

THE JOURNAL

of the

Acoustical Society of America

Vol. 101, No. 5 , Pt. 1

May 1997

ACOUSTICAL NEWS—USA		2401
USA Meetings Calendar		2402
OBITUARIES		2406
ACOUSTICAL STANDARDS NEWS		2408
Standards Meetings Calendar		2408
American National Standards on Acoustics Catalog 16-1997		2412
FORUM		
Room noise criteria—The S12.2-1995 standard raises serious technical and economic considerations	Hsien-sheng (Jason) Pei	2418
REVIEWS OF ACOUSTICAL PATENTS		2421
TUTORIAL REVIEW [10]		
Some hot topics in animal bioacoustics	Whitlow W. L. Au	2433
GENERAL LINEAR ACOUSTICS [20]		
Radiation force on a spherical object in the field of a focused cylindrical transducer	Xucaai Chen, Robert E. Apfel	2443
Double-sum technique for performing a Fourier transformation	G. Maidanik, K. J. Becker	2448
Computation of the modal response of regularly ribbed cylinders	G. Maidanik, K. J. Becker	2452
Transmission and reflection of transient elastodynamic waves at a linear slip interface	Martin D. Verweij, Christopher H. Chapman	2473
Solving ray acoustic problems with a DNA computer	Yang Yang, Yuling Li, Luise S. Couchman	2485
New formulation of the resonance scattering theory	Myoung-Seon Choi	2491
Circumferential creeping waves around a fluid-filled cylindrical cavity in an elastic medium	Waled Hassan, Peter B. Nagy	2496
A study of wave propagation in varying cross-section waveguides by modal decomposition. Part II. Results	N. Amir, V. Pagneux, J. Kergomard	2504
Rational operators for filtering	R. J. Cederberg, Michael D. Collins, Henrik Schmidt, William L. Siegmann	2518
Prediction of the vibro-acoustic transmission loss of planar hose-pipe systems	M. L. Munjal, P. T. Thawani	2524

(Continued)

CONTENTS—Continued from preceding page

NONLINEAR ACOUSTICS, MACROSONICS [25]

- | | | |
|---|--|------|
| On the onset of transient cavitation in gassy liquids | L. Gaete-Garretón, Y. Vargas-Hernández, R. Vargas-Herrera, J. A. Gallego-Juárez, F. Montoya-Vitini | 2536 |
|---|--|------|

UNDERWATER SOUND [30]

- | | | |
|--|---|------|
| Modeling the environmental influence on the vertical directionality of ambient noise in shallow water | T. C. Yang, Kwang Yoo | 2541 |
| Modeling low-frequency reverberation near the Mid-Atlantic Ridge and comparison with ARSRP data | Jerald W. Caruthers, E. J. Yoerger, J. C. Novarini | 2555 |
| Coupled-amplitude approach to solving the Helmholtz equation | Dalcio K. Dacol | 2566 |
| Low-frequency acoustic scattering by a cube: Experimental measurements and theoretical predictions | Paul A. Chinnery, Victor F. Humphrey, Jingdong Zhang | 2571 |
| Simultaneous acoustic and microwave backscattering from the sea surface | Peter H. Dahl, William J. Plant, Bernd Nützel, Anke Schmidt, Heinz Herwig, Eugene A. Terray | 2583 |
| The variability of high-frequency acoustic backscatter from the region near the sea surface | Peter H. Dahl, William J. Plant | 2596 |
| Acoustic measurements of suspended sediments in turbulent currents and comparison with <i>in-situ</i> samples | Peter D. Thorne, Peter J. Hardcastle | 2603 |
| An investigation of the small slope approximation for scattering from rough surfaces. Part II. Numerical studies | Shira Lynn Broschat, Eric I. Thorsos | 2615 |
| The detection of tethered and rising bubbles using multiple acoustic techniques | Timothy G. Leighton, David G. Ramble, Andy D. Phelps | 2626 |
| Numerical simulations with horizontal-refraction-modal tomography. Part I. Adiabatic propagation | A. G. Voronovich, E. C. Shang | 2636 |
| Analytical expressions for high-frequency acoustic fluctuations from suspended particles in the ocean | R. A. Thuraisingham | 2644 |

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

- | | | |
|---|--------------------------------|------|
| Excess attenuation of leaky Lamb waves due to viscous fluid loading | Adnan H. Nayfeh, Peter B. Nagy | 2649 |
|---|--------------------------------|------|

STRUCTURAL ACOUSTICS AND VIBRATION [40]

- | | | |
|---|--------------------------|------|
| Acoustic scattering of a plane wave by two spherical elastic shells above the coincidence frequency | H. Huang, G. C. Gaunaurd | 2659 |
|---|--------------------------|------|

NOISE: ITS EFFECTS AND CONTROL [50]

- | | | |
|--|--|------|
| Noise reduction by a barrier having a random edge profile | Steve S. T. Ho, Ilene J. Busch-Vishniac, David T. Blackstock | 2669 |
| Annoyance caused by sounds of wheeled and tracked vehicles | Niek J. Versfeld, Joos Vos | 2677 |

ARCHITECTURAL ACOUSTICS [55]

- | | | |
|--|------------------|------|
| A parametric error analysis of the backward integration method for reverberation time estimation | Dennis R. Morgan | 2686 |
|--|------------------|------|

ACOUSTICAL MEASUREMENTS AND INSTRUMENTATION [58]

- | | | |
|---|--------------------------------------|------|
| Comparison of methods for determining specific acoustic impedance | Michael G. Jones, Patricia E. Stiede | 2694 |
|---|--------------------------------------|------|

PHYSIOLOGICAL ACOUSTICS [64]

- | | | |
|--|----------------------------------|------|
| Spectral cues for sound localization in cats: A model for discharge rate representations in the auditory nerve | Bradford J. May, Aileen Y. Huang | 2705 |
|--|----------------------------------|------|

CONTENTS—Continued from preceding page

Neural contributions to the perstimulus compound action potential: Implications for measuring the growth of the auditory nerve spike count as a function of stimulus intensity	John R. Doucet, Evan M. Relkin	2720
Is loudness simply proportional to the auditory nerve spike count?	Evan M. Relkin, John R. Doucet	2735
Efferent projections of a physiologically characterized region of the inferior colliculus of the young adult CBA mouse	Robert D. Frisina, Joseph P. Walton, Martha A. Lynch-Armour, Darrel A. Klotz	2741
Sound-pressure measurements in the cochlear vestibule of human-cadaver ears	Sunil Puria, William T. Peake, John J. Rosowski	2754
Click-evoked otoacoustic emissions and the influence of high-frequency hearing losses in humans	Paul Avan, Michel Elbez, Pierre Bonfils	2771
Relations between notched-noise suppressed TEOAE and the psychoacoustical critical bandwidth	Joachim Neumann, Stefan Uppenkamp, Birger Kollmeier	2778
PSYCHOLOGICAL ACOUSTICS [66]		
A psychoacoustic model for the noise masking of plosive bursts	James J. Hant, Brian P. Strobe, Abeer A. Alwan	2789
Spectral weights in level discrimination by preschool children: Synthetic listening conditions	Melodie S. Willihnganz, Mark A. Stellmack, Robert A. Lutfi, Frederic L. Wightman	2803
Spectral weights in level discrimination by preschool children: Analytic listening conditions	Mark A. Stellmack, Melodie S. Willihnganz, Frederic L. Wightman, Robert A. Lutfi	2811
Formant transition duration and speech recognition in normal and hearing-impaired listeners	Christopher W. Turner, Sarah J. Smith, Patricia L. Aldridge, Suzanne L. Stewart	2822
SPEECH PRODUCTION [70]		
An investigation of stop place of articulation as a function of syllable position: A locus equation perspective	Harvey M. Sussman, Nicola Bessell, Eileen Dalston, Tivoli Majors	2826
SPEECH PERCEPTION [71]		
Concurrent vowel identification. I. Effects of relative amplitude and F_0 difference	Alain de Cheveigné, Hideki Kawahara, Minoru Tsuzaki, Kiyoaki Aikawa	2839
Concurrent vowel identification. II. Effects of phase, harmonicity, and task	Alain de Cheveigné, Stephen McAdams, Cécile M. H. Marin	2848
Concurrent vowel identification. III. A neural model of harmonic interference cancellation	Alain de Cheveigné	2857
Intraspeech spread of masking in normal-hearing and hearing-impaired listeners	Van Summers, Marjorie R. Leek	2866
SPEECH PROCESSING AND COMMUNICATION SYSTEMS [72]		
A targeting-and-extracting technique to enhance hearing in the presence of competing speech	Chen Liu, Judith Rosenhouse, Samuel Sideman	2877
A study on robust utterance verification for connected digits recognition	Mazin G. Rahim, Chin-Hui Lee, Biing-Hwang Juang	2892
MUSIC AND MUSICAL INSTRUMENTS [75]		
Acceptance limits for the duration of pre-Helmholtz transients in bowed string attacks	Knut Guettler, Anders Askenfelt	2903
Sound production in recorderlike instruments. I. Dimensionless amplitude of the internal acoustic field	Marc-Pierre Verge, Benoit Fabre, A. Hirschberg, A. P. J. Wijnands	2914
Sound production in recorderlike instruments. II. A simulation model	M. P. Verge, A. Hirschberg, R. Caussé	2925

CONTENTS—Continued from preceding page

BIOACOUSTICS [80]

- | | | |
|---|--|------|
| Inertial cavitation and associated acoustic emission produced during electrohydraulic shock wave lithotripsy | Pei Zhong, Iulian Cioanta, Franklin H. Cocks, Glenn M. Preminger | 2940 |
| The acoustic features of vowel-like <i>grunt</i> calls in chacma baboons (<i>Papio cyncephalus ursinus</i>): Implications for production processes and functions | Michael J. Owren, Robert M. Seyfarth, Dorothy L. Cheney | 2951 |
| A backpropagation network model of the monaural localization information available in the bat echolocation system | Janine M. Wotton, Rick L. Jenison | 2964 |
| Acoustic effects of the ATOC signal (75 Hz, 195 dB) on dolphins and whales | Whitlow W. L. Au, Paul E. Nachtigall, Jeffrey L. Pawloski | 2973 |

LETTERS TO THE EDITOR

- | | | |
|---|---|------|
| Comments on “Flexural wave reduction using a compliant tube baffle” [J. Acoust. Soc. Am. 99(2), 691–699 (1996)] and “Signal pressure received by a hydrophone placed on a plate backed by a compliant baffle” [J. Acoust. Soc. Am. 89(2), 559–564 (1991)] [20] | Ronald P. Radlinski, Sung H. Ko | 2978 |
| Scattering of acoustic waves by intermittent temperature and velocity fluctuations [28] | D. Keith Wilson | 2980 |
| Acoustic scattering by a pair of spheres: Addenda and corrigenda [30] | Guillermo C. Gaunaurd, Hanson Huang, Hans C. Strifors | 2983 |
| Antisymmetric nature of the earliest arrivals of ultrasound propagating in copy paper [35] | Mont A. Johnson, Yves H. Berthelot | 2986 |
| Prediction of sound-pressure level in an occupied enclosure [55] | S. K. Tang, Daniel W. T. Chan, K. C. Chan | 2990 |
| The relationship between localization and the Franssen effect [66] | William A. Yost, Dan Mapes-Riordan, Sandra J. Guzman | 2994 |

ERRATA

- | | | |
|---|---------------|------|
| Erratum: “On the spatial coherence and angular spreading of sound forward scattered from the sea surface: Measurements and interpretive model” [J. Acoust. Soc. Am. 100, 748–758 (1996)] | Peter H. Dahl | 2998 |
|---|---------------|------|

CUMULATIVE AUTHOR INDEX

2999

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 articles can be ordered for \$15 per copy from the AIP/Member Society document delivery service “Articles in Physics,” 75 Varick Street, New York, NY 10013; Fax: 212-301-4060; Telephone: 800-480-PHYS (800-480-7497) (in U.S. and Canada), or 212-301-4000; E-mail: articles@aip.org; URL: <http://www.aip.org/articles.html>

CONTENTS

	page
Technical Program Summary	A8
Schedule of Technical Session Starting Times	A9
Map of Centre County, PA	A10
Hotel Location Map, State College	A12
Map of Meeting Rooms	A14
Calendar–Technical Program	A16
Calendar–Other Events	A22
Meeting Information	A24
Guidelines for Presentations	A36
Dates of Future Meetings	A38
25- and 50-Year Members	A40
NOISE-CON 97 Program Summary	3003
Technical Sessions (1a_), Monday Morning	3015
Technical Sessions (1p_), Monday Afternoon	3030
Tutorial Session (1eID), Monday Evening	3048
Technical Sessions (2a_), Tuesday Morning	3049
Technical Sessions (2p_), Tuesday Afternoon	3067
Technical Sessions (3a_), Wednesday Morning	3092
Technical Sessions (3p_), Wednesday Afternoon	3118
Plenary Session, Business Meeting and Awards Ceremony, Wednesday Afternoon	3133
R. Bruce Lindsay Award Encomium	3133
Helmholtz-Rayleigh Interdisciplinary Silver Medal Award Encomium	3133
Gold Medal Award Encomium	3133
Technical Sessions (4a_), Thursday Morning	3134
Technical Sessions (4p_), Thursday Afternoon	3162
Technical Sessions (5a_), Friday Morning	3185
Technical Sessions (5p_), Friday Afternoon	3199
Author Index to Abstracts	3204
Application Forms	3210
Regional Chapters	3213
Sustaining Members	3216
Index to Advertisers	3220

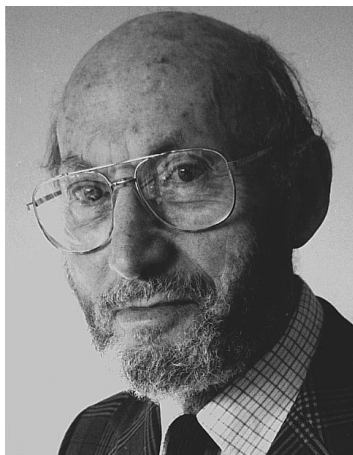
ACOUSTICAL NEWS—USA

Elaine Moran

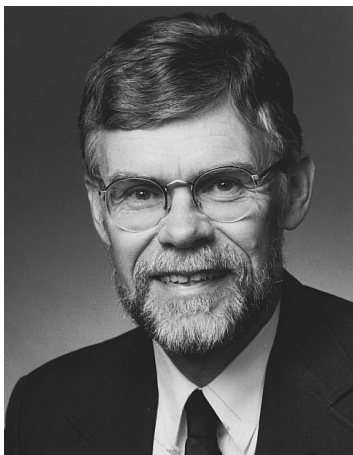
Acoustical Society of America, 500 Sunnyside Boulevard, Woodbury, New York 11797

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

New Fellows of the Acoustical Society of America



Albert Freedman—For contributions to the understanding of echo fields and Lamb modes.



Alfred L. Nuttall—For contributions to understanding the physiology of the cochlea.



Sam H. Ridgway—For contributions to marine mammal acoustics.

Brenda L. Lonsbury-Martin becomes New Associate Editor of the *Journal*



In January 1997 Professor Brenda Lonsbury-Martin of the University of Miami Ear Institute Department of Otolaryngology, became a new Associate Editor of the *Journal* for papers in Physiological Acoustics (PACS 43.64), sharing this responsibility with Professor Robert D. Frisina of the University of Rochester Medical Center. Gradual expansion of physiological acoustics within the *Journal* publication program prompted appointment of an additional associate editor.

Brenda Lonsbury was born in Vancouver, BC, Canada, and received

her B.A. from the University of Victoria in 1967. She received an M.S. degree in 1971 in Neuroscience/Cell Biology from University of Oregon Medical School, and a Ph.D. in Neuroscience/Biochemistry/Physiology in 1975 from University of Oregon Health Sciences Center.

From 1975 to 1978 she was an NIH Postdoctoral Fellow, first at University of California (Irvine) then at University of Washington Medical School where she became an Assistant Professor. From 1984 to 1991 she was at Baylor University Department of Otorhinolaryngology, becoming an Associate Professor in 1987. Since 1991 she has been Professor and Director of Research at University of Miami School of Medicine's Department of Otolaryngology. She is married to Glen K. Martin, Ph.D., Co-Director of Research. Throughout her academic career she has been a lecturer, while tutoring resident physicians, and both pre- and postdoctoral students.

