. The audience understands intuitively that, generally speaking, the closer a seat

is, and the more straight on, the better it is. Imagine an auditorium with open seating and the audience entering one at a time. People will in turn make a selection of what in their opinion is the *best remaining seat*. The order in which the seats are chosen is an indicator of the rank order of the desirability of the individual seats. As the audience makes its seat selection, a geometric pattern of the occupied seats unfolds, and reveals the boundary of the preferred seats. The perimeters describe the equal desirability curves and outlines the auditorium plan with as many good seats as possible. The audience choice of seat is recorded by time lapse photography. The data are analyzed in a combined computer drawing and mathematics program. A mathematical model has been developed that evaluates the desirability of the seating in an auditorium from the audience's point of view.

9:45

5aAA9. Acoustic design and performance of the Bruce Mason Theatre. Joanne M. Valentine (Marshall Day Asociates, P.O. Box 5811, Wellesley St., Auckland 1, New Zealand, mda@bitz.co.nz)

The Bruce Mason Theatre takes an innovative approach to variable acoustics with the use of a variable volume operable ceiling. The overall acoustic design objective was to achieve excellent acoustics with a high level of reverberance for symphony, and good speech clarity with a reduced level of reverberance for theater and drama. The results of commissioning measurements carried out in the auditorium show that a high degree of variability between the acoustic conditions for symphony mode and theater mode has been achieved, and that the operable ceiling/variable volume concept has proved to be highly successful. This paper presents the design concept developed for the auditorium, the findings of a 1:25 scale acoustic model study, and the results of commissioning measurements carried out in the newly opened theater in September 1996.

10:00

5aAA10. Three-dimensional impulse response measurements on S. Maria Church, Florence, Italy. Angelo Farina (Dipartimento di Ingegneria Industriale, Viale delle Scienze, 43100 Parma, Italy, farina@pcfarina.eng.unipr.it) and Lamberto Tronchin (DIENCA—CIARM Univ. of Bologna, Viale Risorgimento, 2 40136 Bologna, Italy)

The measurement of binaural impulse response procedure for acoustic environments has been established and achieved from the International Standard Organization; following the ISO 3382 many acoustical parameters have been already analyzed in the past, ranging from Royal Albert Hall in London, UK, to Teatro la Fenice in Venice, Italy. In this paper a new procedure for measuring 3D impulse response has been utilized, with the aim to reconstruct the 3D sound field especially for virtual reality purposes. Starting from the definition of B-format impulse responses, the pressure gradient has been measured in the Church by using a homemade apparatus and software, and performing measurements on the three axes, gathering seven different impulse responses, without using expensive microphones already developed. From the measures, the relations between the seven different measurements have been found out, and the results have been compared with the binaural impulse responses measurements performed in the same position at the same time.

10:15

5aAA11. Simultaneous measurements of room-acoustic parameters using different measuring equipment? Tor Halmrast (Statsbygg, pb. 8106 Dep. N-0032, Oslo, Norway), Anders Gade (Tech. Univ. of Denmark, DK-2600 Lyngby, Denmark), and Bjorn Winsvold (Norsonic, pb. 24, N-3408, Tranby, Norway)

Often the results from different room-acoustic measurements in the same hall disagree, and the disagreement is just said to be due to different measuring equipment, or different rigging/temperature, etc. The room acoustic of the Oslo Concert Hall was measured simultaneously, using the following different measuring equipment: (1) MLS/MLSSA (Statsbygg), (2) Sweep-Tone (Tech. Univ. Denmark), and (3) Norsonic 840 with MLS +MatLab. For some of the measurements (4) Pistol and (5) Electrical Impulse were also used. The paper will compare the results from the different measuring equipment, for the most known room-acoustic parameters. For the reverberation time parameters RT and EDT, very good agreement was found between the three main measuring equipments. For Ts and C80 the agreement between these three is good/fair for the higher frequencies, but less good for the bass, especially C80. The measurements with Electric Pulse and Pistol as signals (analyzed through Norsonic +MatLab) indicate good agreement for the reverberation times, but EDT is somewhat higher for the Pistol. For Ts and C80 the Electric Impulse and especially the Pistol give less clearness (higher Ts and lower C80), compared to MLSSA, Sweep Tone, and Norsonic/MLS.

10:30

5aAA12. Design of a circular hall improving the subjective preference at each seat. Akio Takatsu, Hiroyuki Sakai, Shin-ichi Sato, and Yoichi Ando (Grad. School of Sci. and Technol., Kobe Univ., Rokkodai, Nada, Kobe, 657 Japan)

The circular type of the medium-sized (400 seats) multipurpose event hall in Kobe Fashion Plaza, which was named Orvis Hall, was designed based on Ando's theory of subjective preference. Acoustic problems caused by such a circular plan were minimized by means of ceiling diffusion panels, side-wall-reflective panels, and a small room at the back wall preventing echo-disturbance created by the long-path echo or ''Whispering Gallery'' effects. One of the most remarkable systems of this hall is controlling the subsequent reverberation time through blending architectural acoustics and electrical acoustics by use of a hybrid system with both a reverberation control room and a digital reverberator. Therefore the total scale value of subjective preference at each seat is maximized by the four acoustic factors, namely, the delay time of first reflection, the listening level, the subsequent reverberation time, and the IACC. After construction of this hall, four factors at each seat were measured, and results are presented.

Session 5aAO

Acoustical Oceanography and Animal Bioacoustics: Acoustics of Fisheries and Plankton IV

John Hedgepeth, Chair
BioSonics, Inc., 4027 Leary Way, NW, Seattle, Washington 98107

Invited Paper

10:00

5aAO1. Vessel avoidance of wintering Norwegian spring spawning herring. Rune Vabø (Inst. of Marine Res., P. O. Box 1870, N-5024 Bergen, Norway), Kjell Olsen (Univ. of Tromsø, School of Fisheries, N-9000 Tromsø, Norway), and Ingvar Huse (Inst. of Marine Res., Bergen, Norway)

Vessel avoidance during acoustic abundance estimation of herring was investigated. Echo energy values measured from a submerged 38-kHz transducer were recorded and values at the time of passage were compared with corresponding echo energy values from a hull-mounted 18-kHz transducer on the passing vessel. A quantification of the loss in echo energy due to avoidance behavior is presented. Experiments are grouped into nighttime and daytime experiments. While a considerable drop in echo energy at the time of passage was observed at night, insignificant loss was found during the day. A general avoidance reaction pattern was identified in all the nighttime experiments. The magnitude of the vessel avoidance reaction was found to be strongly dependent on depth. Between 40 and 100-m depth some nighttime experiments showed loss in echo energy of the order of 90%. The effect of avoidance behavior decreased with depth below 100 m, but was seen down to 150 m at night. Low daytime response is due to deep daytime herring distribution. Body tilting and vertical and horizontal swimming are suggested explanations for the lower echo energy during passage.

Contributed Papers

10:20

5aAO2. FishMASS: ADCP technology adapted to split-beam fisheries echo sounding. R. Lee Gordon (RD Instruments, 9855 Businesspark Ave., San Diego, CA 92131, lgordon@rdinstruments.com) and Len Zedel (Memorial Univ. of New Foundland, Canada)

The objective of FishMASS, a NSF-funded research project, is to adapt elements of ADCP technology into split-beam echo sounding. Improvements come from wide bandwidths (50% bandwidth at 300 kHz), incorporation of velocity (Doppler) data, and addition of autonomous operation (battery and recording capacity to allow operation for up to a year). Wide bandwidths can reduce the uncertainty of single-target echoes from 3–6 dB (characteristic of narrowband signals) down to a fraction of a dB. It enables better differentiation of single-target from multiple-target echoes (when targets are separated either radially or at different ranges). The 50% bandwidth gives access to at least part of the target echo spectrum. These three advantages combine to improve both target discrimination and target classification. Split-beam sonars can sometimes track individual targets through their beams to obtain swimming speed. Doppler processing (identical to the processing of broadband ADCPs) will be added to obtain single-ping swimming speeds, both of individuals and groups. Mean velocities allow observation of migration and flux. Statistical products such as velocity standard deviation enable observation of feeding or predator-avoidance behavior. Correlations of velocity with, for example, TS may further aid species classification.

10:35

5aAO3. The digital transducer, new sonar technology. William Acker (BioSonics, Inc., 4027 Leary Way NW, Seattle, WA 98107)

The new digital architecture in sonar technology designed by BioSonics, Inc. has many advantages when compared with traditional analog technology. The echo signal is digitized inside the transducer at a very low signal level. This removes problems of coupled or induced noise when sending the signal through long lines. The system operates with a 3 to 4 dB noise figure in all modes of operation. Another important feature of the system is very high dynamics (>132 dB continuous). This allows receiving of low signals from small targets (i.e., plankton) simultaneously with high level signals (i.e., bottom) without saturation. The echo signal is sampled at a high rate and stored with high accuracy (0.013%) and high resolution (1.8 cm or less). The sonar is PC controlled for simple operation. The large dynamic range means no adjustment in system gain or transmit power. The data acquisition system is virtually automatic because most decisions are in postprocessing. Raw data is stored on the PC's hard drive or similar device. The sonar is available in a wide frequency range (38 kHz–1 MHz) and single, dual, or split beam transducer configuration depending on the application, i.e., biomass, layering effects, bottom typing, etc.

10:50

5aAO4. An acoustic tag ''radar-type'' tracking system for fish behavior studies near structures. John Hedgepeth, Dave Fuhriman (BioSonics, Inc., 4027 Leary Way NW, Seattle, WA 98107, jhedgepeth@biosonicsinc.com), Robert Johnson, and David Geist (PNNL, Battelle Memorial Inst., Richland, WA)

Recent studies of fish behavior around hydroelectric dams have used acoustics with split-beam methodology. A complementary methodology called the tracking transducer takes advantage of split-beam capabilities for expanding fish behavior investigations. The principle of tracking radar, aligning the antenna beam with a target, was applied with an acoustic transducer and dual-axis rotators for tracking individual fish over long periods of time. Deviation of the target from the beam axis produces a correction to point the axis toward the target. Two of these tracking systems have been used to triangulate the position of a small acoustic transmitter implanted in salmonid fish in the forebay of Lower Granite on the Snake River. This paper describes the system design, development, and early implementation. In October 1997 a 6-mm-diameter 200-kHz acoustic tag was placed in a rainbow trout and fish movement was measured near a surface bypass collector. The tracking equipment consisted of two 201-kHz, 6 deg half-power, full-beamwidth transducers placed 12.2 m

apart and 3.0 m below the water surface. Each split-beam transducer was mounted on a high-speed, dual-axis, motorized armature. These armatures were computer controlled and the system automatically followed the tag by using angle estimates provided by the echosounders.

11:05

5aAO5. Field trials using an acoustic buoy to measure fish response to vessel and trawl noise. Chris D. Wilson (Natl. Marine Fisheries Service, NOAA, Seattle, WA, cdwilson@afsc.noaa.gov)

A freely drifting acoustic buoy was constructed to evaluate the response of fish to vessel and trawl noise. The buoy contains an echosounder and split beam transducer operating at 38 kHz. Transducer heading data from the buoy are collected to assess the directivity of the fish response (i.e., movement relative to vessel and trawl). Geographic position of the buoy is monitored via GPS. Data are stored onboard the buoy and telemetered directly to the support vessel. The radio link between the buoy and vessel is used to control the echosounder, receive buoy positions, and remotely generate echograms in real time. GPS data from the buoy are also transmitted via the Argos satellite system to the vessel to locate the buoy in the event that visual, radar, and direct radio contact are lost. Field trials include repeated passes by the NOAA research vessel Miller Freeman past the buoy at different ship speeds while free-running and trawling to determine the behavioral response of walleye pollock (*Theragra chalcogramma*) to vessel and trawl noise. Performance of the buoy and interpretation of the results will be discussed.

11:20

5aAO6. Noise characteristics of Japanese fisheries' research vessels. Yoshimi Takao, Kouichi Sawada, Yoichi Miyanohana, Tsuyoshi Okumura (Natl. Res. Inst. of Fisheries Eng., Ebidai Hasaki, Kashima Ibaraki, 314-04 Japan, ytakao@nrife.affrc.go.jp), Masahiko Furusawa (Tokyo Univ. of Fisheries, Tokyo, 108 Japan), and Doojin Hwang (Yosu Natl. Fisheries Univ., Yosu, Chonnam, 550-749, Korea)

Noise is one of the important problems for hydroacoustic surveys of fisheries resources, because it may cause a large error in estimated results and it may shorten detectable range. It is necessary to estimate and reduce the contribution of acoustic noise for precise and accurate surveys. Noise is classified into low-frequency noise (audible by fish) and high-frequency noise (affecting echo sounders). High-frequency noise usually results in an overestimate of acoustic return. The average power of the high-frequency noise was measured by the echo-integrator. The echo-integrator output was converted into the equivalent noise spectrum level which can be compared with noises received by other quantitative echo-sounding systems or environmental noises. The noise measurements of several fisheries research vessels, ranging from 29.5-m length (RV ''Taka Maru'') to 93-m length (RV ''Kaiyo Maru''), were conducted to know their characteristics of noise. The dependence of noise upon the ship speed and the screw propeller's setting was investigated for each vessel. The vessel's performance for hydroacoustic survey is discussed from a viewpoint of noise level.

FRIDAY MORNING, 26 JUNE 1998 EAST BALLROOM B (S), 9:15 A.M. TO 12:15 P.M.

Session 5aBVa

Bioresponse to Vibration/Biomedical Ultrasound and Physical Acoustics: Lithotripsy I

Andrew J. Coleman, Cochair
Department of Medical Physics, St. Thomas Hospital, Lambeth Place Roth, London SE1 7EH, England

Robin O. Cleveland, Cochair
Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Invited Papers

9:15

5aBVa1. ESWL—The evaluation of a revolution. C. G. Chaussy (Staedt Krankenhaus München—Harlaching, München, Germany)

After 6 years of experimental research at the Departments of Urology and Surgical Research of the Ludwig—Maximilian University in Munich, extracorporeal shock-wave lithotripsy (ESWL) was introduced for clinical use in 1980. Uniquely successful and increasingly requested by stone patients, the method soon became widespread. Currently more than 2000 lithotriptors are in operation worldwide and over 5 000 000 treatments have been carried out successfully. Clinical experience in all centers has proven the safety, reliability, and reproducibility of the method. Currently approximately 70% of nonselected stone patients are eligible to receive ESWL treatment and, when combined with endourological procedures, more than 95% of patients can benefit from this method and thus avoid open surgery.

9:35

5aBVa2. Extracorporeal shock waves act by shock wave–gas bubble interaction. M. Delius (Inst. for Surgical Res., Univ. of Munich, Klinikum Grosshadern, 81366 Munich, Germany) and W. Eisenmenger (Institut für Physik 1, Univ. of Stuttgart, 70569 Stuttgart, Germany)

Previous animal experimental data had suggested that extracorporeal shock waves acted by the interaction of shock waves with remnant gas bubbles left from cavitational activity of previous shocks. *In vitro* experiments had shown that haemolysis from shock waves were reduced by $>95\%$ by static express pressures from only $1–4\times10^5$ kPa in the exposure vessel. This was interpreted as evidence for a reduced shock wave–gas bubble interaction by compression of the remnant gas bubbles by the excess pressure.

Additional experiments have shown that haemolysis was lower at $0.2-1\times10^5$ kPa excess pressure when the same number of shock waves was administered more slowly. Application of a single strong shock to red blood cells had little effect on haemolysis yet application of two shocks caused an increase of haemolysis by 500%. Static excess pressure abolished the increase. The only interpretation one can think of is that the first shock had caused remnant gas bubbles which acted strongly at the second shock. The term shock-wave gas-bubble interaction is well known from cavitation physics and should be used more often to describe the mechanism of action of extracorporeal shock waves.

9:55

5aBVa3. Can SWL-induced cavitation and renal injury be separated from SWL-induced impairment of renal hemodynamics? Andrew P. Evan (Dept. of Anatomy, Indiana Univ. School of Medicine, MS 259, 635 Barnhill Dr., Indianapolis, IN 46202), Lynn R. Willis, Bret A. Connors, James A. McAteer, James E. Lingeman, Robin O. Cleveland, Michael R. Bailey, and Lawrence A. Crum (Univ. of Washington, Seattle, WA 98105)

SWL to one kidney causes localized tissue damage and impairment of tubular function, but reduces renal plasma flow (RPF) in both kidneys. We examined the effect of SWL voltage (kV) and inversion of the waveform (IW) on these localized and bilateral effects of SWL. Five-week-old pigs were anesthetized for either sham-SWL or SWL (2000 shocks, unmodified HM3) at 12, 18, or 24 kV or 2000 shocks, 24 kV, with a reflector that inverts the waveform. RPF and tubular extraction of PAH (E_{PAH}) were measured 1 h before and 1 and 4 h after SWL. E_{PAH} estimates tubular secretion function. SWL significantly reduced RPF to similar degrees at each kV. E_{PAH} was not significantly reduced in the 12-kV group, but was reduced to progressively greater degrees in the 18 and 24 kV groups. IW eliminated ultrasonic evidence of cavitation, produced minimal tissue damage, and eliminated the reduction of E_{PAH}. It did not eliminate the reduction of RPF. The data suggest that shock-wave voltage and cavitation may be related to the tissue injury and reduced E_{PAH} induced by SWL, but suggests that neither may be directly related to the impairment of RPF. [Work supported by NIH, PO1 DK43881.]

10:15–10:25 Break

10:25

5aBVa4. The potential of lithotripter shock waves for gene therapy of tumors. Douglas L. Miller, Richard A. Gies, Brian D. Thrall (P7-53, Battelle Pacific Northwest Natl. Lab., P.O. Box 999, Richland, WA, doug.miller@pnl.gov), and Shiping Bao (Washington State Univ., Richland, WA)

Tissue destruction by lithotripter shock-wave-induced cavitation can be effective for antitumor therapy. Since DNA transfection can be accomplished through cavitation-induced sonoporation, the potential may exist for an advantageous combination of shock wave and gene therapy of tumors. B16 mouse melanoma cells were cultured by standard methods and a luciferase reporter vector was used as the DNA plasmid. The shock-wave generation system, similar to a Dornier HM-3 lithotripter, had peak pressure amplitudes of 24.4 MPa positive and 5.2 MPa negative. *In vitro* exposures of cell suspensions indicated that results were greatly enhanced by leaving an air space in the exposure chambers to promote cavitation activity. For *in vivo* exposure, cells were implanted subcutaneously in C57BL/6 mice 10–14 days before treatment. DNA at 0.2 mg/ml and sometimes air at 10% of tumor volume was injected intratumorally before exposure. Exposure to 800 shock waves, followed by culture of isolated tumor cells for one day, yielded 1.1 (0.43 SE) pg luciferase production per 10^6 cells, increasing to 7.5 (2.5 SE) pg/10^6 cells for air injection. Significant luciferase production occurred for 200, 400, 800, and 1200 shock waves with air injection. Gene transfer therefore can be induced during lithotripter shock-wave treatment of tumors. [Work supported by NIH Grant No. CA42947.]

10:45

5aBVa5. Effects of lithotripter fields on biological tissues. Diane Dalecki (Dept. of Elec. Eng. and the Rochester Ctr. for Biomed. Ultrasound, Univ. of Rochester, Rochester, NY 14627)

Biological effects resulting from exposure to lithotripter fields include hemorrhage in soft tissues, such as the kidney, lung, and intestine, the production of premature cardiac contractions, malformations in the chicken embryo, and killing of *Drosophila* larvae. Pulsed ultrasound can produce similar bioeffects at comparable pressure thresholds. Tissues that contain gas bodies, either naturally or after the addition of ultrasound contrast agents, are particularly susceptible to damage from low-amplitude lithotripter fields. Lung and intestine contain gas naturally and are hemorrhaged by exposure to lithotripter fields on the order of 1 MPa. After the introduction of an ultrasound contrast agent into the vasculature, many organs and tissues, such as the bladder, kidney, fat, muscle, and mesentery, show extensive hemorrhage after exposure to lithotripter pressures less than 2 MPa. Tissues near developing bone are also selectively susceptible to damage from exposure to low-amplitude lithotripter fields. The thresholds for hemorrhage in tissues near developing bone, such as the fetal head, limbs, and ribs, are all less than 1 MPa for exposures with a piezoelectric lithotripter. Cavitation and purely mechanical forces have been investigated as possible mechanisms for these biological effects of lithotripter fields.

11:05

5aBVa6. Biomechanical effects of ESWL shock waves. Bradford Sturtevant and Murtuza Lokhandwalla (Grad. Aeronautical Labs., Calif. Inst. of Tech., Pasadena, CA 91125)

Impulsive stress in the repeated shock waves administered in ESWL is the mechanical stimulus of injury to the kidney. In order to better understand the mechanical origins of injury, the interaction of focused shock waves with simple, planar polymeric membranes immersed in tissue-mimicking fluids was studied. The ESWL shocks in uniform, noncavitating liquids do not cause damage, but after passing through tissue or simulated tissue they do. This result suggests that the acoustic inhomogeneity of tissue may contribute to injury from ESWL. Shocks with large amplitude and short rise time (e.g., in uniform media) cause no damage in noncavitating fluids, while long-rise-time, dispersed shock waves, though only moderately attenuated, do. A continuum model is

described which incorporates the mechanical properties of the tissue and accounts for the effect on its strength of microscopic inhomogeneities. It is shown that when transient tensile stress is applied it takes a finite time for failure to occur, so if the pulse is not long enough the material survives, in qualitative agreement with the behavior described above. A definition of dose at failure is derived in terms of the stress applied to the membrane by each shock wave, the strain rate, and the material failure stress.

11:25

5aBVa7. Effects of tissue constraining on shock wave-induced bubble oscillation *in vivo*. Pei Zhong (Dept. of Mech. Eng. and Mat. Sci., Duke Univ., Box 90300, Durham, NC 27708)

Using a focused hydrophone (1 MHz), acoustic emission (AE) associated with the rapid oscillaton of cavitation bubbles induced by lithotripter shock waves was measured in kidney pelvis and renal parenchyma of a swine model, and compared with the AE produced in water under the same lithotripsy conditions. At each lithotripter output setting (between 16 and 24 kV), the duration of the primary bubble expansion/collapse was found to decrease as the AE was measured from water to kidney pelvis and renal parenchyma, respectively. As the number of shock waves delivered increased, the ratio of the primary bubble expansion/collapse to the subsequent ringing time was found to be almost unchanged in water, but varied significantly in both pelvis and renal parenchyma. These findings indicate a tissue-constraining effect on lithotripter shock wave-induced bubble oscillation *in vivo*. [Work supported by NIH.]

Contributed Papers

11:45

5aBVa8. SWL cavitation damage *in vitro*: Pressurization unmasks a differential response of foil targets and isolated cells. James A. McAteer, Mark A. Stonehill, Karin Colmenares, James C. Williams, Jr., Andrew P. Evan (Dept. of Anatomy, Indiana Univ. School of Medicine, 635 Barnhill Dr., Indianapolis, IN 46202-5120, mcateer@anatomy.iupui.edu), Robin O. Cleveland (Boston Univ., Boston, MA 02115), Michael R. Bailey, and Lawrence A. Crum (Univ. of Washington, Seattle, WA 98105)

To better understand the role of cavitation in kidney stone comminution and renal injury pressurization (Delius UMB23:1997) to regulate cavitation bubble activity *in vitro* was used. A low acoustic impedance chamber (8×15 cm) that only minimally altered lithotripter shock-wave pressure and waveform was constructed to house targets (aluminum foils or vials of kidney epithelial cells) under pressure (0–1400 psi). Cells exhibited lytic injury and foils sustained pitting at atmospheric pressure, but high pressure (1400 psi) prevented lysis and pitting. The pressure threshold for reduced damage was dramatically different for cells versus foils. Cell lysis was prevented at very low pressure ($>$ 30 psi), while substantially greater pressure ($\sim$600 psi) was needed to prevent pitting. Modest pressure (25–300 psi) actually enhanced pitting. SWL *in vitro* cell lysis and foil pitting are both likely due to cavitation. However, interactions between cavitation bubbles and these two targets appear to be quite different. Our observations are consistent with those of Delius and support his suggestion that it may be possible to reduce cavitation-mediated cell damage by regulating ambient pressure. The finding that slight pressure increased the number of cavitation events on foils suggests that modest overpressure might enhance stone comminution. [Supported by NIH PO1-DK43881.]

12:00

5aBVa9. Effect of overpressure on dissolution and cavitation of bubbles stabilized on a metal surface. Robin O. Cleveland (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215), Michael R. Bailey, Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105), Mark A. Stonehill, James C. Williams, Jr., and James A. McAteer (Indiana Univ. School of Medicine, Indianapolis, IN 46202-5120)

Recent experimental evidence indicates that static overpressure dramatically reduces lithotripsy-induced cell damage or stone comminution [Delius, Ultrasound Med. Biol. **24**, 611–617]. The hypothesis is that the damage and comminution are due to cavitation and that overpressure dissolves the small gas bubbles which act as cavitation nuclei. Delius observed, however, that the overpressure to protect cells (1 atm) was significantly lower than for stones (10 atm). In similar experiments cell lysis and pitting of aluminum foil as indicators of cavitation are used; the thresholds were 1 and 30 atm, respectively. In addition, the authors observed that pitting of foils increased for overpressures up to 20 atm. It is proposed that crevices in the foil stabilize cavitation nuclei whereas the bubbles dissolve in fluids. Foil damage increased because overpressure is small relative to the shock wave pressures that drive bubble expansion but is the dominant driving force at bubble collapse. Calculations using the Gilmore equation predict that overpressures up to 30 atm hasten and intensify cavitation collapses and over 60 atm eliminate cavitation, in qualitative agreement with our foil data. The calculated collapse time was confirmed experimentally by passive cavitation detection. If bubbles are stabilized in crevices in kidney stones, overpressure may provide a safer, more effective lithotripsy treatment. [Work supported by NIH-PO1-DK43881.]

Session 5aBVb

Bioresponse to Vibration/Biomedical Ultrasound: Medical Ultrasound II—Propagation, Media Characterization and Miscellaneous Topics

Timothy G. Leighton, Chair
Institute for Sound and Vibration Research, University of Southampton, Highfield, Southampton SO17 1BJ, England

Contributed Papers

9:15

5aBVb1. Are blood clots Biot solids? Pierre D. Mourad and Steven G. Kargl (Appl. Phys. Lab., UW, 1013 NE 40th St., Seattle, WA 98105, pierre@apl.washington.edu)

Blood clots are saturated porous solids. Their unwanted presence in veins and arteries is the source of a variety of medical problems. Their destruction (thrombolysis) via ultrasound and various enzymes (ultrasound-enhanced thrombolysis, or UET) is still a poorly understood phenomenon. The current view is that cavitation is the mechanism behind UET. This presentation begins with a review of the structure of blood clots and natural thrombolysis. Following that is a review of the literature on UET pointing out that cavitation is at best part of the story, at least *in vitro*, and more suspect *in vivo*. Next is a motivation of the Biot theory (a theory for how sound propagates through and interacts with porous solids) and its application to blood clots. The results of the work thus far predict that blood clots are Biot solids, which has implications for understanding UET. [Work sponsored by DARPA.]

9:30

5aBVb2. A stratified model for ultrasonic propagation in cancellous bone. Elinor R. Hubbuck, Timothy G. Leighton, Paul R. White (Inst. of Sound and Vib. Res., Univ. of Southampton, Southampton SO17 3BJ, Great Britain), and Graham W. Petley (Southampton Univ. Hospitals NHS Trust, Southampton, Great Britain)

A new model for ultrasonic wave propagation in cancellous bone is presented here. The model treats cancellous bone as a stratified medium of periodically alternating solid–fluid layers, and considers propagation of two compressional modes, analogous to the fast and slow waves of Biot theory, in terms of slowness surfaces. The behavior of the two modes is examined *in vitro* using bovine bone samples over a range of incidence angles. Two modes are seen when the cancellous structure is parallel to the propagation direction, but only one mode propagates normal to the structure, and the velocity of the fast wave increases with angle of incidence away from the normal. This behavior can be explained in terms of an angular dependent inertial coupling effect, in keeping with the theory of propagation in stratified media of periodically alternating fluid–solid layers. Quantitative agreement with theory is good. Although the structure being analyzed is essentially an oversimplification of the structure of cancellous bone, the agreement with theory suggests the stratified model offers potential for future research.

9:45

5aBVb3. Effects of the structural anisotropy and the porosity on ultrasonic wave propagation in bovine cancellous bone. Takahiko Otani and Atsushi Hosokawa (Dept. of Elec. Eng., Doshisha Univ., 610-0321 Kyotanabe-shi, Japan)

Ultrasonic wave propagation in water-saturated bovine cancellous (spongy) bone has been experimentally studied *in vitro* by a pulse transmission technique. Fast and slow longitudinal waves have been clearly observed in the earlier investigation when the acoustic wave propagates to the direction of the trabeculae orientation [A. Hosokawa and T. Otani, J. Acoust. Soc. Am. **101**, 558–562 (1997)]. In the present study, propagation speeds of the fast and the slow waves were measured as a function of the propagation angle to the trabeculae orientation. Experimental results show that the propagation speed of the fast wave and the amplitude of the slow wave depend greatly on the angle to the trabeculae orientation, and the speed of the slow wave and the amplitude of the fast wave are little affected by the angle of the trabeculae orientation. The propagation speeds of both waves were also experimentally examined as a function of the porosity. Measured results are discussed in relation to the structural anisotropy and the porosity using Biot's theory.

10:00

5aBVb4. Wideband laser-acoustic spectroscopy of proteins. Alexander A. Karabutov and Natalia B. Podymova (Dept. of Phys., Moscow State Univ., Moscow, Russia 119899, karabut@gpwp1.phys.msu.su)

The absorption of ultrasound in a frequency range of 2–50 MHz in proteins was investigated with a wideband laser-ultrasonic spectrometer. It was found that for fresh chicken proteins ultrasonic absorption is proportional to the squared frequency. Denaturation of proteins transforms the spectrum of ultrasonic absorption to the first power of frequency dependence. In the hole frequency range the ultrasonic absorption in denaturated protein is higher than in the fresh one. The possibility of laser ultrasonic diagnostics of proteins is discussed.

10:15

5aBVb5. Wideband acoustics spectroscopy of liquid phantoms of biological tissues. Valery G. Andreev, Yury A. Pischalnikov (Dept. of Acoust., Phys. Faculty, Moscow State Univ., Moscow 119899, Russia, andreev@na.phys.msu.su), Alexander A. Karabutov, and Natal'ya B. Podymova (Moscow State Univ., Moscow 119899, Russia)

Biological tissues have a specific frequency dependence of the ultrasonic absorption and sound-speed dispersion. It makes fruitful the diagnostics of tissues by studying these parameters. It is necessary to carry out the spectroscopic investigation in a wide frequency band in a real time. Pulsed laser ultrasonic spectroscopy is the most useful technique to solve this problem. The method of wideband acoustic spectroscopy with laser thermo-optical source is proposed. A Q-switched YAG-Nd laser was explored to excite the short acoustic transients (pulse duration 12 ns, amplitude up to 8 MPa) by an especially produced thermo-optical generator. The acoustic transients transmitted through the medium under the testing were detected by a wideband transducer (frequency band 2–350 MHz). The absorption of ultrasound was calculated by FFT. The absorption in glycerol, acetic acid, and butanediol was investigated in a frequency band 2–50 MHz. The spectra of relaxation times were calculated. [Work supported by RFBR and CRDF.]

10:30

5aBVb6. The necessity for acoustics in the biomedical engineering program. Daniel R. Raichel (Dept. of Mech. Eng., Steinman Hall, The City College of The City Univ. of New York, New York, NY 10031) and Latif Jiji (The City College of CUNY, New York, NY 10031)

Biomedical engineering constitutes the fastest growing segment of all engineering fields. In order to effectively deal with a wide variety of engineering problems, the biomedical engineer needs to be a ''jack of all trades'' insofar that he or she should be able to meet challenging problems in mechanics, electronics, thermal analyses, fluid mechanics, materials, manufacturing techniques, etc. In the medical field, acoustics (particularly ultrasound) is used for both diagnostic and therapeutic purposes, and new devices are continuously being developed to provide better diagnostic imaging or more effective therapy. As a part of the biomedical engineering program, undergraduate and graduate students need to gain better understanding of acoustical physics, the effects of sound and ultrasound on living organisms, the principles of measurements, imaging processes, and data analysis concepts. Laboratory experience is essential, and practical experience in a teaching medical hospital would definitely be salutary. The time constraint imposed by a four-year undergraduate program, which also needs to meet ABET criteria, effectively limits the expository scope, and it follows that additional, more intensive training may be better achieved through post-baccalaureate (master's degree and doctoral levels) programs.

FRIDAY MORNING, 26 JUNE 1998 EAST BALLROOM A (S), 9:15 TO 11:30 A.M.

Session 5aEA

Engineering Acoustics: Sonar Transducers

James M. Powers, Chair
Naval Undersea Warfare Center, Code 2131, Newport, Rhode Island 02841-1708

Contributed Papers

9:15

5aEA1. Finite-element simulation of piezoelectric transformers. Takao Tsuchiya, Yukio Kagawa, and Hiroki Okamura (Dept. of Elec. and Electron. Eng., Okayama Univ., Okayama, 700 Japan)

A three-dimensional finite-element approach to the characteristic prediction of piezoelectric transformers is presented. For a numerical example, a Rosen-type piezoelectric transformer is considered. The electrical input admittance and the transmission or transfer characteristics are demonstrated for the fundamental longitudinal mode operation. Based on these characteristics, the lumped parameters are extracted to form the electrical equivalent circuit. The effect of the loading or termination condition on the output voltage and the conversion efficiency are also examined. The numerical results are then compared with the experimental results. Examination is then extended to the optimum design in terms of efficiency and the output voltage.

9:30

5aEA2. A tubular piezoelectric actuator and its characteristic analysis by finite-element modeling. Naoto Wakatsuki, Takao Tsuchiya, Yukio Kagawa, and Kazumichi Hatta (Dept. of Elec. and Electron. Eng., Okayama Univ., Okayama, 700 Japan)

Three-dimensional motion is sometimes required in an actuator. A tubular configuration is proposed to realize the three-dimensional motion at the tip. A cylindrical shell piezoceramic, polarized in the radial direction, is provided with several pairs of electrodes on both surfaces of the shell. The combination of the voltages with proper sense makes it possible to control the motion. The displacements or the forces at the tip are expressed in terms of the linear combination of the electrical input voltages on each pair of electrodes. The coefficient of each term is to be determined by the experiment or numerical simulation. Here, the determination of the coefficients is made by finite-element modeling for which the code previously developed is utilized.

9:45

5aEA3. Substructuring in the finite-element analysis of sonar transducer arrays. J. R. Dunn (School of Electron. & Elec. Eng., Univ. of Birmingham, Edgbaston, Birmingham B15 2TT, England, j.r.dunn@bham.ac.uk), C-L. Chen (6F-3, Taipei, Taiwan), and B. V. Smith (Univ. of Birmingham, Edgbaston, Birmingham B15 2TT, England)

Finite-element (FE) analysis of arrays of sonar transducers becomes costly in memory and computation time for arrays with many elements. This becomes even more severe a problem if the individual transducers have complex structures. However, because in many cases the individual transducer elements are identical, it is possible to treat them as repeated substructures; each substructure may then be viewed as a superelement which is defined by only those nodes which couple into the overall structure of the array. This technique is analogous to the use of Thevenin equivalent sources in the analysis of complex electrical circuits. By the use of this technique it is possible to analyze the electroacoustic performance of larger arrays than hitherto possible within the limitations set by the computer memory. This paper describes this superelement technique and results are presented for a simple two-element transducer array analyzed for its free (unloaded) vibration behavior both by full FE analysis of the complete structure and by the substructuring technique. The results of the two methods are shown to agree very closely, and measurements on an experimental model compare favorably with the theoretical predictions.

10:00

5aEA4. Sonar simulation and display techniques. Yuhong Guo, Hong Liu, and Junying Hui (Dept. of Underwater Acoust. Eng., Harbin Eng. Univ., 150001 Harbin, PROC, guoyh@public.hr.hl.cn)

This paper describes techniques used in the simulation and display of a passive sonar system. The sonar and its environment models, including target, channel, noise background, and array, and a complete sonar signal processing chain are developed in the simulation system. Based upon the simulation system, different display techniques are studied. Modeling for performance assessment of a colored brightness mode sonar display is difficult because of its relation to human visual psychology and physiology. A concept of ''visual threshold'' is proposed in the paper; based on it, as well as the statistical model of display data, detection performance of

the sonar display system can be analyzed quantitatively and theoretically. The second-decision method, based on the statistical property of LOFAR display, and a four-dimensional display technique are presented in the paper.

10:15–10:30 Break

10:30

5aEA5. A new method of multibeamforming for a swath bathymetry sonar system and its MIMD parallel architecture. Haisen Li, Lan Yao, and Xinsheng Xu (Dept. of Underwater Acoust. Eng., Harbin Eng. Univ., Harbin 150001, PROC, euipc@public.hr.hl.cn)

In a swath bathymetry sonar system, real-time multibeamforming is required in order to increase the coverage and the efficiency. To determine the direction of seafloor echo accurately, traditional beamforming must work at a high sample rate, the interpolation is used to reduce computing amounts, but when there are tens of receiving channels and beams, the interpolation filter algorithm will rapidly increase computing amounts to result in the real-time multibeamforming, this is very difficult. Combining with a swath bathymetry sonar system, a new method of quadrature beamforming is described in this paper. It combines time domain beamforming with phase shifting beamforming to process the available bandpass signal. To reduce sample rate, a mixing frequency technique is used. A quadrature component is obtained from a single-channel A/D converter utilizing Hilbert transformation and its coefficient is weighted with Hamming window function, therefore, amplitude and phase is highly accurate. Comparing with the traditional beamforming method, the method presented in the paper is characterized by high accuracy, less sample rate, and less computation complexity, easily realized with a high speed DSP device. The parallel pipeline architecture is given based on TMS320C30. Simulating and lake trial results show that the method is right and efficient. The method is universal for active sonar.

10:45

5aEA6. Specific acoustic impedance of the ultrasonic field by the square flat transducers. Tohru Imamura (Natl. Res. Lab. Metrology, 1-1-4, Umezono, Tsukuba, Ibaraki, 305 Japan, imam@nrlm.go.jp)

Sound pressure, particle velocity, and specific acoustic impedance are investigated, using the case where square flat transducers are tilted from a parallel condition. Specific acoustic impedance is a quotient of sound pressure by particle velocity. Particle velocity is computed by numerical quadruple integration with Huygens's principle, and compared with that of the plane wave. The ratio of the half side length (a) of the square flat transducer to the wavelength (λ) of the ultrasonic wave is set at 2.5, and the distance between the transducers is set at 5a or 10a. The feature of the phase delay leap of the specific acoustic impedance on the central axis of the circular flat transducer does not appear for the square case.

11:00

5aEA7. Estimation of equivalent circuit parameters for piezoelectric transducer using least-squares method. Byung-Doo Jun (R & D 3 Gr., LG Precision Co. Ltd., 148-1 Mabuk-ri, Gusung-Myun, Yongin-shi Kyunggi-do, 449-910, Korea) and Koeng-Mo Sung (Seoul Natl. Univ., Seoul 151-742, Korea)

In general, the resonance method has been used for estimating the equivalent circuit parameters for the underwater piezoelectric transducer. Sometimes, however, it is difficult to estimate the equivalent circuit parameters when the measured electrical impedances have an imaginary part at resonant and antiresonant frequency. In this paper, to estimate the equivalent circuit parameters of the piezoelectric transducers, equivalent circuit equation is modified and the least-squares method is applied. Equivalent circuit equation is composed of real and imaginary equations, respectively. With these equations, matrix form Yule–Walker equations are formulated. The least-squares method is applied with measured electrical impedance data over the effective frequency band to solve these equations. The procedure is as follows: First, four variables which are functions of five equivalent circuit parameters (R, L, C) are obtained by the least-squares method. Then, initial equivalent circuit parameters can be extracted by solving four equations which contain five parameters, one of which is assumed. Finally, by iteration and update, all the equivalent circuit parameters can be estimated with considerably small error. The impedance characteristics of the piezoelectric transducer with estimated parameters are in good agreement with measured ones.

11:15

5aEA8. Study of compound structured ultrasonic transducer made of PZT/PVDF. Moo Joon Kim, Dong Hyun Kim, and Kang Lyeol Ha (Dept. of Phys., Pukyong Natl. Univ., 599-1 Daeyon-dong, Nam-gu, Pusan 608-737, Korea)

When the piezoelectric ceramic, PZT is used in the water it needs matching layers for efficient energy transmission. PZT is known as the proper piezoelectric material for the transmitter by reason of the large electromechanical coupling factor. So actually PZT has been used for the nondestructive testing or ultrasonic diagnostic instrument as good material. In the meantime, there is a kind of polymer by the name of PVDF. Its acoustic impedence is similar to water, so PVDF does not need matching layers but it is not a proper material for the transmitter because of the small electromechanical coupling factor. However, the receiving sensitivity of PVDF is very flat through the wide frequency range. So it is known as a proper material for the receiver like wideband hydrophone. Therefore in this paper it is suggested that the transducer which transmits and receives signals by PZT and PVDF, respectively, could be constructed compositively, and the transmitting and receiving feature of the transducer was investigated theoretically and experimentally.

Session 5aMU

Musical Acoustics: Pastiche

Vladimir Chaloupka, Chair
Physics Department, University of Washington, Box 351560, Seattle, Washington 98125-1560

Contributed Papers

9:00

5aMU1. Physics of the Music Laboratory at the University of Washington. Vladimir Chaloupka and Keith Hughes (Phys. Dept. 351560, Univ. of Washington, Seattle, WA 98125-1560)

The Physics Department at the University of Washington recently completed the acquisition of new equipment for teaching the Physics of Music course, as well as for performing various independent study projects in musical acoustics at both undergraduate and graduate levels. The lab setup will be described, as well as the experience after the first year of operation. Many aspects of the Physics of Music provide excellent opportunities for involving undergraduate students in small, but meaningful and interesting, research projects. Specific examples will be discussed.

9:15

5aMU2. Dynamic optimal tuning of electronic keyboards as they are being played. James E. Steck (Dept. of Mech. Eng., Wichita State Univ., Wichita, KS 67260-0133, steck@me.twsu.edu) and Dean K. Roush (Wichita State Univ., Wichita, KS 67260)

The 12-tone equally tempered musical scale can produce noticeably imperfect matching of acoustic harmonics when some combinations of tones are played together. The lower partials (harmonics) which should match, do not. For a 12-tone octave, if a small number of tones in the octave are to be sounded simultaneously, a tuning can be found which produces harmonic errors that are much less than that of equal temperament. Historically, these tunings have been discovered, explored, and abandoned in favor of equal temperament because they generally have much worse harmonic error for tone combinations other than those for which the tuning was optimized. Here, a mathematical description of the total harmonic error of a musical instrument is presented along with a mathematical method which solves for a tuning (tone frequencies) that minimize this total error. The optimal tuning calculations have been implemented in ''C'' code on a 166-MHZ pentium desktop personal computer equipped with a Creative Labs Sound Blaster AWE 32 soundcard, with a MIDI keyboard attached. The ''C'' code reads the keys (tones) being played on the keyboard, calculates the optimal tuning, and sounds the tones on the PC soundcard at the optimally tuned fundamental frequencies.

9:30

5aMU3. On the use of elements of piano notes to improve the identification of polyphonic piano sounds. Lucile Rossi and Gerard Girolami (Univ. of Corsica, Quartier Grosseti, 20250 Corte, France, lrossi@univ-corse.fr)

An original algorithm for identifying polyphonic piano sounds signals, based only on the use of frequential positions of note partials, has been developed by the authors [Rossi and Girolami, J. Acoust. Soc. Am. **100**, 2842(A) (1996)]. This paper presents a study of the evolution of the amplitude of piano notes' partials and their energy, carried out to see if these elements can be used to improve the identification of notes in cases like short notes, long notes, repeated notes, and polyphonic sounds. In the case of monophonic sounds, Doval [Ph.D. thesis, University of Paris VI (1994)] has shown that the evolution of the amplitude of piano notes' partials along with their frequency variations can be used to identify the fundamental frequency of signals. The case of polyphonic piano sounds (several notes played simultaneously on the same piano) is more difficult, but if a particular behavior of partials belonging to a given note could be detected, for example, synchronous variations or typical amplitude evolution, the sorting of partials would be more reliable, the identification of notes more efficient, and the detection of notes' onsets more precise.

9:45

5aMU4. A new accurate model-based music synthesis technique by using recurrent neural networks. S. F. Liang (Dept. of Control Eng., Natl. Chiao-Tung Univ., Hsin-Chu, Taiwan) and Alvin W. Y. Su (Dept. of CSIE, Chung-Hwa Univ., Hsin-Chu, Taiwan)

Reproduction of the tones generated by playing a particular musical instrument electronically remains an open and difficult problem. An accurate model-based analysis/synthesis approach with respect to plucked-string instruments by using recurrent neural networks is proposed. The procedure is described as follows. A recurrent neural network corresponding to the physical model of a musical string is first configured. The vibration of a plucked string is measured and used as the training data of the neural network. The backpropagation-through-time method is used to train the network such that output of this network can be matched to the measured data as close as possible. The well-trained neural network is then used as our synthesis model for this particular string. Since the characteristics of a string keep changing over the period of vibration, it is also necessary to update the parameters of the neural network during the synthesis process. In our experiments the synthesized tones and the real tones sounded almost identical. These results will also be demonstrated during the presentation in ICA/ASA 98 if this paper is accepted. [This Work is supported by both NSF, Taiwan and Computer/Communication Lab. ITRI, Taiwan.]

10:00

5aMU5. Musical instrument synthesis by nonregenerative nonlinear processing. James W. Beauchamp (School of Music and Dept. of Elec. and Comput. Eng., Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801)

When a sinusoid is distorted by a nonlinear function, a complex tone is produced. An nth degree polynomial nonlinear function produces an output waveform containing n harmonics. Moreover, when the amplitude of the input sine wave changes, the harmonic structure of the output changes in a way which can be predicted from the polynomial coefficients. Also, the coefficients can be set to provide any arbitrary magnitude spectrum when the sine wave has a particular amplitude, say unity. Generally, the spectral centroid, a measure of the spectral bandwidth, increases monotonically as the input amplitude increases. This possibility is enhanced by placing a high-pass filter at the output. An acoustic instrument sound can be imitated by (1) performing time-variant harmonic analysis on it, (2) selecting a representative spectrum from it to determine the polynomial coefficients, (3) measuring the sound's time-varying spectral centroid and rms amplitude, and (4) coercing the sine wave amplitude to change so that the spectral centroid of the output matches that of the original sound. The amplitude of the original sound is matched by a multiplier at the nonlinear output. This method has been successfully used to synthesize instruments such as the trumpet, saxophone, clarinet, oboe, and piano.

10:15

5aMU6. Time-domain modeling and numerical simulation of timpani. Leïla Rhaouti, Patrick Joly (Inria-Rocquencourt, BP 105, 78153 Le Chesnay Cedex, France), and Antoine Chaigne (Enst, Paris Cedex 13, France)

Timpani are made of a circular elastic membrane stretched over a enclosed air cavity and set into vibrations by the impact of a mallet. The motion of the membrane is coupled with both the external and internal sound pressure field. A time-domain model of this instrument has been developed in order to investigate the influence of the main geometrical and physical quantities on the resulting sound. This model consists of a set of partial differential equations which govern the displacement of the membrane and the acoustic pressure inside and outside the cavity, respectively. These equations are coupled with a nonlinear differential equation which governs the excitation by the mallet. A numerical scheme has been derived from this model using three-dimensional finite element methods. Absorbing conditions have been implemented to simulate the free space. The validity of the model is illustrated by successive snapshots showing both the pressure field and membrane displacement. In addition, time histories of energetic quantities are presented for a better understanding of energy balance between membrane, cavity, and external space in real instruments. Sound examples are obtained by simulating the sound pressure at specific positions corresponding to the player's ears.

10:30

5aMU7. Modeling Chinese musical instruments. Andrew B. Horner and Lydia Ayers (Dept. of Comput. Sci., Hong Kong Univ. of Sci. and Technol., Clear Water Bay, Kowloon, Hong Kong, horner@cs.ust.hk)

Chinese instruments include a range of colorful instruments with distinctive characters. For example, the dizi is the most common Chinese flute, and it has a bright buzzing quality produced by a rice paper membrane glued over a hole. This paper describes models for more than 20 Chinese traditional and folk instruments using group additive synthesis with genetic algorithm-optimized parameters. Other types of Chinese flutes modeled include the xiao (vertical flute), paixiao (panpipes), and xun (ocarina). The sheng is a mouth organ with a ring of bamboo pipes attached to a wind chamber, and dates back to 1500 BC. A folk version of the sheng, called the lusheng, has also been modeled, as well as the bawu and hulusi, folk instruments that sound similar to the clarinet. The player's mouth completely covers the bawu's blowhole, which has a vibrating reed cut into a copper strip covering the hole. The hulusi has a playing tube and multiple drone tubes. Several pitched percussion and plucked string instruments have also been modeled. The presentation will show the acoustic instruments, their sound, their spectra, and resynthesized excerpts of music written for these instruments.

FRIDAY MORNING, 26 JUNE 1998 CASCADE BALLROOM I, SECTION A (W), 9:00 TO 10:15 A.M.

Session 5aNSa

Noise: Aircraft Noise: General

Colin G. Gordon, Cochair
Colin G. Gordon and Associates, 411 Borel Avenue, Suite 425, San Mateo, California 94402

Mei Q. Wu, Cochair
Colin G. Gordon and Associates, 411 Borel Avenue, Suite 425, San Mateo, California 94402

Contributed Papers

9:00

5aNSa1. Low-frequency noise around airports. Michel Vallet and Jean Claude Bruyere (INRETS, 25 avenue Francois Mitterrand, Case 24 69675, Bron Cedex, France)

The sound frequency weighting PNdB has been proposed by Kryter (1959) to assess the noise from jet engine aircrafts. Due to the technical evolution of aircraft engines, and to choose a frequency weighting common to all transport systems, the dB(A-weighted) has begun to be used as the acoustical unit around airports; the PNdB is still used for noise certification purposes. A study of noise measurements that was carried out around French airports allows one to see some noticeable differences between acoustical levels, respectively, expressed in dB(A) and dB(C). The difference varies between 4 dB for the landing of recent aircrafts to 14 dB at take-off for 50/70-seat propeller aircrafts. If one considers the additional frequency decrease due to the walls and windows in housing, the pertinence of dB(A) weighting could be questioned for aircraft noise received by people living a long distance from the airport. This question is of interest in light of the new ANSI method for assessing combined noises (Schomer, 1997): this method suggests using the C frequency weighting for sounds with strong low-frequency content.

9:15

5aNSa2. Subjective response to aircraft flyover noise. Sherilyn A. Brown (Structural Acoust. Branch, NASA Langley Res. Ctr., M.S. 463, Hampton, VA 23681, s.a.brown@larc.nasa.gov) and R. David Hilliard (Wyle Labs., Hampton, VA 23666)

In order to obtain increased understanding of community response to aircraft noise, a comprehensive subjective study was designed and implemented to realistically simulate the aircraft noise exposure experienced by residents living both near and far from airports. The test was designed to present a wide range of daily aircraft noise exposures ranging broadly in both quantity and amplitude. The In-Home study was supported by a computer controlled sound playback system which exposed test subjects to aircraft sounds during a daily 14 h test period. This system provided a degree of control over the noise exposure not found in community situations and a degree of situational realism not available in the laboratory. Each day the system played from 0 to 448 flyovers as the test subject went about his or her normal activities. At the end of the day, the test subjects rated their annoyance to the flyovers they had heard. The noise exposures presented during 9 weeks simulated different combinations of runway activity, aircraft flyover altitude, rush hour conditions, and occasional loud aircraft among many low sound level flights. Subsequent data analysis will

seek to provide a basis of understanding of annoyance response to enhance predictive algorithms for a wide range of aircraft noise exposure conditions.

9:30

5aNSa3. Low-flying military aircraft noise—Operational flying. Ian H. Flindell (Inst. of Sound and Vib. Res., Univ. of Southampton, UK) and Ralph J. Weston (Royal Air Force Inst. of Health, Aylesbury, Buckinghamshire, UK)

The distribution of military low-flying training across Great Britain is essentially random within the permitted low-flying areas, although there are certain areas where above-average concentrations of flights occur. However, there are very little data available on actual flight tracks as flown, and training operations are normally conducted so as to encourage as much flexibility as possible. This means that calculations of noise exposure in different areas have to be based to a considerable extent on assumptions about typical flight track distributions, using aircraft noise source reference levels extrapolated from data collected under specially arranged measurement trials. This paper reports the results of a number of recent field surveys carried out in representative overflown areas which show generally good agreement between the assumptions made about typical flight distributions, the special trials data, and actual operational flying.

9:45

5aNSa4. Converting noise level into number of noisy events: A new method to quantify *k* factors reveals area-dependent over- and underenergetic responses around the same airport. Karl Th. Kalveram (Univ. of Duesseldorf, Dept. of Cybernetical Psych., 40225 Duesseldorf, Germany, kalveram@uni-duesseldorf.de)

For aircraft noise, caused by N flight movements within observation time T, with average maximum noise level $L_{\max}$ and duration d_t, the energy equivalent noise level L_{eq3} is approximated by $Q(k,d)=L_{\max}+k^*\log N+d^*\log(dt/T)$, if $k=d=10$. This seems to enable a trade-off between $L_{\max}$, N, and dt. However, the problem is not to keep the L_{eq3} constant, but the subjective annoyance, and this in turn depends on whether people process alterations of $L_{\max}$, N, and dt corresponding to energy equivalence ($k=d=10$) or not. Therefore, valid measurements of public noise responses by $Q(k,d)$ require the estimation of k and d from empirical data. However, multiple correlation techniques mostly used for this purpose lead to contradictory and invalid estimates for k. This paper presents an alternative method, utilizing dose-response curves based on regression of annoyance with $L_{\max}$ at fixed numbers of overflights. Results are independent from the selected annoyance ratings and furthermore also valid for nonlinear dose–response curves. Based on curves characterizing small ($N<5000$), medium ($50000<N<200000$), and large ($200000<N$) airports [Rice (1980)], the proposed procedure predicts a crossover effect also for different areas of the same airport, i.e., $k>10$ for low k and $k<10$ for high noise levels, if this airport changes from medium to large.

10:00

5aNSa5. Relational anomalies in A weighted and linear scale sound level measurements associated with jet aircraft departures. Errol Nelson P.E. QEP (Optimum Environment, P.O. Box 114, Issaquah, WA 98027) and Allan Furney (Regional Commission on Airport Affairs, Des Moines, WA 98198)

Noise levels from jet aircraft operations at major airports are often perceived by nearby residents as being louder and more intrusive than claimed by airport proprietors. This can cause heated opposition by citizens to airport development and expansion plans. To further clarify the issue, a noise study was conducted using four noise meters; two Type 1 and two Type 2. Simultaneous sound level measurements were taken on the A-weighted and linear scales in airport and nonairport noise environments; and included octave band analyses. In both noise environments the average differential between full spectrum A-weighted and linear scale measurements is approximately 15 dB, except when influenced by jet aircraft departures. During jet aircraft departures, this A/linear differential was as small as 4 dB; and the smaller the difference between simultaneous measurements the louder the perceived noise from the departing plane. This measurement anomaly was unique to departing jet aircraft. Propeller driven aircraft operations, or jet aircraft landings, showed little or no change in the A/linear noise differential. The measurements suggest that a significant amount of noise is not being measured by the sound level meters during jet aircraft departures: either deducted by weighting filters or due to instrument limitations. However, by conducting simultaneous measurements of A weighted and linear scale sound levels, this relational anomaly can provide a simple, realistic indicator of jet aircraft loudness in the airport noise environment.

FRIDAY MORNING, 26 JUNE 1998 CASCADE BALLROOM I, SECTION C (W), 8:30 TO 9:30 A.M.

Session 5aNSb

Noise: Source Noise Analysis and Control

Daniel A. Quinlan, Chair
Lucent Technologies Acoustics Research, 600 Mountain Avenue, Murray Hill, New Jersey 07974

Contributed Papers

8:30

5aNSb1. The use of frequency-dependent velocity scaling as a diagnostic tool. Daniel Quinlan (Bell Labs., Lucent Technologies, 600 Mountain Ave., Murray Hill, NJ 07974, dbq@lucent.com)

This paper will discuss a method for investigating broadband aeroacoustic noise sources. The method is based upon the use of frequency-dependent velocity scaling. The dependence of sound power level upon a characteristic velocity is determined experimentally, and that dependence is used as an indicator of primary aeroacoustic processes. The procedure yields an identification of distinct frequency bands within which source trends are observable. Each band is presumed to be controlled by different processes, and the set of likely processes is fixed according to the average velocity exponent value obtained. In principal, the method is applicable to low-speed aeroacoustic source identification problems where a characteristic flow speed can be measured and systematically varied. Measurements of sound radiated by flow over a flat plate were used to evaluate the performance of the method. The procedure was then applied to a small axial fan typical of those used to cool electronic systems. For the fan, the frequency-dependent velocity scaling demonstrated that shifts in dominant aeroacoustic processes occur near 2 kHz and above 4 kHz. The results added insight to a related study which demonstrated that tip gap flows were the primary source of broadband noise above 2 kHz.

8:45

5aNSb2. Noise diagnostic and reduction of a scooter engine motorcycle. Cho H. Lu (Mech. Industry Res. Labs., Industrial Technol. Res. Inst., Hsinchu, Taiwan 310, R.O.C., 830562@mirl.itri.org.tw) and Jen C. Cheng (Natl. Huwel Inst. of Technol., Huwel, Taiwan, R.O.C.)

This paper presents an appropriate way to reduce the noise of a scooter engine motorcycle. In order to investigate the noise contribution of each engine component, the lead covering method was used to rank the noise source. With the noise spectra from measuring the pass-by noise and releasing the lead cover on each component sequentially, the correlations between specific frequencies and each component of the engine were found. Furthermore, the noise characteristics of all components were cataloged into the airborne noise and the structure-borne noise. In this research, it was shown that the airborne noise was usually coming from the air cleaner or muffler. Additionally, the power transmission mechanism like CVT or crankcase generated the structure-borne noise. Based on the diagnostic results above, some modifications of the motorcycle had been done, for example, isolating of the CVT-cover and the crankcase-cover, fixing of high density material at the air cleaner box and adding another layer around the exhaust pipe. For the pass-by noise measurement, it was shown that the noise level of the scooter was reduced by 2.5 dB(A). [Submitted For Noise ''Outstanding Young Presenter'' Paper Award.]

9:00

5aNSb3. Noise in non-premixed turbulent syngas flames. Sikke A. Klein and Jim B. W. Kok (Dept. of Mech. Eng., Univ. of Twente, P.O. Box 217, NL 7500 AE Enschede, The Netherlands, s.a.klein@wb.utwente.nl)

A turbulent syngas flame may generate acoustic noise of high acoustic intensity in a combustion chamber. This may lead to the failure of construction components in a gas turbine engine in periods of the order of 1–100 hours. The research as described in the literature has almost exclusively been performed on the generation of noise in premixed methane or propane flames. Syngas fuel is a mixture of hydrogen and carbon monoxide, and the burners used are of the non-premixed type. In this research, the effect of turbulence and syngas composition on noise generation is investigated. A laboratory is set up to test syngas flames of a thermal power of 50 kW in a cylindrical air-cooled combustion chamber. Experiments are performed at several fuel compositions and burner inlet conditions. The flame sound intensity is measured in the combustion chamber equipped with acoustic dampers. The paper discusses the measured sound spectra. A model is derived for the generation of sound in a turbulent non-premixed flame. In this model it is shown that the sound generation is related to the dependence of density on mixture fraction in a flame with fast chemistry. A fluctuation in mixture fraction will lead to sound generation.

9:15

5aNSb4. Prediction and reduction of centrifugal blower noises. Christopher L. Banks and Sean F. Wu (Dept. of Mech. Eng., Wayne State Univ., 5050 Anthony Wayne Dr., Detroit, MA 48202)

This paper presents the results of an investigation of centrifugal blower noises. Experiments on three different types of blowers were conducted to determine the major noise generation mechanisms. The information thus gained has led to an innovative noise reduction method utilizing a stationary, noncontacting transition mesh. Experiments showed that an overall noise reduction of 3 to 5 dB was achieved for all the blowers running at impeller speeds in the range of 10 to 30 m/s. In addition, a semianalytical model was developed to predict noise spectra from dimensionally similar centrifugal blowers. This model yielded an acoustic power proportional to the fourth power of the blower speeds. Comparisons of the calculated and measured noise spectra from different blowers running under various speeds were demonstrated. Good agreements were obtained in all cases. [Work supported by Ford Motor Company.]

FRIDAY MORNING, 26 JUNE 1998 CASCADE BALLROOM I, SECTION C (W), 9:45 TO 10:45 A.M.

Session 5aNSc

Noise and Architectural Acoustics: Soundscapes—''Acoustical Landscapes'' in Natural and Built Environments

Joseph Pope, Chair
Pope Engineering Company, P.O. Box 236, Newton Center, Massachusetts 02159

Contributed Papers

9:45

5aNSc1. Supportive sound environments in building research. Birgitta Berglund (Inst. of Environ. Medicine and Dept. of Psych., Stockholm Univ., Stockholm, Sweden), Edward Hojan, and Anna Furmann (Inst. of Acoust., Adam Mickiewicz Univ., Poznan, Poland)

The aim of this research is to create good sound environments in rooms of ordinary buildings (schools, dwellings, offices, hospitals) for people with hearing deficits. If the deterioration of hearing with age is considered, this group of people may represent a majority rather than a minority of the population. From the many evaluation criteria of natural sounds, both stationary and nonstationary and in room environments, used by subjects, the authors have identified independent ones. These criteria may be named attributes and primarily define the ''quality'' of the sound-perception space in rooms. New experiments involving perceptual evaluation of sound and environments by 40 randomly chosen subjects and the use of multidimensional data analysis confirm in part our earlier results regarding the particular evaluation criteria applied [Furmann *et al.,* J. Acoust. **4**, 535–551 (1991)]. According to Stevens' argument, it seems that the relationship between the values of the evaluation scale and the perceptions are identical for people of normal and impaired hearing although their perceptions may differ.

10:00

5aNSc2. Features of Japanese soundscapes recognized by foreigners: A questionnaire survey on sounds for the foreigners living in Fukuoka. Shin-ichiro Iwamiya (Dept. of Acoust. Design, Kyushu Inst. of Design, 4-9-1, Shiobaru, Minami-ku, Fukuoka, 815-8540 Japan, iwamiya@kyushu-id.ac.jp)

A questionnaire survey on odd sounds for foreigners living in Fukuoka was done to clarify the feature of Japanese soundscapes recognized by those whose cultural background is different from the Japanese. The foreigners were asked to point out the sound which they usually heard in Japan but heard less in their home countries and their impressions, and the

sound which they usually heard in their home countries but heard less in Japan and their impressions. The traffic sound signal for the blind, the exhaust noise by wild bike riders, the peddlers' cries, and the election campaign broadcast are the typical sounds only heard in Japan. The foreigners usually have a good image for the traffic sound signal for the blind and the peddlers' cries. In contrast, they have a negative impression on the muffler noise by wild riders and speeches concerning election. The Japanese soundscape is characterized by these sounds which we hear in daily life. Comparison between the sounds usually heard in Japan and those heard less in Japan shows that there are too many public sounds, such as announcements using public-address systems and music at the shopping centers, and few natural sounds and human voices in Japan.

10:15

5aNSc3. The acoustic and sound design for a museum exhibition on indigenous music and dance. Peter G. Coulter (Museum of Appl. Arts and Sci., P.O. Box K346, Haymarket, Sydney, Australia 2000)

Museum exhibitions commonly use visual design to orientate and inform the visitor. At the Powerhouse museum in Sydney, Australia, an exhibition (entitiled Ngaramang Bayumi) was recently developed on indigenous music and dance. This exhibition was designed to orient the visitor using aural cues, in addition to the visual. This has been achieved by the use of acoustic design and through the use of a specialized sound systems that presents a 3-D sound (ambisonics), that calls and guides the visitor through the space. The experience, along with audiovisuals, objects, text, and graphics uses unusual geometries, layouts, and building materials. Using this design as a case study, this presentation will explain the research and development of the acoustic elements within the exhibition. These are the elements of cross talk, sound quality, sound focusing, sound isolation, sound ambience, localization, dynamic range, and reverberation times. Visitor reaction to the exhibition will also be discused.

10:30

5aNSc4. Survey on the actual condition of the sound environment in commercial spaces. Michiko So and Sho Kimura (Dept. of Architecture, College of Sci. Technol., Nihon Univ., 1-8 Kanda-Surugadai, Chiyoda-ku, Tokyo, 101 Japan, czf01070@niftyserve.or.jp)

Many people use a great variety of commercial spaces in their daily life, and are exposed to such a peculiar sound environment both consciously and unconsciously. In order to gain a comprehension of the user's awareness of the sound environment in commercial spaces, a survey was conducted on residents of the Tokyo Metropolitan Area (1198 effective responses were received). The result shows a great majority of those questioned regard the sound environment as one of the features that enhance the atmosphere or the character of commercial spaces. Accordingly, the design of the sound environment plays a role as one of the important factors in creating the atmosphere of individual commercial spaces. On the other hand, it was also revealed that under current conditions, the loudness, articulation, quality, the manner of presentation of background music and announcements, etc. (especially how to use PA systems in noisy surroundings), in individual commercial spaces were not adequately adjusted or designed in many cases. This is associated with the limits of preference and the potential question of the awareness of the users.

FRIDAY MORNING, 26 JUNE 1998 GRAND BALLROOM A (S), 9:15 A.M. TO 12:15 P.M.

Session 5aPAa

Physical Acoustics: Sonochemistry and Sonoluminescence: SL I

Thomas J. Matula, Cochair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

R. Glynn Holt, Cochair
Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Chair's Introduction—9:15

Invited Papers

9:20

5aPAa1. Nonlinear bubble dynamics and light emission in single-bubble sonoluminescence. Felipe Gaitan (Natl. Ctr. for Physical Acoust., University, MS 38677) and Glynn Holt (Boston Univ., Boston, MA 02215)

The paradigm system for studying light emission from bubbles has been a single, acoustically levitated air bubble in water. Such a bubble, whose volume pulsations are forced by the time-varying portion of the levitating acoustic field, has a rich dynamic life over a large range of the driving pressure-bubble size parameter space. Measurements have been made of bubble dynamics over a large part of a bubble's parameter space, including the light emission regime. In the context of these measurements, the relationship between observed light emission and the activity of the oscillating bubble will be discussed. In particular, it will be explained how the constraints of mechanical stability and mass flux stability dictate what all experiments in SBSL have seen, namely a very small window of parameter space where light is observed to be emitted. Experimental evidence will be presented to make a clear distinction between features of SL which are due to the mechanism for light emission and those that are due to the dynamics of gas bubbles. [Work supported by NASA and the Office of Naval Research.]

9:40

5aPAa2. Pulse width and shape in single bubble sonoluminescence. Bruno Gompf, Gerald Nick, Rainer Pecha, and Wolfgang Eisenmenger (1.Physikalisches Institut, Universitaet Stuttgart, Pfaffenwaldring 57, D-70550 Stuttgart, Germany)

To test the different theoretical models describing the light emitting process in SBSL measurements of the width and shape of the emitted light pulses are essential. The light pulses have been characterized by two independent methods: with time-correlated single photon counting and with a streak camera. Both methods lead to the same results. The pulse width strongly depends on the parameter's driving pressure, gas and gas concentration and temperature and varies between 60 ps and more than 300 ps, and there is no difference between the red and uv part of the spectrum. The streak camera results additionally show that the pulse shape is slightly asymmetric with a steeper ascent and a slower descent. This asymmetry increases with decreasing temperature.

10:00

5aPAa3. Predictions for upscaling sonoluminescence. Sascha Hilgenfeldt and Detlef Lohse (Dept. of Phys., Univ. of Marburg, Renthof 6, 35032 Marburg, Germany, lohse@stat.physik.uni-marburg.de)

Detailed comparison between the hydrodynamical/chemical approach towards single bubble sonoluminescence (SBSL) and recent experimental data is offered. In particular, the water temperature dependence is focused on. Moreover, detailed predictions of how to upscale SBSL are presented.

10:20

5aPAa4. Single-bubble sonoluminescence and liquid fracture. A. Prosperetti (Dept. of Mech. Eng., Johns Hopkins Univ., Baltimore, MD 21218)

Attention is drawn to the fact that the very mechanism that traps a bubble in a standing acoustic wave causes it to execute an oscillatory translational motion in the direction of gravity. A bubble driven below resonance will move up during the collapse phase and down during the expansion phase. A bubble translating and collapsing develops a jet in the direction of motion. It is hypothesized that sonoluminescence is due to the collision of this jet with the other side of the bubble surface. The mechanism of light emission is a ''fracture'' process of the liquid that initially cannot respond by flowing due to the very short rise time of the applied pressure. The picosecond duration of the light flash is the time it takes for the microjet overpressure to be relieved by reflection from the microjet free surface. Several other observed features of sonoluminescence (such as noble gas and temperature sensitivity, anomalous mass loss process, effect of surfactants) can also be explained, at least qualitatively. [Work supported by the Office of Naval Research.]

10:40

5aPAa5. Conditions during multibubble sonoluminescence. Kenneth S. Suslick, William B. McNamara III, and Yuri Didenko (Dept. of Chemistry, Univ. of Illinois, 601 S. Goodwin Ave., Urbana, IL 61801, ksuslick@uiuc.edu)

Acoustic cavitation results in extraordinary transient conditions inside the collapsing bubble. In addition to interesting chemical effects (sonochemistry), cavitation also produces light emission. Such sonoluminescence from cavitating clouds of bubbles [''multi-bubble sonoluminescence'' (MBSL)] in room-temperature liquids closely resembles flame emission. Effective emission temperatures have been obtained for MBSL from excited state metal atom emission (from sonolysis of several volatile metal carbonyls). From the relative intensities of multiple line emissions from Cr, Mo, and Fe, emission temperatures have been calculated and are all in close agreement with each other. The effects of bubble contents can alter the observed temperatures and this has been now directly observed for the first time. In addition, effective pressures can be estimated from line broadening and line shifts of the metal atom emission. The effective transient conditions formed during cavitation of bubble clouds are roughly 5000 K, 500 atm, which implies heating and cooling rates in excess of 1e10 K/s. Temperatures reached during single bubble sonoluminescence (SBSL) are likely to be very much higher.

11:00–11:10 Break

11:10

5aPAa6. Sonoluminescence and its neighbor cavitation bubble luminescence. Werner H. Lauterborn and Thomas Kurz (Drittes Physikalisches Institut, Universität Göttingen, D-37073 Göttingen, Germany)

The field of sonoluminescence, the making of light from sound, and the related field of cavitation bubble luminescence are reviewed. The single-bubble acoustic trap invented by Crum has started great activities through the enhanced possibilities of observing the dynamics of a single bubble in great detail. The question of spherical versus aspherical collapse in light emission and problems of enhancing the light output are addressed. For sonochemistry the case of many bubbles versus a single bubble is important. Some results on the dynamics of multi-bubble systems and their light emission are given. One neighbor of sonoluminescence, i.e., cavitation bubble luminescence from laser-induced bubbles, may be of great help for clarifying some aspects of light emission from collapsing bubbles.

11:30

5aPAa7. Predictions for stable sonoluminescence in nonaqueous liquids. Michael P. Brenner (Dept. of Math., MIT, Cambridge, MA 02139), Sascha Hilgenfeldt, and Detlef Lohse (Univ. of Marburg, Marburg, Germany)

The oscillation behavior of single sonoluminescing bubbles, and therefore their stability and the intensity of light emission, depends on a multitude of material parameters, including viscosity, surface tension, and liquid vapor. It is shown in analytical and numerical computations how substitution of water by nonaqueous fluids affects the domain for stable and unstable sonoluminescing bubbles in parameter space.

11:45

5aPAa8. Novel techniques for manipulating single-bubble sonoluminescence. Kirk Hargreaves, Thomas J. Matula, Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105), and William C. Moss (Lawrence Livermore Natl. Lab., Livermore, CA 94550)

Various acoustical methods are currently being utilized to manipulate the radial oscillations of single bubbles undergoing sonoluminescence. Previous measurements have shown that the addition of higher harmonics to the drive frequency [J. Holzfuss, R. Mettin, and W. Lauterborn (private communication)] can result in an increase in the light intensity by as much as 50%. The radial response is also dramatically affected by the addition of higher harmonics. Nonlinear bubble dynamics equations agree well with the measured radial response. This and related work will be described. [Work supported by NSF and under the auspices of the U. S. Department of Energy at Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.]

12:00

5aPAa9. Abrupt drive pressure variations for probing the extinction threshold mechanism for single-bubble sonoluminescence. Thomas J. Matula and Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

A novel method utilizing abrupt drive pressure amplitude changes has been used to probe transient conditions in single-bubble sonoluminescence (SBSL). This technique has been used to probe the incipient luminescence threshold [Matula and Crum, Phys. Rev. Lett. (submitted)], where it was found that sonoluminescence from air bubbles depend on the time the bubbles spend in the sonoluminescing state. The abrupt pressure change technique has also recently been used to probe SBSL behavior near the extinction threshold. Simultaneous measurements of the radial response and light emission may be used to better understand the mechanism for SBSL extinction.

FRIDAY MORNING, 26 JUNE 1998 ASPEN ROOM (S), 9:15 TO 10:45 A.M.

Session 5aPAb

Physical Acoustics: Acousto-Optics and Opto-Acoustics

Harry Simpson, Chair
Naval Research Laboratory, Code 7136, 4555 Overlook Avenue, SW, Washington, DC 20375-5000

Contributed Papers

9:15

5aPAb1. Planar backward projection of transient fields obtained from optical imaging methods. Gregory Clement, Ruiming Liu, Stephen Letcher (Dept. of Phys., Univ. of Rhode Island, Kingston, RI 02881, gcle0008@uriacc.uri.edu), and Peter Stepanishen (Univ. of Rhode Island, Kingston, RI 02881)

A polarimetric fiber-optic detection system is used to reconstruct the field in front of a planar source. The field is then back-projected toward the source using a filtered wave vector/frequency domain method. This method is compared with a second temporal projection technique that records data over space at a specific time and projects the field to an earlier time. Numerical results will be presented to illustrate the two methods. Experimental results from a pulsed ultrasonic transducer will be presented and discussed.

9:30

5aPAb2. Measurement of ultrasonic displacement vector with an *F*-*P* interferometer. Menglu Qian, Yongdong Pan (Inst. of Acoust., Tongji Univ., Shanghai 200092, PROC), and Zhongxian Zhang (Zhejiang Univ., Hangzhou 310027, PROC)

The laser interferometer technique for receiving ultrasound has potential for a wide range of noncontact ultrasonic measurements and has therefore received considerable interest in recent years. Several interferometric methods for detecting the normal displacement components of ultrasound have been developed and applied to nondestructive evaluation for material characterization. In order to study the propagation of ultrasound in a medium and to characterize materials more accurately, detecting both normal and tangent components of ultrasonic displacement simultaneously is necessary. Therefore, an *F*-*P* interferometer, which receives scattering lights at scattering angles $+\theta$ and $-\theta$ from sample surfaces and measures the ultrasonic displacement vector simultaneously, is introduced in this paper. In order to increase the scattering light intensity and determine the scattering angle θ precisely, a new method for modification of a sample surface is introduced and some experiments are presented in this paper.

9:45

5aPAb3. Two-frequency degeneracy of an acousto-optic interaction in paratellurite. Alexander Yurchenko and Vladimir Moskalev (Dept. of Quantum Radiophys., Kiev Taras Shevchenko Univ., 64, Vladimirskaya str., 252033 Kiev, Ukraine, alyur@boy.rpd.univ.kiev.ua)

Earlier unknown effects of an acousto-optic interaction degeneracy in paratellurite have been studied. Its essence compared to described phenomena lies in the fact that a light beam can be diffracted twice by acoustic waves with different frequencies instead of being typically diffracted twice by the same acoustic wave. It happens due to strong acoustic anisot-

ropy of paratellurite. Bragg conditions resulting in the degeneracy can be satisfied under a certain frequency correlation of two acoustic slow shear waves. A light beam diffracted by one spectral component of a complex acoustic signal can be effectively diffracted again by another spectral component. Thus, a physical mechanism of this effect is the same as it is in the ordinary case but it is observed under different conditions (different frequencies). Necessary calculations of such interactions have been carried out. The frequency of the ''unfavorable'' spectral component was calculated depending on the frequency of the ''useful'' one for the case of nonaxial acousto-optic interaction in paratellurite. The found effect was shown to take place and to be important for real acousto-optic devices. In the studied case, the frequencies of acoustic waves interacting with a light beam were equal to 44.7 and 43.2 MHz, respectively.

10:00

5aPAb4. Visualization of a phase structure of sound field in a Bragg cell. Vadim Goncharov, Leonid Ilchenko, Vladimir Moskalev, Eugene Smirnov, and Alexander Yurchenko (Dept. of Quantum Radiophys., Kiev Taras Shevchenko Univ., 64, Vladimirskaya str., 252033 Kiev, Ukraine, gva@boy.rpd.univ.kiev.ua)

A phase structure of a sound field in a Bragg cell was investigated experimentally using a dual-beam, dual-frequency laser system. The technique used allows the recording of a spatial phase structure of diffracted light field to visualize phase nonuniformities of a sound field. The underlying physical principle of this technique is to probe an analyzed phase object using two slightly separated in space light beams with different optical frequencies. A difference frequency signal, which is obtained as a photodetector output in an optical heterodyne system, contains information about a phase difference of the adjacent areas of the sound field. When the probing beams are diffracted by a Bragg cell, their optical frequencies are shifted by the same acoustic frequency, so that the difference is the same at any Bragg cell operating frequency. A Bragg cell on paratellurite operating around 100 MHz was investigated. Focused laser beams with a difference of optical frequencies equal to 1 MHz were used as probing beams. Pictures were obtained which demonstrated dependencies of a phase structure of diffracted light on frequency and sound power. Phase nonuniformities across the sound field up to three wavelengths have been estimated. [Work supported by EOARD.]

10:15

5aPAb5. Radiation acoustics—photoacoustics and acousto-optics of penetrating radiation. Leonid M. Lyamshev (N. N. Andreev Acoust. Inst., Shvernik st. 4, 117036 Moscow, Russia, bvp@asu.acoins.msk.su)

Photoacoustics is a science about optical generation of sound in a substance. The study of light diffraction by acoustic waves is the subject of acousto-optics. Light is a beam of particles, i.e., photons. In this connection the study of sound generation by beams of electrons, protons, neutrons, and other particles in a substance, as well as the study of matter waves' (particle beams) diffraction by ultrasound may be considered correspondingly as photoacoustics and acousto-optics of penetrating radiation. Combined together, this is the field which got the name of radiation acoustics. It develops at the interface of acoustics, nuclear physics, solid state physics, and physics of high energy and elementary particles [see L. M. Lyamshev, *Radiation Acoustics* (Nauka-Fizmatgiz, Moscow, 1996), in Russian]. Some universal laws characterizing generation of sound by beams of penetrating radiation are discussed. The problem of acoustic–radiational interaction is considered using the examples of diffraction of x rays and thermal neutrons by ultrasound in solids. [Work supported by RFBR.]

10:30

5aPAb6. Photoacoustic signal in the system of a thin oil layer on the water surface. Antoni Sliwinski, Stanislaw Pogorzelski, and Janusz Szurkowski (Inst. of Experimental Phys., Univ. of Gdansk, 80-952 Gdansk, Poland, fizas@univ.gda.pl)

Thin layers of different kinds of oils spread on the water surface (using standard techniques) were examined by photoacoustic spectroscopy (PAS) methods. Both techniques of excitation were applied: the classic one with a modulated light beam as well as with a laser pulse beam. It was demonstrated that the mechanism of molecular interaction at the oil–water interface represents a good model for determination of the correlation between PAS signal characteristics (amplitude and phase) and physicochemical properties of the thin layer system. The results obtained have confirmed the existence of a very sharp separation of the oil and water phases. Variations in physical properties of layers of different thickness like surface and interface tension thermal conductivity and heat capacity and optical absorption coefficients influenced by different agents (e.g., surfactants) could be detected with the PAS methods applied. [Work supported by KBN.]

Session 5aPP

Psychological and Physiological Acoustics: A Microbrew of Speech Perception and Hearing Impairment (Poster Session)

Kathryn H. Arehart, Cochair
Department of Speech Language and Hearing Science, University of Boulder, Box 409, Boulder, Colorado 80309

Shari L. Campbell, Cochair
University of Georgia, 514 Aderhold Hall, Athens, Georgia 30602

Contributed Papers

All posters will be on display from 7:45 a.m. to 10:45 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 7:45 a.m. to 9:15 a.m. and contributors of even-numbered papers will be at their posters from 9:15 a.m. to 10:45 a.m. To allow for extended viewing time, authors have been requested to place their posters on display on Thursday afternoon beginning at 1:00 p.m. in Grand Ballroom II.

5aPP1. Perceptual consequences of amplitude perturbation in the wavelet coding of speech. Nicolle H. van Schijndel, Tammo Houtgast, and Joost M. Festen (Dept. of Otolaryngol., Univ. Hospital VU, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands, NH.vanSchijndel@azvu.nl)

Wavelet coding of sound is a frequency–time representation that may be able to mirror the properties of auditory coding at a peripheral level, provided that we use the proper so-called mother wavelet. Previous experiments based on intensity discrimination [van Schijndel *et al.*, J. Acoust. Soc. Am. (submitted)] suggested a Gaussian-shaped wavelet with a bandwidth of 1/3 oct. Based on this function, a wavelet coding and reconstruction algorithm has been developed. The frequency–time plane is split up: one wavelet per three periods along the time axis and nine wavelets per octave along the frequency axis. In this study the amplitude of the wavelet coefficients is manipulated before the stage of reconstruction. The perceptual effect of amplitude perturbation on speech perception is studied. A (partial) answer to the following questions will be presented: (1) what degree of amplitude perturbation is just noticeable to the subject, and (2) how is speech intelligibility affected by suprathreshold amplitude perturbation. The results may provide insight into the acuity of the auditory amplitude coding and its effect on speech intelligibility. [Work supported by the Netherlands Organization for the Advancement of Pure Research (N.W.O.).]

5aPP2. The effect of spectral resolution on speech perception in a multitalker babble background. Bom Jun Kwon and Christopher W. Turner (Dept. of Speech Pathol. and Audiol., Univ. of Iowa, Iowa City, IA 52242)

Does finer spectral resolution always have an advantage for speech recognition? In this experiment, speech was divided into separate bands within which spectral information was removed by replacing the fine structure with noise; the temporal envelope cues of each band remained intact. Normal-hearing subjects identified consonants in a/C/a for speech and babble processed with 1, 2, 4, and 8 bands, in addition to the unprocessed speech and babble, for a range of signal-to-babble ratios. For high S/B ratios our results agree with earlier results, in that performance improved with increases in the number of bands, and higher transmission of information was observed for voicing and manner than place of articulation [Shannon *et al.*, Science **270**, 303–304 (1995)]. However, as S/B decreased, the advantage of better spectral resolution was reduced. This suggests that for speech in a competing babble background, increased spectral resolution not only increases the saliency of speech, but at the same time makes background babble a more effective distracter. Thus the benefit provided by fine spectral resolution depends upon the task and S/B ratio. [Work supported by NIDCD.]

5aPP3. Application of auditory models to discrimination thresholds of voicing parameters. Shari L. Campbell (Dept. of Commun. Sci. and Disord., Univ. of Georgia, 514 Aderhold Hall, Athens, GA 30602), Kathryn H. Arehart (Univ. of Colorado, CB 409, Boulder, CO 80309), Ronald C. Scherer (Bowling Green State Univ., Bowling Green, OH 43403), and Amy N. Butler (Univ. of Georgia, Athens, GA 30602)

A recent study [Scherer *et al.*, J. Voice (to be published)] reported discrimination thresholds for synthesized complex tones emulating glottal flow waveforms and output pressure waveforms for the vowel /a/. Each signal type was varied along both an open quotient and a speed quotient continuum using standard values associated with normal voice production. jnds for the output pressure signals were significantly larger than for the glottal flow signals. The current experiment compared quantities derived from ''auditory-perceptual'' representations of the acoustic signals. These quantities (centers of gravity and Euclidean distances) were computed for both the raw amplitude spectra and for modeled excitation patterns and specific loudness patterns. Overall loudness and overall sharpness were also estimated. The comparisons indicated ways in which auditory modeling can enhance our understanding of basic issues in voice and speech perception. [Work supported by UGA College of Education.]

5aPP4. Audio signals in domestic appliances evaluated in terms of the hearing ability of older adults. Kenji Kurakata, Yasuyoshi Kuba, Yasuo Kuchinomachi (Natl. Inst. of Bioscience and Human-Technol., 1-1 Higashi, Tsukuba, Ibaraki, 305 Japan, kurakata@nibh.go.jp), and Kazuma Matsushita (Natl. Inst. of Technol. and Evaluation, Tokyo, 151 Japan)

The audio signals used in domestic appliances currently available on the market in Japan were recorded to identify suitable signals for the hearing ability of older adults. The results of the analysis indicated the following three problems: (1) Some appliances use high-frequency tones around 4000 Hz. Since these sounds are hard for older adults with presbycusis to hear, it would be better to use signals with lower frequencies. However, the problem here is that if the frequency is lowered, then the signal might be masked by domestic sounds whose power would be greater relative to the low-frequency signals. (2) The signals used by some

appliances are too soft. The intensity of some signals should be adjusted to compensate for the hearing loss among older adults. (3) The sounds used in these appliances are often very similar in terms of both timbre and temporal ringing patterns. That may cause confusion because it is difficult to identify which appliance is signaling.

5aPP5. The effects of age and hearing impairment on the time course of backward masking. Sara Elizabeth Gehr and Mitchell Sommers (Washington Univ., Dept. of Psych., 1 Brookings Dr., Campus Box 1125, St. Louis, MO 63130)

The present study was designed to assess the effects of age and hearing loss on the time course of backward masking. Thresholds for detecting a 10-ms, 500-Hz tone were measured as a function of the temporal separation between the signal and a 50-ms broadband masker in normal-hearing young listeners (18–25), normal-hearing older adults (65–82), and hearing-impaired older adults. Younger subjects exhibited less initial masking, had steeper recovery functions, and asymptoted closer to their unmasked thresholds than either group of older adults. Comparisons between the normal-hearing and hearing-impaired older individuals indicated a slower time course of recovery from backward masking for listeners with hearing loss. Differences between the two groups of normal-hearing listeners (young and old) suggest that age, independent of hearing loss, affects the temporal course of backward masking. Differences between the normal-hearing and hearing-impaired elderly listeners, however, indicate that age-related hearing impairment may also alter the rate of recovery from backward masking. [Work supported by the Brookdale Foundation and NIA.]