Lonsbury-Martin's principal research interest has been in Cochlear

Otoacoustic Emissions, the subject of an invited review paper which she co-authored for this *Journal* in 1991. She has authored or co-authored over 65 research publications, and a book "An Introduction to the Practical Applications of Otoacoustic Emissions," appearing in 1997.

The Editor-in-Chief and his colleagues welcome Professor Lonsbury-Martin to membership on the Editorial Board.

DANIEL W. MARTIN
Editor-in-Chief

Call for Nominations

The American Institute of Physics (AIP) Corporate Associates program announces the 1997–1998 competition for its *Prize for Industrial Applications of Physics*. The prize is intended to recognize outstanding contributions to industrial applications of physics.

The Prize. The winner of the prize will receive \$10,000, a certificate of recognition, and a travel allowance to receive the prize from AIP. The prize is sponsored by General Motors Corporation and the AIP Corporate Associates program.

Eligibility. Nominations are open to U.S. residents and individuals who have worked primarily in U.S. industry. The contribution for which individuals are nominated must have led to an advance in industrial physics applications no more than ten years ago. Examples of contributions to applied industrial physics include products, processes, or tools providing new practical applications of physics. Prizes are typically awarded to individuals, but nominations for collaborative work are considered.

Nominations. Nominations should include a two-page statement containing the nominee's vita, a description of the nominee's relevant work, and lists of the nominee's publications and patents.

Nomination Deadline:
31 May 1997

Send nominations and supporting documentation to:

Office of the Director, American Institute of Physics, One Physics Ellipse,
 College Park, MD 20740. Attn: Committee for the IAP Prize. For more
 information, contact: assoc@aip.org

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.

1997

- 13–16 April 23rd International Symposium on Acoustical Imaging,
 Boston, MA [Sidney Lees, Bioengineering Dept., For-
 syth Dental Ctr., 140 Fenway, Boston, MA 02115;
 FAX: 617-262-4021; E-mail: slees@forsyth.org]. 7/96
- 9–10 May Fifth Annual Robert Monzon Memorial Conference,
 Washington, DC. Topic: Wireless Telephones and
 Hearing Aids: New Challenges for Audiology [Univer-
 sity Conference Management, Gallaudet Univ., 800
 Florida Ave., N.E., Washington, DC 20002-3695, Tel.:
 202-651-6060 or 6053; FAX: 202-651-6074 or 6038;
 E-mail: conference@gallua.gallaudet.edu].
- 12–14 May Third AIAA/CEAS Aeroacoustic Conference, Atlanta,
 GA [Dr. Stephen Engelstad, Lockheed Martin Aeronau-
 tical Systems, D/73-47, Z/O-685, Marietta, GA 30063,
 Tel.: 770-494-9178; FAX: 770-494-3055; E-mail:
sengelstad@fs2.mar.1mco.com].
- 15–20 June Eighth International Symposium on Nondestructive
 Characterization of Materials, Boulder, CO
 [Debbie Harris, The Johns Hopkins University,
 Ctr. for Nondestructive Evaluation, 102 Maryland Hall,
 3400 N. Charles St., Baltimore, MD 21218,
 Tel.: 410-516-5397; FAX: 410-516-7249, E-mail:
cnde@jhuvms.hcf.jhu.edu].
- 15–17 June NOISE-CON 97, State College, PA [Institute of Noise
 Control Engineering, P.O. Box 320, Arlington Branch,
 Poughkeepsie, NY 12603, Tel.: 914-891-1407; FAX:
 914-463-0201].
- 16–20 June 133rd meeting of the Acoustical Society of America,
 State College, PA [ASA, 500 Sunnyside Blvd., Wood-

9–13 July

7–11 Sept.

21–27 Aug.

22–24 Sept.

1–5 Dec.

22–26 June

13–17 Sept.

12–16 Oct.

bury, NY 11797, Tel.: 516-576-2360; FAX: 516-576-
 2377; E-mail: asa@aip.org, WWW: <http://asa.aip.org>].
3/97

International Clarinet Association, Texas Tech Univ.,
 Lubbock, TX [Keith Koons, Music Department, Univ.
 of Central Florida, P.O. Box 161354, Orlando, FL
 23816-1354, Tel.: 407-823-5116; E-mail:
kkoons@pegasus.cc.ucf.edu].

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].

1997 Conference on Implantable Auditory Prostheses,
 Pacific Grove, CA [Alena Wilson, Conference Coordi-
 nator, House Ear Inst., 2100 W. 3rd St., Los Angeles,
 CA 90057; Tel.: 213-353-7086; FAX: 213-413-0950;
 E-mail: alena@hei.org; WWW: <http://www.rti.org/ciap97>].

Second Biennial Hearing Aid Research and Develop-
 ment Conference, Bethesda, MD [National Institute
 of Deafness and Other Communication Disorders,
 Tel.: 301-970-3844; FAX: 301-907-9666; E-mail:
hearingaid@tascon.com].

134th meeting of the Acoustical Society of America,
 San Diego, CA [ASA, 500 Sunnyside Blvd., Wood-
 bury, NY 11797, Tel.: 516-576-2360; FAX: 516-576-
 2377; E-mail: asa@aip.org, WWW: <http://asa.aip.org>].

1998

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>].

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].

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>].

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.

W. Dixon Ward • 1924–1996



W. Dixon Ward, one of the most active and distinguished Fellows of the Society, died of heart failure at his home in St. Paul, Minnesota on 19 December 1996. He was a past President of the Society and the 1991 recipient of the Silver Medal in Psychological and Physiological Acoustics, Musical Acoustics, and Noise. His substantial service to the Society also included: Executive Council (1979–1981), Associate Editor for Psychological Acoustics (1984–1986, 1992–1993), Vice President (1986–1987). His Silver Medal in 1991 was for “...furthering knowl-

edge of auditory perception in psychological and musical acoustics and increasing the understanding of the etiology of noise-induced hearing loss.”

Dix was raised in South Dakota and received his B.S. in Physics from the South Dakota School of Mines and Technology in 1944. (He was awarded an Honorary Doctorate from that institution in 1971.) After Navy service and graduate work at the University of Minnesota, Dix followed the example of Walter Rosenblith, his former college teacher, and went to Harvard University to study psychological acoustics. There he was part of the legendary Psychoacoustics Laboratory and received his Ph.D. in 1953 under the mentorship of S. S. Stevens.

At Harvard he collaborated with Ira Hirsh on their classic work on auditory fatigue. This was the first of Dix's many contributions to knowledge about the aftereffects of sound exposure on hearing. His Ph.D. dissertation on the subjective octave began his many contributions to the understanding of pitch and the perception of music.

In 1949 Dix married Edith Bystrom (Bunny), who has attended many meetings of ASA with Dix, and has many lasting friends in the Society. They raised four daughters: Marnie, Laurie, Chris, and Holly, to whom Dix was a loving father.

After his Ph.D. Dix continued studying music perception as a Research Engineer at Baldwin Piano Company. He then became a Research Associate in a Navy-sponsored project conducted by the Central Institute for the Deaf on the auditory and nonauditory effects of high intensity noise. In 1957 Dix joined Aram Glorig at the Research Center of the Subcommittee on Noise in Los Angeles, where he did an extensive series of experiments on temporary threshold shift. These experiments provided the foundation for much of our knowledge about the phenomenon of auditory fatigue.

Dix returned to the University of Minnesota in 1962 and held appointments in the departments of Otolaryngology, Communication Disorders, Psychology, and Environmental Health. During 31 years at Minnesota, Dix continued research on the effects of noise on hearing, and on musical perception. He supervised many graduate students, several of whom continued his legacy of distinguished contributions to hearing research. Dix was also a dedicated teacher. In November he expressed frustration that health problems made him miss many classes in a course that he had volunteered to teach. He was grading exams for that course just a few days before his death.

Dix was a prolific contributor to many areas of acoustics and achieved international recognition. His experimental contributions on auditory fatigue and on pitch were truly fundamental. The former also had great practical importance because they provided much of the scientific basis for the development of criteria for assessing the risk of hearing loss produced by noise exposure, and for the implementation of hearing conservation programs. Dix's many review chapters were models of scholarship; they were lively, provocative yet balanced and fair, and they were beautifully written.

His service to other professional and governmental organizations was also remarkable. Dix was president of two audiological societies and was active on various committees of these organizations. He also served in an editorial capacity for several other journals. Dix was a member and Chairman of CHABA, and was a major contributor to the development of the well-known damage-risk contours published by that organization.

Dix had many personal interests including hunting, fishing, bridge, singing, and playing the piano. He approached these activities with a verve and joy that was contagious. Dix was also a great and colorful character. He was never one to abide foolishness, and could be a curmudgeon when he disagreed with something or someone. He was famous for provocative questions at meetings, and for lively letters to local newspapers and to the *Journal*. There was never malice or ill-will in these exchanges—he always had a memorable smile, even in writing, and was simply being Dix, using his great intellectual energy and integrity to challenge constructively. Dix had many dear friends in the Society and we shall miss him.

There will be a memorial session for Dix at the State College ASA meeting.

NEAL F. VIEMEISTER
WILLIAM MELNICK

Manfred Heckl • 1930–1996

Manfred Heckl, a Fellow of the Society, who served as the director of the Institute for Technical Acoustics at the Technical University of Berlin for over two decades, died on 16 August 1996 after an operation for a brain tumor. He is widely known for his work on structure-borne sound; and particularly for the book, originally co-authored with Lothar Cremer, which appeared in English under the title of “Structure-Borne Sound.” It had been published in German under the title of “Körperschall,” and is considered worldwide as the classic and most authoritative publication on this topic. His “Taschenbuch der Technischen Akustik,” co-authored by H. A. Müller, is the most widely used noise control handbook in German speaking countries.

Manfred Heckl was born on 15 July 1930 in the small Bavarian town of Rennertshofen and der Donau. He studied physics at the University of Munich, and received his doctorate at the Technical University of Berlin in 1959 under Lothar Cremer. He then worked at Bolt Beranek and Newman in Cambridge, MA, where he led and made major contributions to several research and consulting projects until he returned to Munich in 1962. There he, Lothar Cremer, Leo Beranek, Helmut Müller, and others founded Müller-BBM, which has since grown to be Europe's foremost acoustical consulting company. After 11 years at that organization he was invited to succeed Lothar Cremer at the Technical University of Berlin. Until entering

emeritus status in October 1995 he served there as Professor of Technical Acoustics and Managing Director of the Institute for Technical Acoustics. He also served for various intervals as dean of the department of environmental technology.

He was loved by his students, particularly by the 90 doctoral candidates he supervised, and was highly respected by his colleagues and peers. He was known for his insightful lectures and publications, which reflected his uncanny ability to present the substantive core of his subject matter in its simplest form. Although he was an outstanding theoretician, he never pursued theory for its own sake; his view was that the primary purpose of theory is to provide understanding of physical phenomena and to enable their quantitative description. His theoretical understanding, coupled with his experimental skills and his knowledge of practical engineering, earned him his reputation as a world-class consultant.

Although his academic, administrative, and consulting activities demanded much of his time, he was sole author or major contributor to more than one hundred technical papers and a like number of reports, and he co-authored five books and monographs. Widely sought as a lecturer, he presented invited lectures, seminars, and short courses at various locations in Great Britain, Russia, and the United States. He spent two summers as visiting scientist at the Indian Institute of Technology at Madras, and one at the Indian Institute of Science in Bangalore. He also served on the editorial boards of the *Journal of Sound and Vibration*, the *Zeitschrift für Lärmbekämpfung*, and the *Journal of Technical Acoustics* (St. Petersburg), and he was Associate Editor for structural acoustics of *Acta Acustica*. He was a Corresponding Member of the Academy of Science, Cordoba, Argentina, and the 1992 recipient of the coveted Rayleigh Medal of the Institute of Acoustics (United Kingdom).

Manfred Heckl is survived by his wife, Anna, five children, and six grandchildren. His oldest daughter, Maria, has followed in her father's technical footsteps: she studied physics and, after receiving her doctorate at Cambridge (UK), is working in acoustics at a British University.

With Manfred Heckl's passing the world has lost an outstanding acoustician and a warm and caring human being, who was cherished and respected by his family, friends, and colleagues, and who was liked and admired by all who knew him.

ERIC E. UNGAR
ISTVÁN L. VÉR

Polish Acoustician Halina Ryffert • 1916–1996



Professor Halina Ryffert, a past President of the Polish Acoustical Society and retired Professor of Acoustics at the Adam Mickiewicz University, passed away on 7 August 1996. She was born in Bugulma (Russia) to a Polish family. Three years later they moved to Poznan (Poland) where she lived and worked almost all her life. Shortly after she had graduated in mathematics from Adam Mickiewicz University, World War II broke out. As a result of her conspiratorial activity against the invader, Halina Ryffert was arrested and then sent to a concentration camp, "Ravensbrück," where she stayed until the end of the war. She returned to Poland soon after the liberation in 1945, and became an assistant in the Institute of Mathematics at the Adam Mickiewicz University. In 1952 she was offered a position in Acoustics, and received the Ph.D. degree in 1958. Her dissertation was entitled "Acquisition of instantaneous spectra on the basis of a generalized analysis of vibrations." Further work on methods for spectral analysis of nonstationary acoustic signals, conducted both in Poland and abroad at the College de France in Paris (1958) and at the Technical University of Stuttgart, Germany (1960) resulted in many publications. In 1968 Halina Ryffert became a Professor at Adam Mickiewicz University.

She headed the Chair of Acoustics (present Institute of Acoustics) for 19 years (1962–1981). For 7 years she was Dean of the College of Mathematics, Physics and Chemistry, and was a member of numerous academic boards. In 1965–1984 Professor Ryffert was a President of the Polish Acoustical Society, bringing together people interested in diverse fields of acoustics. During that time her efforts to set up a collaboration with acoustical centers abroad resulted in a fruitful collaborative work with GALF (French Acoustical Society). After the establishment of an Acoustics Committee of the Polish Academy of Sciences in 1964, Professor Ryffert was its Vice-President for 20 years.

She was a dedicated teacher. Under her supervision, many masters and doctoral theses were completed. As a reward for her scientific and pedagogic work, she received high honors both in Poland and abroad. Her contributions to the development of Polish Acoustics were invaluable. She turned the Chair of Acoustics at the University into a thriving acoustical center—the present Institute of Acoustics in Poznan. Always helpful and understanding, she was a respected scientist and a great woman. She will remain our mentor. Her passing is an irreparable loss.

EDWARD HOJAN

FORUM

Forum is intended for communications that raise acoustical concerns, express acoustical viewpoints, or stimulate acoustical research and applications without necessarily including new findings. Publication will occur on a selective basis when such communications have particular relevance, importance, or interest to the acoustical community or the Society. Submit such items to an appropriate associate editor or to the Editor-in-Chief, labeled FORUM. Condensation or other editorial changes may be requested of the author.

Opinions expressed are those of the individual authors and are not necessarily endorsed by the Acoustical Society of America.

Room noise criteria—The S12.2-1995 standard raises serious technical and economic considerations [43.15.+s, 43.50.Jh]

Hsien-sheng (Jason) Pei

P.O. Box 5327, Cochituate, Massachusetts 01778

(Received 8 November 1996; approved for publication 17 February 1997)

The key issue with the S12.2-1995 standard is not the differences in the assessment of sound quality, particularly at the low frequencies, but rather the use of octave-band data to describe low-frequency equipment misbehavior. Unfortunately, the “NCB vs RC” controversy leads research efforts and technical discussions in the wrong directions, disregarding serious economic and technical considerations. This conflict can be resolved by: (1) avoiding the use of octave-band data to rate noise with abnormal (spectral and/or temporal) characteristics, (2) addressing the widespread need for the room noise criteria as a practical design tool for general use in architectural spaces, and (3) developing adequate descriptor/procedures to qualify, quantify, and determine the impact of “rumble” on human response. A further goal should be to develop integrated criteria that would apply to both HVAC and computer noise sources. © 1997 Acoustical Society of America. [S0001-4966(97)06905-1]

INTRODUCTION

After more than 20 years of effort, American National Standard “Criteria for Evaluating Room Noise” ANSI S12.2-1995 was released.¹ It is not a good standard² since it offers two different possibilities to evaluate the acceptability of the noise in an occupied space (the NCB and RC methods), with large potential differences in the cost of conforming to these different noise criteria, and insufficient guidance to a user on how to choose between them.