5aPP6. Speech intelligibility and localization in a complex environment by listeners with hearing impairments. Monica L. Hawley, Ruth Y. Litovsky, and H. Steven Colburn (Hearing Res. Ctr. and Dept. of Biomed. Eng., Boston Univ., Boston, MA 02215)

Listeners with hearing impairments often complain of difficulties when trying to communicate in a complex environment such as a cocktail party. Explorations of the dependence of the intelligibility and localizability of speech on the number and spatial positions of simultaneous interfering sources are reported on. Target sources are sentences spoken by a female and interfering sources include other sentences or speech-shaped noise. The total number of sources played at any one time varied between one and four. Listeners with normal hearing, profound monaural deafness, and with symmetric sensorineural hearing impairments were tested. Listeners were tested without hearing aids. Previous work in free field and with virtual stimuli has shown that bilaterally hearing-impaired listeners show large intersubject differences in their performance in both speech intelligibility and localization tasks. In the present study an investigation was made to see whether these intersubject differences are related to each listener's sensitivity to specific binaural cues. Therefore, in addition to localization and speech intelligibility, performance in discrimination of interaural time and level, as well as binaural masking level differences in each listener was measured. The goal of this study is to determine the extent to which performance on speech-based tasks can be accounted for by basic binaural abilities. [Work supported by NIH-NIDCD DC00100.]

5aPP7. Monotic and dichotic musical-interval identification in listeners with cochlear-based hearing loss. Kathryn H. Arehart (Univ. of Colorado, Speech, Lang., Hearing Sci., CB 409, Boulder, CO 80309) and Edward M. Burns (Univ. of Washington, Seattle, WA 98195)

This study examined musical-interval identification in musically experienced listeners with cochlear-based hearing loss. Sounds had missing fundamentals ranging from 100–500 Hz. These missing fundamentals were conveyed by two successive harmonics presented either to the same ear or to separate ears. Performance was measured as a function of the average rank of the lowest harmonic for both monotic and dichotic conditions at both 20 dB sensation level (SL) and 30 dB SL. In both monotic and dichotic conditions, listeners with cochlear-based hearing loss demonstrated excellent musical-interval identification at low fundamental frequencies and low average ranks, but abnormally poor recognition at higher fundamental frequencies and higher average ranks. The results will be related to auditory deficits which might affect complex-tone pitch perception in listeners with cochlear-based hearing loss, and will be discussed in the context of modern pitch theories. [Work supported by Deafness Research Foundation.]

5aPP8. Minimum bandwidth required for speech-reception by normal-hearing and hearing-impaired listeners. Ingrid M. Noordhoek, Tammo Houtgast, and Joost M. Festen (Dept. of Otolaryngol., Univ. Hospital VU, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands, im.noordhoek@azvu.nl)

A new adaptive test has been developed to determine the bandwidth of speech (with a center frequency of 1 kHz) required for 50% intelligibility. Measuring this speech-reception *bandwidth* threshold (SR*B*T), in addition to the more common speech-reception threshold (SRT) *in noise*, may be useful in investigating the factors underlying impaired suprathreshold speech perception. A wider-than-normal SR*B*T points to a deterioration in suprathreshold sound processing in the 1-kHz frequency region. Both the SRT in noise and the SR*B*T were measured for 10 normal-hearing and 34 hearing-impaired listeners. To keep the speech signal above hearing threshold for all frequencies, speech and noise spectra were shaped to fit in the dynamic range of individual listeners. The average SR*B*T of the normal-hearing listeners was 1.4 oct. Most hearing-impaired listeners needed a wider bandwidth to reach 50% intelligibility. Some wider SR*B*Ts were predicted by the SII model (speech intelligibility index) from a high presentation level or narrow dynamic range. The remaining differences between the performance of normal-hearing and hearing-impaired listeners are probably caused by an impaired suprathreshold sound processing in the 1-kHz frequency region. [Work supported by the Foundation ''Heinsius-Houbolt Fonds.'']

5aPP9. Loudness of dynamic stimuli in cochlear-impaired listeners. Chaoying Zhang and Fan-Gang Zeng (House Ear Inst., 2100 West Third St., Los Angeles, CA 90057)

In normal hearing listeners, a temporally fluctuating sound with modulating frequencies from 4 to 400 Hz was found to be louder than predicted by its averaged power. The louder sensation due to temporal fluctuation was also found to be overall level-dependent: approaching peak amplitude loudness at a comfortably loud level, becoming less at lower sensation levels, and disappearing at higher levels. If this level-dependent loudness effect were related to the nonlinear cochlear compression, then it would disappear in cochlear-impaired listeners because their nonlinear compression is presumably damaged. Three listeners with clinically diagnosed cochlear impairment balanced loudness between white noise and sinusoidally amplitude-modulated noise. The results showed that the modulated noise was louder for modulating frequencies between 4 and 400 Hz at all three sensation levels, consistent with the hypothesis that the overall level effect on loudness of dynamic stimuli is related to the cochlear compression. These results indicate that loudness of dynamic stimuli is governed by two different mechanisms: a retrocochlear temporal processing mechanism which is presumably preserved in both normal and cochlear-impaired listeners, and a cochlear mechanism which provides amplitude compression at high levels in normal listeners and is damaged in the cochlear-impaired listeners. [Work supported by NIDCD.]

5aPP10. Further effects of cochlear loss on ICP loudness adaptation, and loudness constancy. Rebecca Schnieder Ludwig, Hongwei Dou, Ernest M. Weiler, Laura W. Kretschmer, David E. Sandman (ML #379, CSD, Univ. of Cincinnati, Cincinnati, OH 45221), and Eleanor Stromberg (Cincinnati Speech & Hearing Ctr.)

ICP loudness adaptation depends on an ipsilateral comparison paradigm with a short increasing or decreasing referent value to reveal loudness adaptation. Janson *et al.* [Br. J. Audiol. **29**, 288–297 (1996)] found the intensity-based function for ICP loudness adaptation for those with a high-frequency loss differed significantly from normal listeners. Loss of loudness constancy in normal listeners was consistent, loss in those with a cochlear loss showed a negative progression as intensity increased, as if it were the inverse of recruitment. The purpose of the present study was to compare ICP adaptation to recruitment. ICP adaptation and recruitment were tested in ten high-frequency loss (probably noise induced) and ten age-matched normal hearing individuals. Although the evidence for recruitment was generally found in those with hearing loss, there was no simple relationship to ICP adaptation. Variation among measures of recruitment, implications of abnormal loudness constancy for speech perception, and the possible use of measures of otoacoustic emissions for future studies will be discussed.

5aPP11. Evoked otoacoustic emissions, suppression and comparison to loudness constancy, at speech frequencies. Ernest M. Weiler, Hongwei Dou, Robert S. Tannen, David E. Sandman, William N. Dember, and Joel S. Warm (ML #379, Hearing Lab., CSD, Univ. of Cincinnati, Cincinnati, OH 45221)

Spontaneous otoacoustic emissions (SOAE) reflect spontaneous activity of the outer hair cells of the cochlea. Suppression of transient evoked otoacoustic emissions (TEOAE) bears operational similarity to loudness adaptation determined by the ipsilateral comparison paradigm (ICP). The present study gathered these measures from 41 normal hearing adults (5 male, 36 female, 20–45 years old). There was no significant relationship between spontaneous SOAE, but there was a significant relationship ($r = -0.36$, $p < 0.05$) between ICP loudness adaptation and suppression of TEOAEs. The relationship between these two active measures has implications for failure of adaptation or the apparent loudness constancy in the speech range revealed by simple adaptation measures. It has been previously noted that the ICP technique yields adaptation in the speech frequencies between 40 and 80 dB, where the simple adaptation apparently does not [Weiler, Sandman, and Dou, program abstract 4pPPa10, J. Acoust. Soc. Am. **101**, 3177 (1997)]. Implications will be discussed in terms of the results uncovered in this study.

5aPP12. Virtual midplane localization of subjects with sensorineural hearing loss. Helen J. Simon (Smith-Kettlewell Eye Res. Inst., 2232 Webster St., San Francisco, CA 94115, helen@skivs.ski.org) and Inna Aleksandrovsky (Univ. of California, Berkeley, CA)

This paper reports the results of a study that investigated localization performance of simulated externalized signals in sensorineural hearing loss (SNHL) listeners when signals were presented with either unbalanced [equal sensation level (EqSL)] or balanced [equal sound-pressure level (EqSPL)] hearing aid gain at the two ears. In previous reports from this laboratory, it has been found that when equalizing by SL, the perceived lateral position was essentially linearly dependent on the degree and direction of asymmetry in normal-hearing (NH) and SNHL listeners. Equalizing by SPL showed no such dependency. It was hypothesized that the current practice of fitting ''binaural'' hearing aids in asymmetric SNHL listeners using prescription formulas might result in SPL imbalance which would disrupt the previously adapted system and impair localization. The present study was designed to address this issue by measuring localization using a simulated externalized auditory image combined with head position control of IID and ITD to spatially fix the image. Absolute and signed error deviations from center were analyzed. Significant interactions of group (NH versus SNHL)* frequency and balance (EqSL or Eq SPL)* frequency were found. These data will be discussed in terms of auditory adaptation and strategies of fitting binaural hearing aids. [Work supported by NIDRR.]

5aPP13. Use of temporal information in recognition of amplitude-compressed speech by older adults. Pamela E. Souza and Johanna J. Larsen (Dept. of Speech and Hearing Sci., Univ. of Washington, 1417 NE 42nd St., Seattle, WA 98105, psouza@u.washington.edu)

Amplitude compression alters the speech envelope and under some conditions can reduce the availability of temporal cues for speech recognition. This may be a particular problem for older listeners, who demonstrated temporal resolution deficits in some studies. Older and younger hearing-impaired listeners were tested on their recognition of VCV syllables processed using the signal-correlated noise (SCN) method [Schroeder, J. Acoust. Soc. Am. **44**, 1735–1736 (1968)]. This method results in a time-varying speech envelope modulating a noise carrier, and preserves temporal cues while minimizing spectral information. Recognition scores were obtained for uncompressed SCN syllables and for the same speech tokens after digital processing with a syllabic compression algorithm. In control conditions, recognition was tested for VCVs that included both temporal and spectral information. Younger and older listener groups were matched for mean hearing thresholds up to 4 kHz. A high-pass masker was included to eliminate contributions from the basal end of the cochlea. When compared to younger listeners with similar hearing losses, the older listeners demonstrated poorer performance on the compressed SCN signals. These results suggest that the poorer temporal resolution abilities of older listeners may play a role in their understanding of amplitude-compressed speech.

5aPP14. Apparent effects of the use of digital hearing aid ''CLAIDHA'' on the several hearing functions of the impaired listeners. Hiroshi Hidaka, Naoko Sasaki, Tetsuaki Kawase, Tomonori Takasaka (Dept. of Otolaryngol., Tohoku Univ. School of Medicine), Kenji Ozawa, Yôiti Suzuki, and Toshio Sone (Res. Inst. of Elec. Commun., Tohoku Univ.)

The portable digital hearing aid system CLAIDHA (compensate loudness by analyzing input signal digital hearing aid), which employs frequency-dependent amplitude compression based on narrow-band compensation, has been developed [Asano *et al.* (1991)]. The patients with hearing loss using hearing aids experience a new sound world as amplified by the device. It seems to be important to know the apparent change of hearing function by the use of a hearing aid. In the present study, the apparent effects of CLAIDHA amplification especially on the intensity discriminability and the masking function were examined in the patients with hearing loss. The results obtained indicate that the CLAIDHA amplification often improves not only the apparent thresholds but also the apparent masking relation between two tones as compared to simple linear amplification; i.e., in the CLAIDHA amplification, more signals could be heard under the same masking condition. On the other hand, from the viewpoint of intensity discriminability, when the compression ratios of the amplification were larger than two, the IDL (intensity difference limen) rapidly increased more than expected. The advantage of the compressive amplification of the CLAIDHA system will be discussed.

5aPP15. The effect of frequency-modulation systems on the communication skills of deaf–blind children. Barbara Franklin (Dept. of Special Education, San Francisco State Univ., 1600 Holloway Ave., San Francisco, CA 94132, dick3580@aol.com)

Classrooms and other public places provide a poor acoustical environment for individuals who have a hearing loss, and it is often difficult for them to discriminate between the speaker and background noise with a hearing aid. The use of a frequency-modulation (FM) system can often reduce this problem by providing a constant sound pressure level of the

speaker's voice. The effect of FM systems on communication skills of deaf–blind children was investigated using a single-subject alternating treatment design. There was a marked improvement for a number of the subjects in the FM condition. The results will be discussed along with individual case studies. FM technology has exploded. A relatively new type of FM receiver is now available which combines the hearing aid and FM in a single behind-the-ear unit (BTE/FM). The most recent development is a ''boot'' attachment which changes a standard over-the-ear hearing aid into a FM system. An overview of these FM systems will be presented. [Work supported by OSERS.]

5aPP16. A metric for determining the degree of compression of a processed signal. Jon C. Schmidt[a)] and Janet C. Rutledge[b)] (Dept. of Elec. Eng. and Comput. Sci., Northwestern Univ., Evanston, IL 60208-3118, jschmidt@resound.com)

When attempting to compare different compression algorithms or different parameter settings of a given compression algorithm, there is a need to compare equivalent amounts of compression. Listening tests that have varied only one parameter while fixing the others have obtained the expected results that increasing the amount of compression that occurs creates a more audible difference between the compressed version and the original. Since there are several parameters that influence the actual amount of compression that occurs, it is desirable to know which set of parameters will create to the least perceptual difference from the original signal for a given amount of compression. Unfortunately, there is not a well-defined approach for determining the amount of compression that has been imparted to a signal. The peak/RMS ratio is woefully inadequate for many compression contexts. Two degree of compression metrics are proposed that can be used to determine the amount of compression by direct analysis of the audio signal before and after compression. These metrics give excellent results in predicting the ratings of subject-based testing on the audio quality of several different audio segments across several parameter variations. [a)]Currently with ReSound. [b)]Currently with the University of Maryland Medical Center.

5aPP17. Modeling tactile speech perception with auditory simulations. Albert T. Lash, Janet M. Weisenberger, and Ying Xu (Speech and Hearing Sci., Ohio State Univ., Columbus, OH 43210)

A useful technique for gaining insights into the perception of speech stimuli presented via vibrotactile aids is to model this perception by presenting degraded auditory stimuli to normal-hearing observers. It is likely that spectral, temporal, and intensive properties of the tactile system could all affect tactile speech perception. In the present experiments, tactile perceptual data were collected for a set of nonsense syllables presented via a vibrotactile vocoder to normal-hearing subjects. The same subjects were tested with sets of degraded auditory stimuli that were modified by reducing spectral, temporal, or intensive parameters. Spectral reductions alone, achieved by reducing the number of filter channels and broadening filter bandwidths, did not appreciably affect auditory performance, nor did intensity reductions, achieved by reducing dynamic range and intensity quantization. Although temporal reductions, achieved by low-pass filtering the stimuli, did result in lower perceptual scores, the pattern of stimulus confusions did not correlate highly with the tactile data. Pairwise and three-way combinations of spectral, temporal, and intensive parameters yielded more promising correlations. Implications for constructing a comprehensive model of vibrotactile speech perception are discussed. [Work supported by NIH.]

5aPP18. A design of a tactile voice coder with different tactile sensations for the hearing impaired. Chikamune Wada, Shuichi Ino, Hisakazu Shoji, and Tohru Ifukube (Res. Inst. for Electron. Sci., Hokkaido Univ., N12 W6, Kita-ku, Sapporo, 060 Japan, chik@sense.hokudai.ac.jp)

A tactile voice coder is a device which changes a speech sound into a vibratory pattern and presents it to the tactile sense of the hearing impaired. The authors have been improving the tactile voice coder in order that the hearing impaired can identify consonants easier. The use of a convex stimuli pattern, similar to Braille, in addition to the vibratory stimuli is proposed, as one of the ways of improving the tactile voice coder. The authors have made this proposal based on the hypotheses that the combination of the vibratory stimuli and the convex stimuli could transmit more speech information. From the combination of basic experiments, the authors have found that presenting a vowel using the convex stimuli and a consonant using the vibratory stimuli was better than only using the vibratory stimuli. Therefore, based on our findings, the authors have designed a new tactile voice coder which uses a different tactile sensation.

5aPP19. Sensitivity of transiently evoked otoacoustic emission to a long-term exposure of impulsive noise. Amnon Duvdevany and Miriam Furst (Dept. of Eng., Tel-Aviv Univ., Tel-Aviv, Israel, mira@eng.tau.ac.il)

In spite of the regular use of ear protectors, 30%–40% of the soldiers that are often exposed to impulsive blast noise are suffering from hearing loss. The purpose of the present study is to find in what ears click-evoked otoacoustic emissions (COAEs) were influenced by the noise exposure. The preliminary study included testing of 15 healthy soldiers. They were exposed to blast noise (M-16 riffles) accompanied by replicas from a semi-enclosed firing range in two consecutive sessions, followed by half a year of exposure to continuous blast noise (service period). In each stage COAEs were measured. The blast noise was precisely measured with microphones located next to the shooters' ears. None of the soldiers acquired a hearing loss. In most subjects COAE levels have significantly changed in the frequency range 2.5–4 kHz. It decreased in the postexposure condition relative to the preexposure by ~3 dB. It increased before the second training session by ~8 dB, and it decreased in the last measurement by ~5 dB. A protection and a destruction mechanism are possibly operating. The protection mechanism reduces the cochlear amplification and causes a reduction in COAE level. The destruction mechanism increases the cochlear nonuniformity which causes an increase in COAE level.

5aPP20. The effects of dexamethasone following an acute acoustic trauma. William A. Ahroon, Ann R. Johnson, B. Sheldon Hagar, and Roger P. Hamernik (Plattsburgh State Univ., Plattsburgh, NY 12901)

An acute acoustic trauma (AAT) results when the unprotected ear is exposed to very high sound pressure levels, usually as the result of an industrial or military accident. There is no commonly accepted treatment in the U.S. for an AAT. A study was performed to evaluate the effects of a corticoid steroid (dexamethasone) on the hearing loss from a series of blast waves that would simulate an AAT. Dexamethasone is commonly used as an anti-inflammatory agent in veterinary medicine. Groups of chinchillas were exposed individually to ten 160-dB peak SPL reverberant blast waves from a conventional shock tube at a rate of one blast per minute. Immediately following the exposure, the animals were injected with dexamethasone alone (1.0–2.0 mg/kg IV) or in combination with dimethyl sulfoxide (DMSO), a free-radical scavenger. Individual groups of animals showed the well-known extreme variability in hearing losses as noted in earlier studies [e.g., R. P. Hamernik *et al.*, J. Acoust. Soc. Am. **90**, 197–204 (1991)]. Examining the median permanent threshold shifts (PTS) of the groups of animals showed a distinct dose-response effect with increasing dosages of dexamethasone associated with lower levels of PTS in the frequency region most affected by the noise exposure.

5aPP21. Protective effects of magnesium on noise-induced hearing loss: Animal studies. Fred Scheibe and Heidemarie Haupt (Dept. of Otorhinolaryngology, Charité-Hospital, Humboldt Univ., Schumannstr. 20-21, D-10117 Berlin, Germany, fscheibe@rz.charite.hu-berlin.de)

Magnesium (Mg) deficiency was found to increase noise-induced hearing loss in laboratory animals. This paper reports both prophylactic and therapeutic effects of Mg on the acoustic trauma produced by a high-level impulse noise series (L_{peak} 167 dB, 38 min, 1/s). Hearing loss was tested by auditory brainstem response audiometry at frequencies between 0.5 and 32 kHz. Permanent hearing threshold shifts (PTS) were measured 1 week after the exposure. For the prophylaxis experiments, anesthetized guinea pigs with either a physiologically high or low Mg status, produced by different diets, were used. For the therapy experiments, animals with the low Mg status received (immediately after exposure) either Mg injections combined with a Mg-rich diet or saline as a placebo. Total Mg concentrations of perilymph, cerebrospinal fluid, and plasma were analyzed to test the Mg status of the animals. The PTS was found to be significantly lower in the high-Mg group than in the low-Mg group. This also applies to the PTS found between the therapy and the placebo groups. There was a good correlation between the PTS and the perilymphatic Mg. The intracochlear Mg level seems to play an important role in the acoustic trauma. [Supported by German Defence Ministry.]

5aPP22. Effects of impulse noise on the hearing of fetal sheep *in utero*. Kenneth J. Gerhardt (Dept. of Commun. Processes and Disord., Univ. of Florida, Box 117420 Gainesville, FL 32611-7420, gerhardt@cpd.ufl.edu), Xinyan Huang, and Robert M. Abrams (Univ. of Florida, Gainesville, FL 32610)

Knowledge of the transmission of exogenous sounds to the fetal head and the effects that these sounds have on fetal hearing is incomplete. The purposes of this study were to measure the transmission of impulse noise into the uterus and to evaluate the effects of impulse noise delivered to pregnant sheep on the hearing of the fetus *in utero*. A shock tube produced impulses that averaged 169.7 dB peak sound pressure level (pSPL) in air. In the uterus, the pSPL varied as a function of fetal head location. When the fetal head was against the abdominal wall, peak levels were within 2 dB of airborne levels and the stimulus resembled a Friedlander wave. When the fetal head was deep within the uterus, the duration of the impulse increased and the peak amplitude decreased. In some instances the decrease in pSPL exceeded 10 dB. Slight elevations of evoked potential thresholds were noted but only for low-frequency stimuli. The integrity of hair cells from these animals was assessed using scanning electron microscopy.

FRIDAY MORNING, 26 JUNE 1998 WEST BALLROOM B (S), 9:15 A.M. TO 12:15 P.M.

Session 5aSA

Structural Acoustics and Vibration: Vibrations of Structural Elements

Linda P. Franzoni, Chair
Department of Mechanical and Aerospace Engineering, North Carolina State University, Box 7910, Raleigh, North Carolina 27695-7910

Contributed Papers

9:15

5aSA1. Moment and force mobilities of semi-infinite tapered beams. Eugene J. M. Nijman (CRF, Strada Torino 50, 10043 Orbassano, Italy, e.nijman@crf.it) and Bjorn A. T. Petersson (Loughborough Univ., Leicestershire LE11 3TU, UK)

The implications of spatially varying structural properties for flexural vibrations will be addressed for beamlike systems. Under the assumption of pure bending, closed-form expressions for the mobilities of a semi-infinite wedge have been derived. The influence of tapering on the energy flow is analyzed for the structural system constituted by a finite length taper attached to a semi-infinite, uniform beam. It is found that the main distinction from the uniform case is a comparatively broadbanded transition from flexural vibrations governed by the properties of the deep part of the system to vibrations governed by those of the slender part. This transition region is featured by a stiffness-controlled global behavior in the case of a moment-to-rotational response, whereas a global resonant behavior is found in the case of force-to-translational response. It is also demonstrated that constructive and destructive interferences due to reflections from the discontinuity between the two structural members are superimposed on the global behavior. Based on the complete expressions for the configurations studied, estimation procedures for the mobilities will be proposed. In addition, experimental results will be presented which corroborate the description developed and the parameter influence established.

9:35

5aSA2. Vibrational power transmission in asymmetric framework structures. Jane L. Horner (Dept. AAETS, Loughborough Univ., Loughborough, Leicestershire LE11 3TU, England)

When attempting to control the vibration transmitted from, say, a machine into and through the structure upon which it is mounted, it is desirable to be able to identify and quantify the vibration paths in the structure. Knowledge of transmission path characteristics enables procedures to be carried out, for example, to reduce vibration levels at points remote from the source, perhaps with the objective of reducing unwanted radiation of sound. One method for obtaining transmission path information is to use the concept of vibrational power transmission. Many machines are installed on frameworks constructed from beamlike members. By using the concept of vibrational power it is possible to compare the contributions from each wave type. Wave motion techniques are used to determine the expressions for vibrational power for each of the various wave types present. The results from the analysis show the amount of vibrational power carried by each wave type and the direction of propagation. Consideration is given to the effect on the vibrational power transmission of introducing misalignment of junctions in previously symmetric framework structures. Unlike other techniques using vibrational power to analyze frameworks, the model keeps the contributions from each of the various wave types separate.

9:55

5aSA3. Vibration transmission via a nonideal beam junction: FEM and analytical combined methods. Igor Grouchetski (Krylov Shipbuilding Res. Inst., 196158 St. Petersburg, Russia) and Eric Rébillard (Chalmers Univ. of Technol., 412 96 Gotheburg, Sweden)

The study is devoted to the calculation of the vibration transmission via a nonideal junction; the example considered is an L-beam system including two steel beams and an elastic block between them. A combination of two calculation methods is used. A finite element method (package ABAQUS) is used to determine the stiffness characteristics of the elastic block having a rectangular parallelpiped shape. These stiffness characteristics are substituted into the calculation for the vibrational behavior of the L-beam system. For such a calculation, the transfer matrix method (TMM) is applied. The experimental approach of the study is divided into two different parts. First the tension-compression and the shear stiffnesses of the elastic block are measured, and, second, the transfer mobility through the complete L-beam system with the rubber elastic block between the two beams is measured. For the two cases, calculated and measured results are compared.

10:15

5aSA4. Measurement of flexural intensity using a dual-mode fiber-optic sensor. Bernard J. Sklanka (Aero/Noise/Propulsion Lab., Boeing Commercial Airplane Group, P.O. Box 3707, Seattle, WA 98124-2207, bernard.j.sklanka@boeing.com), Karl M. Reichard, and Timothy E. McDevitt (Penn State Univ., State College, PA 16804)

An experiment was performed which demonstrated the use of distributed optical fiber sensors to measure power flow in a damped cantilevered beam. Dual-mode optical fiber sensors are an implementation of an optical fiber interferometer in which one fiber acts as both the sensing and reference arms. When bonded to the surface of a structure the sensor provides a distributed measurement of the strain along the length of the sensor. The sensors output is proportional to the integral of the axial component of the strain along the length of the fiber. Two optical fiber sensors were used to measure power flow in a damped cantilevered beam. Finite difference and finite sum approximations were used to determine bending moment and rotational velocity from the output of the fiber sensors. These quantities were then used to calculate the far-field structural intensity. The structural intensity measured using the optical fiber sensors was compared to the structural intensity measured using a two-accelerometer method.

10:35

5aSA5. The condition for beam-bending modes to dominate in the vibro-acoustic behavior of a circular cylindrical shell. Chong Wang and Joseph C. S. Lai (Acoust. and Vib. Unit, Univ. College, Univ. of New South Wales, Australian Defence Force Acad., Canberra, ACT 2600, Australia, c-wang@adfa.oz.au)

The acoustic radiation from circular cylindrical shells is of fundamental and applied interest primarily because cylindrical shells are widely used in industries. However, according to previous studies, only a few special cases, for example, a cylindrical shell under the assumption of beam bending, can be solved analytically. Obviously, in practice, the vibration behavior of a cylindrical shell may not be assumed to be beam bending in the whole frequency range of interest. This is because the vibration behavior of a cylindrical shell changes with frequency. Therefore, it is important to determine the condition under which a circular cylindrical shell would behave like beam bending so that the analytical solution could be used correctly. In this paper, an analysis of the basic vibration behavior of circular cylindrical shells using the Love's equations shows that for an infinite length cylindrical shell, the beam-bending wave would always be propagating with other flexural waves of different circumferential mode numbers. However, for a finite-length circular cylindrical shell, below the cutoff frequency of $n=2$ mode, beam-bending modes would dominate the vibration response so that it can be treated as a beam with reasonable accuracy. This analytical result is supported by calculations obtained using the boundary element method.

10:55

5aSA6. The effect of rib resonances on the vibration and wave scattering of a ribbed cylindrical shell. Martin H. Marcus (Naval Res. Lab., 4555 Overlook Ave., S.W., Washington, DC 20375, marcus@astro.nrl.navy.mil)

Wave phenomena have previously been observed for cylindrical shells with periodically spaced ribs. The understanding of these waves, such as Bragg and Bloch waves, requires knowledge of the periodicity of the ribs and how they interact with each other. In this paper, the understanding of the individual rib resonances gives explanation to the resonance bands that are seen in both experiment and numerical simulations. For these resonances, periodicity becomes less important in computing the frequencies of resonances on the shell and in the far field. However, lack of periodicity in rib spacing and size affects the width and strength of these resonance bands. [Work supported by ONR.]

11:15

5aSA7. A simple model for structure-borne sound transmission based on direct wave and first reflections. Jian Liang and Bjorn A. T. Petersson (Dept. AAETS, Loughborough Univ., Loughborough, Leicestershire LE11 3TU, England)

In finite structures, the total wave field is composed of the direct wave, the first reflections as well as multiple, higher-order reflections. When the structural damping is significant and the divergence of the wave is considered, it is reasonable to assume that those waves reflected more than once only contribute marginally to the total field, and thus can be ignored to a first approximation. This suggests a simple description of the field which takes the direct wave and the first reflections into account. The semi-infinite system realizes an example of such models. The question addressed herein is the performance of those simple descriptions. This contribution first presents an investigation of the performance for one-dimensional systems. It is shown that the corresponding semi-infinite system can be used to: (1) predict the nodal lines in the spatial domain and the locations of deep troughs in the frequency domain; (2) obtain accurate motion transmissibilities. Second, the extrapolation to two-dimensional systems is presented, accompanied by results for platelike structures.

11:35

5aSA8. Estimating the vibration energy of an elastic structure via the input impedance. Yuri I. Bobrovnitskii (Lab. of Structural Acoust., Mech. Eng. Res. Inst. of the Russian Acad. of Sci., M. Kharitonievsky Str., 4, 101830 Moscow, Russia, imash@space.ru)

The following problem is posed and solved: given an elastic structure vibrating under the action of a harmonic point force, find the time-average potential and kinetic energy, loss factor, and other energy characteristics of the structure if known are only the complex amplitudes of the force and velocity response at the driving point, f and v. The structure is assumed to be linear, with viscous and hysteric damping not necessarily proportional. It is shown that the problem has a mathematically exact solution only for lossless systems. For structures with losses, there exist rather accurate estimates valid at low and middle frequencies. The main result is the equation relating the total vibration energy of the structure to the derivative of its input impedance, f/v, or mobility, v/f, with respect to frequency. Computer simulation examples with typical structures (rod, plate) illustrate the accuracy and the range of validity of the estimates. The results presented allow one to obtain the energy characteristics of a forced vibrating structure by the most economic way: without measuring or computing the response all over the structure and using only the data measured at one singular point.

11:55

5aSA9. Improving convergence in scattering problems by smoothing discrete constraints using an approximate convolution with a smoothed composite Green's function. Rickard C. Loftman and Donald B. Bliss (Mech. Eng. and Mater. Sci., Duke Univ., Box 90300, Durham, NC 27708)

A method for efficiently treating discrete structural constraints using an approximate Green's function approach to smooth the constraint influence is presented. A method called analytical/numerical matching is used to formulate a composite Green's function that captures the local high-resolution content associated with the discrete force in a separate analytic local solution. This local solution in turn implies a smooth replacement for the original discrete force. This replacement is approximately integrated following the Green's function formalism modified by a Taylor expansion of the forcing strength function. A smooth replacement for the discrete constraint results. This new smoothed problem converges more rapidly than the original discrete one. However, the difference between the smooth solution and the original is retained in the analytic local solution. The result is a composite solution of the original system that is more accurate for a given computational resolution compared with treating the discrete constraint influence directly. In the present example, a modal method is used. Therefore, a given computational resolution is measured by number of modes calculated. The method is demonstrated in the case of beam incident scattering from an infinite thin elastic shell periodically constrained by rings. [Work supported by ONR.]

FRIDAY MORNING, 26 JUNE 1998 FIFTH AVENUE ROOM (W), 8:15 TO 10:45 A.M.

Session 5aSC

Speech Communication: Production and Other Topics

Ingo R. Titze, Chair
Speech Pathology and Audiology, University of Iowa, 330 WJSHC, Iowa City, Iowa 52242-1012

Contributed Papers

8:15

5aSC1. Vocal tract shape estimation using three noninvasive transducers. Russel Long, Brad Story, and Ingo Titze (WJ Gould Voice Res. Ctr., Denver Ctr. for the Performing Arts, 1245 Champa St., Denver, CO 80204)

Sentence-level speech simulation requires a description of vocal tract shape parameters as a function of time. Vowel-like configurations can often be derived from a microphone signal; however, the influence of consonants may be more easily detected with the use of additional signals. The electroglottograph (EGG) and an articulatory transducer [McGarr and Lofqvist, J. Acoust. Soc. Am. **72**, 34–42 (1982)] are employed to increase the accuracy of vocal tract shape estimation. The microphone signal is analyzed on a frame-by-frame basis for spectral features, energy, and zero-crossing rate. The EGG provides information regarding voicing onset and offset, and the articulatory signal is used to indicate closure and lip rounding. These data are combined to select an appropriate area function for each frame of the sentence. The resulting temporal map of area functions is interpolated and sent to a wave reflection simulator to reconstruct the speech. [Work supported by NIH Grant No. DCO2532.]

8:30

5aSC2. Simulation of sentence-level speech based on measured vocal tract area functions. Brad Story, Ingo Titze, and Russel Long (WJ Gould Voice Res. Ctr., Denver Ctr. for the Performing Arts, 1245 Champa St., Denver, CO 80204)

An inventory of vocal tract area functions, acquired from magnetic resonance imaging experiments, is used as the basis for creating simulations of sentence-level speech. A first approach is a ''brute-force'' method in which a phonetic transcription of a sentence is time aligned with the recorded acoustic signal. Each element of the transcription is assigned one of the stored vocal tract area functions in the acquired inventory, thus serving as a ''target'' through which the time-varying area function should pass. This method has produced intelligible speech but maintains a machinelike quality. A second approach uses a parametric representation of the MRI-acquired area function inventory based on a principal components analysis to generate area functions not in the original inventory. The first three formants of each generated area function were computed and stored in a database, thus creating a mapping between formant frequencies and vocal tract parameters (area functions). Except for the edge of the formant space, the mapping between formants and area functions is one-to-one. This method produces connected speech with natural sounding vowels when only a microphone signal is used but requires additional signals to derive the consonantal contribution to the vocal tract shape. [Supported by NIH grant RO1 DC02532.]

8:45

5aSC3. Effects of naturally occurring external loads on jaw movements during speech. Douglas M. Shiller, Paul L. Gribble, and David J. Ostry (McGill Univ., Montreal, QC, Canada, ostry@motion.psych.mcgill.ca)

The achieved positions of speech articulators are determined by a combination of intrinsic forces due to muscles and extrinsic forces due to gravity and to head and body motion. If not compensated for, changes in these extrinsic forces will affect articulator positions in behaviors such as speech. In this paper, the extent to which speakers adjust control signals to jaw muscles to compensate for the effects on the jaw due to changes in speaker orientation relative to gravity is assessed. Whether control signals are adjusted to compensate for loads arising during locomotion is also assessed. In one study, subjects produce speech utterances where head and body orientation relative to gravity are varied. In another study, subjects produce speech sequences during locomotion such that the timing of the step cycle applies the maximum load force during either the consonant related (jaw elevated) or vowel related (jaw lowered) phase of movement. The empirically observed patterns of jaw motion are compared with predictions of a model of jaw and hyoid motion which includes muscle properties, dynamics, and modeled neural signals. The hypothesis that during speech, control signals to jaw muscles remain unchanged with variation in these extrinsic loads is specifically tested.

9:00

5aSC4. Contrasting chest and falsettolike vibration patterns of the vocal folds. David A. Berry and Douglas W. Montequin (Univ. of Iowa, Dept. of Speech Pathol. and Audiol., 334 WJSHC, Iowa City, IA 52242)

Using excised larynges, a hemilarynx methodology was used in order to view vocal fold vibrations from a medial aspect. This was accomplished by removing one fold, inserting a glass plate at the glottal midline, and

mounting the remaining fold against the glass plate. A stereoscopic imaging system was used to perform 3D tracking of small markers (microsutures) placed on the folds. Chest and falsettolike vibrations patterns were imaged and quantified. These two vibrations patterns will be contrasted with each other, and discussed in terms of previous qualitative comparisons of these phonation types in the literature. [Work supported by Grant No. R29 DC03072 from the National Institute on Deafness and Other Communication Disorders.]

9:15–9:30 Break

9:30

5aSC5. The effect of shear thinning in vocal fold tissues—or why we can phonate over two octaves in pitch. Ingo Titze and Roger Chan (Dept. of Speech Pathol. and Audiol. and The Natl. Ctr. for Voice and Speech, Univ. of Iowa, Iowa City, IA 52242)

The shear modulus and viscosity of vocal fold tissues was measured with a rheometer in the frequency range of 1–30 Hz. Extrapolation to phonation frequencies was possible. As in most biological tissues (and for polymers in general), the shear modulus was found to be roughly constant and the viscosity varied inversely with frequency. This makes the damping ratio only a weak function of frequency in phonation. Since self-oscillation of the vocal folds is critically dependent on this damping ratio, it is reasoned that large *F*0 ranges (on the order of two octaves) are possible only because of this decrease in viscosity, an effect known as shear thinning. A molecular interpretation is given for this effect. [Work supported by NIDCD, Grant No. P60 DC00976.]

9:45

5aSC6. Pressure-flow relationship in a biophysical model of phonation. Fariborz Alipour and Ingo Titze (Dept. of Speech Pathol. and Audiol., The Univ. of Iowa, Iowa City, IA 52242)

Simulation data are presented from a computer model that combines vocal fold tissue mechanics, laryngeal aerodynamics, and vocal tract acoustics. These simulations are based upon the laws of physics that govern airflow, vibration of the vocal folds, and wave propagation in the vocal tract. The tissue mechanics was modeled with the finite element method with 100 nodes and 166 elements in each layer of the 15-layer model. The laryngeal aerodynamics was modeled by the solution of two-dimensional unsteady Navier–Stokes equations with a finite volume method with a 64 $\times$82 nonuniform staggered grid. The model has been used to demonstrate self-oscillatory characteristics of the vocal folds and the flow pattern in the larynx. In this study, simulations were performed for lung pressure ranging from 4 to 24 cm of water. Adduction control was simulated with activation of thyroarytenoid (TA) muscle that caused increased tension and bulging of the vocal folds. Results are reported for glottal waveform, nodal coordinates, and velocity distributions in the larynx. Preliminary data suggest that the relationship between mean pressure and mean flow is almost linear. Frequency and amplitude contours of the model as functions of the lung's pressure and TA activation level reveal their increasing pattern with these parameters.

10:00

5aSC7. Perception of coarticulated vowels by prelingual infants: Formant transitions specify vowel identity. Ocke-Schwen Bohn (Engelsk Inst., Aarhus Univ., DK-8000 Aarhus C, Denmark, engosb@hum.aau.dk) and Linda Polka (School of Commun. Sci. and Disord., McGill Univ., Montreal, Canada)

German-learning infants were tested for discrimination of four German vowel contrasts in the Silent Center paradigm to examine which acoustic properties of coarticulated vowels (target-spectral, dynamic-spectral) define vowel identity in prelingual infants. Acoustic analyses of naturally produced original syllables (/dVt/ with V=/i, e, I, E, o, U/) revealed no vowel inherent spectral change but only formant movement associated with the gestures for the initial and final consonants. The original syllables were electronically modified to obtain test syllables which contained target-spectral information (the vowel nucleus only), only initial or only final formant transitions, or initial and final transitions in their appropriate temporal relationship, with the vowel nucleus attenuated to silence. Four groups of ten infants each (aged 7–11 months) were tested in the conditioned headturn procedure for discrimination of the test syllables. Each group was tested on one of four contrasts (/i/–/e/, /e/–/I/, /E/–/I/, /o/–/U/) after they had discriminated that contrast successfully in the unmodified condition. Results indicate that target spectral information is not needed to discriminate native contrasts; rather, dynamic spectral information spread over syllable onsets and offsets is sufficient to specify vowel identity in infants. [Work supported by grants of the Deutsche Forschungsgemeinschaft to O.-S. Bohn.]

10:15

5aSC8. Some observations of the tense–lax distinction theory with reference to English and Korean. Dae W. Kim (Dept. of English, College of Education, Pusan Natl. Univ., Gumjong-Gu, Pusan 609-735, South Korea, dwokim@hyowon.cc.pusan.ac.kr)

A series of experiments with 13 subjects in total and with /VCV/, /CVC/, /CVCVCV/, and /CVCVC/ words, using various techniques, was undertaken to verify the hypothesis that the speaker may select at least one of muscles involved in the articulation of a phoneme so that either timing or amplitude variable and/or both of the selected muscle(s) could distinguish /p/ from /b/ in English and /ph, p′/ from /p/ in Korean. The English /p/ and the Korean aspirated /ph/ and unaspirated /p′/ were tense, and the English /b/ and the Korean unaspirated /p/ lax. Findings: (1) The EMG data, obtained from the *orbicularis oris* superior muscle and the depressor *anguli oris* muscle, supported the tense–lax opposition, except for one English subject in stressed syllables; (2) The leveled peak intraoral air-pressure, an output of the synchronized respiratory muscle activities during bilabial stops, backed up the tense–lax distinction, except for one English subject in unstressed syllables, and there was intermuscle compensation between the labial muscle activities and the respiratory muscle activities; (3) The data of the maximum lingual contact in alveolar stops also supported the tense–lax distinction theory. Thus it can be said that the hypothesis has been verified in both languages.

10:30

5aSC9. Speech intelligibility is highly tolerant of cross-channel spectral asynchrony. Steven Greenberg (Intl. Comput. Sci. Inst., 1947 Center St., Berkeley, CA 94704, steveng@icsi.berkeley.edu) and Takayuki Arai (Intl. Comput. Sci. Inst., Berkeley, CA)

A detailed auditory analysis of the short-term acoustic spectrum is generally considered essential for understanding spoken language. This assumption is called into question by the results of an experiment in which the spectrum of spoken sentences (from the TIMIT corpus) was partitioned into quarter-octave channels and the onset of each channel shifted in time relative to the others so as to desynchronize spectral information across the frequency plane. Intelligibility of sentential material (as measured in terms of word accuracy) is unaffected by a (maximum) onset jitter of 80 ms or less and remains high (>75%) even for jitter intervals of 140 ms. Only when the jitter imposed across channels exceeds 220 ms does intelligibility fall below 50%. These results imply that the cues required to understand spoken language are not optimally specified in the short-term spectral domain, but may rather be based on some other set of representational cues such as the modulation spectrogram [S. Greenberg and B. Kingsbury, Proc. IEEE ICASSP (1997), pp. 1647–1650]. Consistent with this hypothesis is the fact that intelligibility (as a function of onset-jitter interval) is highly correlated with the magnitude of the modulation spectrum between 3 and 8 Hz.

Session 5aSP

Signal Processing in Acoustics: Signal Processing for Medical Ultrasound I

Stergios Stergiopoulos, Chair
DCIEM, P.O. Box 2000, North York, Ontario M3M 3B9, Canada

Chair's Introduction—9:15

Invited Papers

9:20

5aSP1. Beyond current medical ultrasonic imaging: Opportunities for advanced signal processing. James G. Miller (Dept. of Phys., Washington Univ., St. Louis, MO 63105)

This keynote address will illustrate some aspects of current medical ultrasonic imaging and attempt to identify areas in which advanced signal processing may be able to contribute to the enhancement and extension of clinical ultrasound. Examples will be drawn primarily from echocardiography, in part because of challenges specifically associated with imaging the beating heart. The presentation will address the goals and needs of the (medical) user in order to avoid approaches which are technically elegant but clinically irrelevant. One feature peculiar to medical imaging, which limits the effective resolution to far less than the expected theoretical limit for the few hundred micrometer wavelength imaging systems in current use, is the subtle variation of elastic properties of soft tissue and the corresponding local variation in the speed of sound. Both random and systematic (for example, anisotropic) variations contribute to degradation in image quality, including the effect known as speckle. Opportunities for advanced signal processing may include not only approaches designed to enhance image resolution but also contributions to areas such as contrast agents and tissue characterization. The talk is designed to provide a broad overview which might serve as a common reference point for subsequent presentations.