As the ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers) liaison to the ASA S12 committee since 1981, the author is fully aware of the proprietary underlying issues, and sees the need for inter-society partnership to resolve the conflict. Accordingly the author organized two panel discussion sessions (ASA Boston Chapter meeting on 21 November 1995; NOISE-CON 96 Conference on 30 September 1996), and one technical session (NOISE-CON 96 “Room Noise Criteria—the next 25 Years”).²⁻⁷ The inputs from the acoustical and engineering communities have served to show that there are enough data to develop a roadmap for consensus.

After requests from attendants at the NOISE-CON 96 sessions, the author prepared a report entitled “Room Noise Criteria—The Next 25 Years” which summarized the findings at these sessions. That report and this article may facilitate more productive discussions at the forthcoming ASA/

NOISE-CON 97 meetings scheduled for Penn State in June 1997. There will be one panel discussion session and two technical sessions: (Part 1) Room Noise Criteria—Blind Common Sense? and (Part 2) Room Noise Criteria—Status Report & Integration.

The ASA/NOISE-CON 97 room noise criteria sessions are co-sponsored by two ASA technical committees and four INCE technical committees. All six committees are very supportive of an inter-society effort at this joint meeting. ASHRAE and other relevant technical societies have been invited to be co-sponsors and contributors.

I. KEY ISSUES FOR ROOM NOISE CRITERIA SESSIONS

A. S12.1-1995 standard raises serious economic considerations

In addition to HVAC equipment, other types (e.g., computers) have emerged as important noise sources in office spaces. It is not possible to optimize “room noise” without the integration of computer cooling and building environmental control systems. There is growing evidence⁸ that the future evolution of computers will require increasing emphasis on cooling systems.⁹ The need for integrated noise solutions for buildings and computers is urgent.

Tables are given in the standard that list preferable criterion levels in different spaces. For rooms that demand better acoustical environments, the RC-25 and NCB-25 curves should be chosen. Examples are private residences, executive offices, conference rooms, hospital private rooms, operating rooms, school lecture and classrooms, and legitimate theaters. The RC-25 and NCB-25 curves are plotted in Fig. 4 of Ref. 3. The major differences between the RC and NCB curves are in the 31.5- and 63-Hz bands. Also plotted in that figure are measured background noise spectra in eight unoccupied concert halls, including the famous Philadelphia Academy of Music and the Rotterdam De Doelen halls. In those halls, the levels in the 31.5-Hz band lie 10 to 15 dB higher than the RC-25 curves, but are nearly in conformance with the NCB-15 to NCB-25 curves. These halls are judged to be acoustically superior, and there is no reason why noise levels in the spaces named above should be restricted to grossly lower levels (as required by the RC curves) than those measured in the great concert halls of the world.

The NCB curves have been validated as predictors of unacceptable noise conditions where complaints were severe

enough for owners to call for acoustical advice. The information was furnished by 14 consulting firms.^{3,10}

The principal argument by proponents of the RC method is that it guards against low-frequency rumble in the 16- and 31.5-Hz bands. Octave-band data are not suitable as descriptors for low-frequency misbehavior. It is true that if the RC method is employed, the abnormal cases of fluctuating levels in the 16- and 31.5-Hz bands will be largely handled, but in normal cases the resulting overdesign leads to needlessly costly acoustical treatments for venues with noise levels in the RC-25 to RC-40 range. It is important to observe that the RC-50 and NCB-50 curves are nearly alike.

Since its introduction, the presentations of the RC method have been revised several times, and the RC ratings for interior building spaces have been inconsistent among various ASHRAE publications (1991 Handbook, A Practical Guide for HVAC Systems, 1993 Pocket Guide for AHVR, and 1995 Handbook). Validation of the RC curves is still underway.^{11,12} There has been comment that the result of Ref. 11 can be misleading because earphones were used in the data evaluation of low-frequency rumble. However, earphones were not used in the study of Ref. 12. The RC criterion has not been adopted by ASHRAE as a standard for a very simple reason. ASHRAE is still conducting ongoing research to modify the RC method. In order to avoid conflict of interest, ASHRAE has very stringent requirements on the approval of an ASHRAE standard. The committee designated to prepare an equipment standard is required to have the following distribution: the manufacturers, the users, and the general interests. No one of these three groups can have a majority control. For a room noise standard the manufacturers would be the acoustical consultants; the users would be the building owners, the architects, and the HVAC engineers; and the general interests would be representatives from government and academia.

B. Goal of working group S12/WG18 has not been met

For more than 20 years the scope of the working group S12/WG18 (Criteria for Room Noise) has been “to develop a method for rating the noise encountered in various types of rooms in terms of its effect upon people by means of: (1) a weighting network and (2) a computational procedure to account for temporal pattern of the noise and the need for tone correction as appropriate. The purpose of this method is to arrive at a common yardstick for evaluating environmental noise indoors. In addition to this main charge, also charged with the development of measurement procedures to be utilized in various types of rooms including methods for translating the results of such measurement in the common yardstick.”

It is clear that neither NCB nor RC has met this goal. The NCB curves only work well for steady-state continuous noise. In order to address the “rumble” issue, the RC method treats all noises as though they needed spectral correction for rumble. However, both methods have failed since octave-band data are not suitable to describe most noises with abnormal characteristics such as prominent discrete tones, rumble, hiss, roars,...etc.

Without a separate descriptor and procedure to qualify, quantify, and determine the impact of “rumble,” the RC criteria results in prohibitively expensive treatments for quieting steady-state/continuous noise.

C. Room noise criteria as a practical design tool

ANSI S12.2 is not user friendly, even with the help of today's computers. In order for the standard to be practical, the most important thing is to separate the “exception” from the “norm.” One should note that “HVAC noise” and “HVAC noise problem” are not the same. With proper design most HVAC systems do not create noise problems. Some HVAC noise can even be desirable (i.e., providing masking sound for speech privacy). HVAC noise covers noise from very broad ranges of systems, equipment and components. Like prominent discrete tones, low-frequency rumble is certainly a bad problem. However, it is not the “norm.”

For the engineering community the need is a practical design tool for the “norm.” Although these engineers do design HVAC systems for noise control, very few of them have the opportunity to evaluate the rooms for conformance. When there is a noise problem, they retain the experts.

In the computer industry, the design criterion can be very simple: noise shall not exceed 35 dB A-weighted sound-pressure level, without prominent discrete tone, without impulsive noise, and without high-frequency noise. Similarly, for HVAC equipment, there is no reason for the specification to be longer than: noise shall not exceed NCB-35, without rumble, without prominent discrete tone, and without high-frequency noise.

For structure engineers, resonance can be avoided by following good engineering practice. For mechanical engineers designing large industrial ventilation systems, the beating of fans can be very bad—a situation which could shake walls and floors and even crack buildings (low-frequency noise, of course). Again, the practical solution is not to rate the beating. The solution is to follow good engineering practice so that the problems can be eliminated at the design stage.

Many of the low-frequency noise problems in buildings are introduced by roof-top units located on lightweight, thin slab, and very long spans; by fans with very bad inlet/outlet connections; and/or by fans operating in an unstable region. Again, the basic solution is not to develop a rating scheme. The solution is to recommend good engineering design practice. ASHRAE publications have provided guides to eliminate these problems.

D. Diagnostic tools, the challenges and opportunities for S12/WG18

For some acoustical consultants busy in fixing poorly designed HVAC systems, a diagnostic tool to evaluate these sick systems is required.

From a technology viewpoint, the development of diagnostic tools for noise with abnormal characteristics certainly is a challenge as well as an opportunity for ASA working group S12/WG18.

To develop diagnostic tools, ASA S12/WG18 can draw experience from various industries. For example, in the computer industry: (a) A-weighted sound level is used as the basic rating scheme; (b) a procedure is also available¹³ to determine objectively whether a discrete tone is prominent; and (c) procedures are available to adjust sound level for the presence of prominent discrete tones. It is important to note that this procedure and adjustment are only developed to provide correction to the A-weighted sound level measured, and are not intended to “replace” it.

E. Computers are now major noise sources in rooms

Both RC and NCB were developed with HVAC equipment as the dominant noise source in rooms. However, in the computer age, non-HVAC equipment has emerged as a major noise source, and must be considered when developing any room noise criterion. There should be room noise criteria which can apply to computers as well as to HVAC equipment/systems/components.

F. Low-frequency noise problems—the urgent needs

A demonstration at the ASA Boston Chapter’s panel discussion on Room Noise Criteria (November 1995) showed that low-frequency noises are not all equal (for room noise criteria). In a demonstration by Jacek Figwer, a 75-dB, 31.5-Hz pure tone was generated at a spot in the meeting room. It barely caught people’s attention when they listened there. There has been criticism that use of a pure tone was misleading since low-frequency rumble is more annoying and troublesome. However, it did demonstrate the need of another descriptor and procedure so that “rumble” can be (a) qualified, (b) quantified, and (c) evaluated for its impact on human response.

The use of octave-band data to describe noises with abnormal characteristics (such as rumble) have been the key technical deficiency within the ANSI S12.2 standard. It is clear that the annoyance caused by noise with abnormal sounds is usually greater than predicted by these simplified methods. Since both methods are limited to the assessment of continuous, steady-state background noise, they do not apply to abnormal noise such as rumble. Rumble is not the only problem that these methods face. Similar findings are evident in almost every type of noise with abnormal characteristics.

II. CONCLUSION

The ASA/NOISE-CON 97 is a great opportunity for inter-society partnership on room noise criteria. A pooling of the wisdom/resources from the HVAC and computer industries; the spirit of NC, NCB, RC, and the A-weighted sound level could be brought together into a single criterion for the “norms.” It is hoped that, working closely with INCE, the ASA S12 committee can provide the needed leadership to pool the resources/wisdom from the HVAC and the computer industries (and other concerned parties) to develop the needed room criterion for the computer age.

¹ ANSI S12.2-1996, American National Standard “Criteria for Evaluating Room Noise” (ANSI, New York, 1996).

² A. Marsh and H. Pei, “Room Noise Criteria—A Call for Consensus,” Proc. NOISE-CON 96, Seattle, Washington, 681–686, 1996.

³ L. Beranek, “Applications of NCB and RC Noise Criteria Curves to Specification and Evaluation of Noise in Buildings,” Proc. NOISE-CON 96, Seattle, Washington, 669–674, 1996.

⁴ W. Blazier, Jr., “Room Noise Criteria: The Importance of Temporal Variations in Low-Frequency Sounds from HVAC Systems,” Proc. NOISE-CON 96, Seattle, Washington, 687–692, 1996.

⁵ G. Tocci, “Comparison of NC, NCB, and RC Methods for Evaluating Room Sound Level Spectra,” Proc. NOISE-CON 96, Seattle, Washington, 675–680, 1996.

⁶ D. Sturz and K. Milligan, “A Vote for NC,” Proc. NOISE-CON 96, Seattle, Washington, 665–668, 1996.

⁷ R. Hellweg, Jr., “Room Sound Pressure Levels from Computer and Business Equipment Meeting Sound Power Level Criteria,” Proc. NOISE-CON 96, Seattle, Washington, 79–84, 1996.

⁸ H. Pei, “An Integrated Approach to Office Acoustics in the Computer Age,” 111th Meeting of the Acoustical Society of America, Cleveland, Ohio [J. Acoust. Soc. Am. Suppl. 1 **79**, S1 (1986)].

⁹ H. Pei, “Integrated Cooling in an Electronic Era,” Proc. 8th International Electronics Packaging Conference, Austin, Texas, 548–558, 1992.

¹⁰ L. Beranek, “Applications of NCB and RC Noise Criteria Curves to Specification and Evaluation of Noise in Buildings” (unpublished 1993).

¹¹ ASHRAE Research Project RP-74, “Determination of the Relationships between Low-Frequency HVAC Noise and Comfort in Occupied Spaces—Objective Phase,” 1994.

¹² ASHRAE Research Project RP-879, “Determination of the Relationships between Low-Frequency HVAC Noise and Comfort in Occupied Spaces—Psycho-Acoustic Phase” (in progress).

¹³ ECMA 74 European Computer Manufacturers Association Standard, “Measurement of Airborne Noise Emitted by Computer and Business Equipment, Annex D, Identification of Prominent Discrete Tones,” 3rd edition, 1992.

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:

GEORGE L. AUGSPURGER, *Perception Incorporated, Box 39536, Los Angeles, California 90039*

RONALD B. COLEMAN, *BBN Acoustic Technologies, 70 Fawcett Street, Cambridge, Massachusetts 02138*

HARVEY H. HUBBARD, *325 Charleston Way, Newport News, Virginia 23606*

D. LLOYD RICE, *11222 Flatiron Drive, Lafayette, Colorado 80026*

HAROLD C. ROBINSON, *NUWC, Code 2131, New London, Connecticut 06320*

ROBERT C. WAAG, *University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, New York 14642*

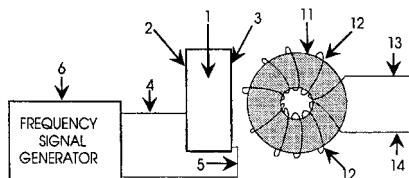
5,568,005

43.38.Ar ACOUSTI-MAGNETIC POWER GENERATOR

Dan A. Davidson, Sierra Vista AZ

22 October 1996 (Class 310/328); filed 24 January 1995

“What is new in this invention is the discovery of the ability of an acoustic field to stimulate the nuclear structure of a material to cause the electrons to wobble under the influence of the acoustic field. If the material is magnetic...then the magnetic field will vibrate under the influence of the



acoustic field.” Acoustic transducer 1 can be piezoelectric, magnetostrictive, or some other type. It is placed “...in close proximity to the permanent magnet 11 such that the acoustic field...can radiate into the permanent magnet material.” With the transducer in operation, “...a voltage and current is (sic) generated at the output 13 and 14.”—GLA

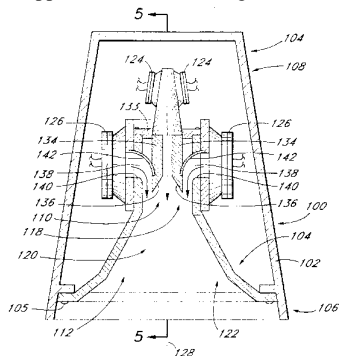
5,526,456

43.38.Ja MULTIPLE-DRIVER SINGLE-HORN LOUD SPEAKER

Ralph D. Heinz, assignor to Renkus-Heinz, Incorporated

11 June 1996 (Class 381/156); filed 30 January 1995

One or two high-frequency drivers 124 are coupled to a short, vertically expanding waveguide. Horn throat 110 is driven by the high-frequency waveguide and by two low-frequency waveguides 138. “The use of a common horn causes the apparent source of the high and low frequency sound to



be coincident in both the horizontal and vertical planes since the single horn acts as the waveguide for all sound produced.”—GLA

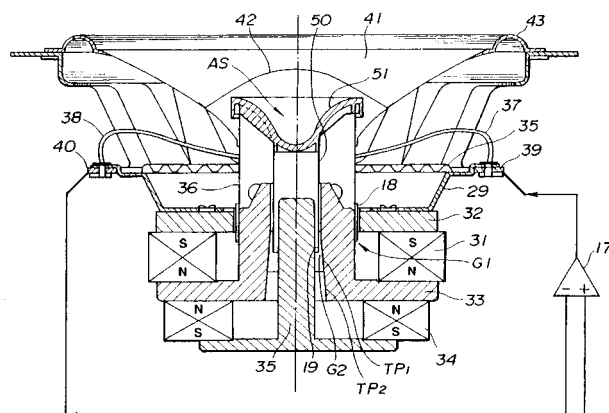
5,533,134

43.38.Ja MOTIONAL FEEDBACK LOUDSPEAKER APPARATUS HAVING A COUPLING MEMBER FOR CONNECTING A VOICE COIL BOBBIN WITH A DETECTING COIL BOBBIN