10:00

5aSP2. The evolution of medical ultrasonic imaging systems. John M. Reid (Consultant 16711 254 Ave. SE, Issaquah, WA 98027)

A limit to effective soft tissue imaging is set by the acoustic properties of tissue. The attenuation is roughly proportional, and the size of a resolution element inversely proportional to frequency. Thus the number of resolution elements per image is fixed. Accordingly, acoustic frequencies from about 1 to 50 MHz are required. Effective imaging has required the development and application of innovative high-frequency hardware; particularly the transducers. These have progressed from single element types to 512 element arrays in production, with two-dimensional arrays having 1282 elements in development. The electronics now require many more channels with a dynamic range exceeding 100 dB and, with the introduction of digital methods, even multiple very fast A/D converters in the beamformer. Display devices have gone from WWII CRT's with 200 spots per radius to digital image stores on CRT's with thousands of spots available. These can now do image processing with software. The earlier systems survive in some applications that require them. The single element types are still used in high-frequency catheters, for example. The wideband transducers have opened up new methods, enhanced Doppler, contrast and tissue imaging that use a wider range of frequencies than previously possible. The single element types are still used in high-frequency catheters, for example.

10:30–10:40 Break

10:40

5aSP3. 3-D ultrasound imaging of the prostate. Aaron Fenster and Donal Downey (Robarts Res. Inst., London, ON N6A 5K8, Canada)

An important aspect that needs improvement in medical ultrasound systems is related to the 2-D imaging of the prostate. 2-D viewing of 3-D anatomy limits our ability to quantify and visualize prostate disease and is partly responsible for the reported variabilities. This occurs because: (i) the diagnostician must integrate multiple 2-D images in his mind during the procedure, leading to inefficiency and variability; (ii) The 2-D ultrasound images represent a thin plane at some arbitrary angle in the body, making it difficult to localize the image plane. To overcome these difficulties, we have developed a 3-D ultrasound system to image the prostate. Our 3-D ultrasound imaging system consists of: a conventional ultrasound machine and transducer; a custom-built assembly for rotating the probe under microcomputer control; a microcomputer with a video grabber; and software to reconstruct and display 3-D images. A typical scan of 200 2-D B-mode images takes only 13 s, and the reconstruction less than 1 s. This paper will detail the 3-D imaging approach and its use for imaging the prostate in 3-D. Various applications will be discussed related to prostate cancer diagnosis, prostate volume measurements, and 3-D ultrasound-guided therapeutic procedures such as cryosurgery.

11:10

5aSP4. The analysis and classification of small-scale tissue structures using the generalized spectrum. Kevin D. Donohue (Dept. of Elec. Eng., Univ. of Kentucky, Lexington, KY 40506), Flemming Forsberg, and Ethan J. Halpern (Thomas Jefferson Univ. Hospital, Philadelphia, PA 19107)

Conventional ultrasonic imaging systems primarily use backscattered signals for creating qualitative images that reveal large-scale structures, such as tissue boundaries. Efforts to extract additional quantitative information have resulted in limited success. The nonstationarities of the scatterers comprising biological tissue often violate conditions for applying common signal characterization and estimation methods. In addition, ultrasonic tissue properties (such as attenuation, velocity, scatterer density, size, and structure) ambiguously encode information into the backscattered signal, making it difficult or impossible to extract and quantify a single property. This paper presents the generalized spectrum (GS) as a method for analyzing and quantifying the properties of small-scale resolvable structures (on the order of 1 to 4 mm), which result from tissue structures such as lobules, ducts, and vessels. The GS extends the capabilities of power spectral density to include meaningful phase information that results from small-scale structure. The relevant properties of the GS include its ability to reduce the effects' diffuse scatterers (speckle), and permit normalization schemes that significantly limit the effects of system response and attenuation from the overlying tissue. An implementation of the GS in an analysis and classification of normal and metastic liver tissues is also described.

11:40

5aSP5. A Wold decomposition-based autonomous system for detecting breast lesions in ultrasound images of the breast. Georgia Georgiou and Fernand S. Cohen (Imaging and Comput. Vision Ctr., Dept. of Elec. and Comput. Eng., Philadelphia, PA 19104)

This paper presents an autonomous system for detecting lesions in the breast. The Wold decomposition algorithm described is used to decompose the RF echo of the breast into its diffuse and coherent components. The coherent component is modeled as a periodic sequence and the diffuse component is modeled as an autoregressive time series of low order. The parameters of the model are estimated from selected regions of the RF image and used as detection features. The database of images that was used contained 370 B-scan images from 52 patients, obtained in the Radiology department of the Thomas Jefferson Hospital. The pathologies of interest are carcinoma fibrocystic and stromal fibrosis disease and fibroadenoma. Empirical ROC techniques were used to evaluate the detection rate on single parameters of the model, such as the residual error variance and the autoregressive parameters of the diffuse component of the RF echo. The area under the empirical ROC curve for detecting lesion regions versus normal RF regions is 0.901. The area under the ROC curve for detecting carcinoma versus normal regions is 0.904. The corresponding areas for normal regions versus stromal fibrosis/fibrocystic regions and fibroadenoma regions are 0.942 and 0.899, respectively.

Contributed Papers

12:10

5aSP6. Tracking and correcting for organ motion artifacts in ultrasound tomography systems. Amar C. Dhanantwari (Dept. of Elec. and Comput. Eng., Univ. of Western Ontario, London, ON N6A 5B9, Canada) and Stergios Stergiopoulos (Defence and Civil Inst. of Environ. Medicine, P.O. Box 2000, North York, ON M3M 3B9, Canada)

Motion artifacts have been identified as a problem in medical tomography systems. This problem, however, is well known in other types of real-time imaging systems such as radar satellites and sonars. In this case it has been found that the application of an overlap processing scheme [J. Acoust. Soc. Am. **86**, 158–171 (1989)] increases the resolution of a phased array imaging system and corrects for motion artifacts as well. Reported results have shown that the problem of correcting motion artifacts in synthetic aperture applications is centered on the estimation of a phase correction factor. This correction factor is then used to compensate for the temporal phase differences between sequential sensor-array measurements in order to synthesize the spatial information coherently. This paper describes an approach which tracks the organ motion and allows the artifacts to be isolated in ultrasound tomography systems. Two sources are utilized so that two sets of projections are generated that are identical in space but separated in time. Then the spatial overlap correlator processing scheme is used to synthesize the 2-D projection data coherently. This provides the desired phase correction factor, which compensates for the phase fluctuations caused by the subject's organ-motion effects and tracks the organ motion.

12:30

5aSP7. Wavefront distortion measurements in the human breast. R. C. Gauss, G. E. Trahey (Dept. of Biomed. Eng., Duke Univ., 136 Eng. Bldg., Durham, NC 27708, rgauss@acpub.duke.edu), and M. S. Soo (Duke Univ. Medical Ctr., Durham, NC 27710)

Published wavefront distortion (phase aberration) measurements for the human breast have been inconsistent, ranging from mild phase aberrations (8 ns rms) to severe phase aberrations (67 ns rms). These measurements are required to specify arrays and assess the potentials of and appropriate algorithms for adaptive imaging. They require high interelement uniformity so that receive signal variations between elements can be attributed to wave propagation effects. Array elements must be small to minimize the integration of the arriving wavefront across the face of the element. The array aperture must be large enough to allow target visualization and to provide high correlations between neighboring elements for pulse-echo measurements. An apparatus to make concurrent pulse-echo and pitch-catch (through transmission) waveform measurements in a clinical setting has been developed. The breast is stabilized between opposing transducers with light compression. After fixing the position of the transducers, pulse-echo and pitch-catch snapshots are captured sequentially with tissue and reference phantom targets. Data were collected from four views of the left breast in 12 volunteers. The rms phase error was significantly smaller for pulse-echo measurements (25 ± 14 ns rms) than for pitch-catch measurements (60 ± 23 ns rms). [Work supported by NIH.]

12:45–1:15 Panel Discussion

Session 5aUW

Underwater Acoustics: Bistatic Scattering from Seabed and Surface: Models and Measurements I

Paul C. Hines, Chair
Defence Research Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada

Chair's Introduction—9:15

Invited Papers

9:20

5aUW1. A practical model for high-frequency seabed bistatic scattering strength. Darrell R. Jackson (Appl. Phys. Lab., College of Ocean and Fishery Sci., Univ. of Washington, Seattle, WA 98105) and Anatoliy N. Ivakin (Andreev Acoust. Inst., Shvernika 4, Moscow 117036, Russia)

The authors have previously reported a model for bistatic scattering by elastic seabeds. This model is extended here to cases of larger roughness by replacing the rough-interface perturbation approximation with the small-slope approximation [Yang and Broschat, J. Acoust. Soc. Am. **96**, 1796–1804 (1994)]. Consistency with older models is shown by comparisons of calculations for silt and sand examples. Correlations between volume fluctuations of different types (e.g., density and compressional wave speed) are considered, and it is shown that the effects of such correlations are more readily observed in bistatic scattering measurements as opposed to backscattering measurements. Finally, the importance of elastic effects in scattering by rocky seabeds is illustrated. [Work supported by ONR.]

9:40

5aUW2. High-frequency bistatic scattering from sediments—experiments and modeling. Kevin Williams and Dajun Tang (Appl. Phys. Lab., College of Ocean and Fishery Sci., Univ. of Washington, Seattle, WA 98105, williams@apl.washington.edu)

Typically, for frequencies in the 10–100 kHz range, the interaction of sound with the ocean bottom is a scattering process. Use of bistatic geometries can give insight into the scattering process not possible via traditional monostatic experiments. The trade-off is that such experiments are more difficult and more data are needed to get statistically meaningful results to compare with models. During the last 5 years, the Coastal Benthic Boundary Layer (CBBL) program allowed the opportunity to develop and fine-tune one method of obtaining bistatic data. That evolution will be briefly described and the data from three sites (one mud, one sand, one sand-silt-clay) shown. The sand and sand-silt-clay site data are compared with a model based on surface roughness and volume perturbation scattering. The mud site had methane bubbles embedded within it and the data are compared with a model based on bubble scattering. The physical insight derivable from the model/data comparison in the mud case will be discussed. [Work supported by ONR.]

Contributed Papers

10:00

5aUW3. Models of volume and roughness scattering in stratified seabeds. Anatoliy N. Ivakin (Andreev Acoust. Inst., Shvernika 4, Moscow 117036, Russia, aniva@glasnet.ru)

At high frequencies, two mechanisms of seabed scattering are usually taken into account which are due to roughness of the seabed surface (sediment–water interface) and volume inhomogeneities of near-surface sediment. The mean parameters of bottom medium are usually considered as independent from spatial coordinates. This corresponds to the assumption that sound penetration into the sediment is sufficiently small because of strong absorption. In this paper, more general models are considered where seabeds are assumed as statistically stratified (plane-layered on the average) with the mean parameters taken as various functions of depth. In particular, they can be step functions corresponding to different types of sediments in different layers. Another type of stratification corresponds to continuous depth-dependences and involve gradients of mean parameters. Effects of stratification are shown to be important at lower frequences due to scattering from internal interfaces and volume inhomogeneities of deep layers as well as the influence of regular refraction and reflection within seabed medium. As numerical examples, angular dependencies of the scattering strength in both monostatic and bistatic cases at different frequencies are calculated for seabeds of various types. Comparison with available data is presented and discussed. [Work supported by ONR.]

10:12

5aUW4. Bottom volume scattering—Modeling and data analysis. Dan Li, George V. Frisk (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Dajun Tang (Univ. of Washington, Seattle, WA 98105)

As part of the Acoustic Reverberation Special Research Program (ARSRP), low-frequency bottom scattering from a deep ocean sediment pond was measured using an omnidirectional source and a vertical line-receiving array deployed near the bottom. Sediment volume heterogeneities were found to be the major contributor to the measured scattered fields [Tang *et al.*, J. Acoust. Soc. Am. **98**, 508–516 (1995)]. Taking advantage of the experimental geometry, a model is developed which can generate realizations of three-dimensional random inhomogeneities in sediments and can simulate scattered fields based on these realizations using a perturbation approach. In this model, the broadband propagator (the Green's function) is obtained using an exact numerical method.

Therefore, the model can handle any layered sediment environments with scatterers distributed within any of the layers. The model predictions are compared to the experimental data across the available frequency band, and the results are favorable. [Work supported by ONR.]

10:24

5aUW5. High-frequency bottom backscattering from a homogeneous seabed. Kevin B. Briggs (Marine Geosciences Div., Naval Res. Lab., Stennis Space Center, MS 39529, kbriggs@zephyr.nrlssc.navy.mil) and Steve Stanic (Acoust. Div., Naval Res. Lab., Stennis Space Center, MS 39529)

High-frequency acoustic backscattering was measured from a sandy site located in very shallow water (8 m) off Tirrenia, Italy, at a range of frequencies and grazing angles. Concomitant measurements of the seabed indicated that the very fine sand was well sorted, relatively uniform with respect to vertical and horizontal distribution of compressional wave velocity and attenuation, and depleted with respect to biological activity. Bottom scattering strengths were very low and variability of sediment parameters was exceptionally low. Coefficients of variation determined for sediment compressional wave velocity and attenuation were 0.29% and 7.2%, respectively. These data compare with previously recorded minima of 0.30% and 14.8% for the same respective parameters, in nearshore, mobile fine sands from the Gulf of Mexico and the Pacific Ocean. Lack of variability in sediment sound velocity and density concurrent with low acoustic bottom scattering strength measurements would endorse recent modeling approaches that predict scattering from sediment impedance fluctuations. Acoustic data are compared with predictions from the composite roughness model (modified for inclusion of sediment volume scattering parameters) at various frequencies and grazing angles in order to test the model's response at minimal perturbation of seabed properties.

10:36

5aUW6. A porous medium with porosity variations. Nicholas P. Chotiros (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029)

The BiotUs equations of acoustic propagation in porous media have no inherent scattering mechanism, because the pores are modeled as smooth-sided tubular passages with constant cross section. They predict two compressional and one shear wave that propagate independently. Real porous media usually have pores with varying cross section, which may be considered as variations in porosity on a microscopic scale. Although any single variation on such a scale would have a negligible effect on the acoustics of the medium, the cumulative effect of a field of such variations is unknown. A one-dimensional structure with varying porosity was modeled numerically to study its properities as an acoustic medium. It was found that, not only does it possess an inherent volume scattering strength, but it also allows energy transfer between the three types of waves as they propagate. [Work supported by ONR 321 OA.]

10:48–11:03 Break

11:03

5aUW7. A comparison of bistatic scattering from two geologically distinct mid-ocean ridges. Chin Swee Chia (MIT, 77 Massachusetts Ave., Cambridge, MA 02139) and Laurie T. Fialkowski (Naval Res. Lab., Washington, DC 20375)

A comparison is made of bistatic scattering from two distinct deep-ocean ridges. The ridges are located along a segment valley on the western flank of the Mid-Atlantic Ridge (MAR). The acoustic data sets were acquired during the Main Acoustics Experiment of the Acoustic Reverberation Special Research Program. Detailed analysis of bistatic scattering from one of these ridges, named B′, has been previously presented by some of the authors. The purpose of this paper is to present a similar analysis for the second ridge, C′, for comparative purposes. This comparison is important because the ridges span the two geologically distinct classes found in the MAR. There is a possibility, therefore, that their scattering characteristics may be extrapolated to ridges throughout the MAR. For example, B′ is an ''outside corner'' and therefore is composed of highly lineated scarps and terraces, while C′ is an ''inside corner'' that is characteristically amorphous with irregular large-throw normal faults. While major scarps on these ridges can be easily resolved by the towed-array system, smaller-scale anomalies such as gullies and trellises are often underresolved by nearly an order of magnitude. These anomalies are found to introduce large variances in the bistatic scattering characteristics of a given scarp.

11:15

5aUW8. Multistatic scattering from anisotropically rough interfaces in horizontally stratified waveguides. Jaiyong Lee (MIT, Cambridge, MA 02139) and Henrik Schmidt (SACLANT Undersea Res. Ctr., La Spezia, Italy)

A model of multistatic scattering and reverberation from an anisotropically rough seabed in shallow water has been developed by combining a scattering theory based the method of small perturbations [W. A. Kuperman and H. Schmidt, J. Acoust. Soc. Am. **79**, 1767–1777 (1986)] with a seismo-acoustic propagation model [H. Schmidt and J. Glattetre, J. Acoust. Soc. Am. **78**, 2105–2114 (1985)]. This new computational model provides full 3-D wave theory simulations of multistatic rough bottom reverberation in horizontally stratified waveguides. The seabed and sub-bottom roughnesses are represented by anisotropic roughness patches covering the actual sonar footprint. The scattering theory convolves the incident sonar field with the roughness to provide a distribution of virtual source panels that are then fed into the propagation model to provide a numerically efficient azimuthal Fourier series of wave number integrals for evaluating the 3-D reverberant field. The efficiency of the new hybrid model enables the modeling of 3-D field statistics by *Monte Carlo* simulation, as well as time-domain field simulations by Fourier synthesis. The model has been verified by comparison with analytical solutions for the simple problem of determining the 3-D multistatic scattering strength of a randomly rough interface separating two fluid half-spaces. Here the new model is used to investigate the dependence of the multistatic reverberant field on environmental parameters such as anisotropic roughness statistics, and bottom stratification and composition. [Work supported by ONR.]

11:27

5aUW9. Integral equation method for bistatic volume scattering from the seafloor. Christopher D. Jones and Darrell R. Jackson (Appl. Phys. Lab., College of Ocean and Fishery Sci., Univ. of Washington, Seattle, WA 98105)

In most theoretical developments of sediment volume scattering, the medium is assumed to have weak fluctuations in sound speed and density and is assumed to be statistically homogeneous. An exact integral equation method has been developed and is used to test these assumptions. Small perturbation theory results for single frequency bistatic scattering from the seafloor are compared with Monte Carlo simulations. Half-space and multiple scattering effects not included in most previous treatments are observed.

11:39

5aUW10. Time/frequency spread for bistatic geometries. R. Lee Culver and Christopher J. Link (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804)

The overall goal of this work is to develop a functional characterization of, and empirical models for, high-frequency, broadband acoustic propagation through a shallow ocean channel. One important aspect of this characterization is the time/frequency spread, which quantifies how much signals spread in time and frequency as they propagate. Because time and frequency spread are random quantities which vary from pulse to pulse, time/frequency spread must be characterized using a second-order statistic, e.g., the scattering function, which is the expected value of the mean-square spreading function. The scattering function has been measured in the following ocean environments: sandy gravel bottom, isovelocity sound

speed profile (SSP); rough rock bottom, slightly downward-refracting SSP; and mud bottom, upward-refracting SSP. A statistical model for the scattering function is being developed for these environments based upon radial basis function decompositions of the scattering function estimates. The ultimate use of a scattering function model will be with simulation to assess and optimize acoustic receiver performance. [Work is sponsored by the Office of Naval Research Ocean Acoustics Program.]

11:51

5aUW11. Sound waves attenuation in shallow water with rough boundaries. Boris G. Katsnelson, Sergey A. Pereselkov (Dept. of Phys., Voronezh Univ., 1 University Sq., Voronezh 394693 Voronezh, Russia), Venedict M. Kuz'kin, and Valery G. Petnikov (General Phys. Inst., 137333 Moscow, Russia)

Here the object is to estimate the effect of scattering by randomly rough seabed and surface waves on attenuation of low-frequency sound waves in shallow water. The experiment was pursued with CW signals (100 and 230 Hz) in Barents Sea. The transmission loss and interference pattern of a sound field were measured within 100 km of the acoustic sources. Mesoscale relief of the bottom with spatial resolution 10 m and wind parameters were recorded as well. The experimental results (the average sound intensity) were compared with theoretical ones calculated with and without taking into account seabed and surface roughness. The calculation was carried out on the base of modal theory of sound scattering by rough boundaries in a waveguide. It was demonstrated that the random roughness substantially affects the transmission loss for long distances. In this paper the method is suggested for bottom parameter estimation using the statistical attributes of sound field in shallow water with random rough boundaries as well. [The work was supported by Russian Foundation of Basic Research (Grant 96-02-17194).]

12:03

5aUW12. Modal theory of sound propagation in a random irregular shallow-water waveguide (numerical calculation and diffusion approach). Boris G. Katsnelson and Sergey A. Pereselkov (Dept. of Phys., Voronezh Univ., 1 University Sq., Voronezh 394693 Voronezh, Russia)

Sound field changing due to scattering by random irregularities in shallow water is considered. Shallow water is a randomly inhomogeneous media bounded by absorbing bottom and surface, which are randomly irregular (rough) as well. Longitudinal scales of all types of irregularities are greater than cross scales. Sound field is the sum of modes with random amplitudes. A system of stochastic differential equations for modal amplitudes is derived. Asymptotic approximation is used to obtain equations for statistical moments of modal amplitudes. The behavior of average modal intensity is described by diffusion equation. Comparative analysis of scattering and absorbing effects on sound propagation is carried out. Scattering leads to two processes: intensity transformation from a coherent (average) field component into a noncoherent (random) one, and intensity redistribution between modes. The next problem is interference structure alteration due to scattering and absorption. This alteration leads in turn to space redistribution of the field intensity. It is shown that for the ranges from the source, a much greater than ''automodel range'' situation takes place, and that depth dependence of intensity is constant in range. It is the result of equilibrium between two phenomena: modal intensity absorbing and modal intensity redistribution due to scattering. Some examples are considered. [Work was supported by RFBR (Grant 97-05-64878).]

FRIDAY MORNING, 26 JUNE 1998 GRAND BALLROOM III (W), 11:00 A.M. TO 12:00 NOON

Session 5aPLb

Plenary Lecture

Thomas D. Rossing, Chair
Department of Physics, Northern Illinois University, DeKalb, Illinois 60115

Chair's Introduction—11:00

Invited Paper

11:05

5aPLb1. Nonlinearity, complexity, and the sounds of musical instruments. Neville H. Fletcher (Res. School of Physical Sci., Australian Natl. Univ., Canberra 0200, Australia, neville.fletcher@anu.edu.au)

Musical instruments can be treated to a first approximation as linear harmonic systems, but closer examination shows that they are, almost without exception, nonlinear and inharmonic. Sustained-tone instruments such as violins and clarinets can be characterized as ''essentially nonlinear'' and owe their deceptively simple behavior to mode locking due to strong nonlinearity. When the nonlinearity is weakened, the behavior becomes complex, with inharmonic and multiphonic sounds that are sometimes exploited in modern music. In contrast, impulsively excited instruments, such as guitars, pianos, gongs, and cymbals, may be described as ''incidentally nonlinear.'' They would function quite acceptably if the nonlinearity were to be eliminated but, when it becomes strong, they exhibit physically and aurally interesting behavior such as period doubling, frequency multiplication cascades, and chaotic oscillation. This paper will explore some of these phenomena and demonstrate both their physical origins and musical utility.

Session 5pAAa

Architectural Acoustics: Classroom Acoustics

Gary W. Siebein, Chair
Department of Architecture, University of Florida, 231 Arch, P.O. Box 115702, Gainesville, Florida 32611-5702

Chair's Introduction—1:30

Contributed Papers

1:35

5pAAa1. Classroom acoustics I: The acoustical learning environment: Participatory action research in classrooms. Mary Jo Hasell, Philip Abbott, Gary W. Siebein, Martin A. Gold, Hee Won Lee, Mitchell Lehde, John Ashby, Michael Ermann, and Carl C. Crandell (Architecture Technol. Res. Ctr., 336 ARCH, Univ. of Florida, P.O. Box 115702, Gainesville, FL 32611-5702)

This pilot study used participatory fieldwork in a number of kindergarten through eighth grade classrooms to evaluate the acoustic setting that supports learning. Dynamic interactions among administrators, teachers, students, parents, and research team members makes a difference in learning. It was found that classroom interaction depends on the social make-up, anticipated behavior, intellectual level, teaching method and theories, and hearing capacity of participants, as well as the physical characteristics of the classrooms. Participatory action research (PAR) resulted in the development of appropriate classroom observation tools utilizing checklists, summary scales, and hierarchial rating scales. The learning climate, lesson clarity, instructional variety, task orientation, and student engagement were used to identify variables for acoustical studies. The PAR was instrumental in determining the protocol for the acoustical measurement studies described in Classroom Acoustics II. Field observations began with an understanding of the adaptive behaviors of the classroom participants such as how the teachers modified their location and sometimes even the furniture to maintain short speaker to listener distances. The findings indicate a range of solutions are needed to improve the acoustical learning environment including innovative teaching methods, improving room acoustical design, reducing background noise levels, and, in some cases, using amplification systems.

1:50

5pAAa2. Classroom acoustics II: Acoustical conditions in elementary school classrooms. Martin A. Gold, Hee Won Lee, Gary W. Siebein, Mitchell Lehde, John Ashby, Michael Ermann, Mary Jo Hasell, Philip Abbott, and Carl C. Crandell (Architecture Technol. Res. Ctr., 231 ARCH, Univ. of Florida, P.O. Box 115702, Gainesville, FL 32611-5702)

Acoustical measurements of speech transmission index, reverberation time, early reverberation time, early-to-late energy ratios, loudness (or relative strength), articulation index, background noise levels, and signal-to-noise ratios were made in a number of elementary school classrooms in one school district to see how many rooms actually had acceptable acoustical conditions. The source and receiver locations for the acoustical measurements were determined through the participatory action research (PAR) described in Classroom Acoustics I. Measurements were made using a TEF analyzer with custom software to compute additional acoustical measurements. Both omni-directional and directional loudspeakers were used for the source signals. The acoustical measurements were conducted in a variety of source receiver conditions as observed in the PAR studies in the classrooms. General observations are related regarding the relationship between the measurements and the PAR. For example, the STI was always greater than 0.75 for the conditions under which the teachers actually taught. Speaker-to-listener distances of 4 m or less were observed for many classroom activities. Likewise the teachers employed many creative ways to control the behavior of the students to reduce background noise levels while they actually spoke.

2:05

5pAAa3. Classroom acoustics III: Acoustical model studies of elementary school classrooms. Gary W. Siebein, Mitchell Lehde, Hee Won Lee, John Ashby, Michael Ermann, Martin A. Gold, Mary Jo Hasell, Philip Abbott, and Carl C. Crandell (Architecture Technol. Res. Ctr., 231 ARCH, Univ. of Florida, P.O. Box 115702, Gainesville, FL 32611-5702)

A 1:4 scale model was constructed with a series of interchangeable wall and ceiling panels to allow a battery of acoustical tests to be conducted for a variety of classroom designs. The walls and ceiling of the room could be changed from sound absorbent to sound reflective materials quickly. Scale furniture was also constructed for the room. The walls of the model could be adjusted from approximately 7 to 10 meters in length and from 2.5 m in height to 7 m in height. Acoustical measurements of speech transmission index, reverberation time, early reverberation time, early-to-late energy ratios, loudness (or relative strength), and articulation index were made in a number of simulated elementary school classrooms using both a TEF analyzer and impulse spark techniques. Classrooms where field measurements were taken as described in Classroom Acoustics II were modified to attempt to improve acoustical conditions in the rooms. The location and amount of absorbent material, the location and amount of sound reflective material, and the room volume were the major variables considered. In general rooms that followed good architectural acoustics design principles produced favorable acoustical measurements.

2:20

5pAAa4. Classroom acoustics IV: Speech perception of normal-hearing and hearing-impaired children in classrooms. Carl C. Crandell, Gary W. Siebein, Martin A. Gold, Mary Jo Hasell, Philip Abbott, Mitchell Lehde, and Hee Won Lee (Architecture Technol. Res. Ctr., 231 ARCH, Univ. of Florida, P.O. Box 115702, Gainesville, FL 32611-5702)

The present investigation examined the speech-perception abilities of normal-hearing and hearing-impaired children in a number of classroom environments. Speech perception was assessed as a function of teacher location, background noise levels, early sound reflections, reverberation time, and signal-to-noise ratio. Specific acoustical measurement procedures are described in Classroom Acoustics I and II. Speech perception was assessed by nonsense syllables, monosyllabic words, and sentences. Normal-hearing populations included children, aged 5–17 years, who are progressing normally in school; learning-disabled children; children with central auditory processing deficits; articulatory- and/or language-disordered children; children with developmental delays and/or attention deficits; and children for whom English is a second language. Hearing-impaired populations consisted of children with minimal-to-severe degrees of bilateral and unilateral, sensorineural or conductive hearing loss. Data will be discussed in view of developing appropriate classroom acoustics for normal-hearing and hearing-impaired pediatric listeners.

2:35

5pAAa5. Assessing speech intelligibility in classrooms at the University of Washington. Dean Heerwagen (Dept. of Architecture, Univ. of Washington, Box 355720, Seattle, WA 98195-5720) and Tarik Khiati (Dept. of Urban Design and Planning, Univ. of Washington)

The University's Committee on Accessibility has initiated an investigation of classroom acoustical conditions. This investigation has been undertaken in response to complaints made by University faculty members who state that they are experiencing difficulties in understanding student's speech during class sessions, when students ask questions or engage in discussion. A survey of University faculty identified 75 problematic classrooms (out of the 300+ classrooms on the University campus). A sample of generally representative classrooms has been selected: these classrooms (and class sessions held in the classrooms) are being systematically measured to determine their acoustical properties. Seven parameters have been anticipated as the causes of inferior speech intelligibility: to strong background noise levels, too weak signal-to-noise ratios, excessive reverberation, too great speak-to-listener distances, various speaker idiosyncrasies (e.g., accenting, mumbling, rapidity, using unfamiliar verbiage, and so forth), acoustical defects of the classroom, and faulty amplification systems. To date, measurements show that the principal acoustical faults causing these intelligibility difficulties are high background noise levels, inadequate amounts and locations of absorptive materials and assemblies, and student speech characteristics including too-long distances separating faculty and students. In selected classrooms corrections are being undertaken.

FRIDAY AFTERNOON, 26 JUNE 1998 GRAND FOYER (W), 1:00 TO 5:00 P.M.

Session 5pAAb

Architectural Acoustics: Room Acoustics Measurement Techniques, Absorption and Diffusion (Poster Session)

Julie Wiebusch, Cochair
Greenbusch Group, 919 NE 71st Street, Seattle, Washington 98115-5634

John Greenlaw, Cochair
Greenbusch Group, 919 NE 71st Street, Seattle, Washington 98115-5634

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.

5pAAb1. Characterizing scattering from room surfaces. Tristan J. Hargreaves, Trevor J. Cox, Y. W. Lam (Dept. of Acoust. and Audio Eng., Univ. of Salford, Salford M5 4WT, UK), and Peter D'Antonio (RPG Diffusor Systems, Inc., Upper Marlboro, MD 20772)

There is a need to characterize the scattering from room surfaces in terms of a diffusion parameter. Such a parameter is required for many geometric room acoustic models and to enable the relative diffusion performance of surfaces to be evaluated. In addition, it is hoped that the characterization of scattering will aid designers of rooms. This paper examines the derivation of diffusion parameters from the distribution of scattered energy, which is one possible route to characterizing surface scattering. Previous investigations have concentrated on obtaining diffusion parameters from two-dimensional polar distributions—either measurements or predictions using BEM techniques. The advantages and disadvantages of these parameters are outlined and an alternative parameter of this type which overcomes some of the disadvantages is described. A new capability to perform measurements in three dimensions, which has broadened the applicability of the results from this study, is described. Extending the use of existing two-dimensional diffusion parameters to characterize hemispherical scattering will be discussed. Some solutions to the problems which arise when using these parameters to characterize scattering from large surfaces or those mounted in baffles will be presented.

5pAAb2. Numerical study on sound absorption characteristics of brick/block absorbing walls. Shinichi Sakamoto, Dong-Jun Joe, Hideki Tachibana (Inst. of Industrial Sci., Univ. of Tokyo, 7-22-1 Roppongi, Minato-ku, Tokyo, 106 Japan), and Hikari Mukai (Ono Sokki Co., Ltd., 1-16-1 Hakusan, Midori-ku, Yokohama, 226 Japan)

The brick/block absorbing walls with openings and backing air space with porous materials are often used for absorption in low frequencies. The absorption in low frequencies can be explained by the well-known theory of Helmholtz resonance. In some measurement results of absorption coefficient, however, high absorption coefficients are seen in high frequencies. Regarding this problem, the authors have been investigating theoretically and experimentally as presented at the last ASA meeting (at Penn. State). Following these studies, further numerical analysis was carried out on simple 2-D models. The wave phenomena in the cavity and backing space/material were formulated by combining the FEM with the

scattering wave theory. Comparison between the calculated result and the result of a 1/10 scale model experiment has indicated the validity of the numerical analysis.

5pAAb3. Absorption characteristics of a double-leaf membrane with an absorptive layer in its cavity. Kimihiro Sakagami, Toru Uyama (Environ. Acoust. Lab., Faculty of Eng., Kobe Univ., Rokko, Nada, Kobe, 657 Japan), Masakazu Kiyama (Kobe Univ., Rokko, Nada, Kobe, 657 Japan), and Masayuki Morimoto (Kobe Univ., Rokko, Nada, Kobe, 657 Japan)

Absorption characteristics of double-leaf membranes with an absorptive layer in the cavity are theoretically analyzed. The double-leaf membrane is modeled as two infinite membranes with a cavity between them, which consists of three layers of arbitrary media. The theory also includes the effect of permeability of the interior leaf. The effects of parameters of the absorptive layer on the absorption characteristics are discussed through numerical examples. The absorption characteristics of an impermeable double-leaf membrane are characterized by a mass-spring resonance peak at low frequencies. This peak becomes larger and shifts to lower frequencies as the thickness of the absorptive layer increases. The absorption characteristics are, in permeable cases, similar to those of the conventional porous absorbents; the absorption is low at low frequencies and higher at high frequencies. As the thickness of the absorptive layer increases, the absorption becomes higher at all frequencies. The change in position of the absorptive layer affects the absorption characteristics of the double-leaf membrane. This effect is related to the standing-wave sound field in the cavity. The flow resistance of the absorptive layer has an optimum value to maximize the absorption of a double-leaf membrane.

5pAAb4. Numerical simulations of the modified Schroeder diffuser structure. Antti Jrvinen, Lauri Savioja (Helsinki Univ. of Technol., Lab. of Acoust. and Audio Signal Processing, P.O. Box 3000, FIN-02015 HUT, Finland), and Kaarina Melkas (Nokia Res. Ctr., FIN-33721 Tampere, Finland)

Schroeder diffusers are widely used in architectural acoustics because they offer unique diffusion characteristics. It has been recently shown that the Schroeder diffusers are also efficient low-frequency absorbers. In this paper a new modified Schroeder diffuser structure is presented. The new folded structure allows one to extend the low-frequency limit of the Schroeder diffuser so that the cutoff frequency is not directly proportional to the depth of the structure. Other important advantage of the new structure is that the total volume of the structure is fully utilized as a diffuser. The modified structure is also easier to manufacture than the traditional Schroeder diffusers. The modified structure is compared with the traditional ones by numerical simulations and measurements. Both the radiation and absorption characteristics are modeled. Modeling techniques, such as finite element and boundary element method and finite difference time domain methods, are applied and compared with each other.

5pAAb5. Evaluation of spatial information from artificial head and four-microphone array measurements. Joerg Becker and Markus Sapp (Insitut für Elektrische Nachrichtentechnik, RWTH Aachen, 52056 Aachen, Germany)

Examinations on the acoustic qualities of enclosures are mostly restricted to the evaluation of single-number quantities like ''reverberation time'' and ''gravity time'' or criteria which qualify the perceptual qualities, for example, ''definition'' or ''clarity index.'' Most of these criteria are measured monaurally, although the spatial distribution of the reflections is essential for the acoustic quality of rooms. Spatial information can be registered with multimicrophone arrays or artificial heads. On the one hand it is easier to get the spatial information by using microphone arrays and on the other hand aurally adequate recordings can only be made with artifical heads. This paper focuses on the further development of algorithms for detecting virtual sound sources and compares evaluations of artifical head recordings with those made by a 4-microphone array.

5pAAb6. Combined beam tracing and Biot transfer-matrix model for predicting sound fields in enclosed spaces. Murray Hodgson, Andrew Wareing, and Callum Campbell (UBC Occupational Hygiene Prog. and Dept. of Mech. Eng., 3rd Fl., 2206 East Mall, Vancouver, BC V6T 1Z3, Canada, hodgson@mech.ubc.ca)

Numerous methods exist for predicting sound fields in enclosed spaces such as rooms, aircraft cabins, and the sea. These are generally energetic approaches which ignore phase and associated wave effects, and which describe sound reflection from surfaces by the energy absorption coefficient. Many, such as ray tracing, require long calculation times. In the work reported here, a new model is presented which is computationally efficient, includes phase, and decribes surface reflection by the angularly varying surface impedance. It is a triangular-beam-tracing algorithm, combined with a Biot-theory transfer-matrix model for laterally homogeneous, multilayer, poroelastic surfaces. The Biot model has been validated in the case of plates, porous materials, and seabeds. The combined approach is being validated in comparison with predictions by analytic and numerical approaches, and with experiment. The model is being used to study the effect of surface extended reaction on enclosed sound fields.

5pAAb7. Broadening the frequency range of panel absorbers by adding an inner layer of microperforated-plate. Jian Kang, Xueqin Zha, and Helmut V. Fuchs (Fraunhofer-Institut fur Bauphysik (IBP), Nobelstr. 12, D-70569 Stuttgart, Germany, kang@ibp.fhg.de)

A disadvantage of panel absorbers, namely a panel backed by an airspace, is that the effective frequency range is usually rather narrow. It has been theoretically demonstrated that the absorption range can be broadened by introducing a microperforated-plate between the panel and the back wall [J. Kang, IBP Report 1993]. In this paper, a series of experimental results on the effectiveness of this method is presented. The experiments were carried out using the single microphone FFT technology in an impedance tube. It is shown that with an inserted microperforated-plate the acoustic resistance of the whole structure is greater than that of the outer panel alone. When the parameters of the microperforated-plate are carefully selected, the absorption range can be effectively extended in both low- and high-frequency directions. Typically the improvement is around 1 oct. The effectiveness of the microperforated layer is much less when its resonance frequency is considerably higher than that of the outer panel alone. [Work supported by AvH.]

5pAAb8. Reverberation time predictions using neural network analysis. Joseph Nannariello and Fergus Fricke (Architectural and Design Sci. Dept., Univ. of Sydney, NSW 2006, Australia)

The study aims to utilize neural network analysis to develop a method to predict the reverberation time for enclosures—in the low and mid frequencies—by using neural networks that have been trained with constructional and acoustical data. The study begins with the hypothesis that, in practice, reverberation time predictions are too difficult to undertake using existing computer models, and too inaccurate when using other methods. Specifically, the study aims at providing an expeditious and accurate method of predicting the reverberation time of enclosures at the initial design stage. To substantiate the hypothesis, and to bring into effect the aims of the study, assessments are made of the predictive powers of the trained neural networks. The results of the investigations have indicated that there is a good basis for using trained neural networks to predict the reverberation time for enclosures. The results have also shown that neural network analysis can identify those variables that have an effect on the predicted reverberation times.

5pAAb9. Sound fields caused by diffuse-type reflectors with periodic profile. Daiji Takahashi (Grad. School of Eng., Kyoto Univ., Kyoto, 606-01 Japan, tkhs@archi.kyoto-u.ac.jp)

For the purpose of providing highly diffuse reflections, the reflection surface is often periodically corrugated. The degree of diffusion of this periodic type diffuser can be evaluated by an objective criterion named $DNSD$ [D. Takahashi, J. Acoust. Soc. Jpn. (E) **16**, 51–58 (1995)], the value of which varies in the range 0–1; $DNSD=0$ for a perfect diffusion, and $DNSD=1$ for a specular reflection. Optimum diffuse reflectors with periodic profiles can be designed by using this criterion, but this type of diffusion might have some aspect of misgivings due to its periodicity, which may be a kind of so-called ''tone coloration.'' In other words, the highly diffuse reflections might have a problem of some particular effect on the spectrum caused by the interference between the direct and the reflected waves, and the effect might be different from an ordinary case of the plane surface. In this report, for the three types of reflectors, the calculated results of the SPL characteristics of the reflected field as well as the diffusion performances are presented and mutually discussed in comparison with the measured data from $\frac{1}{8}$-scale model experiments.

5pAAb10. Directivity patterns of small explosions. Miguel Arana (Phys. Dept., Public Univ. of Navarra, 31006, Pamplona, Spain)

There are methods and equipment for the generation, amplification, and emission of acoustic impulses in order to evaluate the acoustic characteristics of a room by means of the Fourier transform from its impulse response. However, the responses of this equipment (especially the inertia of the loudspeakers and its lack of linearity) produce acoustic signals that do not reproduce reliably the necessary acoustic impulses. A notable dispersion was found in the measurements accomplished with these systems. The directivity patterns of small explosions have been measured and analyzed with the purpose of using them as impulsive acoustic sources. The typical durations of these impulses are between 0.1 and 0.2 ms. The sound power spectrum is relatively smooth from 1 to 10 kHz. The mean acoustic power is around 200 W (143 dB, Ref. 1 pW). For this an experimental device with 16 channels and software based on LabView were used. This communication will show the experimental device as well as the obtained results (mainly the dispersion of both the acoustic power and directivity patterns). [Work supported by Education Department of Government of Navarra.]

5pAAb11. Sound absorption measurements in diffuse field: A study of parameters. Marco Nabuco, Paulo Massarani (Acoust. Testing Lab., DIAVI/INMETRO), and Roberto Tenenbaum (Acoust. and Vib. Lab., COPPE/ Federal Univ. of Rio de Janeiro)

Many researchers have been studying the sound absorption coefficient measurement technique in reverberation chambers, since Sabine presented his formulation during the end of the last century, in order to find reasons for data frequently obtained bigger than 100%, normally correlated to the edge effects of the sample. Some researchers have studied the implications of the lack of diffuseness on the sound absorption data. This paper presents the data obtained for the sound absorption coefficient measured in a 200-m^3 reverberation chamber at the National Institute of Metrology, Standardisation and Industrial Quality-INMETRO, in Brazil, for different areas and configurations materials, from 1 to the standardized 12-m^2 sample. Also presented is the steady-state standard deviation, the cross-correlation between microphones, and the shape of the decays curves to qualify the diffuseness both in steady as well as during the transient state. At last is presented the data obtained using a steady-state measurement of the sound absorption using a standard reference sound source calibrated in semianechoic chamber.

5pAAb12. Dependence of accuracy of reverberation time measurement on input filters. Dragana S. Sumarac (Faculty of Elec. Eng., P.O. Box 35-54, 11001 Belgrade, Yugoslavia, esumarad@ubbg.etf.bg.ac.yu)

In this paper, attention was paid to the standard method for measuring frequency characteristics of reverberation time in rooms. The influence of transfer function characteristics of 1/1 and 1/3 digital octave input filters on differences in measuring values of reverberation time was studied. Procedures for complete digital analysis of recorded impulse response of the room were designed. All procedures were designed by the MATLAB software package. The verification of that measuring software was performed on a great number of recorded impulses in real rooms. The several classes of 1/1 and 1/3 digital octave FIR and IIR filters were designed according to the international standard IEC 225. The complete analysis was performed using the recorded reverberation decay curve in 20 rooms significantly different in volume and reverberation time values. The rooms were excited with broadband pink noise and short duration impulses. The values of measured reverberation times were statistically analyzed to show the expected variance that can appear due to differences of applied 1/1 and 1/3 digital input filters.

5pAAb13. Reverberation time directly obtained from a squared impulse response envelope. Fumiaki Satoh (Dept. of Architecture, Chiba Inst. of Technol., Tsudanuma 2-17-1, Narashino-shi, Chiba, 275 Japan, f_satoh@cc.it-chiba.ac.jp), Yoshito Hidaka (Dept. of Elec. Eng., Tohwa Univ., Chikushigaoka 1-1-1, Minami-ku, Fukuoka-shi, 815 Japan), and Hideki Tachibana (Tokyo Univ., Tokyo, 106 Japan)

For the measurement of reverberation time, a method of directly regressing the envelope of a squared impulse response (direct method) was examined using 100 data obtained in 14 auditoria. The results were compared with those obtained by the integrated impulse response method (Schroeder method) and a very high coincidence was found between them. It has also found that the direct method is more robust to background noise than the Schroeder method and in the former method relatively short-time data will do compared with the latter method.

5pAAb14. Effects of flow resistance on acoustic performance of permeable elastic-plate absorbers. Manabu Tanaka (Bldg. Res. Corp., Osaka-Suita, 5650873 Japan) and Daiji Takahashi (Kyoto Univ., Kyoto, 6068317 Japan)

In general vibration theories of thin elastic plate, the plates have been assumed to be nonpermeable. However, previous work on permeable membranes revealed that the acoustic properties of a membrane are strongly affected by its permeability [Takahashi *et al.*, J. Acoust. Soc. Am. **99**, 3003–3009 (1996)]. In this study, permeable elastic-plate absorbers are proposed and their acoustic performance is examined by means of numerical analysis. For the simplest model, a theory for sound transmission through and sound reflection by a single permeable plate absorber with the oblique incidence of steady-state plane wave is developed. In this theory, permeability of the plate is represented by unit-depth flow resistivity [MKS-rayl/m]. Subsequently, for more complex models, structured absorbers composed of the facing of a permeable plate and several layers of air and/or absorptive materials are investigated theoretically. With these theoretical models, parametric surveys are executed to evaluate the acoustic performance of permeable elastic-plate absorbers. As a result of numerical calculations, it becomes clear that the flow resistance of the permeable elastic plate has a remarkable influence on its acoustic characteristics, especially at high frequencies.

5pAAb15. Qualitative evaluations of the acoustics of rooms: Real room studies and headphone studies. Martin A. Gold (Architectural Acoust. Res. Group, 231 ARCH, Univ. of Florida, P.O. Box 115702, Gainesville, FL 32611-5702)

This pilot study attempts to qualify the use of headphones to evaluate the listening conditions of existing and proposed architectural spaces. Qualitative listening evaluations of speech and music were conducted in multiple locations in three different listening environments: a 2000-seat multipurpose concert hall, a 600-seat lecture room, and a 120-seat lecture room. Groups of college age students were asked to evaluate loudness, clarity, reverberance, spatial impression, and background noise using a questionnaire with a seven-point bipolar rating scale [Cervone (1990)]. The speech and music sources were played through two adjacent loudspeakers at the front of the respective rooms while the students completed the evaluation forms. The speech and music sources were binaurally recorded on a DAT at the listener locations in the respective rooms. At a later time, students were asked to evaluate the acoustical qualities of the binaural recordings made in the rooms using the same questionnaire. The recordings were presented to the listeners in a random order in a blind subjective evaluation. Data from real room evaluations and headphone studies are compared and conclusions are drawn regarding the potential for headphone listening experiments as a means to study qualitative aspects of room acoustics.

5pAAb16. On the acoustical characteristic of a balloon. Anthony P. Nash (Charles M. Salter Associates, 130 Sutter, Ste. 500, San Francisco, CA 94104, tnash@cmsalter.com)

It is common to use a transient acoustical signal when testing the properties of large architectural spaces. Some of the traditional signal sources include pistol shots, pulses from loudspeakers, etc. Assuming the background noise is reasonably low, an inflated balloon can be a very effective signal source in this application. The published literature contains little documentation regarding the sound power, directional properties, and repeatability of an exploding balloon. A carefully controlled experiment was conducted involving a series of balloons exploded in an environment where the closest reflecting surface was located more than 15 m from the source. Several microphones ringing the source were used to measure the free-field acoustical transient on a multi-channel tape recorder. The measured data will demonstrate that a properly inflated balloon is a very satisfactory (and lightweight!) sound source.

5pAAb17. Sound absorption by a large-size isolated acoustic resonator with a cylindrical cavity. V. J. Stauskis (Vilnius Gediminas Tech. Univ., Saulėtekio al. 11, 2040 LT, Vilnius, Lithuania)

An acoustical structure consisting of a large-scale isolated resonator with a large-diameter cylindrical cavity may be used. This resonator differs essentially from the classical Helmholtz resonator whose cavity is only several millimeters in diameter and lined with a sound-absorbing material. Formulas are presented with which the impedance of the cavity, the impedances both inside and outside the cavity, and the impedance of the volume of the resonator are calculated. A radiation directivity pattern expressed by Bessel's function is used to find the radiation impedance. Calculations show that the sound energy is absorbed by resonators made of sound-reflecting materials. The larger the cavity diameter is, the larger the sound absorption. Absorption is of a resonant character with the resonant frequency at 60 Hz. A resonator measuring 200×200 cm, with the cavity diameter 50 cm and the distance to the rigid surface being 30 cm, absorbs 3.5 m^2 of sound energy at the resonant frequency. At very low frequencies, changes in the imaginary parts of both cavity and radiation impedances occur along with the increase in the cavity diameter and frequency.

5pAAb18. Influence of distance and sound recording system on intelligibility and depth localization in highly reverberant conditions. Ruiz Robert (Laboratoire d'acoustique, Univ. de Toulouse-le Mirail, 5 allees Antonio Machado, 31058 Toulouse Cedex 1, France) and Ballet Isabelle (Univ. de Toulouse-le Mirail, 31058 Toulouse Cedex 1, France)

Speech intelligibility degrades when the listener moves from the source. In a highly reverberant and empty church, scores are measured at different distances from a loudspeaker and compared with those obtained from recordings of the speech at the same location. The phonetic material is a set of phonetically balanced lists of triphonemic French words. Two monophonic (omnidirectional, cardiod) and one stereophonic (ORTF) systems are used. Listening of recorded lists is performed with the same loudspeaker(s) as in the church. Results of the experiment indicate smaller variations of scores with the distance. An important improvement of scores is obtained with the stereophonic system. A second experiment consists of using a sound material composed of two voices performing a singing exercise edited in monophony in order to create two distinct sound planes. Listening tests allow the quantiication of the depth localization variations with distance and sound recording system. Results show a good respect of distance and reveal a sudden change in depth localization for stereophony. Acoustical properties of the church, distance factors of the microphones, properties of the stereophony against monophony, and real listening conditions are considered to discuss the experiments.

5pAAb19. Sound field in long enclosures with diffusely reflecting boundaries. Judicaël Picaut, Laurent Simon (LAUM UMR 6613, Univ. du Maine, 72017 Le Mans Cedex 9, France), and Jean-Dominique Polack (Univ. Paris VI, Case 161, 75252 Paris Cedex 05, France)

The sound field modelization by a diffusion equation is used to predict the sound propagation in long enclosures with diffusely reflecting boundaries. The diffusion equation is solved for time-varying sources and particularly in the steady state. Analytical expressions of sound attenuation and reverberation in infinite enclosures are given. Their simplicity makes them easy to apply to realistic diffusing long enclosures. For wide cross sections of enclosures and when absorption at the beginning and at the end walls is high, solutions for finite enclosures are also proposed. The source location and the absorption of each wall are taken into account in the model. Comparisons between analytical solutions, numerical simulations, and classical theory of diffuse sound field show good agreements. [Work supported by the PIR-Villes CNRS.]

5pAAb20. Reverberance of an existing hall in relation to subsequent reverberation time and SPL. Shigeo Hase (Grad. School of Sci. and Technol., Kobe Univ., Rokkodai, Nada, Kobe, 657-8501 Japan), Akio Takatsu (Showa sekkei, Inc., 1-2-1-8000, Benten, Minato-ku, Osaka, Japan), Shin-ichi Sato, Hiroyuki Sakai, and Yoichi Ando (Kobe Univ., Rokkodai, Nada, Kobe, 657-8501 Japan)

Psychological tests in relation to reverberance of sound fields were conducted in an existing hall which is a medium-sized multipurpose hall with 400 seats. Music and speech as sound sources were radiated from an omnidirectional dodecahedron loudspeaker on the stage. Paired-comparison tests were applied for the sound fields changing two orthogonal factors. These factors were the subsequent reverberation time T_{sub} and the SPL. The T_{sub} were adjusted by a hybrid system involving the reverberation control room and an electroacoustic system. 21 subjects took part in this test. The results indicate that the reverberance is influenced independently on the T_{sub} and the SPL. The reverberance increases with increasing both the T_{sub} and the SPL.

Session 5pAO

Acoustical Oceanography and Animal Bioacoustics: Acoustics of Fisheries and Plankton V

Manell E. Zakharia, Chair
CPE Lyon, LISA (EP 92 CNRS) LASSO, 43 Bd. du 11 Novembre 1918, BP 2077, 69616 Villeurbanne Cedex, France

Invited Paper

1:00

5pAO1. Bioacoustic resonance absorption spectroscopy. Orest Diachok (Naval Res. Lab., Washington, DC 20375)

Absorption losses at the resonance frequencies of sardines up to about 18 dB at 1.3 kHz at night, 15 dB at 1.7 kHz during the day and 35 dB at 2.7 kHz at dawn were observed at a range of 12 km during Modal Lion, a multidisciplinary experiment designed to isolate absorptivity due to fish from other effects on long-range propagation at a relatively shallow (83 m) site in the Gulf of Lion. Comparison of transmission loss measurements with a numerical sound propagation model that incorporates absorption layers in the water column permitted estimation of the average absorption coefficient, depth, and thickness of absorption layers. The depths and thickness of layers estimated from sound propagation measurements were in good agreement with echo sounder data. The measured resonance frequencies of dispersed fish at night were within 13% of theoretical computations, based on swim bladder dimensions which were derived from near coincident samples of adult (~16 cm long) sardines. A smaller absorption line at 3.9 kHz at night is consistent with the resonance frequency of dispersed juvenile sardines, which are known to be ~6 cm long, based on historical data. Measured resonance frequencies of sardines in schools were about 0.59 times the resonance frequency of dispersed sardines. The observed frequency shift is analogous to previously observed frequency shifts associated with ''clouds'' of bubbles. The results presented here suggest the possibility of long-term tomographic mapping of fish parameters over large areas using transmission loss measurements. [This work was supported by The Office of Naval Research.]

Contributed Papers

1:20

5pAO2. The application of wideband signals in fisheries and plankton acoustics. John E. Ehrenberg and Thomas C. Torkelson (Hydroacoustic Technol., Inc., 715 NE Northlake Way, Seattle, WA 98105-6429)

Over the last 30 years, there have been considerable improvements in techniques developed for acoustic assessment of fisheries and plankton populations. Techniques have evolved from analog echo integrators used to obtain crude abundance estimates to digital multiple-beam techniques for measuring *in situ* target strength, location, and velocity of individual targets. These advancements have been largely due to digital signal processing techniques applied to signals at the output of an echo sounder. Constant-frequency sine wave (CW) pulses have been used with nearly all plankton and fisheries assessment systems. The same digital signal processing technology used to improve post-processing is now finding its way into biological assessment echo sounders that use FM slide/chirp signals. Systems with time-bandwidth products of up to 50 have been developed and implemented to provide a signal-to-noise advantage in excess of 15 dB when compared to CW pulse signals with the same spatial resolution. This paper describes the implementation and application of FM slide/chirp signals for echo integration and target strength measurement. Practical implementation issues, such as the use of pulse shaping for range side-lobe control, are discussed. Recent data collected with a multifrequency acoustic assessment system developed for WHOI are presented.

1:35

5pAO3. Wideband fisheries sounder. From individual echoes analysis to classification of schools at sea. Manell E. Zakharia (CPE Lyon, LISA (EP92 CNRS)/LASSSO, 43 Bd du 11 Novembre 1918, BP 2077 69616 Villeurbanne Cedex, France, zakharia@cpe.fr)

Several works have been carried out in the last decade concerning wideband fisheries systems and their ability to discriminate fish species. The paper will first review previous work showing the relevance of wideband systems for fisheries acoustics (analysis of individual echoes). A short description of a wideband prototype (20 to 80 kHz) will then be given showing the technological issues. An emphasis will be made on the analysis of fish schools either in a controlled situation (lakes) or at sea. Wideband echo processing was based on autoregressive modeling of individual pings (modeling of spectrum resonances). This modeling concentrates the information contained in the wideband echo in a reduced set of parameters that could be used for classification. A neural network was trained to recognize several fish species. Ground truth was achieved by species identification, at sea, using associated trawling. For the construction of an echo data base, only monospecific schools were considered. Classification results at sea using only the spectral signature of the echoes were as high as 75% from a single ping. Lake tests have shown that the use of several pings (ten and above) could increase the recognition rate to more than 90%.

1:50

5pAO4. Searching for species identifiers in multifrequency target strength variability. James Dawson (BioSonics, Inc., 4027 Leary Way NW, Seattle, WA 98107)

Acoustic research has shown that the species of some schooling fish can be assigned based solely on measurement parameters from acoustic returns. For single targets, this process is complicated by the high variability of *in-situ* target strength (TS) measurements. Attempts at speciation based on the patterns of TS variability were confounded by interactions between acoustic frequency and fish directivity. Initial attempts at defining speciation parameters based on multifrequency target strength measures look promising. This paper reports on the results of simultaneously measuring the *in-situ* target strength of fish at 38-, 120-, and 420-kHz acoustic frequencies, and proposes the definition of several speciation parameters.

2:05

5pAO5. A broadband acoustic fish identification system. Gerald Denny and Patrick Simpson (Sci. Fishery Systems, Inc., P.O. Box 242065, Anchorage, AK 99524, scifish@alaska.net)

A broadband fish identification system has been built and tested in the Great Lakes, Prince William Sound, and the Bering Sea. Unique signatures have been obtained on free swimming and captive animals. These signatures include jellyfish, euphausiids, Pacific Rockfish, and Walleye Pollock. Acoustic characteristics and utilization of the system will be discussed, as well as results of *in situ* data collection.

2:20–2:35 Break

2:35

5pAO6. The use of two freqencies to interpret acoustic scattering layers. Denise R. McKelvey (Natl. Marine Fisheries Service, NOAA, 7600 Sand Point Way NE, Seattle, WA 98115, denise.mckelvey@afsc.noaa.gov)

The use of multiple frequencies during acoustic stock assessment surveys allow methods to be used, which can discriminate between different species groups based on their scattering properties and assist in target species identification, thus improving estimates of biomass. Research, using multiple frequencies to study zooplankton, have used the difference in backscatter between several frequencies to identify the species or species groups comprising the scattering layers. This paper presents the results of using a dual-frequency differencing algorithm on 120- and 38-kHz mean volume backscatter strength (MVBS) data collected during a 1995 United States west coast acoustic survey of Pacific hake (*Merluccius productus*) as a means to acoustically separate hake from zooplankton. Scatterers were sampled using various nets with mouth areas ranging from 5 to 1000 m^2. Backscatter data from both frequencies were corrected for background noise. The effect of species composition, time of day, depth, and density of scatterers on the (delta)MVBS data were evaluated. Results indicate that the dual-frequency differencing algorithm shows a marked difference between fish and zooplankton (delta)MVBS and can be used to distinguish between fish and zooplankton scattering layers.

2:50

5pAO7. Fish'n krill: A 38/120-kHz acoustic separation. Yvan Simard (Dept. of Fisheries and Oceans, Maurice-Lamontagne Inst., P.O. Box 1000, Mont-Joli, QC G5H 3Z4, Canada, simardy@dfo-mpo.gc.ca)

Volume backscattering data (Sv) at 38 and 120 kHz were collected in the Gulf of St. Lawrence, during daylight, in regions where krill scattering layers (SL) and pelagic fish (mainly capelin) were present. The SLs and fish aggregations were sampled with the Bioness zooplankton samples and with small-mesh pelagic trawls equipped with acoustic monitoring systems. Echointegration was performed on small bins of ~5 m vertically and ~50 m horizontally. Comparisons of Sv at the two frequencies indicated that krill SLs had stronger backscattering at 120 than at 38 kHz, as expected. The ratio was relatively constant within the same SL, but varied from one SL to the other in parallel with the size of krill. The backscattering from fish schools or fish aggregations or fish layers was stronger at 38 kHz than at 120 kHz. This permitted the separation of fish from krill on the echograms, even for complex cases where fish were embedded in krill SL or had structural patterns that mimic the signature of krill SL. The 38/120-kHz backscattering ratio appears therefore useful to separate krill from fish, or vice-versa, for better interpretation and estimation of these taxa individually.

3:05

5pAO8. Multifrequency acoustic assessment of fisheries and plankton resources. Thomas C. Torkelson (Hydroacoustic Technol., Inc., 715 NE Northlake Way, Seattle, WA 98115-6429), Thomas C. Austin, and Peter H. Wiebe (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

The range of target species encountered in marine biological resource assessment and the inability to obtain a broad view of these resources by direct physical sampling techniques has lead to the utilization of various indirect measurement techniques. Among these techniques is the use of acoustics (sonar) to determine biomass distribution through integration and target strength measurements. WHOI is using a multifrequency acoustic system to obtain more accurate measurements of target species on the Georges Bank. The system uses advanced acoustic techniques to enhance detectability at the lower trophic levels and uses an advanced towing technique to enable sampling over the water column to depths in excess of 200 m. This presentation discusses the challenges encountered with indirect acoustic sampling and the techniques implemented to meet them. The system to be discussed operates at frequencies between 38 kHz and 2 MHz, using either FM slide (chirp) or traditional CW acoustic pulses.

3:20

5pAO9. Pulse compression processing of zooplankton echoes. Joseph D. Warren (MIT/WHOI Joint Prog. in Oceanogr. and Ocean Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, jwarren@whoi.edu), Timothy K. Stanton, Dezhang Chu (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Duncan E. McGehee (Tracor Appl. Sci., San Diego, CA 92123-4333)

A pulse compression process was used to analyze echoes from individual animals of three distinct morphological groups of zooplankton. The groups studied were fluidlike animals (Euphausiids, decapod shrimp), elastic-shelled (Gastropods), and animals containing a gas inclusion (Siphonophores). A broadband chirp signal was used to insonify freshly collected animals that were then tethered in a seawater tank either in the laboratory or on the deck of a ship. The decapod shrimp was tethered to a stepper motor and was rotated in 1-deg increments. The other experiments had some animals fixed and others moving freely. The pulse compression processing of the echoes from the animals temporally resolved multiple returns from an individual. Using existing scattering models [Stanton *et al.*, J. Acoust. Soc. Am. (to be published)], the temporal and amplitude differences of the returns were used to estimate size and orientation. Estimates of the size of the animal were made for the Euphausiids, Gastropods, and Siphonophores, while orientation estimates were made for the decapod shrimp. These results agreed well with the actual dimensions of the animals.

3:35

5pAO10. Multistatic, multifrequency scattering from zooplankton. C. F. Greenlaw, D. V. Holliday, and D. E. McGehee (Tracor Appl. Sci., Inc., 4669 Murphy Canyon Rd., San Diego, CA 92123)

Very satisfactory results have been obtained in estimating size abundances of small zooplankton such as copepods from inversion of multifrequency backscattering measurements. Application of this method has become almost routine in many situations. Fluctuations in the scatterer physical properties (density and compressibility) tend principally to cause errors in the abundance estimates whereas size estimates are largely unaffected. The model commonly employed for small crustacean scatterers, a truncated-mode version of the Anderson fluid sphere model, predicts bistatic scattering as well as monostatic backscattering. The behavior of bistatic scattering at angles approaching 90 deg predicted by this model shows that scattering spectra at these angles are quite sensitive to physical properties. This suggests that simultaneous measurements at multiple frequencies and one or more off-axis angles might permit estimation of the physical properties of zooplanktonic scatterers as well as their sizes. Preliminary modeling has been done for a single fluid scatterer to illustrate

the potential for simultaneous estimation of size, density, and compressibility of zooplanktonic organisms in a laboratory setting. The character of the bistatic scattering may argue for nontraditional methods for inverting this sort of data.

3:50

5pAO11. Laboratory studies and theoretical modeling of bistatic scattering of fish. Li Ding (VITech Innovative Res. and Consulting, Victoria, BC V8P 3MB, Canada) and Zhen Ye (Natl. Central Univ., Taiwan, R.O.C.)

Bistatic scattering of fish, in particular forward scattering (scattering in the forward direction), has been identified as a potential means of detecting fish, supplementary to the traditional backscattering. A series of laboratory experiments has been conducted over the past 2 years to measure bistatic scattering characteristics of fish, including forward scattering strength and scattering patterns around the forward direction. It is observed that the forward scattering strength increases rapidly with acoustic frequency and is much stronger than the backscattering strength. The measured scattering patterns (at 120 and 200 kHz) suggest that the scattering is strongest in the forward direction when the incident wave is normal to the fish, and drops by 10–20 dB when the scattering angle deviates from the forward direction by a few degrees. In the latter case, the scattering strength is still much stronger than backscattering strength. These results are compared with a recently developed deformed cylinder model, and with a shelled cylinder model currently under development. [Work supported by the Science and Technology Agency of Japan, the Department of Fisheries and Oceans of Canada, and the National Central University of Taiwan.]

FRIDAY AFTERNOON, 26 JUNE 1998 EAST BALLROOM B (S), 1:30 TO 4:50 P.M.

Session 5pBV

Bioresponse to Vibration/Biomedical Ultrasound and Physical Acoustics: Lithotripsy II

Robin O. Cleveland, Cochair
Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Andrew J. Coleman, Cochair
Department of Medical Physics, St. Thomas Hospital, Lambeth Palace Road, London SE1 7EH, England

Invited Papers

1:30

5pBV1. Shock wave measuring techniques in liquids. Wolfgang Eisenmenger (Physikalisches Inst., Univ. of Stuttgart, Pfaffenwaldring 57, 70550 Stuttgart, es@pi1.physik.uni-stuttgart.de)

The parameters of shock waves in liquids as used for lithotripsy have been measured with needle-, membrane-, and fiber optic probe hydrophones. The true rise times of planar shock waves are determined by surface detection or by the transient response of piezoelectric crystals. The different measuring techniques are compared with respect to their suitability for the determination of shock wave parameters, such as risetime, pulse width, positive and negative pressure amplitudes, and also with respect to aging, calibration sensitivity lifetime, etc. The fiber optic probe hydrophone appears as an optimum high precision standard for calibrated shock wave measurements in lithotripsy. Finally, aspects of the shock parameters with respect to the stone destruction efficiency are discussed.

1:50

5pBV2. Full-wave modeling of lithotripter fields. Eckard Steiger (Inst. für Hoechstfrequenztechnik und Elektronik/Akustik, Univ. of Karlsruhe, D-76128 Karlsruhe, Germany, eckard.steiger@etec.uni-karlsruhe.de)

The propagation of extracorporeal shock wave lithotripter pressure pulses is considered. A computational modeling must account for the effects of diffraction, nonlinear steepening and generation, and propagation of shock waves. Some commercial lithotripters are reflector focusing systems. Therefore, reflections at curved boundaries have to be considered. For specific interests in the propagation through and the interaction with human tissue structures, refraction, absorption, and scattering must be included. A full-wave model based on a set of nonlinear acoustic time domain equations is applied. This system is discretized to obtain a finite difference representation of broad bandwidth and low dispersive character. To enable an accurate shock wave computation an algorithm that modifies the temporal and spatial resolution adaptively is developed. Exemplary a reflector focusing lithotripter is modeled. To validate the model resulting computed wave profiles are compared with measured ones. The measurements are performed using a fiber-optic probe hydrophone. Excellent agreements show the method to provide an accurate and flexible tool for field predictions of complex devices. Characterizing parameters like −6-dB focal regions and energy quantities can be obtained, too. A human tissue model is applied to obtain predictions of the relevant parameters inside the human body.

2:10

5pBV3. Acoustic shock-wave induced cavitation: A comparison of theory and experiment. Andrew J. Coleman and Mark D. Cahill (Medical Phys. Directorate, Guys' & St. Thomas' Hospital Trust, Lambeth Palace Rd., London SE1 7EH, UK, a.coleman@umds.ac.uk)

A calibrated membrane hydrophone has been used to make measurements of pressure in a cavitating field generated in water by a single 30-μs duration ultrasound pulse with a 0.2-MHz center frequency and a peak negative pressure (p_-) variable from 0.5 to 5 MPa as employed in clinical lithotripsy. The acoustic emission from cavitation in the vicinity of the hydrophone is identified as the high-frequency ($>$1 MHz) component of the waveform. This ''cavitation component'' is compared with theoretical predictions of the pressure radiated by a bubble made using the Gilmore model [C. C. Church, J. Acoust. Soc. Am. **86**, 215–227 (1989)]. The model was run for initial bubble radii of 1–100 μm using the measured pressure waveform to drive the radial dynamics of the bubble. The theory predicts a transition from completely driven dynamics at p_- below 3 MPa to largely undriven dynamics above 3 MPa. This transition, which is substantially independent of bubble size, is experimentally observed and occurs at a measured p_- of 3 ± 0.4 MPa. The time between the pressure pulse and the first undriven inertial collapse is predicted to increase approximately twice as rapidly with increasing amplitude as is measured. A mechanism resulting in shorter than predicted collapse times is discussed. [Work supported by the Medical Research Council, UK.]

2:30

5pBV4. A new shock pressure waveform to amplify transient cavitation effect. Dominique Cathignol (INSERM U281, 151 Cours Albert Thomas, 69424 Lyon Cedex 03, France, cathignol@lyon151.inserm.fr), Jahangir Tavakkoli, Françgoise Chavrier, Alain Birer, Alex Arefiev, and Mathieu Prigent

Shock-wave-induced tissue damage and/or stone disintegration are greatly dependent on cavitation effect. In order to control this effect in tissue, a new piezocomposite shock-wave generator (200 mm in diameter and focused at 190 mm) has been developed with the capability of producing two different kinds of shock pressure-time waveform at its focus (type 1: a tensile wave followed by a compressive wave, and type 2: a compressive wave followed by a tensile wave). The cavitation effects produced by these two pressure waveforms were compared in the following experiments: (1) measurement of cavitation bubble collapse time (defined as the time difference between the shock wave arrival and the main bubble collapse), (2) study of cavitation-induced destructive effects in the agar gel, (3) study of disintegration rate on plaster balls (a kidney stone-mimicking material), and (4) study of cavitation-induced destructive effects on rabbit liver *in vivo*. All of these experiments revealed the amplified effect of transient cavitation produced by the type 1 pressure-time waveform. Comparisons with a computed simulation using the Gilmore Akulichef model, taking into account the rectified diffusion, confirm the experimental results. It is concluded that the temporal characteristics of the shock pressure waveform have a major influence on transient cavitation.

2:50

5pBV5. Stone tracking with time-reversal techniques. J. L. Thomas and M. Fink (Laboratoire Ondes et Acoustique, Ecole Superieure de Physique et de Chimie, Industrielles de la Ville de Paris, Universite Denis Diderot, 10 rue Vauquelin, 75231 Paris Cedex 05, France)

Time reversal of ultrasonic fields allows a very efficient approach to focusing pulsed ultrasonic waves through lossless inhomogeneous media. Time reversal mirrors (TRM) are made of large transducer arrays, allowing the incident acoustic field to be sampled, time reversed and reemitted. Time reversal processing permits a choice of any temporal window to be time reversed, allowing operation in an iterative mode. In multitarget media, this process converges on the most reflective target, i.e., the dominant scatterer. The time reversal process is applied to track, in real time a moving gall bladder or kidney stone embedded in its surrounding medium. The feasibility of adaptative beam forming techniques to track the stone during a lithotripsy treatment is investigated. It is shown that TRM allows sharp focusing on one bright point of the stone. Once the bright point is selected, a time of flight profile is determined and used in a least-mean-square method to calculate the spatial coordinates of the stone. Stone trajectories can be tracked by this technique at 30 Hz.

3:10–3:25 Break

3:25

5pBV6. On a mechanism of target disintegration at a shock wave focusing in ESWL systems. Valery K. Kedrinskii (Lavrentyev Inst. of Hydrodynamics, Lavrentyev prospect 15, Novosibirsk 630090, Russia)

It is well known how the lithotriptors operate and the discussion concerning the mechanisms of target disintegration are continued. Meanwhile, to understand these mechanisms means to obtain a possibility to influence on the fracture process. The spall mechanism, as a result of acting with the shear forces, the crack growth inside a stone, and an erosion mechanism (impacts of cumulative microjets forming at bubble collapse in the cavitation zone near a target wall), were suggested by Sturtevant, Kuwahara, Delius, Takayama, Groenig, Crum, and others to explain the disintegration effects observed. At the same time, all authors do not deny a rise of the cavitation zone in the vicinity of a focus under the action of a negative pressure phase in ''a tail'' of a shock wave. However, the role of the cavitation cluster is interpreted in different ways assuming that this effect is a secondary one. This paper proposes an approach to the estimation of the disintegration mechanism, which is based on the development and dynamics of the cavity cluster in a focus vicinity. Within the framework of a two-phase model of cavitating liquid the problem of bubbly cluster development and its collapse on a solid wall are considered. The experimental data on a simulation of cavity cluster pulsation on a target and the idea that hydraulic impacts generated by cluster can determine the mechanics of target disintegration in a focusing zone of shock wave and its rarefaction phase are discussed.

3:45

5pBV7. First *in vitro* experiments using a new reflector to concentrate shock waves for extracorporeal shock wave lithotripsy. Achim M. Loske and Fernando E. Prieto (Inst. Física, UNAM, México D.F., México)

Electrohydraulic lithotripters to perform extracorporeal shock wave lithotripsy (ESWL) treatments use symmetrical reflectors in order to concentrate the energy generated in one of their foci ($F1$). The new reflector presented here was designed in an attempt to increase the efficiency of ESWL. It is obtained by combining sectors of two rotationally symmetric ellipsoidal reflectors with different separations between their foci $F1$ and $F2$. These sectors are joined together in such a way that the $F1$ foci coincide, creating a separation between the $F2$ foci. The purpose is to temporally and spatially dephase the shock wave generated at $F1$. This is achieved because the initial shock wave is divided, by reflection, into two shock fronts converging toward two, rather than one, $F2$ foci. Pressure measurements, and *in vitro* experiments using kidney-stone models and dog kidneys, indicate that the new design could be more efficient in breaking up renal stones, without producing more tissue damage. The performance of the new reflector was compared with one having the geometry of a Dornier HM3 (nonmodified) reflector. Both reflectors were installed separately in an electrohydraulic experimental shock wave generator (MEXILIT II). [Work supported by DGAPA IN502694, UNAM.]

Contributed Papers

4:05

5pBV8. Use of two pulses to control cavitation in lithotripsy. Michael R. Bailey (Appl. Phys. Lab, 1013 NE 40th St., Seattle, WA 98105), Robin O. Cleveland (Boston Univ., Boston, MA 02215), David T. Blackstock (Appl. Res. Labs., Austin, TX 78713), and Lawrence A. Crum (Appl. Phys. Lab., Seattle, WA 98105)

In lithotripsy acoustic cavitation appears to play a role in both comminution of kidney stones and damage to tissue. To investigate the role of cavitation, we used a pair of confocal electrohydraulic lithotripters (ellipsoidal reflectors). A delay in firing of Δt modified the cavitation field without altering the peak positive and negative pressures. Bubble collapse time t_c, measured by passively detecting acoustic emissions generated by the bubble, was used as a measure of cavitation activity. A two-pulse sequence from the source shortened t_c relative to the single pulse value t_{c1}. Collapse times calculated by using the Gilmore equation agreed well with measurements. Relative intensity of bubble collapse was determined by depth of pits created in aluminium foil. For pulse delays $\Delta t < 0.3t_{c1}$ the pit size decreased, a result that implies the second pulse stifled the collapse. Conversely, for delays in the range $0.5t_{c1} < \Delta t < t_{c1}$, cavitation apparently intensified, since pit depth increased by a factor of up to 2. Pit depth correlates well with (calculated) pressure radiated by the bubble. When the two reflectors were pointed at each other and fired simultaneously, pitting was localized to a spot less than 1 cm in diameter. [Work supported by ONR, ARL:UT IR & D, and NIH PO1-DK43881.]

4:20

5pBV9. Detection and control of lithotripsy-induced cavitation in blood. Brenda Jordan (Dept. of Biol., Tennessee State Univ., Nashville, TN 37209-1561), Michael R. Bailey (Appl. Phys. Lab., Seattle, WA 98105), Robin O. Cleveland (Boston Univ., Boston, MA 02215), and Lawrence A. Crum (Appl. Phys. Lab., Seattle, WA 98105)

Monitoring and controlling cavitation offer ways of isolating the role of cavitation in lithotripsy-induced cell damage. Finger cots containing 7 ml of PBS-diluted blood (3% hematocrit) were treated by 100 lithotripter pulses focused by either a conventional rigid reflector or a pressure-release reflector. The rigid reflector produced 7% hemolysis at 18 kV. The pressure-release reflector, which produces an equally strong acoustic pulse but less intense cavitation, produced 1% hemolysis which was indistinguishable from control levels. A passive cavitation detector consisting of two confocal focused transducers operating in coincidence was used to monitor cavitation within each blood sample. With the rigid reflector, each transducer recorded a spike when pre-existing bubbles were collapsed by the positive portion of the lithotripter pulse and a second spike when the bubbles, caused to grow by the negative portion of the pulse, underwent inertial collapse. The time between spikes was 1.25–1.5 times longer in blood than in water. The results appear to indicate that transient cavitation is an important mechanism in lithotripsy-induced hemolysis, that lithotripter-induced cavitation can be monitored in cell samples, and that bubble dynamics in blood and water are not identical. [Work supported by NIH PO1 DK43881, UW STAR, and DARPA/ONR.]

4:35

5pBV10. Experiments on the relation of shock wave parameters to stone disintegration. Thomas Dreyer, Rainer E. Riedlinger, and Eckard Steiger (Inst. für Hoechstfrequenztechnik und Elektronik/Akustik, Univ. of Karlsruhe, Kaiserstr. 12, D-76128 Karlsruhe, Germany, thomas.dreyer@etec.uni-karlsruhe.de)

The sound field of different focusing piezoelectric transducers designed for lithotripsy was investigated. The shock wave parameters according to proposed standards, e.g., FDA Draft 1991, were determined. The parameters achieved are based on measurements using a fiber-optical probe hydrophone. In contrast to PVDF-based hydrophones, this hydrophone is able to give a correct representation also of the tensile components of a complete signal and provides a higher spatial resolution. Thus some shock wave parameters like beam energy can be calculated more precisely. The different lithotripter pulses were applied to model stones *in vitro*, recording the amounts of removed material. Fragmentation results were compared with the different parameters. Attempts were made to arrange these parameters according to their relevance to disintegration. The evaluation of these experiments revealed up to now that there is no significant relation between some of the shock wave parameters and fragmentation efficiency. It is concluded that most of the proposed parameters do not describe well the efficacy of different lithotripter sources on concrements.

Session 5pEAa

Engineering Acoustics: Scattering and Radiation

Sung-Hwan Ko, Chair
Naval Undersea Warfare Center, Code 2133, Newport, Rhode Island 02841

Contributed Papers

1:00

5pEAa1. Scattering by a horizontal strip on a hard side wall. Djamel Ouis and Sven Lindblad (Eng. Acoust., LTH, Lund Univ., P.O. Box 118, S-221 00, Lund, Sweden, djamel.ouis@kstr.lth.se)

The study of the effect of a horizontal hard striplike element on a side wall is presented. This may give an estimate of the usefulness of such elements in reinforcing the early sound components in a listening room. The theoretical calculations are based on geometrical assumptions in combination with a diffraction model to take into account the diffraction by the edge of the scattering element. This latter is based on the time-domain theory of Biot and Tolstoy for the diffraction by a hard wedge applied to the half-plane. An accurate and useful approximation for the diffracted field in the frequency domain using Fresnel integrals is presented. Some measurements on a scale model are also presented.

1:15

5pEAa2. Acoustic wave scattering from a coated cylindrical shell. Sung H. Ko (Naval Undersea Warfare Ctr. Div., Newport, RI 02841)

A theoretical model was developed to evaluate the effect of a coating on a cylindrical shell. The coating is designed to reduce the flexural wave noise generated by the vibration of the cylindrical shell. In underwater applications, this coating is called an inner decoupler. The outer surface of the composite structure, which consists of the cylindrical shell and the inner decoupler, is in contact with water; the core (cavity) of the structure contains air. The analysis of the problem is made for a double-layer cylindrical structure that is immersed in water. The theory of elasticity, the elastic wave propagation, the acoustic wave propagation, and pertinent boundary conditions are used in formulating the problem. The two-dimensional problem is treated in this study. This implies that all the quantities are independent of the axial coordinate. Numerically calculated results for both coated and uncoated cylindrical shells are presented to compare these two directivity patterns.

1:30

5pEAa3. Numerical solution of the linearized Euler equations with high-order centered schemes. John A. Ekaterinaris (Nielsen Eng. and Res., 526 Clyde Ave., Mountain View, CA 94043-2212)

The unsteady, linearized Euler equations governing propagation of acoustic disturbances in the presence of a uniform free stream or sound scattering from solid surfaces are solved numerically. The governing equations are discretized in space using the standard fourth- and sixth-order central-difference schemes, and compact finite-difference schemes with spectrallike resolution are also used. The semidiscrete system of equations obtained after space discretization is marched in time using high-order, Runge–Kutta schemes. Numerical boundary conditions are applied at the far field in order to limit the extent of the computational domain. Solutions are computed first on Cartesian, equally spaced grids. Computed results for propagation of a pulse in a uniform medium are compared with existing analytic solutions. Next, sound scattering from a cylinder surface is computed using a nonuniform body-fitted numerical mesh. The effects of grid spacing, time accuracy, the method used for space discretization, and the boundary conditions on the computed solutions are investigated.

1:45

5pEAa4. Modeling acoustic radiation from end-initiated explosive line charges. William J. Marshall, Jr. (BBN Technologies, Union Station, New London, CT 06320, bmarshall@bbn.com)

Acoustic radiation produced by end-fired explosive line charges has been examined experimentally and theoretically. In terms of directivity effects, line charges are found to behave like beam-steered continuous arrays over moderately wide bands. This talk briefly summarizes measured waveforms and derives a math model which predicts levels and spectrum shapes from such sources over a wide range of view angles and source characteristics. The model is driven by two parameters which may be extracted from experimental data. The result is a practical means of designing explosive line arrays to desired source levels, bandwidths, and beam patterns. [Work supported by DARPA.]

2:00

5pEAa5. Resonant frequencies in degeneration of hybrid longitudinal and torsional vibration system. LiQun Zhang and TieYing Zhou (Dept. of Modern Appl. Phys., TsingHua Univ., Beijing, PROC)

The Langevin vibrator is made of the longitudinal and torsional piezoelectric plates, which can be used for an ultrasonic motor and an ultrasonic welding transducer. Because of the difference between the longitudinal and torsional sound velocity, their resonant frequencies are not equal. To make resonant frequencies of two vibration modes in degeneration, a matching block with a thin neck was connected to the end of the vibrator. Two calculation methods, a dimension network which transmits matrixes and a finite element method have been used. The feasibility of frequencies tuning through a matching block has been analyzed and how to effect the resonant frequency by the radius and thickness of the thin neck and matching block has been summarized. The resonant frequencies of several steel cylinders with variant thin necks and matching blocks were calculated. Some experimental results of measuring the resonant frequencies have been compared with the calculation. According to the calculation and experimental results, a prototype of hybrid longitudinal-torsional ultrasonic motor has been manufactured, and it can be operated very well [Tomikawa *et al.*, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39** (1992)].

2:15

5pEAa6. The estimation of an interior sound field by source methods. Gee-Pinn James Too and Shing Maw Wang (Dept. of Naval Architecture and Marine Eng., Natl. Cheng Kung Univ., 1 Ta-Hsueh Rd., Tainan, Taiwan 70101, R.O.C.)

In an earlier study, three different source methods (SSM, similar source method; IPSM, internal parallel source method; ISM, internal source method) were successfully developed to estimate radiation and scattering sound field. All these are methods of estimating exterior sound fields which are not suitable for estimating interior sound fields. In this study, the similar source method is modified to estimate the interior sound field. The modification is to move the sources to somewhere outside the boundary surface in order to estimate the interior sound field and to prevent the singularity problem when the source point is close to the field point. Finally, the sound field inside a 2D infinite cylinder and in a car will be estimated. The results will be compared with the results obtained by the boundary element method.

FRIDAY AFTERNOON, 26 JUNE 1998 EAST BALLROOM A (S), 3:00 TO 4:30 P.M.

Session 5pEAb

Engineering Acoustics: Acoustic Waveguides

Anthony A. Atchley, Chair
Graduate Program in Acoustics, Pennsylvania State University, P.O. Box 30, Applied Science Building, State College, Pennsylvania 16804

Contributed Papers

3:00

5pEAb1. Insertion loss measurements of an acoustical enclosure by using Intensity and MLS methods. Pedro Cobo, Carlos Ranz, Salvador Santiago, Jose Pons, Manuel Siguero, and Carmen Delgado (Instituto de Acustica, CSIC, Serrano 144, 28006 Madrid, Spain, iacpc24@fresno.csic.es)

The Insertion Loss (IL) is one of the most appropriate descriptors for the acoustical performance of an enclosure. The IL is conventionally measured by comparing either sound power of space-averaged sound-pressure measurements before and after enclosing the sound source. In this paper, a comparison is made between a power-based IL measurement, using a sound intensity meter, and a pressure-based IL measurement, using the MLS method. Special attention is paid to practical and technical aspects of both methods, such as time consuming, low-frequency limit, signal-to-noise ratio, and background noise immunity. [Work supported by CICYT, Project AMB97-1175-C03-01.]

3:15

5pEAb2. Intensity measurements in various rooms: A new intensity probe. W. F. Druyvesteyn (Philips Res. Labs., Eindhoven, The Netherlands) and H. E. de Bree (MESA Res. Inst., The Netherlands)

The main advantage of acoustic intensity measurements is that in a reverberant room the direct and the reverberant sound field can be determined. This property has been tested using the Bruel & Kjaer p-p probe in four different rooms (anechoic, reverberant and two listening rooms). As compared to the results in the anechoic room, small deviations, which can be positive and negative, were found. The deviations can be explained using a model of early reflections. A new intensity probe consisting of a microphone and a sound velocity sensor has been tested and compared with the Bruel & Kjaer p-p probe. The sound (particle) velocity sensor contains the Microflown, a novel micro-machined sensor. This silicon-based velocity microphone operates on a thermal principle. The dimensions of the Microflown are on the order of a cubic millimeter. For symmetry considerations, however, the sensors were packaged in a quarter inch package. Experiments were done in the anechoic and in the reverberation room. A good agreement was found.

3:30

5pEAb3. *In situ* absorption measurements using a transfer function technique and MLS. G. Dutilleux, U. R. Kristiansen, and T. E. Vigran (Norwegian Univ. of Sci. and Technol., NTNU, Acoust. Group, Dept. of Telecommunications, N-7034 Trondheim, Norway, tor.erik.virgran@tele.ntnu.no)

The transfer function (or two-microphone) technique is a tool for measuring materials absorption related parameters as impedance or absorption coefficients. But this method is laboratory bound in its original form due to free field assumption and practical issues. It is shown that after proper modifications the transfer function technique can be used successfully in noisy and reverberating fields and work well at low frequency even in small measurement rooms, becoming therefore a really *in situ* measurement technique. A wide frequency range implementation is presented, based on a two-channel MLS measurement system, ''energy ratio invariant'' time windowing, a basic specular reflection propagation model with spherical decoupling, and a particular geometrical configuration. The measurement method is tested on a wide range of absorbers. For impedance measurements using infinite surface propagation models the uniqueness of the inversion procedure is investigated using the conformal mapping technique. In addition, faster inversion and model computation routines are proposed. Measurement results obtained under realistic *in situ* conditions are provided and compared with reference methods.