Kunihiko Tokura et al., assignors to Sony Corporation

2 July 1996 (Class 381/96); filed in Japan 16 August 1993

The apparatus is an interesting motional feedback loudspeaker design. It is a mystery why the patent document devotes a full page of text to such matters as the construction of a suitable box, the method of mounting an associated tweeter (set screws), and the location of a LED tally light (mid lower region thereof). Voice coil 18 and detection coil 19 are mechanically



coupled by cup-shaped element 51. The shape and material of the coupler are chosen to extend the usable frequency range of feedback control.—GLA

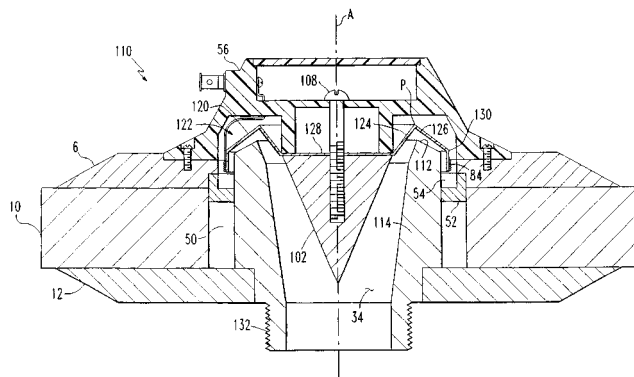
5,537,481

43.38.Ja HORN DRIVER

Alexander G. Voishvillo and Sergei Shurupov, assignors to The AWS Group, Incorporated

16 July 1996 (Class 381/192); filed 5 April 1994

The text of the patent includes a number of incorrect negative assertions about horn driver phasing plugs, but fails to mention their most significant drawback: good phasing plugs are expensive to make. A ring-



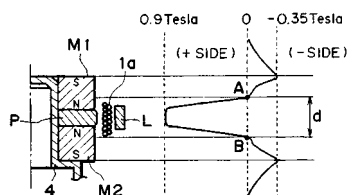
shaped diaphragm, on the other hand, can be narrow enough to drive a circular slot through a very small coupling chamber without significant path length differences. The patent refines this concept by clamping the central area of diaphragm **122** and then driving its outer edge by voice coil **84**. Losses in the diaphragm itself minimize path length differences at high frequencies, which emanate mainly from the region nearest the coil.—GLA

5,550,332

43.38.Ja LOUDSPEAKER ASSEMBLY

Yoshio Sakamoto, assignor to **Kabushiki Kaisha Kenwood**
27 August 1996 (Class 181/148); filed in Japan 20 October 1992

The patent describes a repulsive magnetic circuit loudspeaker design. By properly setting the depth of voice coil **1a** in relation to magnetic assem-



bly **M1**, relatively long linear travel at reasonable efficiency can be obtained.—GLA

5,550,921

43.38.Ja STEREO SOUND SOURCE FOR PORTABLE COMPUTER

Tommyca Freadman, assignor to **Sparkomatic**
27 August 1996 (Class 381/24); filed 15 May 1995

Back chambers for two small stereophonic loudspeakers vent into a mixing chamber which is coupled to the outside air via a third tuned vent. The system functions as a kind of stereo combining double-chamber reflex box. The patent specifications and claims stubbornly insist that it is to be used only for portable computer audio.—GLA

5,568,560

43.38.Lc AUDIO CROSSOVER CIRCUIT

Christopher E. Combest, assignor to **Multi Service Corporation**
22 October 1996 (Class 381/99); filed 11 May 1995

The circuit uses the principle that underdamped filters have response curves with steeper initial skirts. Opposed series-connected, mutually coupled inductors presumably contribute to the desired effect.—GLA

5,530,769

43.38.Lc EQUALIZER AND AUDIO DEVICE USING THE SAME

Hiroyuki Saitoh, assignor to **Rohm Company**
25 June 1996 (Class 381/103); filed in Japan 13 October 1993

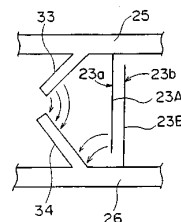
The basic circuit can be used in the design of practical graphic equalizers. Boost and cut functions are voltage controlled, providing flexibility in product design. Moreover, the actual filter for each band is located downstream of associated attenuating circuitry to improve signal-to-noise characteristics.—GLA

5,548,256

43.38.Rh SAW FILTER HAVING ELECTRODES FOR RELAXATION OF ELECTRIC FIELD CONCENTRATION

Toru Watanabe and Kenichi Noto, assignors to **Murata Manufacturing Company**
20 August 1996 (Class 333/194); filed in Japan 14 January 1994

Surface acoustic wave (SAW) filters often use substrate materials such as lithium niobate which are pyroelectric as well as piezoelectric. Thus, excess charges on the electrodes may be generated due to the application of thermal shock. At the ends of the interdigital transducer (IDT), the electric fields are particularly strong due to a discontinuity in the electric potential. These excess charges can discharge abruptly, damaging the end electrodes and altering the characteristics of the filter. This patent proposes adding "electric field concentration electrodes" **33** and **34** to the bus bars **25** and **26**



of each IDT which are oblique to the surface acoustic wavefront and do not overlap. These concentration electrodes are preferably wider than the interdigital electrodes **23a** and **23b**. The electric field lines, denoted by the arrows, now no longer terminate abruptly at **23a** (as they would without the concentration electrodes), but make a much smoother transition, lessening the chance of catastrophic discharge. The concentration electrodes are less susceptible to discharge damage, in turn, due to their larger size. The obliqueness of the concentration electrodes prevents spurious reflections of the SAW, while their separation from each other and from the interdigital electrodes prevents generation of undesired SAWs.—HCR

5,559,481

43.38.Rh SURFACE ACOUSTIC WAVE FILTER

Yoshio Satoh et al., assignors to **Fujitsu Ltd.**
24 September 1996 (Class 333/143); filed in Japan 28 October 1991

This patent describes a ladder-type surface wave filter for use in mobile telephones. For this application, filters with a large bandwidth, a large sidelobe suppression factor, and a small insertion loss are required. In this particular filter, a SAW resonator with a particular resonance frequency and an inductor are placed in the parallel arm of the filter. A second SAW resonator, with a resonance frequency approximately equal to the antiresonance of the first, is placed in the series arm. The two resonators yield a passband centered near the antiresonance frequency of the parallel resonator. Improved sidelobe suppression has typically been achieved by increasing the capacitance of the parallel arm, but the cost of this is a loss in bandwidth. In this patent, the presence of the inductor in the parallel arm leads to improved sidelobe suppression characteristics but also increases the difference between the effective resonance and antiresonance frequencies of the parallel arm, thereby also increasing the bandwidth simultaneously. The insertion loss and passband ripple of this filter design appears to be controllable by controlling the difference between the series resonance and parallel

antiresonance. Several versions of this patent, including cascading filter elements together and varying the inductances of each, are described.—HCR

5,565,725

43.38.Rh SURFACE ACOUSTIC WAVE DEVICE

Hideaki Nakahata *et al.*, assignors to Sumitomo Electric Industries, Ltd.

15 October 1996 (Class 310/313); filed in Japan 10 May 1994

The operating frequency of a surface acoustic wave (SAW) device is determined by the period of the interdigital transducers generating the SAW, which gives the SAW wavelength, and by the surface acoustic wave velocity in the substrate. Because photolithographic techniques provide a lower limit on the wavelength, using SAW devices at higher frequencies requires larger surface acoustic wave velocities. Furthermore, the electromechanical coupling coefficient of many pure SAW substrate materials is quite low, so increasing it is also desirable. This patent proposes depositing a layer of lithium niobate on a layer of diamond, which may or may not be deposited on a layer of silica, to achieve the increase in SAW velocity and coupling coefficient. The device uses the zeroth, first or second layered Rayleigh modes, the velocity of which are determined by the thickness of the lithium niobate layer. If the diamond is deposited on another base material, the thickness of the diamond becomes the controlling SAW velocity parameter. The coupling coefficient increases due to the change in the SAW velocity. The interdigital transducer (IDT) electrodes may be placed on top of the lithium niobate layer or between the lithium niobate and the diamond. Shorting electrodes may also be added, on the face of the lithium niobate layer which does not have the IDT.—HCR

5,524,059

43.38.Si SOUND ACQUISITION METHOD AND SYSTEM, AND SOUND ACQUISITION AND REPRODUCTION APPARATUS

Frederic Zurcher, assignor to Prescom

4 June 1996 (Class 381/92); filed in France 2 October 1991

A common configuration for freestanding teleconference systems is an upward-firing loudspeaker with a coaxial cluster of microphones at its base. The system minimizes sound pickup from the loudspeaker by introducing frequency-independent electrical phase shift between adjacent microphones such that net gain along the axis of symmetry is zero. The idea is clever and the patent is clearly written.—GLA

5,528,694

43.38.Vk AUDIO SIGNAL PROCESSING ARRANGEMENT FOR DERIVING A CENTER CHANNEL SIGNAL AND ALSO AN AUDIO VISUAL REPRODUCTION SYSTEM COMPRISING SUCH A PROCESSING ARRANGEMENT

Leon M. Van De Kerkhof *et al.*, assignors to U.S. Philips Corporation

18 June 1996 (Class 381/27); filed in European Patent Office 27 January 1993

The patent describes circuitry for dynamically deriving a center channel from non-encoded two-channel stereophonic program signals. Left and right channels are compared in discrete frequency bands. Only those bands displaying correlation above a predetermined threshold are combined to form the center channel.—GLA

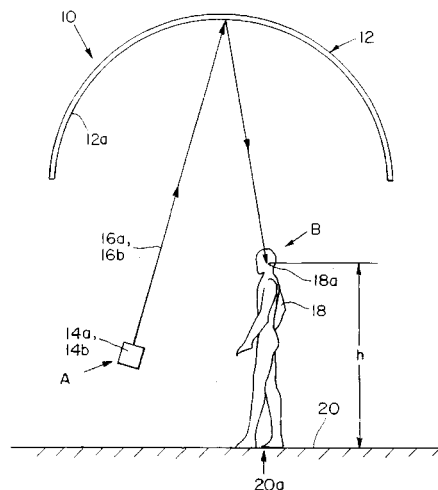
5,532,438

43.38.Vk ACOUSTIC IMAGING SOUND DOME

Kevin Brown, Boston, MA

2 July 1996 (Class 181/155); filed 4 November 1993

Parabolic reflectors have long been used to produce sound seemingly from nowhere, as in a whispering gallery. The patent shows that by using a spherical reflector 12, the location of receiving focal point 18a becomes a function of source location 14a. It follows that sound fields from a pair of stereophonic loudspeakers can be focused to the respective ears of a listener



without disturbing other people nearby. The geometry of the illustration apparently is not to be used as an example, because the addition of a few more rays makes it obvious that they do not really focus anywhere.—GLA

5,572,443

43.38.Vk ACOUSTIC CHARACTERISTIC CORRECTION DEVICE

Naohiro Emoto and Tsugio Ito, assignors to Yamaha Corporation

5 November 1996 (Class 364/550); filed in Japan 11 May 1993

Charts and graphs take up the first 32 pages of this 54-page patent. The patent discloses an elaborate arrangement for equalizing a stereo playback system, involving a waveform memory, convolution operator, frequency conversion, Hilbert transform, and all sorts of other digital goodies. Although some questionable assertions are made, the patent includes a lot of useful information for people interested in this field.—GLA

5,498,115

43.40.Tm ACTIVE SUSPENSION SYSTEM

Seiji Osaki *et al.*, assignors to Honda Giken Kogyo Kabushiki Kaisha

6 February 1996 (Class 280/707); filed in Japan 6 August 1990.

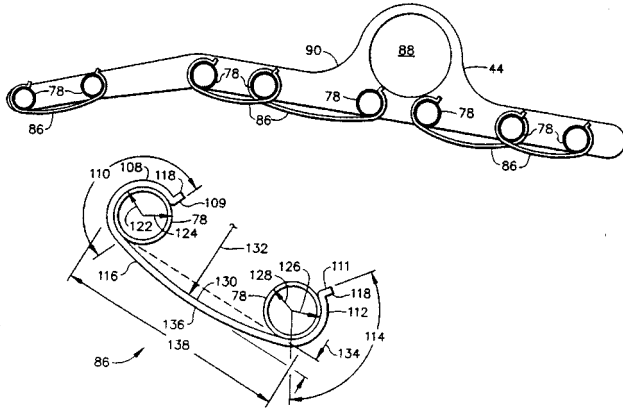
Conventional active isolation systems for automobiles attempt to modify the apparent spring rate and the damping of the passive isolator located between the vehicle body (the sprung mass) and the wheel (the unsprung mass). In addition, the apparent damping rate between the sprung mass and ground can be modified. This later approach is referred to as "sky-hook" damping. This patent augments these conventional approaches by including an additional term in the cost function of the controller to regulate the apparent mass of the wheel. This additional inertial-force constraint, in combination with the sky-hook damping constraint, is said to enhance the ability of the tire to maintain contact with the road surface. This leads to significant driving improvements in handling with little effect on the behavior of the sprung mass.—RBC

5,540,547

43.40.Tm METHODS AND APPARATUS FOR DAMPING VIBRATIONS OF EXTERNAL TUBING OF A GAS TURBINE ENGINE

Lester E. Cole, assignor to General Electric Company
30 July 1996 (Class 415/177); filed 23 June 1994

This patent relates to the damping of vibrations in the external tubing segments of gas turbines. Spring-loaded metal clips are placed at predeter-



mined positions along the tubing as a way of adding structural damping to the system and selectively changing the frequency of vibrations.—HHH

5,540,551

43.40.Tm METHOD AND APPARATUS FOR REDUCING VIBRATION IN A TURBOMACHINE BLADE

Roger W. Heinig, assignor to Westinghouse Electric Corporation
30 July 1996 (Class 416/190); filed 3 August 1994

This patent relates to the reduction of vibrations in the rotating blades of a turbomachine such as a steam turbine. Both inner and outer blade restraints are incorporated in the form of shrouds, welded lugs, or tie wires to tie together groups of blades. The numbers of blades in each group is a function of the blade resonance frequency, the total number of blades, and the radial distances involved.—HHH

5,485,053

43.40.Vn METHOD AND DEVICE FOR ACTIVE CONSTRAINED LAYER DAMPING FOR VIBRATION AND SOUND CONTROL

Amr M. Baz, Rockville, MD
16 January 1996 (Class 310/326); filed 15 October 1993

This patent presents a method and device for an actively controlled constrained layer (ACLCD) treatment which can be used as an effective means for damping out vibrations and sounds from flexible structures. The ACLCD treatment consists of a viscoelastic damping layer which is sandwiched between two piezoelectric layers. The three-layer composite, when bonded to a surface which is subject to vibration, acts as a smart constraining layer damping treatment with built-in sensing and actuation capabilities. The sensing capability is provided by the piezoelectric layer bonded to the

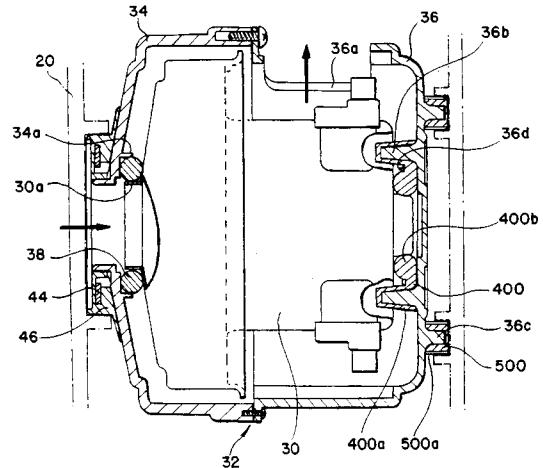
surface of the flexible structure, whereas the actuation or control capability is generated by the other piezoelectric layer which acts as an active constraining layer.—RBC

5,548,867

43.50.Gf VIBROISOLATING APPARATUS OF VACUUM CLEANER

Jin-Sung Hwang and Jin-Seung Sohn, assignors to Samsung Electronics Company
27 August 1996 (Class 15/339), filed in Republic of Korea 3 March 1994

This patent relates to the reduction of noise from a vacuum cleaner which incorporates a suction motor. A vibroisolating ring at the front surface of the suction motor and a rear surface packing are provided to struc-



turally isolate the motor and minimize the vibrations transmitted to surrounding structural elements.—HHH

5,549,454

43.50.Gf HIGH SPEED VACUUM PUMP WITH REDUCED EXHAUST NOISE

Stephen R. Earle, assignor to WABCO Automotive
27 August 1996 (Class 417/312); filed in the United Kingdom 4 March 1993