3:45

5pEAb4. Lumped impedance of a planar discontinuity in an acoustic waveguide. Ralph T. Muehleisen (Dept. of Civil, Environ., and Architect. Eng., Univ. of Colorado, Boulder, CO 80309) and Anthony A. Atchley (Penn State Univ., P.O. Box 30, State College, PA 16804)

When modeling planar discontinuities in acoustic waveguides at low frequencies, the standard assumptions involve continuity of pressure and volume velocity at the discontinuity. While continuity of volume velocity is required by mass conservation, excitation of evanescent higher-order modes and thermoviscous loss suggest that pressure is not continuous across the discontinuity. Such a pressure loss can be modeled by a series acoustic impedance. A new model for a lumped series impedance at a planar waveguide discontinuity has been developed and is presented. The model uses a modal expansion of the acoustic field near the discontinuity and matrix techniques to obtain a closed-form approximation for the re-

sistive and reactive elements of the series impedance. As an example of the accuracy of the model, the prediction of the resonance frequency of a constricted annular resonator is presented. [This work was supported by the Office of Naval Research and the American Society of Engineering Education.]

4:00

5pEAb5. A short water-filled pulse tube for the measurement of the acoustic properties of materials at low frequencies. Debra M. Kenney (Carderock Div., Naval Surface Warfare Ctr., 9500 MacArthur Blvd., West Bethesda, MD 20817, kennde@oasys.dt.navy.mil) and Peter H. Rogers (Georgia Inst. of Technol., Atlanta, GA 30332)

A method of measuring the acoustic properties of materials at low frequencies in a water-filled pulse tube, which was a fraction of a wavelength long, was investigated. The method employed active sound cancellation at the end of the tube opposite from the sample, and signal processing of two hydrophone signals to separate the signal incident on the sample from its reflection. Two analytical models were developed and an experimental tube was built to examine the feasibility of such measurements. The results from the simulation and experiment suggest that the short water-filled pulse tube would be useful for measuring compliant materials at moderately low frequencies (down to 300 Hz), but not for measuring high impedance materials or for making measurements at very low frequencies. The simulation and experiment showed that the thicker the sample, relative to a wavelength, the lower the sensitivity to experimental errors. Fortunately, many of the materials of interest in underwater acoustics are compliant materials with relatively slow phase velocities and consequently shorter wavelengths. The impedance of the backing behind the sample also influences the sensitivity of results to errors in the measurements. [Work supported by ONR and NSF.]

4:15

5pEAb6. The generation of low-frequency acoustic waves in electrolytes. Bronislaw Zoltogorski (Wroclaw Univ. of Technol., Inst. of Telecommunication and Acoust., Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland)

The phenomenon of the generation of acoustic waves in liquids based on interaction of alternating ion current and constant magnetic field is analyzed. The properties of such a magneto-electric transducer having no mechanical vibrating elements is examined theoretically (with the use of an electric equivalent circuit) and experimentally. The following measurements were carried out: (1) amplitude of conductivity current as a function of frequency, (2) amplitude of acoustic pressure at the distance of 0.5 m without constant magnetic field and (3) with magnetic field. For water from a municipal water-pipe network, the upper limit frequency of the current characteristic is 4 to 5 kHz. The same frequency properties exhibit a pressure characteristic. The directional pattern of the source depends on the shape of the electrodes and can be optimized for special aims. The energetic efficiency of the transducer is not great because the majority of the electric power is converted into heat and causes an increase in temperature of the liquid. That process is relative to a big value of resistance of weak electrolytes. That type of transducer can be applied in liquids without ion conductivity by submerging a container with very thin rubber walls and with mounted electrodes and filling with an electrolyte.

FRIDAY AFTERNOON, 26 JUNE 1998 ASPEN ROOM (S), 12:55 TO 6:00 P.M.

Session 5pPAa

Physical Acoustics: Sonochemistry and Sonoluminescence: SL II

R. Glynn Holt, Cochair
Department of Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Thomas J. Matula, Cochair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

Chair's Introduction—12:55

Invited Papers

1:00

5pPAa1. Energy focusing in bubbly flows. Seth Putterman, Keith Weninger, Robert A. Hiller,[a] and Bradley P. Barber[b] (Phys. Dept., Univ. of California, Los Angeles, CA 90095)

Sonoluminescence, cavitation damage at surfaces, and cavitation in accelerating flows are realizations of spectacular levels of energy focusing in nature. In a resonant sound field a single trapped bubble of gas can focus the ambient sound energy by 12 orders of magnitude to generate a clocklike string of picosecond flashes of ultraviolet light. [Barber *et al.*, ''Defining the unknowns of sonoluminescence,'' Phys. Rep. **281**, 65 (1977)]. In more complicated geometries a high level of sound leads to the formation of hemispherical bubbles attached to an exposed surface. These bubbles also emit light and in addition damage the surface. Measurements show that the pulsation of these bubbles maintains the hemispherical symmetry [Weninger *et al.*, ''Sonoluminescence from an isolated bubble on a solid surface,'' Phys. Rev. E **56**, 6745 (1997)], thus raising the question as to whether cavitation damage is due to (micro)jets or imploding (hemispherical) shock waves. Finally, flow through a Venturi tube generates a stream of bubbles which also emit subnanosecond flashes of light [F. B. Peterson and T. P. Anderson, Phys. Fluids **10**, 874 (1967)]. Luminescence from an isolated trapped bubble in water seems to work well with any noble gas, whereas luminescence from cavitating flows and surface bubbles is quite dependent on xenon [argon bubbles appear to give no light at all]. The width of the SL flash [Gompf *et al.*, Phys. Rev. Lett. **79**, 1405 (1997), Hiller *et al.*, Phys. Rev. Lett. **80**, 1090 (1998)] has been found to be independent of wavelength suggesting that

light is emitted from a new high energy phase of matter-probably a cold dense nano-plasma. The key unknowns of SL are the size and temperature of the hot spot from which the light is emitted. Experiments aimed at measuring these quantities will be discussed. [Research supported by the NSF.] [a)]Present address: CMS, Los Alamos National Laboratories, Los Alamos, NM. [b)]Present address: Lucent Technologies, Murray Hill, NJ.

1:20

5pPAa2. Sonoluminescence stability for gas saturations down to 0.01 Torr. Robert E. Apfel and Jeffrey A. Ketterling (Dept. of Mech. Eng., Yale Univ., 9 Hillhouse Ave., New Haven, CT 06520-8286)

Water solutions with very small argon partial pressures were investigated. The experiments were performed in a sealed cell. The argon was used in both a pure form and in mixtures with other gases. A regime of stable sonoluminescence was seen at lower driving pressures for argon partial pressures down to 0.01 Torr. Mie scattering, light pulse measurements, and visual observations over tens of thousands of cycles were used to determine if the bubble was stable. As the drive pressure was increased, the bubble moved into a slow time scale instability on the order of several seconds which sometimes caused break up. These experiments are discussed and compared to other published data involving gas mixtures at saturation levels up to 1 atm.

1:40

5pPAa3. Evolution of a sonoluminescence bubble under a magnetic field. J. B. Young, H. Cho, and W. Kang (James Franck Inst. and Dept. of Phys., Univ. of Chicago, 5640 S. Ellis Ave., Chicago, IL 60637)

Studies of sonoluminescence (SL) in magnetic fields up to 20 T have revealed a striking magnetic field dependence. The intensity of emitted light is suppressed under increasing magnetic fields and vanishes above threshold magnetic field that depends on the applied sound pressure. Further increase in the magnetic field leads to the destruction of the bubble through dissolution. At a constant magnetic field, the light intensity is found to increase roughly linearly with increasing sound pressure until the bubble disappears above a cut-off pressure which depends on the strength of the magnetic field. The cut-off pressure is found to approximately double between 0 and 20 T. Further study of the SL bubble through the measurement of the phase of the emitted radiation relative to acoustic drive shows a large increase in phase under magnetic fields. These results suggest that the bubble grows significantly in size and that there may be a substantial modification of bubble dynamics under magnetic field. The origin of the observed effects, possibly due to magnetic field-induced anisotropies in the SL bubble, will be discussed in light of our recent anisotropy studies under magnetic fields.

2:00

5pPAa4. Computed spectral and temporal emissions from a sonoluminescing bubble. William C. Moss and David A. Young (Lawrence Livermore Natl. Lab., L-200, 7000 East Ave., Livermore, CA 94551, wmoss@llnl.gov)

Although the mechanism of emission in single bubble sonoluminescence (SBSL) has still not been identified conclusively, the ''plasma'' model [Moss *et al.*, Science **276**, 1398–1401 (1997)] appears to explain more features of SBSL than other models. Recent measurements of the light pulses emitted by SBSL [Gompf *et al.*, Phys. Rev. Lett. **79**, 1405–1408 (1997)] show that the pulse widths may be longer than the original UCLA measurements and can vary between 60 and 250 ps. These new results provide information for improving the modeling of the energy loss and emission mechanisms in the plasma model. In particular, the plasma can be described more accurately as a ''strongly coupled plasma'' (a plasma with a small Debye screening length). The main difference between the strongly coupled plasma and our earlier plasma model is that there is less energy loss by electron thermal conduction in the strongly coupled plasma, which gives rise to longer calculated pulse widths that are consistent with the new experimental data. [This work was performed under the auspices of the U. S. Department of Energy by Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.]

2:20

5pPAa5. Gas dynamics of mixtures within sonoluminescence bubbles. Andrew Szeri and Brian Storey (Dept. of Mech. Eng., Univ. of California—Berkeley, Berkeley, CA 94720-1740)

Sonoluminescence experiments are often conducted using mixtures of gases intentionally dissolved in the liquid, and, presumably, resident in the bubble interior. Even when only a single gas is dissolved in the host liquid, vapor is undoubtedly present within the bubble in addition to the intended gas. Hence, all sonoluminescence experiments can be assumed to involve gas mixtures even though the contents of the bubbles cannot be directly assayed. Two constituents of a gas mixture will have differing solubility in the host liquid, which complicates the amount of each species one would expect is resident within a bubble, depending on how the experiment was prepared. Moreover, there is little opportunity for mixing of the bubble contents due to the highly spherically symmetric nature of single-bubble sonoluminescence. Using the methods of computational fluid dynamics, the fate of gas mixtures within a strongly driven bubble is investigated. [Work was supported by a grant from the National Science Foundation.]

2:40–3:00 Break

3:00

5pPAa6. Time-resolved measurements of optical emission from sonoluminescence. D. Froula, R. W. Lee, W. C. Moss, P. E. Young (Lawrence Livermore Natl. Lab.), and T. J. Matula (Univ. of Washington, Seattle, WA 98105)

The results of detailed measurements of sonoluminescence from a gas bubble that is acoustically driven in deionized water are presented. The pulse width and relative jitter of the optical flash and the minimum and ambient radius of the bubble as a function of the driving pressure, temperature, and gas concentration of the water are measured. The optical flash is collected by an f/1.4 lens, and then optically relayed onto the slit of a Hamamatsu C1587 optical streak camera which records the time evolution of the flash with picosecond resolution. Single event streaks are obtained, showing FWHM pulse widths of 200–400 ps, comparable to recently reported results [Gompf *et al.*, Phys. Rev. Lett. **79**, 1405 (1997)]. The jitter of the flash between events has been found to be a minimum of ± 200 ps. Mie scattering has been used to measure the maximum and ambient bubble radius at the same time for comparison to the model [Moss *et al.*, Science **276**, 1398 (1997)]. The consequences of these results on modeling will also be discussed. [This work was performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.]

3:15

5pPAa7. Single-bubble sonoluminescence: Acoustic emission measurements with a fiber-optic probe hydrophone. Bruno Gompf, Zhaoqiu Wang, Rainer Pecha, Wolfgang Eisenmenger (1. Physikalisches Inst., Univ. of Stuttgart, Pfaffenwaldring 57, D-70550 Stuttgart, Germany), and Ralf Guenther (Natural and Medical Sci. Inst., Eberhardtstr. 29, D-72762 Reutlingen, Germany)

In single-bubble sonoluminescence additionally to the short light pulses, the bubble emits in the collapse phase a sound wave which can be measured with a fiber-optic probe hydrophone [Staudenraus and Eisenmenger, Ultrasonics **31**, 267 (1993)]. This type of hydrophone is an absolute ultrasonic wideband reference standard. Compared to piezoelectric hydrophones it allows measurements with higher spatial (0.1 mm) and temporal (10 ns) resolution. The intensity and the width of the emitted sound wave increase with increasing driving pressure, but vary only slightly with water temperature in contrast to the emitted light intensity. The total radiated energy of the sound wave is below 10% of the initial energy of the bubble. The results are compared with earlier measurements on transient cavitation bubbles and with new theoretical results.

3:30

5pPAa8. Luminescence from spherically and aspherically collapsing laser-induced bubbles. Claus-Dieter Ohl, Olgert Lindau, and Werner Lauterborn (Drittes Physikalisches Institut, Univ. of Göttingen, Bürgerstr. 42-44, D-37073 Göttingen, Germany, ohl@physik3.gwdg.de)

Luminescence from laser-induced cavitation bubbles (SCBL) is investigated experimentally. This experimental technique offers several advantages over the acoustical driving technique for luminescence studies: The optical resolution of the imaging system allows the investigation of the shape of the luminescence spot. This technique facilitates the study of the spherical asymmetries during the bubble collapse phase. Finally, complications due to gas diffusion and Bjerknes forces are avoided. Comparison with theory gives good agreement between the size of the luminescence spot and the observed minimum bubble size for spherical symmetry. A rigid boundary greatly influences the light output; SCBL is only observed for a mildly aspherical bubble collapse.

3:45

5pPAa9. On the theory of supercompression of a gas bubble in a liquid-filled flask. Robert I. Nigmatulin, Iskander Sh. Akhatov, Nelya K. Vakhitova (Ufa (Bashkortostan) Branch of Russian Acad. of Sci., 6 K. Marx St., Ufa, 450000, Russia, iskander@ncan.ufanet.ru), and Richard T. Lahey, Jr. (Rensselaer Polytechnic Inst., Troy, NY 12180-3590)

The spherically symmetric problem of the oscillations of a gas bubble in the center of a spherical flask filled with a compressible liquid that is excited by pressure oscillations on the flask wall is considered. A generalization of the Rayleigh–Plesset equation for a compressible liquid is given in the form of two ordinary difference-differential equations that take into account the pressure waves that are reflected from the bubble and those that are incident on the bubble from the flask wall. The initial value problem for the initiation of bubble oscillations due to flask wall excitation and for the evolution of these oscillations is considered. Linear and nonlinear periodic bubble oscillations are analyzed analytically. Nonlinear resonant and near-resonant solutions for the bubble nonharmonic oscillations, which are excited by harmonic pressure oscillations on the flask wall, are obtained. The influence of heat transfer phenomena on the bubble oscillations is analyzed.

4:00

5pPAa10. Observations of single-bubble sonoluminescence in simulated microgravity and hypergravity. Jeremy E. Young, Nathaniel K. Hicks, A. C. Binner, Susan Richardson, Mark J. Marr-Lyon, and Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814)

Single-bubble sonoluminescence (SBSL) in water occurs when a bubble undergoes highly nonlinear volume oscillations in a sound field. In the presence of gravity, the SBSL bubble experiences a buoyancy force described by $\rho g V(t)$. There is also a Bjerknes force from the sound field. The time average of these two forces balances at the equilibrium location. If the SBSL apparatus is in normal gravity the bubble is displaced from the pressure antinode of the acoustic standing wave [Matula *et al.*, J. Acoust. Soc. Am. **102**, 1522–1524 (1997)]. Because of this displacement, the bubble has a vertical oscillation. One method of exploring the effects of the translational oscillations is to compare the light output of SBSL in microgravity to the light output of SBSL in normal or hypergravity. These were examined by an experiment performed aboard NASA's KC-135A. The airplane performs parabolic maneuvers, giving short periods of microgravity and hypergravity (approximately 0 and 2 g, respectively). The results suggest a correlation of higher light output in microgravity than in hypergravity. [Work supported by Flags Up Feed, Inc., Space Grant Consortium, NASA, and W.S.U.]

4:15

5pPAa11. Analytical results for the conditions of shock formation inside SBSL bubbles. Ralf Guenther (Natural and Medical Sci. Inst., Eberhardstr. 29, D-72762 Reutlingen, Germany)

The mechanism of light emission during single-bubble sonoluminescence is still an unresolved question. Of central importance in this respect is whether a converging shock forms inside the bubble. Using an expansion of the fields inside the bubble, a set of dimensionless equations is derived, which allows exact solutions as well as perturbative treatment during different stages of the bubble collapse. The transition from nonshock to shock solutions is obtained as a function of dimensionless con-

stants characterizing the dynamics. The explicit form of spatial distributions of temperature, velocity, and density is derived and consequences for the light emission are discussed.

4:30

5pPAa12. Shock formation in a sonoluminescing gas bubble. Takeru Yano (Mech. Sci. Dept., Hokkaido Univ., Sapporo, 060 Japan, yano@lebes-me.eng.hokudai.ac.jp)

Whether a shock is formed or not in a sonoluminescing gas bubble is a fundamental problem in the sonoluminescence study. The problem is numerically studied by solving Euler equations of gasdynamics in conjunction with an approximate equation of Keller type for the motion of the bubble radius. A high-resolution TVD upwind finite-difference scheme is used for Euler equations and a fourth-order Runge–Kutta method for a Keller-type equation. The result shows that, owing to the strong nonlinearity of the bubble motion, a slight increase of the driving pressure over a critical value makes the maximum collapse speed of a bubble wall supersonic, and hence the gas behavior changes from a nearly incompressible and adiabatic flow to a compressible one including a shock wave. The relation between the critical driving pressure and other parameters (the driving frequency, the initial bubble radius, etc.) is analyzed. The maximum temperature and pressure attained in the bubble are compared with those in earlier theoretical and numerical works.

4:45

5pPAa13. Thermal wave from a sonoluminescing gas bubble. Yoon-Pyo Lee, Sarng-Woo Karng (Korea Inst. of Sci. and Technol., P.O. Box 131, Cheongryang, Seoul, Korea), and Ho-Young Kwak (Chung-Ang Univ., Seoul, 156-756, Korea)

Heat diffusion from the sonoluminescing gas bubble which produces a high temperature and pressure environment adjacent to the bubble wall was considered by solving the heat diffusion equation numerically. The boundary conditions employed in this calculation are the time-varying bubble wall temperature obtained from the analytical solutions around the collapse point. The thermal boundary layer was found to be decreased rapidly during the collapse as expected. The heat flux at the collapse exceed 3 GW/m^2. After the collapse a solitonlike heat wave, which was generated due to the rapid increase in the bubble wall temperature at/near the collapse for the sonoluminescing gas bubble, was found to propagate through the medium. On the other hand, the liquid temperature near the bubble decreases after the collapse for the nonsonoluminescing case. [Work supported by Korea Science and Engineering Foundation.]

5:00

5pPAa14. Effect of quantum radiation pressure on the bubble dynamics in sonoluminescence. Weizhong Chen and Rongjue Wei (Inst. of Acoust. and State Key Lab. of Mod. Acoust., Nanjing Univ., Nanjing 210093, PROC, wzchen@nju.edu.cn)

In the theoretical frame of quantum vacuum radiation (QVR), the effect of the QVR on the gas bubble dynamics in the single bubble sonoluminescence is studied. The dominated hydrodynamic equation of the gas bubble trapped in a liquid, Rayleigh–Plesset equation, has been modified by considering the reaction of the radiation force on the bubble surface and therefore causing a loss of energy. In the modified model the collapse and expansion of the bubble, especially near the turnaround at minimum radius, related to photon emission instead of independence on it. The numerical computation shows that the emission energy due to the QVR seems too low to explain the experimental data, if some present values of parameters, such as the ambient radius of the bubble and the sound pressure in liquid, are used. The radiation energy due to the QVR, however, can arise by changing those parameters. Based on this, it is concluded that the QVR may be a possible candidate for mechanisms of the sonoluminescence. [Work supported by NSF of China.]

5:15

5pPAa15. Sonoluminescence: The effect of magnetic fields in the Planck theory. T. V. Prevenslik (2E Greenery Court, Discovery Bay, Hong Kong)

Sonoluminescence (SL) observed in the cavitation of water may be explained by the Planck theory of SL that treats the bubbles as collapsing miniature masers having a Planck energy intrinsic to the dimensions of the maser cavity. Microwaves at frequencies proportional to the collapse velocity are created in the bubble wall molecules as the Planck energy at optical frequencies is absorbed. Because of the interaction between the microwaves and molecular dipole moments, the Planck energy accumulates in the electronic state of polar molecules with a rotation quantum state. As the bubble wall is polarized, intense electrical fields develop to charge the bubble surface molecules to visible photon levels. By this theory, SL can only be observed if the maser walls close to optical dimensions thereby allowing the bubble wall molecules to accumulate Planck energy to the 3–6-eV levels necessary to produce VIS-UV photons. Closure to optical dimensions is assured by imposing a net positive collision pressure. The consistency of the Planck theory of SL was assessed with data on the effects of applied magnetic fields on SL. The net pressure was taken as the difference between the stagnation pressure in bringing the liquid bubble walls to rest and the pressure generated by the magnetic field in the presence of a charged bubble wall. Consistency with the Planck theory of SL is found as the SL intensity with magnetic field is parabolic and the peak SL intensity is linear with collapse velocity.

5:30–6:00 Discussion

Session 5pPAb

Physical Acoustics: Nonlinear Acoustics II: 1. Solitons; 2. Biomedical Nonlinearities

Mack A. Breazeale, Cochair
National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677

Lev A. Ostrovsky, Cochair
NOAA/ERL/ETL, 325 Broadway, R/E/ET-1. Boulder, Colorado 80303

Chair's Introduction—1:00

Invited Papers

1:05

5pPAb1. Chaotic dynamics in acoustics. Werner H. Lauterborn (Drittes Physikalisches Inst., Univ. of Göttingen, D-37073 Göttingen, Germany)

The theory developed for the description of chaotic and nonlinear dynamical systems also pertains to acoustics. A survey is given of the new ideas and methods that for experimentalists rely on the embedding of experimentally obtained data (time series) into high-dimensional spaces. The embedding is the starting point of nonlinear time series analysis for determining the static (fractal dimensions) and dynamic (Lyapunov exponents) properties of the point sets (attractors) obtained. Acoustical systems that can be investigated with these methods are, for instance, vibrating liquid surfaces (Faraday experiment), vibrating machines, musical instruments, nonlinear oscillators, speech and hearing, bubble dynamics, and acoustic cavitation.

1:25

5pPAb2. Nonlinearities in the bioeffects of ultrasound. Edwin L. Carstensen (Dept. of Elec. Eng. and Rochester Ctr. for Biomed. Ultrasound, Univ. of Rochester, Rochester, NY 14627)

Biological effects of ultrasound involve nonlinear phenomena (1) in the propagation of sound itself, (2) in the generation of acoustic cavitation, and (3) in the biochemistry, physiology, and pathology of the biological system. Most nonthermal bioeffects of interest to users of diagnostic ultrasound require acoustic pressures great enough that the wave can become distorted by nonlinear propagation. Under limiting conditions this process can increase the absorption parameter of weakly absorbing media by orders of magnitude and make the absorption parameter of a material such as water as great as the linear absorption coefficient of liver tissue. At low amplitudes the response of the bubbles is dominated by the acoustic pressure. At a critical acoustic pressure, however, the inertia of the surrounding medium becomes controlling. At this threshold of acoustic pressure a 10% or 20% increase in acoustic pressure leads to an increase in the collapse pressure in the bubble by orders of magnitude. The rates of biochemical processes, including denaturation of biological macromolecules, are exponential functions of the temperature. Whether the physical process of heating by ultrasound is linear or nonlinear, this leads to a very strong nonlinear dependence of thermal tissue damage and teratological effects upon the levels of ultrasound.

1:45

5pPAb3. Solitary waves and solitons in acoustics. Nobumasa Sugimoto (Dept. of Mech. Eng., Fac. of Eng. Sci., Univ. of Osaka, Toyonaka, Osaka, 560 Japan)

This study reports on the propagation of acoustic solitary waves and solitons in ordinary air. In theory, it can be predicted by using a long and rigid tube with an array of Helmholtz resonators (or more generally, compact and closed side branches) connected to it with equal axial spacing. In the framework of nonlinear acoustics, the theory exploits three assumptions: (1) a plane wave over the tube, (2) continuum approximation for the array, and (3) negligibly small dissipation. It can be shown that solitary waves are *compressive*, that their speed is *subsonic* but faster than a threshold value determined by the array, and that the solitary waves tend to solitons of Korteweg–de Vries type asymptotically as the speed approaches the threshold. Thus it may be possible to implement a ''soliton tube.'' As the speed approaches the linear speed of sound, the height of the solitary waves becomes greater, approaching the limiting value. Discussions are given on amplification of solitary waves and also on solitary waves in a tube with a double array of resonators.

2:05

5pPAb4. Some consideration of processes of soliton formation from an initially sinusoidal wave form in a thin fiber of fused silica. Akira Nakamura (Dept. of Elect. Eng., Fukui Inst. of Technol., 3-6-1, Gakuen, Fukui, 910 Japan)

Soliton formation processes in a thin fiber of fused silica by computer simulation are discussed. It has already been found that tone-burst sinusoidal sound successively makes solitons from negative half-periods coming in order in the wave train. In this case, the first soliton made from the first negative period is the most stable with respect to the following solitons, because some disturbances caused by preceding positive pressure periods are delayed and superposed with following negative pressure periods. Then, the pressure profiles of these subsequent negative periods are changed. Therefore, nonlinear distortion is affected. In this paper the intention is to discuss the difference of soliton formation processes between two single sinusoidal pulses with preceding negative and positive periods. A preceding negative period makes a stable soliton without any disturbance generated from the following positive pressure part, while the disturbance from the preceding positive period gives change to the pressure profile of following negative period, then nonlinear distortion of this negative period is affected. As a result, frequency spectra of both are differing with the propagation. Discussion is made by operational treatment using a nonlinear distortion operator and linear propagation operator.

2:20

5pPAb5. Fast spectral algorithm for modeling focused sound beams in a highly nonlinear regime. Vera A. Khokhlova (Dept. of Acoust., Phys. Faculty, Moscow State Univ., Moscow 119899, Russia, vera@na.phys.msu.su), Michalakis A. Averkiou (ATL Ultrasound, Bothell, WA 98041-3003), Steven J. Younghouse, Mark F. Hamilton (Univ. of Texas at Austin, Austin, TX 78712-1063), and Lawrence A. Crum (Univ. of Washington, Seattle, WA 98105)

Numerical modeling of high-intensity focused sound beams is important for medical applications of ultrasound. The KZK equation, which may be modified to include relaxation in the time domain or arbitrary absorption and dispersion in the frequency domain, is widely used to model ultrasound beams in tissue. Numerical solutions are available in either the time or frequency domain, but they become time consuming at high intensities when thin shocks are developed. To speed up calculations, a known high-frequency asymptotic result for shocks can be incorporated in the frequency domain algorithm. In comparison with exclusively numerical schemes, the present approach substantially reduces the number of harmonics and thus the computation time [Yu. A. Pishchal'nikov, O. A. Sapozhnikov, and V. A. Khokhlova, Acoust. Phys. **42**, 362–367 (1996)]. Focused cw beams radiated from Gaussian and piston sources are simulated in water and biological tissue. The accuracy of the asymptotic approach is verified via conventional numerical methods of solution. It is shown that the present asymptotic method provides accurate results with a considerable reduction in computation time. In combination with conventional numerical schemes, the developed asymptotic method provides an effective tool for investigating focused sound beams for a wide range of parameters. [Work supported by NIH PO1-DK43881, Fogarty Research Program, and CRDF.]

2:35

5pPAb6. Acoustic nonlinearity calculations using the Tait equation of state. Bruce Hartmann, Gilbert F. Lee, and Edward Balizer (Naval Surface Warfare Ctr., West Bethesda, MD 20817-5700)

Beyer's B/A parameter of acoustic nonlinearity was calculated using the Tait equation of state supplemented with specific heat results, in the same manner as done earlier with our equation of state [J. Acoust. Soc. Am. **82**, 614 (1987)]. Calculations were made of sound speed, sound-speed temperature and pressure derivatives, the two components of B/A, and the value of B/A for a series of six n-alkane liquids as well as the limiting case of molten polyethylene. All these results were compared with experimental data from the literature. In addition, a comparison of the results with Ballou's rule (linear relation of B/A versus reciprocal sound speed) was made. Overall, values calculated using the Tait equation are found to agree with experiment about as well as our equation of state, though the pressure derivatives of the sound speed and the final value of B/A are generally not as accurate using the Tait equation. [Work supported by the Naval Surface Warfare Center's In-house Laboratory Independent Research Program.]

2:50

5pPAb7. Acoustic streaming near Albunex spheres. Gerard Gormley and Junru Wu (Dept. of Phys., Univ. of Vermont, Burlington, VT 05405)

Acoustic streaming near bubbles in an external standing-wave sound field has been an interesting research topic for decades. It was relatively difficult to visualize the acoustic streaming pattern experimentally, especially for small bubbles due to the instability of the bubbles and difficulties associated with the optical technique. Using Albunex spheres (a contrast agent for ultrasonic imaging), a specially designed microscope device, and a video peak store technique, several acoustic streaming patterns around a single and a pair of Albunex spheres of tens of micrometers radius in a 160-kHz ultrasonic standing wave field have been recorded. Spatial peak pressure amplitude of the standing wave was estimated to be in the order of 0.5 MPa by using optical interferometry. The acoustic streaming velocity was in the range of 50–100 μ/s.

3:05–3:20 Break

3:20

5pPAb8. Acoustical nonlinearity and its images in biomedicine. X. F. Gong, D. Zhang, S. G. Ye, X. Z. Liu, X. Chen, and W. Y. Zhang (Inst. of Acoust., State Key Lab. of Modern Acoust., Nanjing Univ., Nanjing, 210093, PROC)

Some acoustical nonlinear effects in biological media at medical ultrasonic frequencies and intensities were studied experimentally. These phenomena include wave distortion, harmonic generation, energy transformation, extra-attenuation, departure of intensity response curve from linear relation, etc. The influence of these effects on medical ultrasonic diagnosis and therapy was discussed. As a basic parameter in nonlinear acoustics, the acoustical nonlinearity parameter can describe the above-mentioned nonlinear effects and provide some information on the structure and composition of biological tissues. Thus the imaging method of an acoustical nonlinearity parameter using secondary wave detection was developed and investigated. Nonlinearity parameter tomography for normal and a series of pathological porcine tissues was presented. The results obtained in experiments indicate that the difference between normal and diseased tissue in nonlinearity parameter tomography is more sensitive than that in linear parameter tomography. Therefore, nonlinearity parameter tomography may become a novel imaging method for tissue characterization. [Work supported by NSF of China.]

3:35

5pPAb9. Determination of the acoustic nonlinearity parameter B/A using a phase locked ultrasonic interferometer. Jevon R. Davies, Jonathan Tapson (Univ. of Cape Town, Elec. Eng., Private Bag, Rondebosch, 7700, South Africa), and Bruce J. P. Mortimer (Cape Technikon, Cape Town, 8000, South Africa)

The acoustic nonlinearity parameter B/A is an important parameter for characterizing acoustic propagation in fluids. Recent developments in the fields of shock waves and biomedical ultrasonics have necessitated an accurate method for measuring this quantity. This paper presents a new technique based on an ultrasonic interferometer that allows accurate measurements of the change in sound speed within a liquid cavity. The system makes use of a novel continuous wave phase locking principle, which has

the effect of nullifying many of the nonlinearities inherent in the acoustic interferometer. The phase lock circuitry automatically adjusts the frequency so as to keep an exact integer number of waves in the cavity regardless of changes in sound speed. Using this principle, B/A is determined by relating a small change in ambient pressure (less than 2 bar) to a change in frequency under isentropic conditions. B/A values are calculated using this experimental setup for several liquids including two slow sound-speed fluorocarbon liquids, FC-43 and FC-75. These liquids are unique in that they have the lowest range of sound speeds of any organic liquids and in that they have the highest values of B/A recorded thus far.

3:50

5pPAb10. Experimental measurement and numerical modeling of nonlinear propagation in biological fluids. Victor F. Humphrey (Dept. of Phys., Univ. of Bath, Claverton Down, Bath BA2 7AY, UK), Prashant K. Verma (Leicester Royal Infirmary, Leicester LE1 5WW, UK), and Francis A. Duck (Royal United Hospital, Combe Park, Bath BA1 3NG, UK)

The nonlinear pressure fields generated in three biological fluids (amniotic fluid, urine, and a 4.5% human albumin solution) by a circular 2.25-MHz transducer have been investigated experimentally using a flexible measurement facility and compared with numerical predictions. The acoustic field was generated by a single element 38-mm-diameter transducer coupled to a PMMA lens with a focal gain of 12 in water for source pressures from 0.007 to 0.244 MPa. The experimental results illustrate significant harmonic generation in all three fluids which may have safety and imaging implications. In addition, the facility was used to make low level measurements of the acoustic parameters of the fluids, including the frequency dependence of attenuation, at both room and physiological temperature. This data has been used in conjunction with a finite difference model based on the Khokhlov–Zabolotskaya–Kuznetsov equation (the Bergen Code) to model the experiments. The agreement with experiment was found to be generally good, indicating that this type of model may be useful for predicting *in-vivo* pressure distributions given an accurate knowledge of the medium parameters. [Work supported by EPSRC].

4:05

5pPAb11. Asymptotically unified exact solutions of the Khokhlov–Zabolotskaya equation. Yuri N. Makov (Dept. of Phys., Moscow State Univ., Moscow 119899, Russia, makov@acs364.phys.msu.su)

For the large distance traversed by the sonic beam the exact analytical and physically realistic solutions of the Khokhlov–Zabolotskaya (KZ) equation were found for the first time. The procedure of finding these solutions as the self-similar traveling waves identifies their asymptotical generality with respect to the initial arbitrary wave profiles. The resulting solutions consist of the periodically repeating parts of the second-order curvers such as hyperbolas, ellipses, and, in particular, parabolas. The boundaries between these parts are defined with the use of the preserved integrals of the KZ equation. The influence of nonlinearity, diffraction, which is seen in many-dimensional cases, and initial curvature of the phase front of the wave beam on the forming of the typical form of a specific asymptotical wave profile (N shaped or U shaped) was discussed. A comparison was made between analytical solutions with the numerical results for the KZ equation and the experimental data and a full agreement was discovered. These solutions play the same role among all the variations of solutions of the KZ equation, such as the well-known sawtoothlike wave for the equation of the simple waves (Riemann's equation). [Work supported by CRDF.]

FRIDAY AFTERNOON, 26 JUNE 1998 FIFTH AVENUE ROOM (W), 1:00 TO 3:45 P.M.

Session 5pPP

Psychological and Physiological Acoustics: Processing, Location, Distance and Motion

Armin Kohlrausch, Chair
Institute for Perception Research, P.O. Box 513, NL-5600MB Eindhoven, The Netherlands

Contributed Papers

1:00

5pPP1. Binaural coherence and the localization of sound in rooms. William M. Hartmann, Zachary A. Constan (Phys. and Astron., Michigan State Univ., East Lansing, MI 48824), and Brad Rakerd (Michigan State Univ., East Lansing, MI 48824)

The absorption coefficient of materials tends to decrease with decreasing frequency. Therefore, the ratio of direct to reverberant sound is normally least favorable at low frequency, and binaural coherence tends to be a minimum. However, the ability to localize sounds using interaural time differences (ITD) in the steady-state waveform requires binaural coherence, especially at low frequency. Lateralization experiments, using a forced-choice staircase technique, indicate that adequate localization based on ITD requires binaural coherence of at least 0.2. Measurements of binaural coherence in rooms find that binaural coherence is often less than that, which means that ITD cues can only be used on the transient portions of a signal. By contrast, interaural level differences are perceptually useful at high frequencies and increasingly dominate the localization of steady-state sounds as the environment becomes more reverberant. [Work supported by the NIDCD, Grant No. DC00181.]

1:15

5pPP2. Interaural correlation sensitivity. John F. Culling, Matthew Spurchise, and H. Steven Colburn (Dept. of Biomed. Eng., Boston Univ., 44 Cummington St., Boston, MA 02215)

Sensitivity to differences in interaural correlation was measured as a function of reference interaural correlation and frequency (250 to 1500 Hz) for narrow-band-noise stimuli (1.3 ERBs wide) and for the same stimuli spectrally fringed by a broadband correlated noise. d′ was measured for two-interval discriminations between fixed pairs of correlation values, and these measurements were used to generate cumulative d′ versus correlation curves for each stimulus frequency and type. The perceptual cue reported by subjects was perceived as intracranial breadth for narrow-band stimuli (wider image for lower correlation) and loudness of a whistling sound heard at the frequency of the decorrelated band for the fringed stimuli (louder for lower correlation). At low correlations, sensitivity was greater for fringed than for narrow-band stimuli at all frequencies, but at higher correlations, sensitivity was often greater for narrow-band stimuli. For fringed stimuli, cumulative sensitivity was greater at low frequencies than at high frequencies, but listeners produced varied patterns for narrow-band stimuli. The forms of cumulative d′ curves as a function of frequency were interpolated using an eight-parameter fitted function.

Such functions may be used to predict listeners' perceptions of stimuli that vary across frequency in interaural correlation. [Work supported by MRC and NIDCD Grant No. DC00100.]

1:30

5pPP3. Modeling the ''effective'' binaural signal processing. Carsten Zerbs (Graduiertenkolleg Psychoakustik, Univ. of Oldenburg, P.O. Box 2503, 26111 Oldenburg, Germany, cz@medi.physik.uni-oldenburg.de), Torsten Dau, and Birger Kollmeier (Univ. of Oldenburg, 26111 Oldenburg, Germany)

This paper describes a quantitative model for binaural processing in the auditory system. To overcome shortcomings of several proposed models, a monaural stage is employed for signal preprocessing [Dau *et al.*, J. Acoust. Soc. Am. **99**, 3615–3622 (1996)], which is able to simulate a large variety of monaural detection data. A binaural processing stage is added following Durlach's EC model which is also a simple functional unit with a small set of constant parameters unit. The decision algorithm for both the monaural and the binaural channel is realized as an optimal detector in the same way as described in the monaural model. In the simulations the whole adaptive measurement procedure with the actual signals is realized. The model describes both monaural as well as binaural thresholds. The typical binaural masking experiments with the following stimulus parameters and their combinations were examined: signal frequency, signal duration, masker bandwidth, interaural masker and signal phase, interaural masker correlation, and delay in simultaneous and forward masking conditions. The model describes the experimental data with an accuracy of a few dB with some exceptions. Model predictions are discussed and compared to the performance of human observers and to predictions of other binaural models.

1:45

5pPP4. The influence of masker variability on estimates of monaural and binaural critical bandwidths. Armin Kohlrausch, Steven van de Par, and Jeroen Breebaart (IPO Ctr. for Res. on User–System Interaction, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands)

If integration bandwidths are derived from band-widening masking experiments, binaural estimates are usually a factor 2 to 3 larger than monaural estimates, at least at high levels. It is proposed that this difference does not reflect ''different critical bandwidths,'' but that it is a consequence of detection statistics. In monaural experiments, the variability in the masker energy limits detectability of the signal. Therefore with increasing subcritical bandwidths, the S/N ratio at threshold decreases. In binaural NoSπ experiments, the S/N ratio at subcritical bandwidths remains constant because the masker correlation is always 1 without uncertainty. Due to this absence of external variability in the masker correlation, detection must be limited by internal noise. Binaural detection in *narrowband* conditions can gain from the spread of excitation across several critical bands, assuming that the internal noises are uncorrelated. The wider binaural critical band is then caused by masking the off-frequency spread through masker components outside the central critical band (for binaural model simulations, see Breebaart *et al.*). In this study it is shown that this scheme also applies to monaural frozen-noise maskers, which do not have external variability. In agreement with results from binaural experiments, the integration bandwidths increase considerably with masker level.

2:00

5pPP5. Modeling spectral integration in binaural signal detection. Jeroen Breebaart, Steven van de Par, and Armin Kohlrausch (IPO Ctr. for Res. on User–System Interaction, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands)

Estimates of critical bandwidths in binaural experiments vary considerably across experimental conditions. For masking experiments using signals of varying bandwidths (i.e., spectral integration), and experiments with a frequency-dependent interaural masker correlation, bandwidth estimates agree with those from monaural experiments. However, if auditory bandwidths are estimated from binaural band-widening masking experiments, the effective bandwidth is usually a factor of 2 to 3 times larger than monaural estimates. One often ignored detail is that this difference between monaural and binaural estimates decreases with decreasing masker level [Hall *et al.*, J. Acoust. Soc. Am. **73**, 894–898 (1983)]. Using the binaural model described by Breebaart *et al.*, we have simulated the binaural experiments described above. It can be shown that, without any parameter change in the model, basically all relevant conditions can be modeled accurately, including the band-widening and spectral integration experiments. In agreement with the scheme discussed by Kohlrausch *et al.*, detection in binaural narrow-band-noise conditions is improved by analyzing the internal representation in several adjacent filters. It is concluded that the absence of this detection advantage in monaural random-noise conditions is the primary cause for the so-called ''wider binaural critical band.''

2:15

5pPP6. The role of the head related transfer function low-frequency component for sound localization on a median plane. Motoki Yairi, Yasuko Kuroki, Masayuki Morimoto (Environ. Acoust. Lab., Faculty of Eng., Kobe Univ., Nada, Kobe, 657 Japan), and Kazuhiro Iida (AVC Res. Lab. Matsushita Comm. Ind., Saedo, Tsuzuki, Yokohama, Japan)

Asano *et al.* [J. Acoust. Soc. Am. **88**, 159–168 (1990)] demonstrated that localization is not accurate for the simulation of sound sources on the median plane when the low-frequency component of head related transfer functions (HRTFs) are smoothed. In the present localization experiment, the low- and high-frequency components of white-noise stimuli were separated at fc=4.8 kHz (−48 dB/oct.). Those components were independently presented from different directions. Results show that the sound image is localized in the direction from which the high-frequency component is presented, regardless of the direction of low-frequency component, showing that the low-frequency component has no evident effect on the sound localization on a median plane. Based on the results of these two experiments, it is concluded that only the high-frequency component of an HRTF provides spectral cues for sound localization on a median plane. However, changing the low-frequency component of the HRTF, which is unlikely to naturally occur in sound localization on a median plane, may interfere with accurate localization.

2:30–2:45 Break

2:45

5pPP7. The AUDIS catalog of human HRTFs. Jens Blauert, Marc Brueggen (Ruhr Univ., D-44780 Bochum, Germany, blauert@ika.ruhr-uni-bochum.de), Adelbert W. Bronkhorst, Rob Drullman (TNO NL-3769 ZG Soesterberg), Gerard Reynaud, Lionel Pellieux (Sextant Avionique, F-33166 Saint-Medard-en-Jalles), Winfried Krebber, and Roland Sottek (Head Acoust. GmbH, D-52134, Herzogenrath, Germany)

In the European-Union-funded project AUDIS (AUditory DISplay), a multipurpose auditory display for 3D hearing applications is being developed. The main goal is to develop a special sound generator in combination with an auditory-symbol database that can, for example, be used in the avionic and automotive industry. As the project relies heavily on binaural technology and, hence, on the availability of reliable human HRTF data, a special program for collecting data has been undertaken. In the course of this program, round-robin tests have been performed in four different laboratories. Differences in the data across the different laboratories have been analyzed. The analysis resulted in recommendations for common measuring procedures to be applied. A catalog of HRTFs has then been collected and perceptual verification tests have been performed. It is planned to make the catalog publicly available. In the presentation, the measuring methods applied and the contents of the catalog will be discussed.

3:00

5pPP8. A modeling of distance perception based on an auditory parallax model. Yôiti Suzuki, Shouichi Takane, Hae-Young Kim, and Toshio Sone (Res. Inst. of Elec. Comm., Tohoku Univ., Sendai-shi, 980-8577 Japan)

Distance perception of sound from a source (<5 m) situated close to a listener was studied under the state without information on intensity and reflections using a small movable loudspeaker in an anechoic room. Distance perception in such a condition is modeled by use of the difference in HRTF (head-related transfer function) due to the parallax angle calculated from the difference between directions from each ear to the source. This model is called the ''auditory parallax model'' here. Sound signals to realize this model were digitally synthesized for the listening experiment. In addition, an experiment with an actual sound source and one with HRTF synthesized as precisely as possible were also conducted. The distances obtained for sound with the auditory parallax model and that with synthesized HRTF showed very similar tendencies to those with actual sound sources. That is, in all the three conditions, the perceived distance of a sound image monotonically increased with the physical distance of source for distances up to around 1.2 m. If the presented distance was over 1.2 m, the perceived distance increased lower. These results show that the cues for the distance perception derived by the auditory parallax model is almost sufficient.

3:15

5pPP9. Evidence of spatially tuned auditory analyzers. Russell L. Martin (AOD, AMRL, P.O. Box 4331, Melbourne Australia, 3001) and Ester Klimkeit (Monash Univ., Clayton, Australia, 3168)

The primary-plus-probe uncertainty technique was used to explore the possibility that analyzers tuned to different spatial locations exist within the human auditory system. Eight subjects' psychometric functions for detection of a 500-Hz tone when presented from a speaker located at 0 and 90 degrees azimuth were measured in each of two conditions. In the first, the intermixed condition, subjects knew that 80% of signals would be presented from one location (the primary location) and 20% would be presented from the other (the probe location). In the second, subjects knew that all signals would be presented from either the primary or probe location. For each subject the primary location was 0 degrees for half the sessions and 90 degrees for the others. Each function was measured using a two-interval forced-choice task and five signal levels ranging from 7.5 dB above to 4.5 dB below the subject's previously estimated threshold. For primary location signals, average functions for the two conditions were indistinguishable. For probe location signals, however, functions differed, reflecting significantly poorer detection performance in the intermixed condition [$F(1,7)=3D22.38$, $p<0.01$]. This uncertainty effect provides support for the notion that spatially tuned analyzers exist within the auditory system.

3:30

5pPP10. Detection of Doppler-like signals. Mark A. Ericson and Lawrence L. Feth (The Ohio State Univ., Dept. of Speech and Hearing Sci., Columbus, OH 43221)

Moving sounds induce a Doppler shift of frequency in direct proportion to a sound source's velocity. Concurrent with the frequency change is a change in intensity related to the source's velocity and distance from the listener. These simultaneous changes in frequency and intensity may be audible for any moving sound source. Jorasz and Dooley [Arch. Acoust. **21**, 149–157 (1996)] modeled the physical effects of constant velocity motion on a sound source's frequency and intensity as observed at a single point in space. Three linear trajectories of a sound source were simulated over headphones for different minimal distances between the moving sound source and observer. These minimal distances were 1, 10, and 100 m. An adaptive tracking procedure was used to determine detection thresholds for velocity induced changes in frequency and intensity. Intensity and frequency were varied independently as well as together. A series of experiments was conducted to test a perceptual model of these Doppler effects.

FRIDAY AFTERNOON, 26 JUNE 1998 WEST BALLROOM B (S), 1:00 TO 3:00 P.M.

Session 5pSA

Structural Acoustics and Vibration: Vibrations of Complex Structures II

Noureddine Atalla, Cochair
GAUS Mechanical Engineering, Universite de Sherbrooke, Sherbroooke, QC J1K 2R1, Canada

Paul E. Barbone, Cochair
Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Contributed Papers

1:00

5pSA1. Measurement and analysis of the structure-borne sound of a fan installation: Evaluation of source signal filtering techniques and rotational degrees of freedom. Michael A. Sanderson, Lars Ivarsson (Dept. of Appl. Acoust., Chalmers Univ. of Technol., 412 96 Gothenburg, Sweden, mike@ta.chalmers.se), and Andrej G. Troshin (Krylov Shipbuilding Res. Inst., St. Petersburg, Russia)

The structure-borne sound characteristics of a commonly found fan installation were examined through measurements and further analyses in the frequency range up to 500 Hz. The focus was to determine the prediction improvements by including the rotational degrees of freedom. Free velocities of the fan were measured in five degrees of freedom (in-plane rotation or twisting motion were neglected), while fan mobilities were measured and foundation mobilities were predicted in three degrees of freedom: a translational and two rotational, those resulting in bending motion. The two in-plane translations were considered by transforming their motion to rotations resulting in moment excitations on the foundation. Measurements were also made of the corresponding isolators' stiffness values. Comparisons are made with power measurements of the complete system. It was found that an unsatisfactory moment excitation resulted in one of the rotational degrees of freedom while employing a twin shaker arrangement when determining the fans' mobilities. Therefore, a source signal filtering technique was applied to improve the mea-

surement. Comparisons are made with the originally measured mobilities. [This work has been supported by the Swedish Council for Building Research (BFR) through Contract 950192-8 and the Swedish Research Council for Engineering Sciences (TFR) through Contract 95-160.]

1:20

5pSA2. Modeling the vibration behavior of an induction motor. Chong Wang and Joseph C. S. Lai (Acoust. and Vib. Unit, Univ. College, Univ. of New South Wales, Australian Defence Force Acad., Canberra, ACT 2600, Australia, c-wang@adfa.oz.au)

With the advent of power electronics, variable speed induction motors are finding increased use in industries because of their low cost and savings in energy usage. However, the acoustic noise emitted by the motor increases due to switching harmonics introduced by the electronic inverters. Consequently, the vibro-acoustic behavior of the motor structure has attracted more attention. In this paper, considerations given to modeling the vibration behavior of a 2.2-kW induction motor will be discussed. By comparing the calculated natural frequencies and the mode shapes with the results obtained from experimental modal testing, the effects of the teeth of the stator, windings, outer casing, slots, endshields, and support on the overall vibration behavior are analyzed. The results show that when modeling the vibration behavior of a motor structure, the laminated stator should be treated as an orthotropic structure, and the teeth of the stator could be neglected. As the outer casing, endshields, and the support all affect the vibration properties of the whole structure, these substructures should be incorporated in the model to improve the accuracy.

1:40

5pSA3. Model-based shaping of vibroacoustic properties of gears. Stanislaw Radkowski (Vibroacoustic Lab. IPBM PW, 02-524 Poland, Warszawa, Narbutta 84)

The work undertakes an issue of estimating and predicting the level and structure of the vibroacoustic signal generated by a pair of toothed wheels and propagated by the structure of the gear's housing. Assuming that the basic role in transmission of such diagnostic information is played by the phenomena of amplitude and phase modulation of the vibroacoustic signal, it is pointed to the significance of the signal's nonlinear components. For that reason in this work a model of a signal in the form of a Volterra series is assumed, which enables examination of the influence that disturb nonlinear components. On the basis of exemplary analysis of the contact conditions disturbances, the possibility of the generation of a bilinear component is presented. What serves this purpose is the model of nonlinear and instantaneous contact. Simultaneously problems of selection of signal analysis methods that enable realization of assumed aims are presented. Emphasizing the fact that defect development does not always lead to a significant growth of the vibroacoustic signal's power, the importance of methods of higher-order spectral analysis like bispectrum and bicepstrum has been shown.

2:00

5pSA4. Experimental identification of mechanical joint dynamic properties. Olivier Chiello, Franck C. Sgard (LASH, DGCB URA CNRS 1652, ENTPE—Rue Maurice Audin, 69518 Vaulx-en-Velin, Cedex, France, olivier.chiello@entpe.fr), and Noureddine Atalla (GAUS, Univ. de Sherbrooke, QC J1K 2R1, Canada)

Large structures are often composed of substructures connected through mechanical joints. The dynamic behavior of these joints is generally difficult to characterize. In this paper, a method for identifying dynamic joint characteristics is presented. Joint properties are extracted directly from measured frequency responses of the assembled structure. Measurement points are not taken on the interface between the structure components. This feature is of practical importance since it is often difficult to directly attach a transducer on the joint. The method only requires the knowledge of the frequency responses of the substructures taken separately. These responses can be either measured, analytically calculated, or simulated by FEM. The presented approach allows for the determination of the joint properties as a function of frequency. The joint properties extracted from one assembled structure can be used to predict the behavior of another structure assembled with the same joint condition. The applicability of the proposed method is investigated by numerical simulations. Errors in simulated frequency responses are introduced to investigate the robustness of the method. It is shown that the accuracy of the identified results is good.

2:20

5pSA5. The vibroacoustical analysis of the teeming crane. Jerzy W. Wiciak, Ryszard J. Panuszka, and Marek A. Iwaniec (Structural Acoust. and Intelligent Mater. Group, Tech. Univ. of Mining and Metallurgy, Al. Mickiewicza 30, 30-059 Cracow, Poland)

The present article considers the vibration analysis of the teeming crane (lifting capacity $Q=160$ tons) and noise caused by working units of the crane. The vibroacoustical investigations and theoretical analysis have been applied to the main bridge with box-trapezoid cross section. In the article there is shown a comparison of the vibration acceleration and dynamic stress that have been measured in real structure with calculations obtained by means of the finite element method. The correct and incorrect (two-stage) reduction roll gear running also was considered. The significant influence of the incorrect reduction roll gear running on the dynamic parameters was noticed. The noise caused by working units of the crane has been presented for three operating conditions: when the tilting ladle was descending, when it was hoisting, and pass-by load running of the teeming crane. [Work supported by TUMM, Cracow, Poland.]

2:40

5pSA6. Metal and composite aircraft panels: Selection of parameters optimal from the viewpoint of acoustic fatigue lifetime. Nikolay I. Baranov, Sergey N. Baranov, Lev S. Kuravsky, and Konstantin P. Zhukov (''RusAvia'' Res. Group, Mikoyan Design Bureau, 6 Leningradskoe Shosse, 125190 Moscow, Russia)

Optimization techniques to ensure synthesis of stiffened and laminated composite panels subjected to acoustic excitation are presented. The criteria for optimization are established to balance the structure weight and fatigue life. Test spectrum of loading pressure fluctuations is in use, with the parameters of pressure correlation ensuring the greatest response. For stiffened structures, stress spectra and corresponding root-mean-square values in points of maximum response are computed by means of special software to draw the special final diagrams that connect nondimensional structure parameters and optimization criteria. These diagrams are then used to select optimal parameters of reinforcement. For composite structures, correlation of the nondimensional parameters representing structure stress level, acoustic load level, dominant frequency of vibration, and panel anisotropy is computed by the same technique. The functions computed are later used to estimate how the criterion for optimization depends on the number of layers, lay-up sequence, and the directions of fibers. The approach proposed allows aerospace designers to select and motivate their decisions having increased confidence in the development and modifications of these structures.

Session 5pSC

Speech Communication: Segmental Phonetics, L2, and Cross Language Studies (Poster Session)

Joseph S. Perkell, Chair
Massachusetts Institute of Technology, Room 36-591, 50 Vassar Street, Cambridge, Massachusetts 02139-4307

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. Do nasalized fricatives exist? John J. Ohala (Dept. of Linguist., Univ. of California at Berkeley, 2337 Dwinelle Hall #2650, Berkeley, CA), Maria-Josep Sole (Univ. Autonoma de Barcelona, 08193 Bellaterra, Barcelona, Spain), and Goangshiuan S. Ying (Univ. of California at Berkeley, Berkeley, CA)

It has been claimed that nasalized fricatives, even sibilant fricatives, exist in some languages. To assess this claim we sought to discover how much velopharyngeal opening can be present before the acoustic and auditory characteristics of fricatives are affected. Two trained phoneticians produced steady-state voiced and voiceless strong and weak fricatives. The back pressure was intermittently bled with a tube of varying diameter (and thus impedance) inserted into the speaker's mouth via the buccal sulcus and the gap behind the back molars, simulating nasal leakage. Intraoral pressure (Po) was sampled via a catheter inserted into the pharynx through the nose. The changes in amplitude and quality of frication were analyzed acoustically. We found that (1) a vent area of about 18 mm^2 caused auditorily and acoustically noticeable lessening of fricative energy, e.g., sibilants sounded more like nonsibilants, (2) for a given vent aperture, Po was diminished less for voiceless than voiced fricatives (because the voiced fricatives have higher upstream impedance at the glottis) and consequently the same vent aperture impaired the amplitude and quality of frication of voiced fricatives more than voiceless ones. Thus the aerodynamic requirements for fricatives seem to be relatively narrow and unforgiving. [Research supported by Committee on Research, UCB, and by DGICYT, Spain, PB 96-1158.]

5pSC2. Aerodynamic characteristics of trills. Maria-Josep Sole (Dept. of English, Univ. Autonoma de Barcelona, 08193 Bellaterra, Barcelona, Spain, sole@cc.uab.es), John J. Ohala, and S. Goangshiuan Ying (Univ. of California at Berkeley, Berkeley, CA 94720-2650)

What are the aerodynamic conditions required for trills? To find out we had two subjects produce steady-state voiced, voiceless, and ejective alveolar trills. The backpressure during trills was intermittently bled with a tube of varying diameter (and thus impedance) inserted in the speaker's mouth via the buccal sulcus and gap behind the back molars. Intraoral pressure was measured via a catheter inserted into the pharynx through the nose. The variation, impairment, or extinction of trilling as a function of gradual decrease in intraoral pressure was analyzed acoustically. It was found that (1) bleeding the oro-pharyngeal pressure by 2 cm H_2O impaired sustained trilling; (2) the minimum Po required to sustain tongue-tip vibration is lower than that required to initiate it; (3) extinction (and reinitiation) of trilling generally results in a fricative; (4) the range of Po variation for trills is narrower than that for fricatives; (5) voiceless and ejective trills are significantly less affected by venting the backpressure than voiced trills. The behavior of trills in varying aerodynamic conditions accounts for observed phonological patterns: final trill devoicing, alternation between trills and fricatives, co-occurrence of trilling and frication, and limited distribution of trills. [Work supported by DGICYT, Spain, PB 96-1158, and by Committee on Research, UCB.]

5pSC3. Motor equivalence in the production of /ʃ/. Joseph S. Perkell, Melanie L. Matthies, and Majid Zandipour (Speech Commun. Group, R.L.E., MIT, Rm. 36-511, 50 Vassar St., Cambridge, MA 02139)

To explore the idea that speech motor goals are acoustic targets, upper lip protrusion and tongue blade fronting were examined in the sibilant /ʃ/ for evidence of motor equivalence in eight speakers of American English. Positive correlations across multiple repetitions of /ʃ/ (motor equivalence) would occur if the upper lip compensated with more protrusion when the tongue blade was further forward and vice versa. This motor equivalence would serve to maintain an adequate front cavity size for good acoustic separation from /s/ and enhance the acoustic stability of /ʃ/. It was hypothesized that motor equivalence would be found among less prototypical tokens, as elicited in a ''clear+fast'' speaking condition. Acoustic spectral analyses indicated excellent acoustic separation between the two sibilants, even in the ''clear+fast'' condition. There were significant positive correlations of the tongue-blade and upper lip movements for 28% of all possible cases, with considerable individual variation. When motor equivalence could be compared across the speaking condition, the tokens produced with the motor equivalent pattern were acoustically less prototypical. No occurrences of negative movement correlations were found. One possible source of individual variation may be the use of saturation (quantal) effects, which could also enhance acoustic stability. [Work supported by NIH.]

5pSC4. Further results on the perception of place coarticulation in Taiwanese stops. Shu-hui Peng and Terrance Nearey (Dept. of Linguist., Univ. of Alberta, 4-32 Assiniboia Hall, Edmonton, AB T6G 2E7, Canada, speng@gpu.srv.ualberta.ca.)

Previous study of place coarticulation using logistic discrimination [Peng and Nearey, J. Acoust. Soc. Am. **101**, 3110(A) (1997)] showed that 87% of Taiwanese word-final dental or velar stops in stop consonant clusters [CVstop1-stop2V(C)] spoken in various speech rates were correctly classified based on formant transitions from the preceding vowel. This automatic classification was sigificantly (though only moderately) correlated with categorization by Taiwanese speakers in a concept formation test. An earlier study [Peng (1996)] also showed that there was not a consistent correlation between Taiwanese listeners' categorization of the stops and patterns of gestural overlapping shown in EPG measurements. To avoid complication involving listeners' difficulty with the concept formation task and the influence of native speakers' lexical knowledge, the present study examined the perception of the stops by phonetically trained

listeners in a different experimental paradigm. Listeners were asked to identify the consonant clusters (tt, kk, tk, kt) from the excised first word in a four-way forced choice identification test. The phonetic judgments of these listeners will be compared with the categorization by Taiwanese speakers, the results of discriminant analysis, and EPG measurements.

5pSC5. Ejectives in Tanana Athabaskan. Siri G. Tuttle (Univ. of Washington, Dept. of Linguist., Box 354340, Seattle, WA 98195, iris@u.washington.edu)

Ejectives in Athabaskan languages have been investigated phonetically by Lindau (1984), McDonough and Ladefoged (1993), in Navajo, and by Davis and Hargus (1994, 1996) in Babine/Witsuwit'en. In the latter two investigations, three different types of production were found associated with glottal consonants, only one of which showed the classic ejective pattern with a silent period. Examination of samples of Tanana Athabaskan recorded in 1961, 1962, 1974, and the early 1990s found similar variation in samples of this Alaskan Athabaskan language. There appeared, however, to be no examples of ejectives without silent periods in the oldest samples (1961 and 1962). This suggests the hypothesis that the ejective without silent period is an innovation in the last generation of Tanana speakers. Recordings analyzed include a set of standardized word lists recorded by McKennan in 1962 with speakers of Minto, Chena, and Salcha dialects; texts of Teddy Charlie (Minto) recorded by Michael Krauss in 1962; and 3keywords2 recordings of Isabel Charlie (Minto) and Eva Moffit (Salcha) recorded by James Kari and Siri Tuttle in 1991. [This work has been supported by NSF.]

5pSC6. Perception of synthesized Hindi geminate and cluster sounds. Nisheeth Shrotriya and Shyam S. Agrawal (Speech Technol. Group, CEERI Ctr., CSIR Complex, Hillside Rd., New Delhi-12, India)

This paper describes the various differences in acoustic parameters of geminate, nongeminate, and cluster sounds of Hindi, which can be used for their perceptual distinction. Words containing these sounds were synthesized using a PC-based Klatt synthesizer, which uses a set of about 60 parameters. For distinction between geminate and nongeminate sounds, the parameters such as duration of the plosive gap and preceding vowel, and the formant frequencies and intensity of the burst are important but the closure duration plays the most significant role. This difference is of the order of 1:1.5. However, if the closure duration is reduced to 50% or less, the unvoiced sounds are perceived as voiced sounds. In the case of the cluster (-C1C2-) words containing stop consonants, when C1 and C2 have the different places of articulation, a burst is observed in the middle of the closure. It was found that this burst has no perceptual effect when it was moved to the left and right positions adjacent to the vowels. Even the removal of the burst has no perceptual effect.

5pSC7. Integration of acoustic cues in Spanish voiced stops. Sergio Feijóo, Santiago Fernández, and Ramón Balsa (Dpto. Física Aplicada, Fac. de Física, 15706, Santiago, Spain, fasergio@uscmail.usc.es)

In Spanish, voiced stops in initial position usually show the presence of a voice bar before the release burst. In this study, the perception of word-initial voiced stops when the voice bar is removed is examined. Nine men and nine women pronounced 15 Spanish words in which the initial consonant was /b/, /d/, or /g/ in combination with the vowels /a/, /e/, /i/, /o/, /u/. Seventeen subjects with normal hearing heard the stops in two conditions: (1) whole voice bar + 102.4 ms of signal from burst onset; (2) 102.4 ms from burst onset. In condition (2), tokens were perceived as the correct stop, as its voiceless counterpart, or as an ambiguous signal. Removing the voice bar caused 37% of /b/ tokens to be classified as /p/; 41% of /d/s to be classified as /t/s ; and 19% of /g/s to be classified as /k/s. The change in voicing characteristics was significantly more marked for women. The results indicate that the integration of the voice bar with the burst and the vowel plays a relevant role in the perception of voicing in voiced stops, especially in women's speech, smoothing the dynamic frequency characteristics of burst and vowel.

5pSC8. Acoustic properties of English fricatives. Allard Jongman, Joan Sereno, Ratree Wayland, and Serena Wong (Cornell Phonet. Lab., Dept. of Modern Lang., Cornell Univ., Ithaca, NY 14853)

A fundamental issue in speech research concerns whether distinctions in terms of place of articulation are more succesfully captured by local (static) or global (and/or dynamic) properties of the speech signal. While most studies of place of articulation have investigated stop consonants, it is uncertain whether these properties can successfully classify fricatives. In the present study, both static and dynamic metrics were used to investigate place of articulation in fricatives. Static metrics include spectral peak location, noise duration, and noise amplitude; dynamic metrics include relative amplitude, locus equations, and spectral moments. Twenty speakers produced the fricatives /f, v, θ, ð, s, z, ʃ, ʒ/ followed by the vowels /i, e, æ, ɑ, o, u/. Preliminary results based on ANOVA and discriminant analysis suggest that all metrics could distinguish nonsibilant /f, v, θ, ð/ from sibilant /s, z, ʃ, ʒ/. In addition, spectral moments separated /s, z/ from /ʃ, ʒ/ with about 95% accuracy and /f, v/ from /θ, ð/ with about 65% accuracy while locus equations only served to separate labiodental /f, v/ from all other fricatives. Accuracy of the metrics will also be compared to human perception data. [Work supported by NIH.]

5pSC9. Role of *F*1 in the perception of voice offset time as a cue for preaspiration. Jörgen Pind (Dept. of Psych., Univ. of Iceland, Oddi, IS-101 Reykjavík, Iceland, jorgen@rhi.hi.is)

Preaspiration, an [h]-like sound inserted between a vowel and a following stop consonant, can be cued by voice offset time (VOffT), a speech cue which is a mirror image of voice onset time (VOT). Previous research has shown VOffT to be an effective cue for preaspiration and also that the perception of VOffT differs from the perception of VOT in that the former cue is much more sensitive to the duration of a preceding vowel than is VOT to the duration of a following vowel [J. Pind, Q. J. Exp. Psychol. 49A, 745–764 (1996)]. The hypothesis has been put forward that this is due to the fact that preaspiration in Icelandic follows a phonemically short vowel, whereas aspiration can precede either a long or a short vowel. Vowel quantity will therefore affect phoneme boundaries for VOffT, not for VOT. This hypothesis was confirmed in experiments where vowel spectra were used to manipulate quantity rather than vowel duration. The present paper reports experiments investigating whether this effect could possibly be mediated through the effect of *F*1 on perceived voicing. [Supported by Icelandic Science Foundation and University of Iceland.]

5pSC10. Vowel and consonant durations as cues for quantity in Icelandic. Jörgen Pind (Dept. of Psych., Univ. of Iceland, Oddi, IS-101 Reykjavík, Iceland, jorgen@rhi.hi.is)

Icelandic distinguishes short and long consonants and vowels in stressed syllables. Phonologically, these show complementary quantity in that a short vowel is followed by a long consonant and vice versa. Perceptual experiments have shown that the ratio of vowel to consonant duration is the major cue for quantity, although the vowel spectrum can override the relational cue in those cases where there is a marked difference in the spectral character of the long and short vowels [J. Pind, Scand. J. Psychol. **37**, 121–131 (1996)]. Measurements will be reported of vowel and consonant durations in stressed syllables in words spoken repeatedly in sentence frames and in longer connected texts. The utterances were read at different speaking rates. Two questions are of major interest in this study: (1) To what extent do vowel and consonant durations in different environments support a relational cue for quantity? and (2) Is the dura-

tional cue weaker in those cases where the spectra of short and long vowels provide another cue for quantity? [Work supported by the Icelandic Science Foundation and University of Iceland.]

5pSC11. Timing effects on successive alveolar consonants in Swedish. Robert Bannert and Peter E. Czigler (Dept. of Phonet., Umeå Univ., S-90187 Umeå, Sweden)

The duration of a consonant is potentially affected by various factors related to the phonetic nature of the segment and its phonetic environment. The present study addresses the concurrent effects of prominence, syllable structure, and manner of articulation in intervocalic /s/+alveolar consonant /t, n, l/, and alveolar consonant+/s/ clusters in Swedish. Six real words, imbedded in a carrier sentence, were repeated five times with, and five times without, focal accent by ten native speakers of standard Swedish. Results show that the duration of a cluster, and the duration of its constituent consonants, are longer with focal accent than without focal accent. Consonant clusters with /s/ as the first member do not differ in duration from consonant clusters with /s/ as the second member. Clusters containing an alveolar stop are shorter than clusters with an alveolar nasal or alveolar liquid. The duration of /s/ is always longer than the duration of any adjacent alveolar stop, nasal or liquid. A number of interactions was found between prominence, cluster structure, and manner of articulation, indicating clear differences between the patterns of cluster internal timing for the investigated cluster pairs. These differences will be presented, compared with data on American English, and discussed.

5pSC12. Acoustic and perceptual effects of vowel length on voice onset time in Thai stops. Chutamanee Onsuwan and Patrice S. Beddor (Prog. in Linguist., Univ. of Michigan, Ann Arbor, MI 48109, beddor@umich.edu)

In Thai, vowel length (e.g., [pan]-[paan]) and unaspirated–aspirated stop (e.g., [pan]-[p^han]) distinctions are contrastive; this study examines possible acoustic and perceptual interactions between these temporal properties. VOT and vowel duration were measured for five repetitions of 40 (near-) minimal word pairs beginning with aspirated or unaspirated bilabial, alveolar, or velar stops followed by long or short vowels, produced by three native Thai speakers. Although VOT for velars was slightly longer (about 5 ms) before long than before short vowels, the overall acoustic picture across places of articulation showed no systematic effect of vowel length. To test for perceptual influences of contrastive vowel length on VOT, a 15-step 0–90 ms VOT continuum was created by editing the aspiration portion of naturally produced [p^haan]. The VOT continuum was spliced onto the vowel portion of four different natural tokens, [pan], [paan], [p^han], and [p^haan], creating four continua, two with long vowels and two with short. Identification tests were presented to 18 native Thai listeners. Responses to the continua created from the originally unaspirated tokens ([pan], [paan]) showed robust effects of vowel length: the long vowel continuum elicited more aspirated ([p^h]) responses. A compensatory account of the perceptual findings is offered. [Work supported by NSF.]

5pSC13. Perception of glottalized consonants in Babine/Witsuwit'en. Katharine Davis and Sharon Hargus (Dept. of Linguist., Univ. of Washington, Box 354340, Seattle, WA 98195-4340)

Previous studies of Babine/Witsuwit'en have shown speaker differences in the production of glottalized stops [Davis and Hargus, J. Acoust. Soc. Am. **100**, 2690 (1996)]. The present study investigates native listeners' perception of the glottal/nonglottal contrast. Ten subjects were presented with 16 naturally produced minimal pairs contrasting glottalized and voiceless unaspirated stops or affricates. These tokens, interspersed with 18 control minimal pairs, were presented in random order. Listeners' judgments were recorded on a forced-choice answer sheet. The task was performed twice, once with a male voice and once with a female voice. Subjects showed greater confusion when identifying members of the glottal/nonglottal contrast than when identifying members of the control contrasts (84% correct versus 93% correct). Among control items, voiceless unaspirated and voiceless aspirated pairs were correctly distinguished 97% of the time. ANOVA results considering all experimental factors (subject, speaker, place of articulation, etc.) will be presented. In addition, acoustic analyses of correctly identified versus confused glottal/nonglottal pairs will be presented. Particular attention will be paid to significant factors (VOT, *F*0) from earlier production studies.

5pSC14. A perceptual experiment on Thai consonant types and tones. Rungpat Roengpitya (Univ. of California, Berkeley)

Tonogenesis in Thai arose historically from neutralization of voicing contrast prevocalic consonants: new higher tonal variants after what had been voiceless consonants and new lower variants after the earlier voiced consonants. Today, Thai has again acquired a series of voiced stops and the same allophonic variation in *F*0 may be observed on the tones following voiced and voiceless stops. How important are these *F*0 cues to Thai listeners? Can one stage of the tonal multiplication process that shaped Thai tonal systems be demonstrated in the lab? A perceptual experiment was conducted. Six pairs of Thai words with voiced and voiceless initial stops were recorded, spliced at six different points of time, and put into contexts. These tokens were played to 11 native-Thai subjects for judgment as voiced or voiceless. The result showed that listeners tended to identify those syllables with gated-out bursts and closures as voiceless. However, there was a significant success at differentiating voiced and voiceless stops at gates past these points. It remains to be determined whether the cues that facilitate this are *F*0, vowel amplitude, or voice quality. If *F*0 plays a role, then this result mirrors the rise of distinctive tones in the Thai history.

5pSC15. The nasal vowels of Iberian Portuguese. Maria Joao C. Galvao (Univ. of Washington, Dept of Linguist., Box# 354340, Seattle, WA 98195-4340, mgalvao@u.washington.edu)

Previous analysis [Mateus, ''Aspectos da Fonologia Portuguesa'' (1982)] of Iberian Portuguese nasal vowels has suggested that there are three degrees of nasality, predictable from syllable structure and from the presence or absence of an underlying nasal consonant following the nasal vowel. This study addresses the question of whether vowel quality affects the degree of phonetic nasality in Portuguese. Nasalance measurements of seven native speakers of Portuguese were obtained via a Kay Nasometer 6200, while spectrographic analysis was provided by a Kay CSL 4300B. The present analysis confirms that a trace of the nasal consonant is phonetically present which provides a peak target for the nasal vowel, but the degree of nasality of the vowel itself is solely determined by vowel height: as occurs in other languages with nasal vowels, high vowels have a more constricted oral air passage and hence higher nasalance percentage.

5pSC16. EMG evidence for the automaticity of intrinsic *F*0 of vowels. D. H. Whalen, Bryan Gick, Masanobu Kumada (Haskins Labs., 270 Crown St., New Haven, CT 06511), and Kiyoshi Honda (ATR Human Information Processing Res. Labs, 2-2 Hikaridai Seika-cho Soraku-gun, Kyoto, 619-02 Japan)

Although vowels can be produced with many *F*0 values, high vowels' *F*0s tend to be higher than low vowels', a tendency found for every language examined. Nonetheless, this ''intrinsic *F*0'' (or *IF*0) has been called a deliberate ''enhancement'' of the speech signal, aiding vowel height perception. Since *IF*0 remains constant with the number of vowels in a language, this enhancement seems unlikely. The only positive evidence for it is that activity of the cricothyroid (CT) muscle, which primarily raises *F*0, is larger for high vowels than for low. The present experiment explores this by having subjects produce vowels at slightly different *F*0s, with the differences being similar to the *IF*0 magnitude for each

subject. The first subject (of four planned) showed no vowel height difference in CT; further, an analysis factoring out $F0$ as a covariate showed that high vowels had higher CT activation. Thus different vowels involve different levels of CT activity for a given $F0$; previous findings were likely due not to $F0$ control but to $F0$/vowel quality interactions. Even the EMG evidence, then, renders $IF0$ an unlikely speech enhancement. Rather, $IF0$ appears to be an automatic consequence of vowel production. [Supported by NIH Grant DC-02717.]

5pSC17. Interacting spectral and temporal properties in Jamaican Creole and Jamaican English vowel production. Alicia Beckford (Program in Linguist., Univ. of Michigan, 1076 Frieze Bldg., Ann Arbor, MI 48109-1285)

The vowel system of Jamaican Creole has been described as exploiting a contrastive length distinction, wherein long vowels and falling diphthongs contrast with single vowels in identical phonological environments [Wells (1973); Mead (1996)]. For example, /pan/ on, pan, can /tan/ stand /pen/ pen /paan/ to grasp /taan/ turn /pien/ pain. However, the nature of the relative contributions of spectral and temporal differences remains unclear. The interactions between spectral ($F1/F2$) and temporal properties, and their roles in shaping a three-dimensional vowel space ($F1\times F2\times$length) were examined in an acoustic study of the vowel production of 24 speakers at different points along the creole continuum from Jamaican Creole to Jamaican English. Data suggest that Jamaican Creole and Jamaican English speakers utilize vowel duration differently. It will be demonstrated in this paper that, for the Creole speakers, vowel duration functions to enhance slight spectral distinctions among low central and upper-mid-back vowels /a,o,u/. [Mead, ''On the phonology and orthography of Jamaican Creole,'' J. Pidgin Creole Languages (1996); Wells, *Jamaican Pronunciation in London* (Blackwell, London, 1973).]

5pSC18. The effect of $F3$ on the /ʊ–ɪ/ boundary. Matthew R. Buehler and Anna K. Nábělek (Dept. of Audiol. and Speech Pathol., Univ. of Tennessee at Knoxville, 457 S. Stadium Hall, Knoxville, TN 37996)

Boundary locations (50% response points of identification) were tested for /ʊ–ɪ/ continua for two different values of $F3$. All formants were steady-state. $F1=250$ Hz, $F2$ was varied in 20 steps from 1000 to 2000 Hz, and $F3=2500$ or 3000 Hz. All stimuli were 200-ms long. The boundaries for the two continua were determined with ten subjects. The boundary for $F3=2500$ Hz was 1652 Hz and for $F3=3000$ Hz was 1526 Hz. The difference between the boundaries was significant. These results indicate that $F3$ can affect the perception of $F2$. When $F3$ was increased the stimuli were identified as /ɪ/ at a significantly decreased value of $F2$.

5pSC19. The effect of transition slope and transition duration on vowel reduction in V_1V_2 complexes. Pierre L. Divenyi (Experimental Audiol. Res. (151), V. A. Medical Ctr., Martinez, CA 94553) and René Carré (Ecole Nationale Supérieure de Télécommunications, Paris, France)

Digital V_1V_2 vowel pairs were generated using a cascade-parallel formant synthesizer. The duration of V_1 was 100 ms and that of a completed V_2 150 ms, while the duration of the transition was varied between 40 and 100 ms, depending on the condition. For each of six English V_1V_2 pairs (/iɑ/, /ɑi/, /iu/, /ui/, /uɑ/, and /ɑu/) 8–12 tokens were obtained by progressively cutting back the V_2 portion and the vowel transition in 10-ms steps. Separately for each vowel pair, the cut-back tokens were presented to trained listeners who had to indicate their confidence level (using a 4-point scale) that the final portion of the token was, indeed, the vowel V_2. Since high-confidence V_2 responses were given not only to tokens that actually reached V_2, we observed vowel reduction; the degree of vowel reduction (VR) was measured as the severity of the cut-back at which such a misperception occurred at a certain percent of the time. Using this measure, we determined the degree of VR for different durations and velocities of transition. In general, VR was larger for steeper transition slopes and for durations around 80–100 ms. [Work supported by NATO Office of Scientific Research, a grant from N.I.H., and the V.A. Medical Research.]

5pSC20. Acoustic correlates of ''devoiced'' vowels in standard modern Japanese. J. Kevin Varden (Meiji Gakuin Univ., Dept. of English, Meiji Gakuin Univ., 1-2-37 Shirokane-dai, Minato-ku, Tokyo, 108 Japan, varden@ltr.meijigakuin.ac.jp)

This paper documents the acoustic characteristics of what have been called devoiced vowels in Tokyo Japanese, traditionally attributed to the phonological rule of high vowel devoicing [T. J. Vance, Introduction to Japanese Phonology (1987)]. More recently they have been attributed to articulatory gesture overlap [S-A. Jun and M. E. Beckman, Paper, Annual Meeting of the LSA (1993)]. Current research by the author indicates both processes are consistent with the data. The vowels which have undergone phonological processing, however, are not simply devoiced; they have undergone a process of obstruentization similar to that found in the Turkish language of Uyghur discussed both pedologically [R. F. Hahn, Spoken Uyghur (1991)] and phonologically [E. M. Kaisse, Language **68**, 313–332 (1992)]. Specifically, the vowel sites surface as voiceless fricatives at the place of articulation of the preceding voiceless consonant. This is evidenced acoustically by high-frequency frication in spectrograms and a high number of zero crossings. Additionally, formant frequency information is preserved, as evidenced by formant bars in spectrograms and spectral sections, allowing the perception of the original vowels in the face of heavy loss of vocality.

5pSC21. Nasal vibration spectra in vowels. Mechtild Tronnier (Dept. of Linguist. and Phonet., Lund Univ., Helgonabacken 12, S-22362 Lund, Sweden, mechtild.tronnier@ling.lu.se)

An accelerometer microphone attached to the upper part of the nose for measuring pathological nasality has been applied. The advantage of this technique is the easy application, the minor degree of invasiveness, and the minimal speaker distraction. The analysis of the obtained signal has been subject to discussion. Some researchers judged from signal intensity whether speech was produced on a hypernasal level or not. Others calculated the intensity ratio of the accelerometer signal to a voice source signal to eliminate intensity variation. The analysis of the spectral pattern of the accelerometer signal is suggested here. Spectral information in such a signal is a good indicator to reflect the velar port opening, which results in nasal resonance and nasal vibration at the nasal surface. Different vowels induce a different degree of velar opening, which does not necessarily reflect the degree of perceived nasality and measured acoustic resonance. This is due to different nasal and oral impedance of the vowels. Therefore, the spectral patterns of individual vowels in the accelerometer signal are described here. Simultaneously, a fiberendoscopic film has been recorded, showing the velar opening to confirm visually the presence or absence of velopharyngeal coupling.

5pSC22. Perceived vowel quantity in Swedish: Effects of postvocalic voicing. Dawn M. Behne (Norwegian Univ. of Sci. and Technol., 7034 Trondheim, Norway), Peter E. Czigler, and Kirk P. H. Sullivan (Umea Univ., 90187 Umea, Sweden)

Swedish is described as having a distinction between phonologically long and short vowels. This distinction is realized primarily through the duration of the vowels, but in some cases also through resonance characteristics of the vowels. In Swedish, like many languages, vowel duration is also longer preceding a voiced postvocalic consonant than a voiceless one. This study examines the weight of vowel duration and the first and second formant frequencies F_1-F_2 frequencies when distinguishing phonologi-

cally long and short vowel before a voiceless consonant (experiment 1) and before a voiced consonant (experiment 2). For three pairs of Swedish vowels ([iː]-[ɪ], [oː]-[ɔ], [ɑː]-[a]) 100 /kVt/ (experiment 1) and 100 /kVd/ (experiment 2) words were resynthesized having ten degrees of vowel duration and ten degrees of F_1 and F_2 adjustment. In both experiments listeners decided whether presented words contained a phonologically long or short vowel. Reaction times were also recorded. Results show that vowel duration is a more dominant cue than the first and second formant for distinguishing phonologically long and short vowels. However, for [ɑː] and [a] the perceptual influence of vowel duration and spectral attributes appears to be more complex. These findings are discussed for vowels preceding postvocalic voiceless and voiced consonants.

5pSC23. Isolating the critical segment of AE /r/ and /l/ to enhance non-native perception. Rieko Kubo, John S. Pruitt, and Reiko Akahane-Yamada (ATR Human Information Processing Res. Labs., 2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-02 Japan, rkubo@hip.atr.co.jp)

Brief synthetic stop-consonant+vowel syllables as short as 20 ms are reliably identified for place of articulation. Further, removal of the steady-state vowel in synthetic stop-consonant+vowel series improves discrimination performance, but does not affect identification. Subsequently, vowel truncation of consonant-vowel syllables has been used to improve native English speaker's perception of Hindi dental and retroflex stop-consonants. Whether such truncation techniques are effective for native speakers of Japanese on the AE liquids /r/ and /l/ was examined. The /r/ or /l/ segment from 50 minimal word pairs, produced in five phonemic positions, was isolated by truncating the surrounding phonemic environment. Six durations of the segment were created: 60, 100, 140, 180, 220 ms, and full-length (unedited productions). Native Japanese speakers were asked to identify the target phoneme for each segment, which was presented in a random sequence but blocked by speaker and segment duration. Segment duration significantly affected performance. Interestingly, while longer segments were perceived better than full-length stimuli, shorter segments were not perceived as well. Identification training with these stimuli confirmed the effectiveness of such manipulations. The implications of these findings are discussed with regard to theories of cross-language speech perception and methods of training non-native contrasts.

5pSC24. English vowel production by native speakers of Beijing Mandarin. Xinchun Wang and Murray J. Munro (Dept. of Linguist., Simon Fraser Univ., Burnaby, BC V5A 1S6, Canada, jxwang@sfu.ca)

Ten English vowels produced in a controlled elicitation paradigm by 15 Beijing Mandarin speakers and 15 native English speakers were randomly presented to trained English listeners who identified them in a forced-choice listening task. Although the non-native productions of the vowels with Mandarin counterparts were generally as intelligible as those of the native English speakers, the remaining vowels were produced with varying degrees of intelligibility. The results suggested a strong influence of the $L1$ vowel system on the production of $L2$ vowels. In particular, acoustic analyses revealed that the acccented English vowels tended to be ''pulled'' toward their Mandarin counterparts so that they fell between the English and Mandarin vowels in an $F1-F2$ space. Spectral overlap in the tense–lax pairs accounted for some intelligibility problems, while the dense distribution of second-language ($L2$) vowels in the area of the vowel space where the $L1$ vowel system is less crowded apparently increased the level of difficulty for the $L2$ learners in producing some vowels. The results are discussed in terms of error prediction models for $L2$ speech learning.

5pSC25. The effect of immersion on second language productions: The acquisition of American English /r/ and /l/ by Japanese children. Jesica C. Downs-Pruitt, Reiko Akahane-Yamada (ATR-HIP Labs, 2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-02 Japan, jcpruitt@hip.atr.co.jp), Reiko Mazuka (Duke Univ., Durham, NC 27708), and Akiko Hayashi (Tokyo Gakugei Univ., 4-1-1, Nukui-Kita, Koganei-Shi, 184 Japan)

Many studies have shown that the American English (AE) /r/ and /l/ are difficult for adult native speakers of Japanese to perceive and produce, sometimes even after substantial experience and/or intense training. In fact, there are many reports of first language ($L1$) effects on second language ($L2$) learning. Common lore promotes immersion as the best method of second language learning, particularly for overcoming $L1$ biases. The current study examined the effect of English-language immersion on Japanese children's productions of AE /r/ and /l/. We studied productions from children in three age groups, 4-, 6-, and 8-year-olds. Native Japanese children, who had been immersed in an English-language environment, nonimmersed Japanese children, and native English children produced English-language words and nonwords which included /r/, /l/, and /w/ in three positions. The productions were evaluated by categorizing them into one of seven categories, which included a ''correct'' and ''distorted'' category for each intended gesture, as well as an ''other'' category. Significant differences in intelligibility were found dependent upon the age and language-experience group. Furthermore, the children in the three language-experience groups showed different substitution patterns. The results are discussed in the context of recent theories of $L1$ development and $L2$ learning.

5pSC26. Modification of $L2$ vowel production by perception training as evaluated by acoustic analysis and native speakers. Reiko Akahane-Yamada (ATR HIP Labs., Seika, Soraku, Kyoto, 619-02 Japan, yamada@hip.atr.co.jp), Winifred Strange (Univ. of South Florida, Tampa, FL 33620-8150), Jesica C. Downs-Pruitt (ATR HIP Labs., Seika, Soraku, Kyoto, 619-02 Japan), and Yasushi Masuda (Grad. School of Information Sci. NAIST, 8916-5 Takayama, Ikoma, Nara, 630-01 Japan / ATR HIP Labs.)