This patent relates to the reduction of exhaust noise from high-speed vacuum pumps. Interior design details provide a primary exhaust passage between the outer wall of the liner and the inner wall of the body to an exhaust port or ports. The exhaust passage includes restrictors which permit the interior spaces to be used as silencer chambers.—HHH

5,550,333

43.50.Gf REDUCED NOISE TRAILERABLE ENGINE

Marvin E. Whiteman, Jr., assignor to Whiteman Industries, Incorporated
27 August 1996 (Class 181/204); filed 10 January 1995

This patent applies to trailerable housings for gasoline or diesel engines used as power sources at remote locations. Unibody engine enclosures of sheet metal construction are provided which allow cross flow ventilation

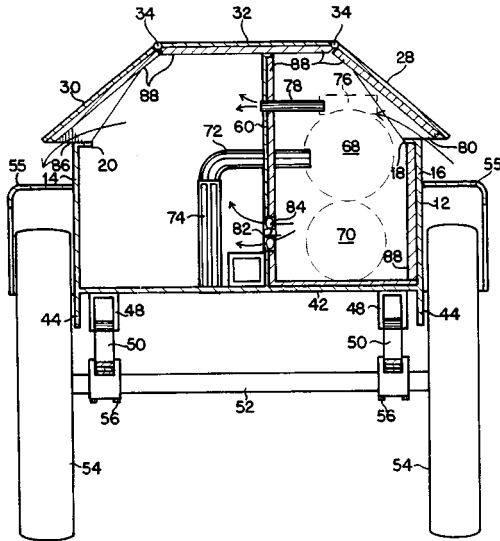
5,485,523

43.50.Ki ACTIVE NOISE REDUCTION SYSTEM FOR AUTOMOBILE COMPARTMENT

Manpei Tamamura *et al.*, assignors to Fuji Jukogyo Kabushiki Kaisha

16 January 1996 (Class 381/71); filed in Japan 17 March 1992

This patent discusses an active noise control system to reduce the engine noise reaching an automobile compartment. The basic control approach is a filtered-X least-mean-square (FX-LMS) controller. Several reference sensor options are proposed to obtain a reference signal which is well correlated with the engine noise reaching the interior. The preferred reference signal is the ignition pulse signal, which is said to be better than using accelerometers mounted to the engine. Several schemes for system identification are proposed to determine the plant transfer function between the loudspeakers and the residual microphones. One approach uses seat sensors to determine the number of occupants. Gain scheduling is then used to select the appropriate plant filter through which the reference signal is to be passed.—RBC



but which obstruct direct noise radiation. Sound absorbing materials are attached selectively to interior surfaces to maximize noise reduction while minimizing the obstruction to cooling air flows.—HHH

5,555,726

43.50.Gf ATTENUATION OF FLUID BORNE NOISE FROM HYDRAULIC PISTON PUMPS

Robert J. Huebner, assignor to Caterpillar, Incorporated
17 September 1996 (Class 60/469), filed 31 March 1995

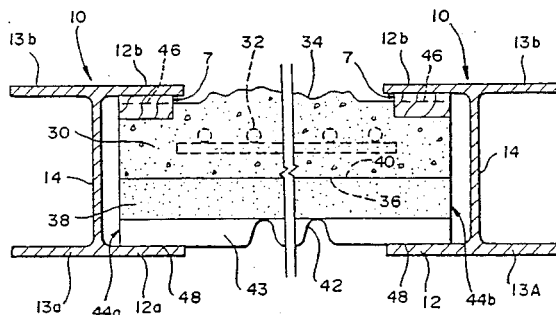
This patent relates to the attenuation of fluid borne noise in systems which incorporate a hydraulic piston pump. The objective is to reduce the flow ripple produced by the pump to produce a generally uniform flow to the rest of the system. A sensor arrangement is included to supply electrical signals representing the operating parameters to a microprocessor which transmits command signals to the electrically controlled valve and port mechanisms to selectively control the fluid flow.—HHH

5,572,847

43.50.Gf RAPIDLY ERECTABLE, REMOVABLE, REUSABLE AND RAISABLE OUTDOOR ACOUSTICAL WALL SYSTEM

J. Thomas Elmore *et al.*, assignors to JTE, Incorporated
12 November 1996 (Class 52/766); filed 22 February 1995

In order to make this noise barrier wall easily "erectable, removable, reusable and raisable," the precast wall panel is held between the flanges 12b of posts 10 with wedge members 7 in sealing contact with the inside



surfaces of the flanges, in order to prevent noise leakage between the edges of the panels and the posts.—DWM

5,488,667

43.50.Ki VEHICLE INTERNAL NOISE REDUCTION SYSTEM

Manpei Tamamura *et al.*, assignors to Fuji Jukogyo Kabushiki Kaisha

30 January 1996 (Class 387/71); filed in Japan 1 February 1993

This is another system for controlling vehicle interior noise using a combination of engine-related reference sensors, audio loudspeakers, microphones, and a multichannel filtered-X least-mean-square (FX-LMS) control algorithm. Reference sensors on the engine are passed through adaptive filters to provide drive signals to the loudspeaker electronics. In addition, these signals are filtered by estimates of the plant (transfer function from the loudspeakers to the microphones). The novelty is said to be in storing these plant transfer functions as a function of engine RPM and load (known as gain scheduling in the literature). Based on a signal from a fuel-injection sensor, the engine RPM and load are estimated, and the appropriate plant transfer function is extracted for use from a look-up table. An explanation is not given for why the transfer functions from loudspeakers to microphones should depend on engine load.—RBC

5,498,127

43.50.Ki ACTIVE ACOUSTIC LINER

Robert E. Kraft and Karen B. Kontos, assignors to General Electric Company

12 March 1996 (Class 415/119); filed 14 November 1994

This patent describes an active acoustic liner for attenuating noise in an aircraft engine nacelle. The active liner incorporates piezoelectric actuators with collocated pressure transducers to attenuate relatively low-frequency noise (below about 1000 Hz) within the nacelle by regulating the acoustic impedance at the face of the actuator which is in contact with the flow. The desired (i.e., optimum) acoustic impedance is determined analytically. The measured pressure signal is filtered to produce a drive signal to the piezoelectric actuator such that the pressure divided by the resulting velocity response of the actuator face approximates the desired acoustic impedance. Multiple actuators and pressure sensors are envisioned to achieve control at multiple tonal frequencies, for broadband control below 1000 Hz.—RBC

5,499,301

43.50.Ki ACTIVE NOISE CANCELLING APPARATUS

Yuko Sudo *et al.*, assignors to Kabushiki Kaisha Toshiba
12 March 1996 (Class 381/71); filed in Japan 19 September 1991

This patent describes an active noise canceling apparatus for controlling noise generated by a compressor or similar rotating machinery. The premise is that passive or structural enclosures around such machinery will have openings through which sound will "leak" or radiate. A nominally collocated microphone and loudspeaker pair are used in a feedback control system to minimize noise radiated through these openings. The example given pertains to controlling compressor noise in a refrigerator.—RBC

5,511,127

43.50.Ki ACTIVE NOISE CONTROL

Glenn E. Warnaka, assignor to Applied Acoustic Research
23 April 1996 (Class 381/71); filed 5 April 1991

In this patent by a former reviewer for this journal, an active control system is presented for reducing tonal noise from fans radiated out of the inlet ducts. The approach uses loudspeakers in the duct to generate canceling noise and microphones to measure the residual noise in the duct. The approach does not include the traditional tachometer to measure the instantaneous rotational frequency in order to infer the blade-passage frequency and harmonics. Instead, these "offending" frequencies are "embedded" in the controller as *a priori* knowledge. This can eliminate the need for a traditional reference sensor in controlling systems for which the tonal disturbances are highly stable. That is, if you know the frequency to be controlled, there is no need to measure it.—RBC

5,553,150

43.55.Lb REVERBERATION-IMPARTING DEVICE CAPABLE OF MODULATING AN INPUT SIGNAL BY RANDOM NUMBERS

Koishi Kozuki, assignor to Yamaha Corporation
3 September 1996 (Class 381/61); filed in Japan 21 October 1993

Several means, including random modulation and cross-fading, are used to produce a more natural sounding digital reverberator. The patent is not easy to follow but should be informative to people interested in audio signal processing.—GLA

5,555,311

43.58.Ta ELECTRO-ACOUSTIC SYSTEM ANALYZER

Robert W. Reams, assignor to Electronic Engineering and Manufacturing, Incorporated
10 September 1996 (Class 381/58); filed 1 April 1994

The analyzer is a computer-based method and apparatus, "...for analyzing performance parameters of an electro-acoustic system." Readers expecting a menu of FFTs, windowed impulse response calculations, and cumulative decay spectra will be disappointed. Instead, the system provides information about such things as equalizability, spurious vibrations, unambiguous bandwidth, and "thermal limit and/or mass."—GLA

5,500,900

43.66.Qp METHODS AND APPARATUS FOR PRODUCING DIRECTIONAL SOUND

Jiashu Chen *et al.*, assignors to Wisconsin Alumni Research Foundation
19 March 1996 (Class 381/17); filed 23 September 1994

The U.S. Government, among others, is interested in accurately positioning phantom sound sources around the head of a person wearing earphones. The patent describes a method for measuring free-field-to-ear transfer functions at discrete locations and then mapping and filtering these in a manner that permits interpolation between measured locations. The patent document itself is unusual in that most of the heavy math is found in the claims section.—GLA

5,536,171

43.70.Dn SYNTHESIS-BASED SPEECH TRAINING SYSTEM AND METHOD

Hector R. Javkin *et al.*, assignors to Panasonic Technologies, Incorporated
16 July 1996 (Class 434/185); filed 28 May 1993

This articulation training system, designed particularly for deaf or otherwise impaired students, produces a display of vocal tract shapes for a word or phrase based on an articulatory model of phonetic characteristics. The desired utterance is entered by keyboard and a synthesizer rule system generates both articulatory controls for the display and formant parameters for a typical formant synthesizer. The student may then speak the utterance, which is analyzed to recover comparable articulatory parameters. Similarities between the student and the synthesizer utterances are displayed.—DLR

5,533,118

43.72.Ar VOICE ACTIVITY DETECTION METHOD AND APPARATUS USING THE SAME

Claude Cesaro and Gerard Richter, assignors to International Business Machines Corporation
2 July 1996 (Class 379/386); filed in European Patent Office 29 April 1993

This signal classifier determines whether a telephone signal contains voice, a single call-progress tone, a dual-tone pattern, or noise. A ratio is first computed relating the squared maximum amplitude during a 20-ms interval to an energy measure taken during the interval. Three thresholds serve to divide the ratio value into one of the four classes. If the ratio is above 5.2, then a signal level above -43 dB m is taken as speech and a lower signal level as noise. If the ratio is below 5.2 but greater than 3, a dual tone is indicated, while a ratio below 3 would mark a single tone.—DLR

5,537,647

43.72.Ar NOISE RESISTANT AUDITORY MODEL FOR PARAMETERIZATION OF SPEECH

Hynek Hermansky and Nelson H. Morgan, assignors to U S West Advanced Technologies, Incorporated; International Computer Science Institute
16 July 1996 (Class 395/2.2); filed 19 August 1991

This low-noise speech parameterization method is based on the first author's published technique, known as Perceptual Linear Predictive (PLP) analysis. The primary stated goal of the method is to reduce both additive and convolutive noise components from the speech signal. The FFT power spectra are warped to a Bark scale, then convolved with a critical-band masking function which approximates the cochlear response. The resulting critical band power spectra are resampled into 18 bands. An offset log operation then allows spectral filtering to reduce both additive and convolutive noise.—DLR

5,544,277

43.72.Ar SPEECH CODING APPARATUS AND METHOD FOR GENERATING ACOUSTIC FEATURE VECTOR COMPONENT VALUES BY COMBINING VALUES OF THE SAME FEATURES FOR MULTIPLE TIME INTERVALS

Raimo Bakis *et al.*, assignors to International Business Machines Corporation

6 August 1996 (Class 395/2.2); filed 28 July 1993

In this speech analyzer, intended for use in a recognition system, additional spectral feature vectors are generated at each frame interval by combining the spectral analysis results from two or more frames in various weighted combinations. For example, two different spectral weighting vectors would produce two different feature vectors from a single frame of spectral data. These could be further proliferated by combining the previous and present frames with various weightings. The weights could then be made to vary according to noise levels or other factors.—DLR

5,539,859

43.72.Dv METHOD OF USING A DOMINANT ANGLE OF INCIDENCE TO REDUCE ACOUSTIC NOISE IN A SPEECH SIGNAL

Francois Robbe and Luc Dartois, assignors to Alcatel N.V.
23 July 1996 (Class 395/2.42); filed in France 18 February 1992

Two microphones are placed a suitable distance apart to pick up a desired signal in a noisy environment. The first processing stage is a version of channel delay beamforming in which the direction to the source is determined from the delay required to time align the two channels. Here, the delay operation is done independently on each of a set of frequency bands. After determining a best estimated source direction considering all bands, a noise spectrum is updated with the band energy of any bands whose individual direction estimate differs by more than a threshold angle from the overall estimated direction. The estimated noise spectrum is subtracted from the incoming signals, the appropriate delay is applied, and the signals summed, resulting in the clean output.—DLR

5,535,299

43.72.Gy ADAPTIVE ERROR CONTROL FOR ADPCM SPEECH CODERS

Neal K. Riedel, assignor to Pacific Communication Sciences, Incorporated
9 July 1996 (Class 395/2.21); filed 2 November 1993

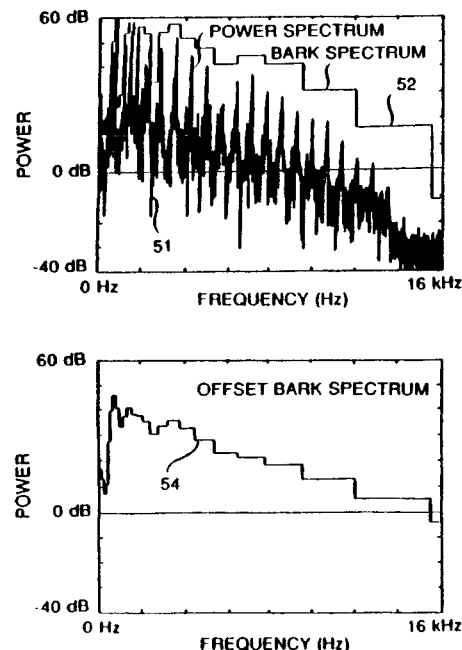
The CCITT standard G.721 for ADPCM speech coding has been adopted for use in certain applications, such as Personal Communication Services, in which the expected transmission error rates may exceed that which the ADPCM system can comfortably tolerate. This patent covers a method of limiting the absolute magnitude of the incoming data stream prior to decoding into speech. In addition, channel error indicators may allow the replacement of doubtful quality data with ADPCM-encoded silence at the receiver.—DLR

5,535,300

43.72.Gy PERCEPTUAL CODING OF AUDIO SIGNALS USING ENTROPY CODING AND/OR MULTIPLE POWER SPECTRA

Joseph L. Hall II and James D. Johnston, assignors to AT&T Corporation
9 July 1996 (Class 395/2.36); filed 30 December 1988

This is a coding technique suitable for speech or music which is said to introduce only minimal perceptible distortion into a high-quality signal while compressing the bitrate from CD rates to 192K or even 128K bit



per s. For a 32K sample rate, 1024-sample frames are overlapped by 128 points. For stereophonic data, sum and difference channels replace left and right. The FFT power spectra are summed across Bark-sized bands, then quantized using an adjustable bit allocation based on a masking noise level estimate. Finally, for stereophonic data, redundancies are reduced by considering the different statistical properties of the difference channel.—DLR

5,544,278

43.72.Gy PITCH POST-FILTER

Leon Bialik and Felix Flomen, assignors to Audio Codes Limited
6 August 1996 (Class 395/2.77); filed 29 April 1994

The patent begins by describing a method of enhancing the quality of speech which has been compressed for transmission or storage. Assuming that the coding technique provides a reasonably good quality pitch estimate, the speech can be enhanced by adding fractions of corresponding samples from the previous pitch period. The authors then note that certain coding techniques which break the frames into shorter subframe segments make available to the receiver, in certain cases, subframe data which is a pitch period ahead of that currently being synthesized. Adding in a weighting of this future information improves the result even more.—DLR