Recent cross-language studies have shown that there is a link between speech perception and speech production when learning second language ($L2$) consonants. This paper reports additional evidence for such a link from a study which examines whether identification training of $L2$ vowels improves the trainees' production ability. Twenty native speakers of Japanese served as subjects. Half of the subjects received a pretest, identification training on English /ɑ, æ, ʌ/ and a post-test. For the other half of the subjects, only the pretest and post-test were administered. In addition to the perception test, productions of the English vowels were recorded in the pre- and post-test. These productions were later examined in two ways; acoustical analysis and evaluation by native English speakers. Identification ability significantly improved from pretest to post-test for the trained group, but not for the untrained group as reported by Strange and Akahane-Yamada [J. Acoust. Soc. Am. **102**, 3137 (1997)]. More importantly, production ability showed a significant improvement from pretest to post-test for the trained group, suggesting that there is a link between vowel perception and vowel production. The results are discussed in the context of the theories of $L2$ learning.

5pSC27. Categorial discrimination of English and Japanese vowels and consonants by native Japanese and English subjects. Susan G. Guion, James E. Flege (Dept. of Rehab. Sci., VH503, Univ. of Alabama, Birmingham, AL 35294, guionsg@uab.edu), Reiko Akahane-Yamada, and Jesica C. Downs-Pruitt (ATR, Kyoto, 619-02 Japan)

Categorial discrimination tests (CDTs) with equal opportunities for ''hits'' and ''false alarms'' assessed native Japanese and English subjects' perception of English and Japanese segmental contrasts. The stimuli consisted of vowels and consonants spoken at two rates by five native speakers each of English and Japanese. The listeners chose the odd item out in

triads, if there was one, or responded ''none'' if they heard three physically different tokens of a single vowel or consonant. A' scores were computed for each contrast. The subjects were ten native English (NE) controls, ten ''experienced'' bilinguals (Japanese speakers who had spent several years in the United States), and ten ''inexperienced'' bilinguals (Japanese speakers who had never lived outside Japan, but had studied English in school). Analysis of the CDT results will help determine (1) if the subjects differ in their ability to discriminate consonants and vowels, (2) if perceptual sensitivity to vowel and/or consonant contrasts is affected by the speech rate, and (3) if the experienced bilinguals more closely resemble the NE controls than the inexperienced bilinguals do. [Work supported by NIH.]

5pSC28. Effects of listener training on intelligibility of Chinese-accented English. Catherine L. Rogers (Dept. of Speech and Hearing Sci., Ohio State Univ., 110 Pressey Hall, 1070 Carmack Rd., Columbus, OH 43210, rogers.140@osu.edu)

Previous research has demonstrated significant improvements in the intelligibility of synthetic speech following listener training [E. Schwab *et al.*, Human Factors **27**, 395–408 (1985)]. In the present study, the effect of listener training on second-language speech intelligibility was investigated. Chinese-accented words and sentences were presented to four groups of native English listeners—one training and one control group for each of two speakers. On word trials, listeners were instructed to select the word presented from two minimal-pair alternatives; on sentence trials, listeners were instructed to write down what they heard. Listener responses were scored as percent of intended words correctly identified. The two training groups participated in three 1-h training sessions. Auditory and visual feedback was provided to listeners during the training sessions only. Performance of the training and control groups was compared across pre- and posttraining test sessions. Improvement in minimal-pair word intelligibility following training was found across speakers for the training group, relative to the control group; a similar but smaller improvement in sentence intelligibility was found. Word results will be discussed in terms of improvement in discriminability following training for the twelve phonemic contrasts tested. [Work supported in part by NIH-NIDCD Grant No. 2R44DC02213.]

5pSC29. An instance-based model of Japanese speech recognition by native and non-native listeners. Kiyoko Yoneyama and Keith Johnson (Dept. of Linguist., Ohio State Univ., 222 Oxley Hall, 1712 Neil Ave., Columbus, OH 43210)

Japanese has a short and long segment contrast (*oji-san* versus *ojii-san*, *kita* versus *kitta*, and *kana* versus *kanna*). Previous research has shown that this durational property is one of the characteristics of moras in Japanese. It has also been reported that Japanese learners generally have difficulty in acquiring this contrast, but after extensive language experience with Japanese, they do learn the contrast. This suggests that the contrasts and similarities are based on remembered instances of linguistic objects in various environments during a lifetime. This paper investigates whether an instance-based approach of speech recognition can replicate performance by native listeners and second language learners of Japanese. An instance-based model was trained with naturally produced utterances of *koma* spin and *komma* comma and then tested on a 24-step nasal duration continuum from *koma* to *komma* [T. Uchida, doctoral dissertation, Nagoya University (1996)]. The identification function produced by the model was very similar to identification responses produced by native Japanese listeners. Additional simulations trained with utterances produced by English native speakers (e.g., *coma*, *comma*) will be reported. These simulations will address the role of linguistic experience in second language acquisition. [Supported by NIDCD R29-016435-06.]

5pSC30. Training American listeners to perceive and produce Mandarin tones. Yue Wang, Allard Jongman, and Joan A. Sereno (Dept. of Linguist. and Modern Lang., Cornell Univ., Ithaca, NY 14853, yw36@cornell.edu)

Auditory training of non-native speech contrasts is based on the assumption that the perceptual system of adults can be modified. Previous research has shown substantial improvements in the identification of segmental distinctions which are absent in the listeners' native language after a simple phonetic laboratory training procedure. However, it is not known whether such a training procedure is applicable to the acquisition of non-native suprasegmental contrasts. In the present study, native speakers of English are trained to perceive Mandarin lexical tones, using the high-variability procedure shown to be efficient in previous studies concerning the generalization and long-term retention of the training effects. In 20 sessions lasting four weeks, ten American learners of Chinese are trained to identify the four tones in natural words occurring in different phonetic contexts, and spoken by different talkers. The results of the pre- and post-tests are compared to determine if suprasegmental contrasts can be obtained and improved by training, if the contrasts obtained can be generalized to novel words and talkers, if contrasts can be maintained long after training, and if learning gained perceptually can be transferred to production. The results will be discussed in terms of nonnative perceptual modifications at the suprasegmental level.

5pSC31. A CALL system for teaching the duration and phone quality of Japanese tokushuhaku. Goh Kawai and Keikichi Hirose (Univ. of Tokyo, Dept. of Information and Commun. Eng., 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113 Japan, goh@kawai.com)

This paper describes a CALL (computer-aided language learning) system that teaches the pronunciation of Japanese tokushuhaku (long vowels, the mora nasal, and mora obstruents) to beginning learners of Japanese as a second language. The system detects mispronounced tokushuhaku durations and quality, including, for instance, shortened or lengthened tokushuhaku, dentalized mora nasals, diphthongized long vowels, and aspirated plosives. Learners are prompted by the system to read tokushuhaku minimal pairs aloud. Tokushuhaku durations are measured using speech recognition and compared with distributions of native speakers. Mistakes in phone quality are detected using a speech recognizer incorporating bilingual monophone models of both the learner's native and target languages. HMMs for the two languages are trained separately on language-dependent speech data, but are bundled together during recognition so that the closest phone recognized indicates the non-nativeness of the utterance should the recognized phone not be Japanese. Knowing the phonetics and phonology of the learner's native language can identify non-native articulatory gestures that result in Japanese pronunciation errors, thus allowing precise corrective feedback to the learner. The paper includes an overview of the system, its reliability, validity, and effectiveness in the foreigen language classroom.

5pSC32. The organization of articulator gestures: A comparison of Swedish, Bulgarian, and Greenlandic. Sidney A. J. Wood (Dept. of Linguist., Univ. of Lund, Helgonabacken 12, S-22362 Lund, Sweden, sidney.wood@ling.lu.se)

This paper reports speech articulator gestures in Swedish, Bulgarian, and Greenlandic to study universal and language-specific components in articulation. This work is based on analysis of movements of individual articulators from x-ray motion films of speech, and continues from previous reports from this study by Wood [J. Phon. **7**, 25–43 (1979); **19**, 281–292 (1991); Proc. 3rd Conf. I.C.P.L.A., 191–200, Helsinki (1994); Proc. 13th I.C.Ph.S. Vol. 1, 392–395, Stockholm (1995); J. Phon. **24**, 139–164 (1996); Proc. 4th Speech Prod. Sem., 61–64, Grenoble (1996); Speech Commun., (in press)], and by Wood and Pettersson [Folia Ling. **22**, 239–262 (1988)]. Common principles concerned utilization and integration of articulator gestures (articulator gestures executed in approach, hold, and withdrawal phases, four tongue body gestures used, all gestures

available for vowels and consonants, gesture conflicts resolved by gesture queuing, coarticulation and assimilations implemented by coproduction). Language-specific principles concerned implementation of assimilations like palatalization of alveolar stops in Bulgarian and uvularization of vowels in Greenlandic. One assimilation, palatalization of velar consonants, is common to all three languages. A model of gesture programming based on these results is proposed.

5pSC33. The relationship between non-native phoneme perception and the MMN. Janet W. Stack and Susan D. Dalebout (Commun. Disord. Prog., Univ. of Virginia, 2205 Fontaine Ave., Charlottesville, VA 22903 jws8n@virginia.edu)

The mismatch negativity (MMN) auditory-evoked potential reflects the detection of minimal stimulus differences at a preattentive level. It has been reported that the MMN reflects acoustic, rather than phonetic processing of speech stimuli. The purpose of the present study was to examine the relationship between discrimination at a neurophysiologic level, as reflected by the MMN, and the behavioral ability of native speakers of American English to discriminate and label pairs of Hindi stop consonants that are contrasted by a place of articulation feature that is not phonemic in English. These pairs were the voiceless unaspirated Hindi stop consonants, dental and retroflex /t/, and the voiced unaspirated Hindi stops, dental and retroflex /d/. The present study addressed two questions: (1) whether the MMN response would be evoked by acoustic differences resulting from differences in place of articulation, even though behavioral discrimination ability was poor; and (2) whether there would be a difference in the presence/absence or other parameters (latency, duration, amplitude, and area) of the MMN response depending on which voicing contrast was presented.

5pSC34. Acquisition of language-specific word-initial unvoiced stops: VOT, intensity, and spectral shape in American English and Swedish. Eugene H. Buder (Univ. of Memphis, Memphis, TN 38105, ehbuder@memphis.edu) and Carol Stoel-Gammon (Univ. of Washington, Seattle, WA 98105-6246)

By 30 months of age, children learning American English (AE) or Swedish (S) have begun to acquire language specific aspects of word-initial /t/s [Stoel-Gammon *et al.*, Phonetica **51**, 146–158 (1994)]. The coronal stop is ostensibly alveolar in AE and dental in S, suggesting a place mechanism for the distinctive acoustic measures obtained from both adults and toddlers from these languages (specifically burst frequency shape as measured by spectral mean and standard deviation, and burst intensity as calibrated to following vowel). However, these measures, as well as voice onset time, are also distinctive in the same directions for other unvoiced stop consonants in adults and children, suggesting a system-wide mechanism for language-specific behaviors in this domain. Current results on word-initial /p/s, /t/s, and /k/s from both 24-month- and 30-month-old children are interpreted as evidence for early acquisition of specific allophonic phonetic aspects of consonants. [Work supported by NICHD.]

5pSC35. The effects of postvocalic voicing on the duration of high front vowels in Swedish and American English: Developmental data. Carol Stoel-Gammon (Univ. of Washington, Seattle, WA 98105-6246) and Eugene H. Buder (Univ. of Memphis, Memphis, TN 38105)

In most languages of the world, a vowel preceding a voiced obstruent is longer than the same vowel preceding a voiceless obstruent. Although the effect of postvocalic voicing on vowel duration is often considered to be phonetically driven, the extent of influence differs considerably across languages. In English, for example, vowels preceding voiced obstruents are nearly twice as long as those preceding voiceless obstruents, whereas in Swedish, the influence is minimal, perhaps because vowel length is phonemically contrastive in this language. The present study examines acquisition data from Swedish and American children, aged 24 and 30 months, to determine developmental patterns of vowel duration associated with context-sensitive voicing. It was hypothesized that the 24-month-old children from both language backgrounds would exhibit the phonetically based tendency for vowels to be significantly longer in the voiced context, but that at 30 months of age, the effects of final consonant voicing would be much stronger in English than in Swedish. Durational measures of high front vowels (tense/long /i/ and its short/lax counterpart) supported the proposed hypothesis. [Work supported by NICHD.]

5pSC36. Perceptual assimilation of Hindi dental and retroflex stop-consonants by native English and Japanese speakers. John S. Pruitt, Reiko Akahane-Yamada (ATR Human Information Processing Res. Labs., 2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-02 Japan, jspruitt@hip.atr.co.jp), and Winifred Strange (Univ. of South Florida, Tampa, FL 33620)

Previous research has shown that English speakers have great difficulty distinguishing the dental and retroflex stop-consonants of the Hindi language. However, native Japanese speakers have somewhat less difficulty perceiving this contrast even though it is not employed in the Japanese language. Best and colleagues have attributed similar differences in the perception of non-native speech sounds to differences in the assimilation of such sounds to native-language speech categories [Best, McRoberts, and Sithole, J. Exp. Psych: Human Percept. Perform. **14** (1988)]. According to Best's Perceptual Assimilation Model (PAM), four patterns of assimilation have been described which are purportedly predictive of perceptual difficulty. To determine if this model could be applied to the present case, native speakers of English and Japanese transcribed multiple instances of the Hindi consonants in different voicing/manner classes, produced with different vowels. Marked differences were found between the responses of the two language groups which were partially dependent upon the voicing/manner class of the contrast. Interestingly, an assimilation pattern emerged which was different from those discussed in PAM. Implications of these findings for PAM and for perceptual training of non-native speech sounds will be discussed. [Work supported in part by NIDCD.]

5pSC37. Comparison of American English vowel production and identification by native speakers of Luso-Portuguese. Christina F. Famoso, Patricia N. Schwartz (Dept. of Speech, Commun. Sci., and Theatre, St. John's Univ., Jamaica, NY 11439), and Adelia DaSilva (Blythedale Children's Hospital, Valhalla, NY)

Learning a new language requires mastering the phonetic as well as lexical and syntactical aspects of the new language. As part of a larger study of second-language learning, the identification of American English vowels by speakers of Luso-Portuguese and the identification of those speakers' productions by native speakers of American English were explored. Four adult bilingual Luso-Portuguese/English speakers (two women and two men) participated in this study. The subjects performed two tasks. First, they identified ten English words (ten repetitions of each, presented in random order, of: heed, hid, head, had, hod, hawed, hood, who'd, hud, heard) produced by a native speaker of American English. Then recordings were made of the subjects as they produced four repetitions each, in random order, of these ten English words. These recordings were digitized using SoundDesigner II software with an audiomedia board on a Macintosh Quadra 800 computer. Listening tapes will be prepared, using SoundDesigner II software, for identification of the Luso-Portuguese speakers' vowels by native American English speakers. The results of the

identification of the Luso-Portuguese speakers' vowels will be compared with the identification of American English vowels by the native Luso-Portuguese speakers. [Work supported by St. John's University.]

5pSC38. Acoustic analysis of American English vowels by native speakers of Luso-Portuguese. Patricia N. Schwartz, Christina F. Famoso (Dept. of Speech, Commun. Sci., and Theatre, St. John's Univ., Jamaica, NY 11439), and Adelia DaSilva (Blythedale Children's Hospital, Valhalla, NY)

The ability of individuals to learn a new language requires mastering the phonetic as well as lexical and syntactical aspects of the new language. As part of a larger study of second-language learning, the acoustical characteristics of American English vowel production by native speakers of Luso-Portuguese were explored, and these characteristics were compared with the identification of these speakers' English vowels by native speakers of American English. Four adult bilingual Luso-Portuguese/English speakers (two women and two men) participated in this study. Recordings were made of the subjects as they produced four repetitions each, in random order, of 10 English words (heed, hid, head, had, hod, hawed, hood, who'd, hud, heard). The recordings were digitized using SoundDesigner II software with an audiomedia board; formant frequency analyses will be carried out using Signalyze 2.0 software implemented on a Macintosh Quadra 800 computer. The acoustical analyses will be examined in light of the identification of the Luso-Portuguese speakers' English vowels by native speakers of American English (reported in another paper at this meeting). [Work supported by St. John's University.]

5pSC39. Comparison of American English vowel production and identification by native speakers of Russian. Fredericka Bell-Berti (Dept. of Speech, Commun. Sci., and Theatre, St. John's Univ., Jamaica, NY 11439), Lisa Jayne Romano, and Eugenia Lorin (St. John's Univ., Jamaica, NY 11439)

Learning a new language requires mastering the phonetic as well as lexical and syntactical aspects of the new language. As part of a larger study of second-language learning, the identification of American English vowels by speakers of Russian and the identification of those speakers' productions by native speakers of American English were explored. Four adult bilingual Russian/English speakers (two women and two men) participated in this study. The subjects performed two tasks. First, they identified ten English words (ten repetitions of each, presented in random order, of: heed, hid, head, had, hod, hawed, hood, who'd, hud, heard) produced by a native speaker of American English. Then recordings were made of the subjects as they produced four repetitions each, in random order, of these ten English words. These recordings were digitized using SoundDesigner II software with an audiomedia board on a Macintosh Quadra 800 computer. Listening tapes will be prepared, using SoundDesigner II software, for identification of the Russian speakers' vowels by native American English speakers. The results of the identification of the Russian speaker's vowels will be compared with the identification of American English vowels by the native Russian speakers. [Work supported by St. John's University.]

5pSC40. Acoustic analysis of American English vowels by native speakers of Russian. Lisa Jayne Romano (Dept. of Speech, Commun. Sci., and Theatre, St. John's Univ., Jamaica, NY 11439), Fredericka Bell-Berti (St. John's Univ., Jamaica, NY 11439), and Eugenia Lorin (St. John's Univ., Jamaica, NY 11439)

The ability of individuals to learn a new language requires mastering the phonetic as well as lexical and syntactical aspects of the new language. As part of a larger study of second-language learning, the acoustical characteristics of American English vowel production by native speakers of Russian were explored and these characteristics were compared with the identification of these speakers' English vowels by native speakers of American English. Four adult bilingual Russian/English speakers (two women and two men) participated in this study. Recordings were made of the subjects as they produced four repetitions each, in random order, of ten English words (heed, hid, head, had, hod, hawed, hood, who'd, hud, heard). The recordings were digitized using SoundDesigner II software with an audiomedia board; formant frequency analyses will be carried out using Signayze 2.0 software implemented on a Macintosh Quadra 800 computer. The acoustical analyses will be examined in light of the identification of the Russian speakers' English vowels by native speakers of American English (reported in another paper at this meeting). [Work supported by St. John's University.]

5pSC41. Is allophonic variation in Japanese /ɾ/ a factor in Japanese listeners' difficulty in perceiving English /ɹ-l/? Mieko Ueno, Patrice S. Beddor (Prog. in Linguist., Univ. of Michigan, Ann Arbor, MI 48109, beddor@umich.edu), and Reiko Akayane-Yamada (ATR Human Information Processing Res. Labs., Kyoto, Japan)

Japanese listeners' difficulty in perceiving the English liquid (/ɹ-l/) contrast has been attributed to various phonological and phonetic factors. To explore allophonic variation in the Japansese flap as a contributing factor, American English listeners' perception of Japanese /ɾ/ was investigated in a range of phonetic contexts. Stimuli were 152 Japanese words containing /ɾ/ in all phonotactically possible contexts produced by two native Japanese speakers. Three stimulus randomizations were presented to 12 American English listeners for identification as ''l,'' ''r,'' ''D'' (=flap), ''d,'' ''t,'' ''no sound,'' or ''other sound.'' Results showed strong effects of syllable structure: Japanese /ɾ/ sounded more stoplike (''D,'' ''d,'' ''t'' responses) in consonantal contexts and more liquidlike (''r,'' ''l'' responses) in vocalic contexts. However, liquid responses were more likely to be ''r'' intervocalically, compared with ''l'' in initial position. Vowel quality also had a strong effect: Japanese /ɾ/ elicited more stop responses in high vowel ([i u]) contexts and more liquid responses in lower vowel ([a o]) contexts. Acoustic measures of the stimuli were generally consistent with the perceptual patterns. The perceptual and acoustic findings are compared with the results of recent training studies [Akayane-Yamada, Proc. 3rd Joint Meeting of ASA-ASJ, 953–958 (1996)] with Japanese listeners on English /ɹ–l/.

5pSC42. The perception of place of articulation in three coronal nasal contrasts. James D. Harnsberger (Prog. in Linguist., Univ. of Michigan, 1076 Frieze Bldg., Ann Arbor, MI 48109, jharns@umich.edu)

Place of articulation contrasts which undergo perceptual assimilation are often difficult to discriminate if absent in a listener's native language, as demonstrated by American English (AE) listeners with Hindi dental-retroflex stops [Polka (1991)] and Salish velar-uvular ejectives [Werker and Tees (1984)]. To see if this pattern generalizes to other manners of articulation, nasal contrasts (dental-alveolar, dental-retroflex, alveolar-retroflex) from three languages (Malayalam, Marathi, and Oriya), varying in talker, syllabic position, and vowel context, were presented to speakers of Malayalam, Marathi, and Oriya (control groups), and Bengali, AE, Tamil, and Punjabi in two tests: AXB discrimination, and identification with goodness ratings. AE and Bengali were chosen because both possess

a single relevant perceptual category for these stimuli, an alveolar nasal. Tamil and Punjabi were selected for their two nasal categories, dental and retroflex. Of the contrasts tested, Malayalam dental-alveolar was the most challenging, with all non-native groups performing at-chance on the AXB test. The discriminability of the other contrasts proved to be highly dependent on the particular talker, syllable position, or vowel context, with performance ranging from at-chance to 98.4%. An acoustic analysis of the stimuli will also be provided.

5pSC43. Vowel perception by formant variation. Byunggon Yang (English Dept., Dongeui Univ., Kayadong Pusanjingu, Pusan 614-714, Korea)

The acoustic parameters of six Korean vowels produced by a healthy male subject with normal hearing were analyzed to synthesize the six vowels by a formant synthesis method. The f_0, amplitude envelopes, and the first six formant values were elaborately matched to the those of the original vowels. Then, F_1 and F_2 values of the synthesis file were modified by 50 Hz over and under the original values but not interfering with adjacent formants. F_3 varied 500 Hz over and under the original values by 100 Hz to obtain 270 stimuli. Twenty male and female subjects in a quiet room listened to the original vowel followed by the corresponding synthesized one 0.5 s later and judged whether they sounded qualitatively ''similar'' or ''different.'' From the experiment, it was found that male subjects responded ''similar'' within the range of variance by an average of 163 Hz for F_1, 415 Hz for F_2; and 843 Hz for F_3. Female subjects showed 40-Hz higher range than males. The ranges supported the nonlinearity of the human auditory scale in the speech sound. This suggests that the validity of exact formant frequency measurement varies according to frequency levels.

5pSC44. Cross-linguistic evidence for the acquisition of accent by the onset of word use. Rory A. DePaolis, Marilyn M. Vihman (Dept. of Psych., Univ. of Wales, Bangor, Bangor LL57 2DG, Gwynedd, UK, r.depaolis@bangor.ac.uk), and Leonard P. Lefkovitch (Carleton Univ., Ottawa, Canada)

An investigation of the acoustic correlates of stress or accent in infant vocalizations was undertaken using five English- and five French-learning participants at two developmental points, the onset of word use (10–13 months) and late in the single-word period (14–19 months). A data base of 555 disyllabic utterances (words and babble) was digitized from naturalistic recordings and quantified acoustically by indices of duration, amplitude, and fundamental frequency (F_0). Cluster analysis as well as the results of perceptual judgments (i.e., both a bottom-up and a top-down approach) revealed a developmental trend toward the adult prosodic system. In addition, acoustic comparisons of French and American infants at both developmental points revealed differences in their use of amplitude, F_0, and duration to accent a syllable. The differences were traceable to the adult patterns for the respective adult languages. These results suggest that infants' perceptual attunement to prosody in the prelinguistic period manifests itself in the beginnings of adultlike production as early as the onset of word use. [Work supported in part by the UK Economic and Social Research Council.]

FRIDAY AFTERNOON, 26 JUNE 1998 GRAND BALLROOM B (S), 1:00 TO 5:00 P.M.

Session 5pUW

Underwater Acoustics: Bistatic Scattering From the Seabed and Surface: Models and Measurements II

Paul C. Hines, Chair
Defence Research Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada

Chair's Introduction—1:00

Invited Papers

1:05

5pUW1. Implications of a bistatic treatment for the second echo from a normal incidence sonar. Gary J. Heald (Defence Evaluation and Res. Agency, Weymouth, Dorset, England) and Nicholas G. Pace (Univ. of Bath, Claverton Down, Bath, England)

Echo sounders have traditionally been used for the measurement of water depth but recent developments have shown how the information contained in the acoustic backscatter, related to the sediment type, may be extracted. As echo sounders operate at normal incidence it is possible to use the first and second backscatter from the seabed for sediment discrimination. This study has investigated why additional information may be contained in the second backscatter. It is shown that the geometry for the second backscatter must be treated as an on-axis bistatic configuration in which the source is effectively 3 water depths above the seabed. The resulting scattering equations are used to show how the range dependence of the scattering places the receiver in the near field for the second echo. The importance of the sensor geometry and the processing of the received signals are also discussed. As a result, it is shown that the combination of information contained in the first and second backscatter may be used to descriminate between different sediments. The result from tank experiments using a 200-kHz transducer over different sediments has allowed comparison with the theory and shows the effect of different water surface roughness on the second backscatter.

1:25

5pUW2. In-plane bistatic calculations of sub-bottom volume reverberation in shallow water. Dale D. Ellis and Paul C. Hines (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada)

A shallow-water reverberation model based on normal modes and ray-mode analogies has been previously developed [J. Acoust. Soc. Am. **97**, 2804–2814 (1995)]. It was first developed for boundary scattering and recently extended to handle volume reverberation in both the water and sub-bottom. The volume scattering is formulated in terms of an empirical plane-wave scattering function, which depends on depth but which can also depend on the incident and scattered angles. Our first approximation for the volume reverberation

is to choose the scattering to be isotropic, but to get a better approximation based on physical parameters, the in-plane bistatic scattering function of Hines is used [J. Acoust. Soc. Am. **99**, 836–844 (1996)]. In this paper, calculations of shallow-water reverberation are made with the Hines scattering function used as input to the normal-mode model. Results are compared with reverberation assuming isotropic scattering in the bottom, as well as with reverberation due to bottom interface scattering using Lambert's rule.

Contributed Papers

1:45

5pUW3. A review of the SSA for rough surface scattering. Shira L. Broschat (School of Elec. Eng. and Comput. Sci., Washington State Univ., P.O. Box 642752, Pullman, WA 99164-2752) and Eric I. Thorsos (Univ. of Washington, Seattle, WA 98105)

The small slope approximation (SSA) for scattering from rough surfaces was introduced by Voronovich in the mid-1980s [A. G. Voronovich, Sov. Phys. JETP **62**, 65–70 (1985)]. Numerical studies have shown that the SSA gives excellent results over a broad range of scattering angles. However, in the high-frequency limit the second-order (in the field) SSA scattering amplitude reduces to the geometrical optics result and, thus, to a single scattering model limited to local interactions with the surface. In this paper, small slope theory will be reviewed briefly, expressions for the bistatic scattering cross section will be presented, and numerical results will be shown for a number of different problems and several different roughness spectra. In addition, in an effort to extend the nonlocal scattering effects, Voronovich modified the SSA. He applied the same approach underlying the SSA, but he began by iterating the integral equation of the second kind [A. G. Voronovich, Waves Random Media **6**, 151–167 (1996)] which led to the nonlocal small slope approximation (NLSSA). While the main focus of this talk will be the SSA, the NLSSA will be compared with the SSA for selected results. [Work supported by ONR.]

1:57

5pUW4. The lowest-order small slope approximation for rough surface scattering. Eric I. Thorsos (Appl. Phys. Lab., College of Ocean and Fishery Sci., Univ. of Washington, Seattle, WA 98105) and Shira L. Broschat (Washington State Univ., P.O. Box 642752, Pullman, WA 99164-2752)

The small slope approximation (SSA) [A. G. Voronovich, Sov. Phys. JETP **62**, 65–70 (1985)] for scattering from rough surfaces leads to a systematic series and improved accuracy is obtained by using higher terms in the series. Numerical implementation of the lowest-order SSA for the scattering cross section has the same level of effort required as for the Kirchhoff approximation (a single integration for one-dimensional surfaces, a double integral for two-dimensional surfaces), while higher orders require higher-order integrations. Thus the accuracy of the lowest-order SSA is of considerable practical interest. In this paper, the accuracy of lowest-order SSA will be illustrated for surfaces with power-law roughness spectra through comparison with exact numerical results. Both Dirichlet and two-fluid boundary conditions will be considered. For typical power-law spectra, the accuracy of the lowest-order SSA is generally quite good. [Work supported by ONR.]

2:09

5pUW5. The Kirchhoff approximation: A new iterative solution. Suzanne T. McDaniel (4035 Luxury Ln., Bremerton, WA 98311)

A new iterative solution to the problem of sea surface scattering is obtained which has the Kirchhoff approximation as its leading term. This solution is obtained by transforming an integral equation of the first kind for the surface potential to one of the second kind, and applying the method of successive approximations. The advantages of this method are that the conditions under which this solution converges may be readily established and the solution may be iterated to sufficiently high order in the interaction to examine the rate at which it converges. A numerical example of scattering from an echelette grating is considered, and the regime of convergence established. It is found that the new iterative solution converges only for such surfaces on which $k\sigma\sigma' < \pi$ where k is the wave number of the scattered radiation and σ and σ' are, respectively, the rms surface height and slope. This finding raises doubts concerning the regime in which the Kirchhoff approximation is traditionally deemed applicable. Finally, numerical studies are presented to illustrate the convergence of the iterative solution.

2:21

5pUW6. Statistical characteristics of bistatic sea surface scatter. Peter D. Neumann (Planning Systems, Inc., 7923 Jones Branch Dr., McLean, VA 22102) and R. Lee Culver (Penn State Univ., University Park, PA 16804)

The characterization of reverberation, from the surface, the bottom, and/or the volume, is important in advancing the understanding of the mechanisms involved in underwater sound scattering. Numerous experiments and much theoretical work have been done in the area of monostatic reverberation, but the area of bistatic reverberation has been much less researched. This certainly is a result of the much more complex geometries involved in bistatic reverberation. This work studies the statistics of bistatic sea surface reverberation data from the FLIP experiment conducted in January of 1992. The data were verified as homogeneous and normally distributed using the Kolmogorov–Smirnov two-sample test with a confidence interval, α, equal to 0.1. Using a technique by which ensembles from different times were combined, meaningful deviations from a normal distribution were observed at the highest wind speed of 7.2 m/s. The bistatic surface scattering strengths were calculated from the data and compared with the predictions of a theory developed by S. T. McDaniel for the prediction of bistatic surface scattering strengths. The comparison showed good agreement with a rms deviation of about 3 dB. [Work supported by ONR.]

2:33

5pUW7. Time dependence in forward scattering: Experimental results and model comparisons derived from sea surface and seabed bistatic cross sections. Peter H. Dahl and Kevin L. Williams (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98195)

In shallow, littoral regions the acoustic transmission channel can often be governed by the process of forward scattering from both the sea surface and seabed. In this paper, measurements of the intensity time dependence of high-frequency O(10) kHz sound that has been forward scattered from both the sea surface and seabed are discussed along with simulated time series for this process. The data are from an experiment conducted near Key West, Florida in waters 25-m deep, for which the seabed was characterized by a carbonate sand-silt-clay mixture. The simulated time series are derived from the bistatic cross sections for sea surface and seabed by convolving the bistatic scattering impulse response with the waveform emitted from the transmitter. For the seabed *in situ* estimates of sediment properties were used as inputs to a model described in Williams and Jackson [J. Acoust. Soc. Am. (submitted)]. For the sea surface, the necessary representation of the two-dimensional surface wave correlation function was derived from estimates of the directional wave spectrum. The latter were measured with a directional wave buoy that operated in the immediate vicinity of the acoustic measurements. [Work supported by ONR Code 321OA.]

2:45

5pUW8. Stacking and averaging techniques for bottom echo characterization. Daniel D. Sternlicht and Christian P. deMoustier (Marine Physical Lab., Scripps Inst. of Oceanogr., 9500 Gilman Dr., La Jolla, CA 92093-0205)

Bottom-looking sonars operating at frequencies greater than 30 kHz transmit acoustic signals which penetrate at most a few meters into seafloor sediments, making them well suited to characterize the water/sediment interface. Acoustic wavelengths at these frequencies are small compared to the rms relief of the interface, resulting in bottom echoes dominated by incoherent energy and varying significantly in amplitude and shape as the sonar translates longitudinally above the interface. Because of this variability, echoes must be treated stochastically and determining the average shape of the bottom echo involves careful selection of stacking and averaging algorithms. Comparisons of a geoacoustic temporal model with 33- and 93-kHz backscatter echoes from sand and silt substrates reveal that a representative echo shape is obtained by averaging echoes along-track over a distance roughly equivalent to the 6-dB footprint diameter of the sonar beam pattern. To be meaningful, such averaging must be performed on echoes that have been aligned in time, thus removing effects due to transducer heave as well as small depth variations over consecutive pings. Echo alignment techniques based on thresholding minima, maxima, or half energy, on low-pass filtering techniques and on carrier frequency phase tracking, have been investigated and their relative merits are discussed. [Work supported by ONR N00014-94-1-0121.]

2:57–3:12 Break

3:12

5pUW9. A comparison of the finite-difference time-domain and integral equation methods for scattering from shallow water sediment bottoms. Frank D. Hastings, John B. Schneider, Shira L. Broschat (School of Elec. Eng. and Comput. Sci., Washington State Univ., Pullman, WA 99164-2752, shira@eecs.wsu.edu), and Eric I. Thorsos (Univ. of Washington, Seattle, WA 98105)

Scattering from sediment bottoms has a significant effect on acoustic propagation in shallow water. When shear effects are negligible, the water–sediment interface can be modeled using a fluid–fluid model. In this paper, finite-difference time-domain (FDTD) and integral equation (IE) results are compared for a fluid–fluid rough interface. Both methods are exact in that they make no physical approximations, and both can be used either to benchmark approximate analytical methods or to study the propagation problem directly. The comparison presented verifies the numerical accuracy of the two methods. Scattering strengths are computed using a Monte Carlo average over a set of 50 surfaces. Results for single surface realizations are presented as well. Both single-scale Gaussian and multiscale modified power law (MPL) roughness spectra results are presented. While the MPL spectrum is more realistic for shallow water problems, the Gaussian spectrum gives low scattering levels in the backscatter direction, permitting a stringent test of numerical accuracy. Several different bottom types are considered for which the shear speeds are negligible, and attenuation has not been included. For all cases, the FDTD and IE results are obtained using the same surface realizations. In general, the results agree well. [Work supported by ONR.]

3:24

5pUW10. Shallow-water in-plane bistatic scattering experiment. D. Vance Crowe, Paul C. Hines, and Patrick J. Barry (Defence Res. Establishment Atlantic, Dartmouth, NS, Canada)

Every reverberation experiment done in shallow water is by its nature, bistatic at long ranges. That is to say, in addition to simple monostatic returns, the received signal is composed of energy arriving along bottom bounce paths for which incident and scattered angles are different. Therefore, it is possible to measure in-plane bistatic scattering strength using a monostatic geometry [Hines, Crowe, and Ellis, J. Acoust. Soc. Am. (in review)]. Noting this, a simple array consisting of a pair of free-flooding-ring projectors and two hydrophones was used to measure monostatic and bistatic scatter in 100 m of water on the Scotian Shelf. A series of linear FM (LFM) pulses was transmitted in four frequency bands over the range 900 to 2100 Hz. The reverberation data between the first bottom return and the first surface return provided estimates of monostatic backscatter strength for grazing angles down to 10°. Data arriving after the first-bottom–first-surface interaction provided estimates of in-plane bistatic scattering strength for pairs of incident and scattered grazing angles (ϕ_i, ϕ_s) down to (7°, 57°). In this paper, the experiment and data are described in light of the monostatic and bistatic arrivals.

3:36

5pUW11. Sediment density heterogeneity spectra estimated from digital X-radiographs. Dajun Tang (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105, djtang@apl.washington.edu) and Robert A. Wheatcroft (Oregon State Univ., Corvallis, OR 97331)

It has been confirmed that sediment volume heterogeneity is one of the major bottom scattering mechanisms, especially in soft-sediment areas. In most cases, scattering by sediment heterogeneity can be modeled by a first-order perturbations technique. The crucial input to such models is the spectrum of sediment spatial heterogeneity. It is difficult to obtain reliable data to estimate such spectra, however, because (1) two-dimensional data sets collected from statistically similar sites are needed, and (2) the data need to cover several orders of magnitude in spatial dimensions. Digital X-radiographs of sediment cores provide an ideal form of data for estimating the spectrum of sub-bottom density variability. Here the data processing procedures will be provided in detail, and physical insights into the resultant spectra discussed. Last, bistatic scattering coefficients resulting from these spectra will be presented.

3:48

5pUW12. Measurement and modeling of bistatic scattering on the Florida Atlantic coastal shelf. Tokuo Yamamoto, Christopher Day, and Murat Kuru (Univ. of Miami, Div. of Appl. Marine Phys., 4600 Rickenbacker Cswy., Miami, FL 33149)

Measurements of bistatic sea floor scattering on the Fort Pierce, Florida shelf are presented. Scattering measurements have been collected using a new 32 element bilinear acoustic array. According to the volume and roughness scattering models of Yamamoto (1996) and Jackson *et al.* (1992), bistatic scattering depends on the bottom sediment characteristics, the interface roughness, and the incident and scattered grazing angles. Parameters for the combined volume and roughness scattering model are estimated using the observed bottom roughness and tomographically measured sediment sound speed and attenuation. Observations of bistatic scattering are compared to the combined scattering model. Good agreement between the observed and modeled scattering pattern permits an accurate mapping of mean sound speeds, attenuation, densities, and their three-dimensional spatial spectra using observations of bistatic scattering. [Work supported by ONR.]

4:00

5pUW13. Analysis of laboratory measurements of sound propagating into an unconsolidated water-saturated porous media. Harry J. Simpson and Brian H. Houston (Naval Res. Lab., 4555 Overlook Ave., SW, Washington, DC 20375-5350, simpson@fugazi.nrl.navy.mil)

A series of measurements in the NRL shallow-water laboratory have been conducted to understand the propagation of sound into and within an unconsolidated water-saturated porous medium. The initial set of measurements was taken to quantify sound penetration as a function of incident angle and interface roughness and to identify the types of waves generated within the water-saturated sandy bottom. These investigations included two-dimensional synthetic array measurements for both a smoothed and a roughened interface. Additionally, above interface and buried source, one-dimensional synthetic array measurements were taken to investigate specifically the types of compressional waves generated within the sandy bottom. To complete the physical characterization of the wave speeds

propagating in the bottom, measurements of the shear wave speeds have been made. These measurements provide a large data base for analysis to understand the physics of this sandy bottom, which consists of a 3.0-m-deep bed of a manufactured 212-μm mean diameter sand. Analysis of these measurements using Biot models and other fluid models is presented. In addition, analysis of enhanced sound penetration at shallow grazing angles due to a roughened interface is discussed. [Work supported by ONR.]

4:12

5pUW14. The low-frequency radiation and scattering of sound from bubble distributions near the sea surface. William M. Carey (Dept. of Ocean Eng., MIT, Cambridge, MA 02139) and Ronald A. Roy (Boston Univ., Boston, MA 02215)

Microbubble plumes are produced when waves break and are convected to depth by the vortical wave motion. The question is, What role is played by these distributions in the production and scattering of sound near the sea surface at low frequency (LF)? Deep- and shallow-water ambient noise show dramatic increases with wave breaking. Individual wave breaking events show radiation down to 30 Hz. Tipping water-filled trough and bucket experiments show LF radiation associated with the entrainment of known bubble size distributions; suggesting that single bubbles cannot be the cause of this radiation. Microbubble distributions with void fractions greater than 0.01% act as collective monopole resonant oscillators excited by the energy of formation. Multipole expansions show the radiation is described by a modified Minnaert expression and a dipole radiation pattern (owing to the presence of the sea surface). Since the plume stays put, scattered sound would have a zero Doppler shift with ample Doppler spread. Acoustic images of the sea surface show discrete scattering events, while experiments show LF scattering from free bubble distributions. Results are shown which indicate that bubble plumes from breaking waves not only radiate but scatter LF sound. [Work sponsored by ONR.]

4:24

5pUW15. Inversion of bistatic reverberation and scattering measurements for seabottom properties. Andrew Rogers, Greg Muncill, and Peter Neumann (Planning Systems, Inc., 7923 Jones Branch Dr., McLean, VA 22102, arogers@plansys.com)

Considerations associated with the inversion of bistatic reverberation and scattering data from a three-dimensional space are discussed. The SCARAB (SCAttering and ReverberAtion from the Bottom) acoustic propagation and scattering model has been the kernel for the study. This model considers three-dimensional, non-plane-wave, range-independent acoustic signals in both mono- and bistatic source receiver configurations. Due to the computationally expensive problem of computing a reverberative field which includes many hybrid acoustic paths and the iterative nature of the inverse process, inversion for seabottom properties requires highly effective search parameter space reduction and algorithm optimization. The parameter space reduction methods and simulated annealing and genetic optimization algorithms utilized are discussed, and their effectiveness measured against several test cases. [Work supported by ONR.]

4:36

5pUW16. High-frequency sound propagation in shallow water with rough surface scattering. X. Tang, M. Badiey (Marine Studies, Univ. of Delaware, Newark, DE 19716), and J. Simmen (Office of Naval Res., ONR 3210A, Arlington, VA 22217)

Sound propagation at mid to high frequencies (1–20 kHz) in shallow-water environments is extremely complicated, variable, and poorly understood due to multiple processes that cover a wide range of temporal and spatial scales. Numerical simulations using a full-wave PE model are carried out to interpret the acoustic data taken during very carefully designed experiments at a shallow-water site inside the Delaware Bay. Records of broadband acoustic wave time series (from hours to days) are numerically reproduced by using the collected environmental data as input to the PE model. The statistics of both measured and simulated multipath signals are calculated and compared in the presence of rough sea-surface scattering and surface bubble layer attenuation. The surface wind effects can be separated from the slowly varying volume fluctuations of sound speed, by comparing the temporal variations; while it can be separated from the bottom effects by examining different multipath arrivals. The dependence of coherence on the parameters of wind strength and acoustic frequency is examined.

4:48

5pUW17. Recent observations of high-frequency acoustic wave propagation in the Delaware Bay. M. Badiey (Marine Studies, Univ. of Delaware, Newark, DE 19716), S. Forsythe (Naval Undersea Warefare Ctr., Newport, RI 02841), J. Simmen (Office of Naval Res., ONR 3210A, Arlington, VA 22217), and X. Tang (Univ. of Delaware, Newark, DE 19716)

Broadband acoustic propagation measurements, as well as complementary environmental measurements including CTD, ADCP, and wind speed, have been conducted in the Delaware Bay during March and September, 1997. With the data from these complete experiments, acoustic pulse travel time and intensity are related to the sound-speed fluctuations in the water column, the sea-surface profile, and the bottom roughness. The direct arrival between acoustic source and receiver contains information about ocean current and temperature, while the arrival corresponding to a single surface bounce contains additional information about the sea-surface roughness. Independent measurements of temperature and current allow the sorting of acoustic fluctuations due to ocean volume fluctuations from those due to sea-surface fluctuations. Coherence of the broadband acoustic signals is calculated from the observed data and is found to be correlated with the environmental parameters.