5,546,498

43.72.Gy METHOD OF AND DEVICE FOR QUANTIZING SPECTRAL PARAMETERS IN DIGITAL SPEECH CODERS

Daniele Sereno, assignor to Sip—Societa Italiana per l'esercizio Delle Telecomunicazioni S.p.A.
13 August 1996 (Class 395/2.38); filed in Italy 10 June 1993

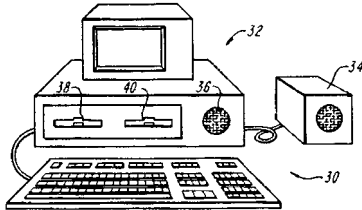
This speech coding system reduces the bitrate for encoding spectral parameters by taking account of correlations among the parameter values, both within frames and between frames in sequences of two or more frames. A parameter indexing scheme is described which is said to simplify the correlation computations.—DLR

5,540,589

43.72.Ja AUDIO INTERACTIVE TUTOR

Richard C. Waters, assignor to Mitsubishi Electric Information Technology Center
30 July 1996 (Class 434/156); filed 11 April 1994

This computer-based tutoring system centers around a speech recognition unit. Study material read from a CD-ROM disk consists, in this example presentation, of a Spanish language course. The primary emphasis of



the patent is on how error rates in factual replies from the user's responses can be combined with speech recognition error rates to evaluate the student's progress, to control the flow of the tutorial session, and, not least, to allow the use of a lower accuracy speech recognizer system.—DLR

5,533,105

43.72.Kb METHOD AND APPARATUS FOR EMULATING A TELEPHONE WITH A MODEM AND HEADSET

Paul M. Brown *et al.*, assignors to Cirrus Logic, Incorporated
2 July 1996 (Class 379/93); filed 15 September 1993

This patent reveals an analog circuit design which may be added to a traditional voice/fax/modem chip to provide full modem, fax, and speech transmission functionality. A communications-type headset plugged into the circuit may be used as a normal telephone connection. The circuit provides electret-style "phantom" power, if required. All standard modem and fax functions remain available via direct phone connection to the modem chip. Alternate version consider the patented circuit as external to or as incorporated into the modem chip.—DLR

5,533,133

43.72.Kb NOISE SUPPRESSION IN DIGITAL VOICE COMMUNICATIONS SYSTEMS

Allan Lamkin *et al.*, assignors to Hughes Aircraft Company
2 July 1996 (Class 381/94); filed 26 March 1993

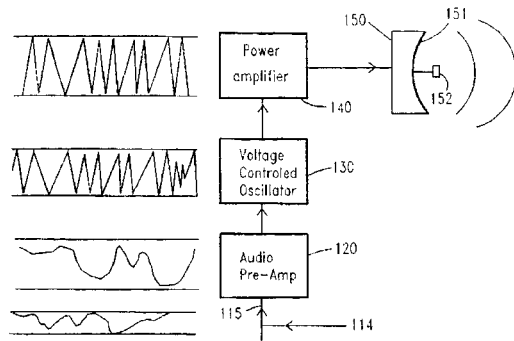
The background noise level presented to the listener during pauses and silences in speech transmission is deemed to be an important parameter in the perceived quality of the connection. This system blocks loud noises but passes a small noise signal during the gaps such that the listener "senses" the continued connection. A slightly lower noise level is presented when the connection is still open but the sender is not speaking.—DLR

5,539,705

43.72.Kb ULTRASONIC SPEECH TRANSLATOR AND COMMUNICATIONS SYSTEM

M. Alfred Akerman *et al.*, assignors to Martin Marietta Energy Systems, Incorporated
23 July 1996 (Class 367/132); filed 27 October 1994

This device allows a speech communication channel to be modulated onto an ultrasonic acoustic carrier for a "local" wireless connection. Two advantages over rf transmission are cited; use where rf emissions are prohibited, such as near explosives, and improved security from eavesdropping. Some prior systems of this type used amplitude modulation, which would



often be demodulated by any of numerous natural nonlinear pathways, compromising the secure aspect of the system. Here, frequency modulation avoids that problem. Various applications mentioned include vehicle-to-vehicle, room-to-room in a large building, or in noisy conditions. Range of up to a few miles is said to be possible with currently available ultrasonic transducers.—DLR

5,539,858

43.72.Kb VOICE CODING COMMUNICATION SYSTEM AND APPARATUS

Seishi Sasaki and Masayasu Miyake, assignors to Kokusai Electric Company
23 July 1996 (Class 395/2.21); filed in Japan 31 May 1991

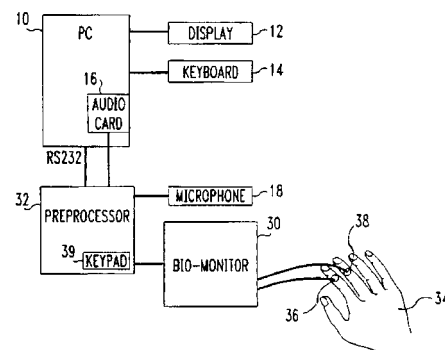
This ADPCM-based speech coding and transmission system detects voice activity at the transmitter and enters a power-saving idle state when there is no speech. The receiving unit detects the loss of transmitted carrier signal and produces a noise signal of the appropriate power level to give the appearance that the channel remains open.—DLR

5,539,860

43.72.Ne SPEECH RECOGNITION USING BIO-SIGNALS

Joseph DeSimone and Jian-Tu Hsieh, assignors to AT&T Corporation
23 July 1996 (Class 395/2.43); filed 22 December 1993

This patent describes how an auxiliary signal, derived from a biosensor attached to or worn by the talker, is used as additional information input to a speech recognizer, hopefully improving the performance of that recognizer. For illustration, a skin conductivity sensor feeds an oscillator circuit,



producing a variable rate pulse which is sampled concurrently with the speech input. Recognizer details are scarcely touched upon, yet it is assumed that the biosignal will vary in frequency in a fixed relationship to variations in certain speech frequencies, all as a function of the speaker's emotional state, and that this correlation will be in a form usable by the recognizer.—DLR

5,539,861

43.72.Ne SPEECH RECOGNITION USING BIO-SIGNALS

Joseph DeSimone, assignor to AT&T Corporation
23 July 1996 (Class 395/2.43); filed 22 December 1993

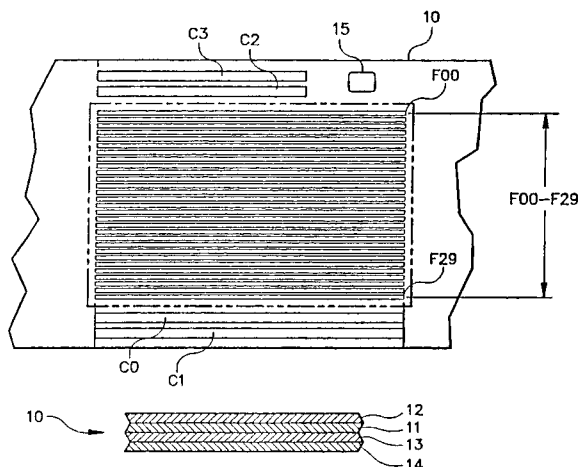
The patent text is nearly identical to Patent 5,539,860 (reviewed above). Several new claims have been added, dealing primarily with various ways to preprocess the biosensor signal.—DLR

5,546,145

43.72.Ne CAMERA ON-BOARD VOICE RECOGNITION

Bryan D. Bernardi *et al.*, assignors to Eastman Kodak Company
13 August 1996 (Class 354/76); filed 30 August 1994

This patent adds a speech recognizer subsystem to a digital camera. Vocabulary management facilities are included by means of a docking mechanism so the camera may be plugged into a supporting host computer. The upshot is that the recognizer is used as a speech compressor, that is, the recognized words or phrases are digitally stored with the image information



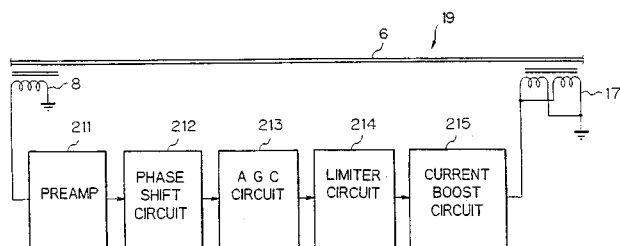
for each frame, as shown in the figure. There is no mention of using the recognition results for camera control. This is ironic, because the free form dictationlike input for annotating images would probably be nearly impossible to recognize with today's technology, whereas a camera control system might be feasible.—DLR

5,585,588

43.75.Gh ELECTRIC STRINGED INSTRUMENT HAVING A DEVICE FOR SUSTAINING THE VIBRATION OF A STRING AND AN ELECTROMAGNETIC DRIVER FOR THE DEVICE

Kenji Tumura, assignor to Fernandes Company
17 December 1996 (Class 84/726); filed in Japan 24 October 1991

Musical string instrument vibrations have been sustained electromagnetically through the use of magnetic pickups in proximity to the string supplying amplified output to magnetic driver units adjacent to the string. In



the present patent, phase shift circuit 212 has been inserted into the amplification system to provide a "well-balanced excitation" for greater control of the sustaining effect.—DWM

5,585,587

43.75.Tv ACOUSTIC IMAGE LOCALIZATION APPARATUS FOR DISTRIBUTING TONE COLOR GROUPS THROUGHOUT SOUND FIELD

Toshihiro Inoue and Hiroyuki Torimura, assignors to Yamaha Corporation
17 December 1996 (Class 84/662); filed in Japan 24 September 1993

In this electronic musical system patent various tone signals representing different types of musical instruments are routed to different loudspeakers in a spaced array in order to simulate more closely the sound of an ensemble of acoustic instruments, such as an orchestra or band. In some cases a specific signal may be routed to more than one loudspeaker to provide images at other than loudspeaker locations. The terminology used in the patent sounds newer than the principles used.—DWM

5,533,510

43.80.Qf REAL TIME ULTRASOUND ENDOCARDIAL DISPLACEMENT DISPLAY

Albert F. Koch III and David M. Prater, assignors to Hewlett-Packard Company
9 July 1996 (Class 128/660.07); filed 22 September 1995

This display uses circuitry that classifies pixels as either tissue or fluid in a sequence of frames. A comparator determines which pixels change classification from one type to another between frames, and a color is assigned to pixels that change so that the changes can be readily seen. Different color values are used from frame-to-frame on both expansion and contraction cycles of the ventricle.—RCW

5,546,946

43.80.Qf ULTRASONIC DIAGNOSTIC TRANSDUCER ARRAY WITH ELEVATION FOCUS

Jacques Souquet, assignor to Advanced Technology Laboratories, Incorporated
20 August 1996 (Class 128/662.03); filed 24 June 1994

Elements of this transducer array are subdiced in the elevation direction to produce subelements with aspect ratios that vary in proportion to distance from the central longitudinal axis of the array. The electromechanical coupling between subelements is consequently changed so that the intensity of the transmitted energy is centered about the longitudinal axis. In the elevation direction, the elements may have an extension that is proportional to the displacement of the element from the longitudinal axis. Extended elements are acoustically separated into subelements that focus in elevation or spatially compound transmitted acoustic energy.—RCW

5,540,656

43.80.Sh ULTRASONIC ANGIOPLASTY DEVICE HAVING SURFACE DISRUPTIONS

Russell Pflueger *et al.*, assignors to Baxter International, Incorporated
30 July 1996 (Class 604/22); filed 24 January 1995

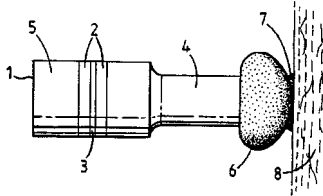
The proximal end of this device attaches to a generator of ultrasound, and a bulbous distal head is used to ablate an occluding lesion. Surface disruptions on the distal head are included to improve the efficacy of treatment. The device may be incorporated into an angioscope.—RCW

5,549,544

43.80.Sh APPARATUS FOR ULTRASONIC THERAPEUTIC TREATMENT

Michael J. R. Young and Brian R. D. P. Bradnock, assignors to Orthosonics Ltd.
27 August 1996 (Class 601/2); filed 10 December 1993

This apparatus contains a piezoelectric vibrator that produces ultrasonic energy. The energy is transmitted to a plastic head. The shape of the



head is chosen to allow control of the ultrasonic frequency and amplitude.—RCW

5,549,638

43.80.Sh ULTRASOUND DEVICE FOR USE IN A THERMOTHERAPY APPARATUS

Everette C. Burdette, Champaign, IL
27 August 1996 (Class 607/97); filed 17 May 1994

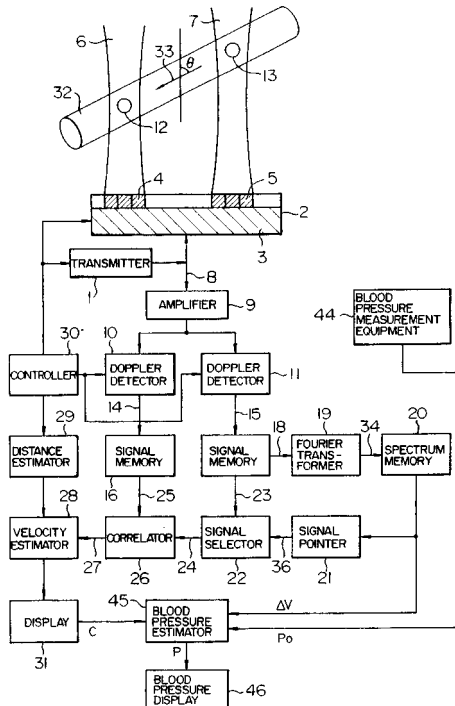
The transducers in this device are arranged in cylindrical shape for patient treatment. A variable number of the transducers can be excited to customize therapy.—RCW

5,535,747

43.80.Vj ULTRASONIC EQUIPMENT

Kageyoshi Katakura, assignor to Hitachi, Ltd.
16 July 1996 (Class 128/660.02); filed in Japan 4 March 1994

This equipment estimates the propagation time of a pulse between a number of measurement points using echoes from the measurement points. The ultrasonic speed is also estimated using the propagation time and the



distance between the measurement points. Additionally, the speed of flow and pressure are estimated, and from these the pressure at a different time is estimated.—RCW

5,540,097

43.80.Vj ULTRASONIC MICROSCOPE APPARATUS

Nahoko Hisata, assignor to Olympus Optical Company
30 July 1996 (Class 73/620); filed in Japan 25 April 1991

In this apparatus, the brightness and contrast of an image signal are adjusted by one controller that varies brightness using the echo signal and an operator setting, and another controller that varies contrast, also using the echo signal and an operator setting.—RCW

5,540,228

43.80.Vj SYSTEM AND METHOD FOR TOPOGRAPHIC MAPPING OF ULTRASONIC IMAGE DISPLAYS

Ming Li, assignor to Siemens Medical Systems, Incorporated
30 July 1996 (Class 128/660.07); filed 29 September 1995

The ultrasonic image of an interrogated region is represented by a sequence of intensity values that are stored in a memory. The values of intensity are displayed using a combination of a geometry-independent function of an ambient light source and a geometry-dependent function of the values representing the original image. The geometry-dependent function weights the input intensity values, and is itself a function of the input intensity values. Noise-floor filtering, lowpass filtering, and surface normal vector averaging of the input intensity are included in the weighting.—RCW

5,540,229

43.80.Vj SYSTEM AND METHOD FOR VIEWING THREE-DIMENSIONAL ECHOGRAPHIC DATA

Antoine Collet-Billon and Raoul Mallart, assignors to U.S. Philips Corporation
30 July 1996 (Class 128/660.070); filed in France 29 September 1993

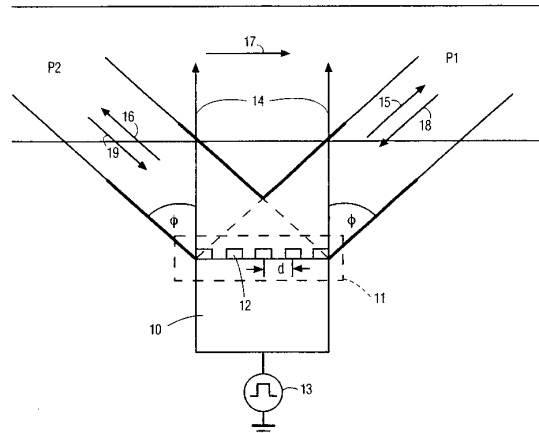
An ultrasonic imaging instrument equipped with a probe that senses position is used with a workstation that includes memory for the ultrasonic data. A figure depicting the subject is used to facilitate viewing the data. The plane to be imaged is specified on the figure.—RCW

5,540,230

43.80.Vj DIFFRACTING DOPPLER-TRANSDUCER

David Vilkomerson, assignor to Echocath, Incorporated
30 July 1996 (Class 188/662.04); filed 24 October 1994

This transducer includes a diffracting structure that permits emission and reception of ultrasonic beams at a number of different angles. The angles are used to determine the velocity of flow through a lumen and also



the flow volume. A determination of the lumen diameter is employed in the calculations.—RCW

5,544,654

43.80.Vj VOICE CONTROL OF A MEDICAL ULTRASOUND SCANNING MACHINE

Sean Murphy and Daniel E. Need, assignors to Acuson Corporation
13 August 1996 (Class 128/660.07); filed 6 June 1995

The control of an ultrasonic imaging instrument is accomplished by a structured vocabulary in which subsets of a large vocabulary are used for speech recognition. A two-way communication link between the voice control system and the imaging apparatus permits the state of the apparatus to be monitored so that subsets of the total vocabulary may be made active. From knowledge of the ultrasonic imager status, commands may be implemented for streamlined control of the imager.—RCW

5,544,655

43.80.Vj ULTRASONIC MULTILINE BEAMFORMING WITH INTERLEAVED SAMPLING

Ronald E. Daigle, assignor to Atlantis Diagnostics International
13 August 1996 (Class 128/661.010); filed 16 September 1994

Echo signals received by each element of a transducer array are sampled using two or more interleaved sampling signal sequences. Each sequence is timed to begin at the arrival of echos from a unique spatial line. The interleaved signal samples from each transducer element are separated according to the sampling signal sequence, and the separated sequences are individually processed to form corresponding spatially separate ultrasonic beams.—RCW

5,544,656

43.80.Vj METHOD AND APPARATUS FOR MYOCARDIAL WALL MEASUREMENT

Koullis F. Pitsillides and John C. Longhurst, assignors to the University of California
13 August 1996 (Class 128/661.04); filed 2 December 1994

A closed-loop, single-crystal ultrasonic system is used to identify and track the myocardial muscle–blood interface throughout the cardiac cycle. During steady state, a reference volume is adjusted to remain within the myocardium while half of an interface volume is in the myocardium and half is in the ventricular chamber. Echos of pulsed ultrasound are processed to extract a Doppler signal. The Doppler decoded signal is sampled at time intervals corresponding to the reference and interface sample volumes, and an amplitude-locked-loop circuit continuously adjusts the position of the sampled volumes so that they track the muscle–blood interface.—RCW

5,544,658

43.80.Vj DOPPLER ULTRASOUND VELOCITY ESTIMATION

Jin H. Kim and Dong-Chyan Liu, assignors to Siemens Medical Systems, Incorporated
13 August 1996 (Class 128/661.09); filed 18 September 1995

This approach to estimate spatial-temporal Doppler velocity is designed to extend the detectable range of Doppler frequencies beyond the Nyquist limit. The approach uses multiple velocity estimation functions. One function is the conventional function typically bound by the Nyquist limit. For additional functions, the $(n-1)$ th echo is spatially shifted with respect to the n th echo. A power function is then calculated for each velocity estimation function. The estimation function having the highest power is selected for use in deriving the Doppler shift for a given sample point.—RCW

5,544,659

43.80.Vj ULTRASONIC DOPPLER IMAGER HAVING A REDUCED HARDWARE ADAPTIVE TISSUE REJECTION FILTER ARRANGEMENT

Zoran B. Banjanin, assignor to Siemens Medical Systems, Incorporated
13 August 1996 (Class 128/661.09); filed 29 December 1994

Tissue signals are suppressed using a rejection filter in which the attenuation characteristic is controlled by an estimate of at least one spectral characteristic of tissue echoes to be removed. In one mode of operation, echo signals are routed to the Doppler processor to form estimates of tissue motion. In another mode of operation, echo signals are fed to the Doppler processor after passing through the rejection filter.—RCW

5,546,807

43.80.Vj HIGH SPEED VOLUMETRIC ULTRASOUND IMAGING SYSTEM

John T. Oxaal *et al.*, Durham, NC
20 August 1996 (Class 73/606); filed 2 December 1994

This system simultaneously displays multiple slices of an object in real time to enable study of relationships between parts of the object. The slices are selectable. The selection may be comprised of *b*-scans or *c*-scans or inclined scans (*i*-scans) in which the plane has a selected orientation. Multiple *i*-scans may be viewed in perspective to visualize three dimensions. Color and spectral Doppler is incorporated to provide additional information about dynamic properties.—RCW

5,549,110

43.80.Vj DEVICE FOR GENERATING SOUND IMPULSES FOR MEDICAL APPLICATIONS

Werner Krauss *et al.*, assignors to Richard Wolf GmbH
27 August 1996 (Class 128/661.01); filed in Germany 11 March 1993

This device is made for the generation of shock waves, and the emitted energy may be particularly high by using a bias potential and an excitation pulse with a polarity opposite that of the bias.—RCW

5,549,111

43.80.Vj METHOD AND APPARATUS FOR ADJUSTABLE FREQUENCY SCANNING IN ULTRASOUND IMAGING

J. Nelson Wright *et al.*, assignors to Acuson Corporation
27 August 1996 (Class 128/742); filed 2 May 1995

In this method and apparatus, the carrier frequency of the imaging pulse is higher at the center of the field of view than at the edges. Modulating appropriately delayed programmable initial waveform samples with a programmable carrier frequency accomplishes this on transmit. The reduction of center frequency with steering angle mitigates grating lobe artifacts and preserves signal energy, because the modulation translates the signal in frequency without substantial modification of the pulse shape. Correction for systematic scan-line-to-scan-line phase variations is accomplished by post-beam-formation predetection remodulation, in order to insure scan-line-to-scan-line phase coherence for subsequent coherence processing across scan lines, or for coherent image formation using the phase and amplitude information from multiple beams.—RCW

5,551,433

43.80.Vj METHOD AND APPARATUS FOR A GEOMETRIC ABERRATION TRANSFORM IN AN ADAPTIVE FOCUSING ULTRASOUND BEAMFORMER SYSTEM

J. Nelson Wright *et al.*, assignors to Acuson Corporation
3 September 1996 (Class 128/660.07); filed 5 August 1994

In this method and apparatus, aberration correction values that correspond to a region of aberration in an imaged subject are determined at a particular transmit focal depth. These values are stored in a geometric aberration transform (GAG™) table. The table is used to obtain aberration correction values at other focal depths for focus correction on both transmission and reception.—RCW

5,555,886

43.80.Vj APPARATUS AND METHOD FOR DETECTING BLOOD VESSEL SIZE AND DIRECTION FOR DOPPLER FLOW MEASUREMENT SYSTEM

Lee Weng and William H. Phillips, assignors to Siemens Medical Systems, Incorporated
17 September 1996 (Class 128/661.10); filed 28 September 1995

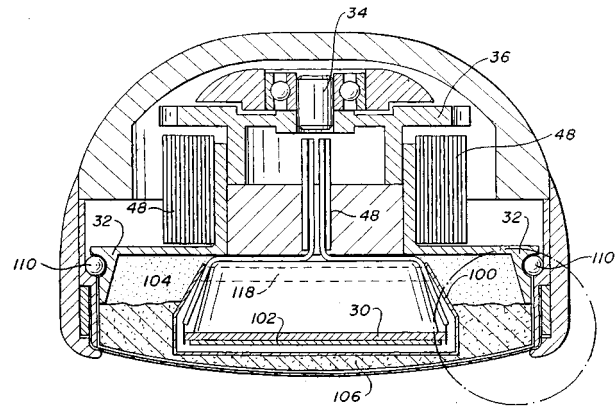
The volumetric flow rate in a displayed blood vessel is determined by estimating the vessel size and direction and using Doppler signals from fluid flowing through a region of interest. The vessel is selected on the display by an operator. A search is then performed in a number of directions that extend radially outward from within the vessel for an upper and lower set of edgepoints that correspond to inner wall surfaces of the vessel. Once upper and lower sets of edgepoints are identified, curves are fitted through the respective sets of points to locate inner wall surfaces. Next, the difference in angle between the curves and the direction of the beam continued with Doppler-shifted signals is determined. Finally, a volumetric flow rate is calculated.—RCW

5,555,887

43.80.Vj ULTRASONIC TRANSDUCER PROBE WITH HEAT DISSIPATING LENS

John D. Fraser *et al.*, assignors to Advanced Technology Laboratories, Incorporated
17 September 1996 (Class 128/663.01); filed 9 January 1995

Heat dissipation is accomplished in this transesophageal echocardiographic probe by an aluminum sheet **100** embedded in the acoustic lens **106**



on the face of the transducer **30**. Heat that develops in the lens is conducted away from the probe surface that contacts the patient to a heatsink **118** behind the transducer. A mylar sheet with an aluminized inner surface covers the front of the transducer to serve a like purpose.—RCW

5,556,372

43.80.Vj APPARATUS FOR ULTRASONIC BONE TREATMENT

Roger J. Talish *et al.*, assignors to Exogen, Incorporated
17 September 1996 (Class 601/2); filed 15 February 1995

This apparatus contains an ergonomically constructed ultrasonic transducer head module fabricated with a conductive plastic material. A portable operating unit is constructed to fit in a pouch worn by the patient. The transducer head is positioned adjacent to the area of injury and excited for a predetermined period of time. To ensure the transducer module is properly positioned and in compliance with the treatment protocol, a safety interlock is provided.—RCW

5,557,047

43.80.Vj METHOD AND APPARATUS FOR ULTRASONIC WAVE MEASUREMENT

Makoto Koide, assignor to Fuji Ultrasonic Engineering Company
17 September 1996 (Class 73/597); filed in Japan 10 February 1993

In this method and pparatus, the period of a transmitted ultrasonic wave is controlled to equal the time interval between like-order echoes produced in a test chamber or sample of known dimension.—RCW

Some hot topics in animal bioacoustics

Whitlow W. L. Au

Marine Mammal Research Program, Hawaii Institute of Marine Biology, P.O. Box 1106, Kailua, Hawaii 96734

(Received 16 November 1996; accepted for publication 19 February 1997)

This paper is derived from a "Hot Topics in Animal Bioacoustics" presentation at the 130th meeting of the Acoustical Society of America in St. Louis, Missouri. Six bioacoustics studies on a wide variety of species are discussed. Two of the studies are concerned with insects, the parasitoid fly, and cotton bollworms. The remaining bioacoustics studies are on aquatic animals including the West Indian manatee, elephant seals, and dolphins. © 1997 Acoustical Society of America. [S0001-4966(97)07005-7]

PACS numbers: 43.10.Ln, 43.80.Ka, 43.80.Lb [DWM]

INTRODUCTION

The field of animal bioacoustics is diverse, involving a wide variety of species varying from insects and birds to terrestrial and aquatic animals. Interest in this area of acoustics has been growing, and the Animal Bioacoustics Special Technical Group will become a full committee of the Acoustical Society of America following the fall 1996 meeting in Honolulu. This paper is derived from a "Hot Topics in Animal Bioacoustics" presentation at the 130th meeting of the ASA in St. Louis, Missouri, in which interesting research being performed with different species was highlighted. Results of some of the studies that will be discussed have not been published so they can be discussed only in general terms without compromising the publication plans of the various investigators. The objective of this paper is to present the reader with a general flavor of the type of animal bioacoustics research that is currently being conducted. The seven specific projects highlighted deal with subjects as varied as insects and mammals of aquatic and marine origin.

I. DIRECTIONAL HEARING IN THE PARASITOID FLY

Humans and most large animals generally localize sounds by using two mechanisms: time of arrival difference of sounds arriving at the two ears (which can also be considered as a phase difference cue), and the difference in the intensity of sounds arriving at the two ears caused by diffraction of sounds around the head and body. However, the ears of insects are often fractions of a wavelength apart, making these two cues relatively ineffective. Yet insects and other small animals can and do localize sounds using very different mechanisms such as pressure gradient receivers in some crickets (Michelsen, 1992).

The female Parasitoid fly, *Ormia ochracea* must find and deposit her parasitic larvae on a live field cricket. It locates the host cricket at night, apparently using acoustics cues. However, the fly's ears are not physically separated but are contained within a common air-filled chamber, only about 450 to 520 μm apart (Miles *et al.*, 1995). The small separation distance between tympanum makes the encoding of time difference and intensity difference cues impractical. A group from Cornell University (Robert *et al.*, 1992, 1994) has been studying the hearing processes of the parasitoid fly

and has discovered a new mechanism of sound localization involving a mechanical coupling between the tympanal membranes of the fly's hearing organ (Miles *et al.*, 1995).

Figure 1 is a schematic showing the external anatomy, location of the ears, and an electron micrograph of the tympanal membranes. The bulbae acoustica are attached to the membranes through a stiff cuticular rod that connects to the tympanum at the tympanal pit (TP). The cuticular structure connecting the tympanum pits to each other and to the pivot point is known as the intertympanal bridge. Each bulba acoustica contains about 70–75 auditory reception cells that are innervated by the auditory nerve.

Miles *et al.* (1995) measured the response of the tympanal membranes to sounds arriving at different angles by using a laser vibrometer, as depicted in Fig. 2. The laser light could be focused to an area of approximately 5 mm in diameter allowing Miles and his colleagues to measure the membrane vibration at a large number of locations. The incident sound was also measured using a small microphone, as shown in Fig. 2. The acoustic stimulus was a burst of band-limited white noise lasting 0.01 s and having a frequency range of 1 to 30 kHz. The signals from the laser vibrometer and the microphone were digitized and the transfer function of the different locations on the tympanum were measured.

Miles and colleagues found that the tympanal pit (TP) which was closest to the sound source (ipsilateral) responded with as much as 20 dB greater amplitude than the tympanal pit farthest from the sound source (contralateral). The tympanal structure rocks about the central pivot point with the two ears moving in nearly opposite phase. A mechanical model of the tympanal structure on a pivot is depicted in Fig. 3. When a sound first arrives at the ipsilateral tympanal pit, the structure deflects downward and this motion produces forces on the contralateral ear via the intertympanal bridge, which tend to cancel the effects of the external acoustic pressure. This mechanical action essentially produces an interaural intensity difference that would normally not exist since the ears are closer than 1/200th of a wavelength at 5 kHz.

II. ACOUSTIC DETECTION OF PINK BOLLWORMS

Pink bollworm infestation of cotton bolls is a serious problem for cotton growers. The present methods for exam-

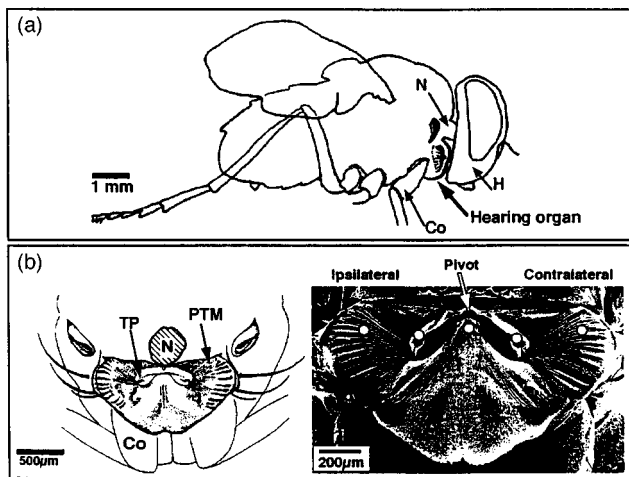


FIG. 1. (a) External anatomy of the *Ormia orchrcea* showing the location of the ears. (b) Frontal view of the ears with the head removed in left panel and a frontal scanning electron micrograph view of the ears. The prosteral tympanal membranes (PTM) show radial corrugations which converge upon the tympanal pit (TP) to which the internal sensory organ is attached (from Miles *et al.*, 1995).

ining samples of cotton bolls to determine if a field has been infested are laborious and prone to error. In order to solve this problem and devise a more expedient method of examining cotton bolls for bollworms, Robert Hickling and colleagues from the National Center for Physical Acoustics have turned to acoustic detection of larvae. Larvae constantly eat and move, generating sounds that can be detected by a sensitive acoustic sensor.

Boll cutting is the traditional method to determine if boll infestation has occurred. The cotton boll are pulled from the plant and are cut open and examined for pink bollworm lar-

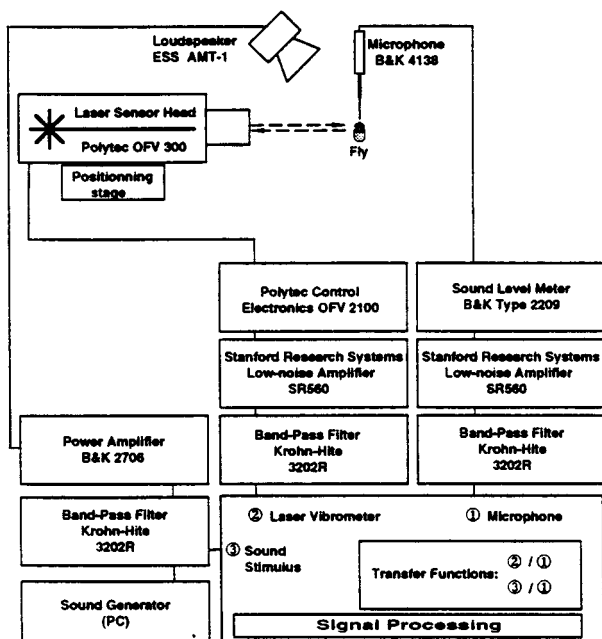


FIG. 2. Schematic of the experimental geometry (adapted from Miles *et al.*, 1995).

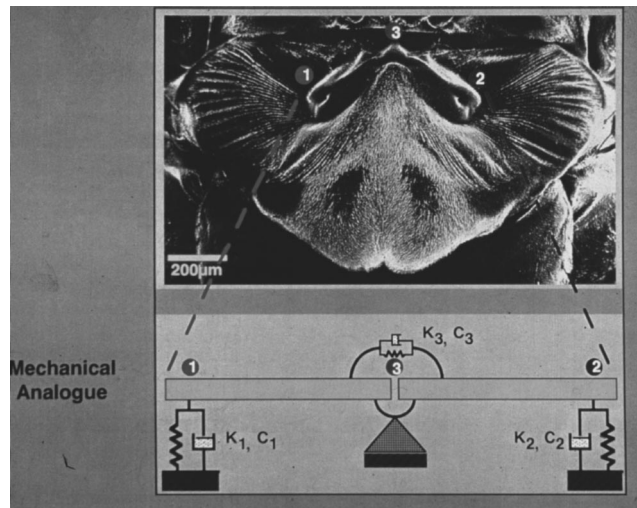


FIG. 3. Mechanical model of the parasitoid fly ears. The intertympanal bridge is assumed to be two rigid bars connected at the pivot through a coupling spring k_3 and dashpot c_3 (from Miles *et al.*, 1995).

vae. However, the larvae are usually very small, and so this technique is not only error prone but also time intensive. In a second method, cotton bolls are stored and left in boll boxes at the appropriate temperature and daylength regime for a time. Larvae within the bolls finish feeding and cut out of the boll. The number of damaged bolls and larvae/pupae produced are then counted. However, this technique requires up to 2 weeks and therefore has limited value.

The prototype multiple acoustic sensor developed by Hickling *et al.* (1994) to detect the sounds of pink bollworm larvae eating and moving in cotton bolls has four major components: the sound isolation box, sensors, amplification and filtering components, and power supply. The individual sensors are based on low cost fetal monitors consisting of adapted electret microphones connected to stethoscope headphones that were developed at the National Center of Physical Acoustics. A photograph of the prototype multiple acoustic sensor consisting of 48 sensors is shown in Fig. 4. The sensor signals are amplified and bandpass-filtered so that only a narrow band of low-frequency sound is detected.

Bolls from a Pima cotton field in Coolidge, Arizona were used in a test of the acoustic detection system. Three hundred bolls were cut using standard procedures, and were put into boll boxes and held in an insectary. Three hundred other bolls were used with the multiple acoustic sensor and then were carefully cut to verify sensing results. Binocular microscopes were used in several cases to find "suspected" larvae for the bolls tested with the acoustic unit. The bolls were warmed to a temperature of 38 °C prior to sensing. This increased audible larval activity thereby increasing the probability of sensing the larvae.

Results of the test are shown in Fig. 5. The acoustic sensor found as many or more larvae than did the other methods. Furthermore, acoustic sensing took less time and labor. As with the standard boll cutting and storing techniques, the acoustic sensor unit was not error free. Cases in which a larva was heard, but could not be found (false positive) occurred as shown in Fig. 5, along with cases in which the sensors did not detect larvae but one was found upon



FIG. 4. Photograph of the multiple acoustic sensor system (courtesy of R. Hickling).

cutting (false negative). The results indicate that use of a multiple acoustic sensor may be a good way to examine infestation of bollworms.

III. HEARING IN MANATEES

Manatees were once very plentiful in the warm waters of Florida but were placed on the endangered species list in 1973. A variety of causes could account for the decrease in population: destruction of habitat, loss of sea grasses, and collisions with motorboats. Boating accidents involving manatees have been particularly disturbing to Florida officials as the number of boats using Florida's waterways continues to increase. It seems strange that manatees do not

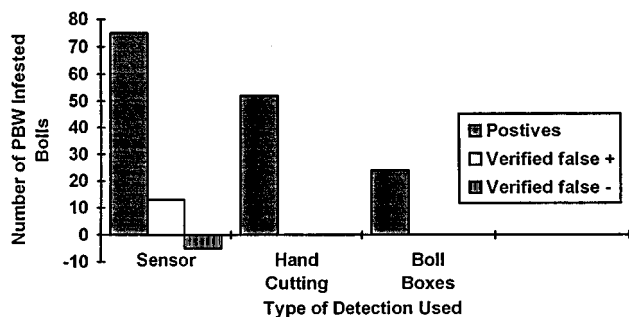


FIG. 5. Results of test one using Pima cotton (from Hickling *et al.*, 1994).

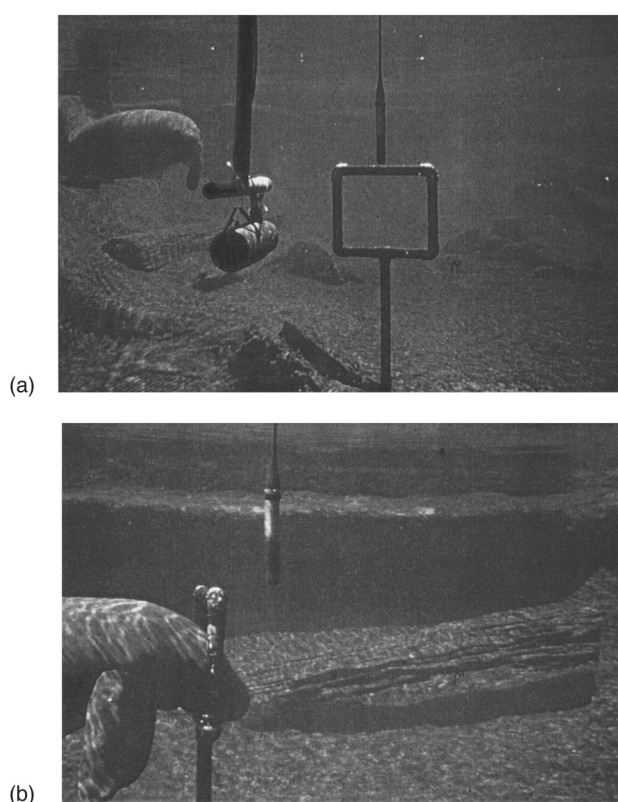


FIG. 6. (a) Picture of the experimental apparatus showing the J-9 sound projector, H-56 measuring hydrophone, a hoop station and a manatee approaching the hoop. (b) Manatee in the hoop station awaiting an acoustic signal (courtesy of Edmund Gerstein).

detect the approach of on-coming boats and avoid collision by swimming away. Lack of knowledge on the hearing capabilities of manatees has made it difficult to solve the boat collision problem without placing heavy constraints on the use of Florida's waterways by boaters.

A team of scientists led by Edmund Gerstein has been performing auditory experiments with the West Indian manatee (*Trichechus manatus*) in order to address the boat collision problem. They have performed experiments to measure the underwater hearing sensitivity as a function of frequency (audiogram) for two 7-year-old captive born males (Gerstein *et al.*, 1993), and their masked hearing threshold and sound localization capabilities (Gerstein *et al.*, 1995). The hearing studies are being conducted at the Lowry Park Zoo in Tampa, Florida.

The hearing sensitivity of the manatees was determined by a controlled psychoacoustic experiment in which they were trained to place their head into a hoop facing a J-9 transducer. The J-9 sound projector, H-56 measuring hydrophone, a hoop station, and a manatee approaching the hoop are shown in Fig. 6(a); and the manatee in the hoop station is depicted in Fig. 6(b). A pure-tone signal was then played to the animal. If it could hear the tone, the animal would respond by touching the signal-present paddle, and if a signal could not be heard, the animal would respond by touching the signal-absent paddle. The signal level was systematically decreased after correct detection trials and increased after the animal missed or failed to respond to a sound. Thus, the

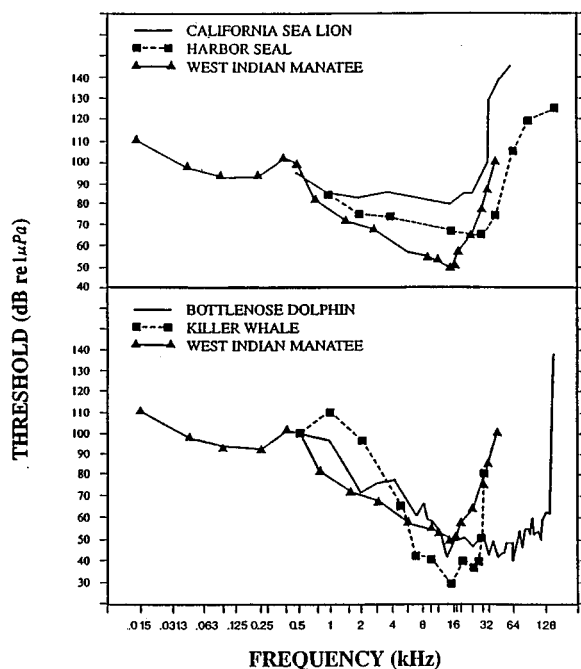


FIG. 7. (top panel) Audiogram of a manatee compared with a California sea lion and a harbor seal. (bottom panel) Manatee audiogram compared with that of a bottlenose dolphin and a killer whale (from Gerstein *et al.*, 1997).

signal level was varied in an up-down staircase fashion depending on the response of the subject.

The audiogram of one of the subjects (Gerstein *et al.*, 1997) is shown in Fig. 7 along with the audiograms for a California sea lion and for a harbor seal in the top panel; and for a killer whale and a bottlenose dolphin in the bottom panel. The manatee audiogram has a typical u shape in the hearing range from 500 Hz to 38 kHz. The data also indicated that a manatee can hear low-frequency sounds (below 16 kHz) as well or better than other marine mammals.

Manatees are agile enough to avoid oncoming boats (Gerstein, 1994, 1995), if they are aware of impending collision, but yet they do not. However, the low-frequency nature of boat noise may preclude detection at sufficient range because of elevated hearing sensitivity at low frequencies (below 1 kHz). It is also possible that a manatee may hear an oncoming boat but not be able to localize the direction of the sound. Finally, their hearing in noise may also not be good. Gerstein and colleagues are continuing to study the hearing of manatees by conducting a masked hearing threshold study and a sound localization study.

IV. ACOUSTIC ENVIRONMENT OF DIVING ELEPHANT SEALS

Many marine animals rely on acoustics to capture prey, avoid predators, reproduce, and navigate. Yet the ocean is a very noisy environment, especially at low frequencies. Buney Le Boeuf and his colleagues at the University of California at Santa Cruz have spent many years studying the natural behavior and physiology of the northern elephant seal on the Año Nuevo rookery off the coast of California. They

have recently expanded their research to consider the effects of low-frequency sounds, including the ATOC signal on these animals (Fletcher *et al.*, 1996).

Special electroacoustics instrument packages have been developed that can be attached to elephant seals to study the acoustic environment of these animals when they are out at sea. Adult female elephant seals spend 83% of the year at sea, and 90% of the time they are submerged (Le Boeuf *et al.*, 1988). They are deep divers, averaging 500-m depth on each dive with maximum dives to 1500-m depth. They are one of few marine mammals that can dive down to the axis of the deep sound channel in temperate waters. Therefore, the kinds of sounds that these animals encounter, especially when they are in the deep sound channel are of much interest.

The objectives of Le Boeuf's project are to determine (a) if diving seals make active sounds, (b) the frequencies and levels of sounds diving seals encounter in their environment, and (c) the best package design and attachment process for minimizing flow noise. Instrument packages were attached to juvenile seals, on their return to Año Nuevo Island from Monterey Bay. The first model of the attachment package contained a Sony digital audio tape recorder with a hydrophone, along with time depth recorder and swim speed detectors and the appropriate datalogger. Later models use a Tattletale data logger which also performs the analog-to-digital conversion of the incoming acoustic signals and stores the data on a hard disk. A picture of the latest electroacoustic package is shown in Fig. 8. The hydrophone is mounted on the back of the package to minimize the effects of flow noise. The top panel of Fig. 9 shows a seal being released from a boat with an electroacoustics package attached to its back. The bottom panel shows the seal on the beach at Año Nuevo Island.

An example of the acoustic signal received by a seal is shown in Fig. 10, with the depth of dive shown above the color sonogram. Most of the received signals had frequencies in the range of 20–200 Hz. Snapping shrimp, cetacean sounds, boat noise, and seal swim strokes and heart beats are clearly audible in some of the data. Flow noise correlated with swim speed, suggests that optimal time for acoustic

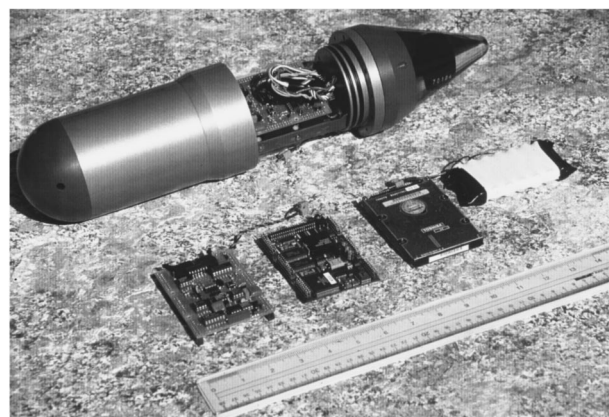


FIG. 8. Picture of the electroacoustic package showing the hydrophone located on the back of the package to minimize flow noise (courtesy of W. Burgess).

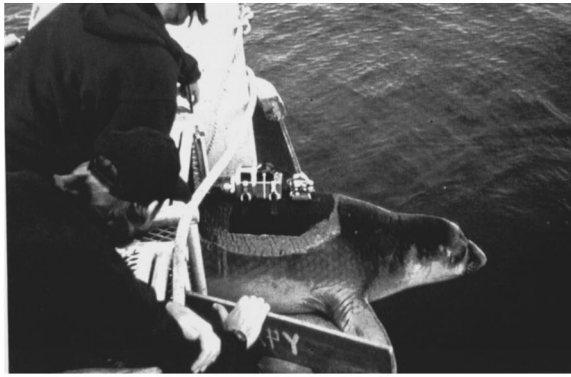


FIG. 9. The top panel shows a seal with an instrumentation package being released at sea and the bottom panel shows a seal on the beach at Año Nuevo Island (courtesy of B. LeBeauf).

sampling would be when the seals are swimming slowly. Results of several deployments have indicated that it is also feasible to obtain long-term, reliable, quantitative, and non-invasive cardiac monitoring of elephant seals and other ma-

rine mammals (Burgess *et al.*, 1996). This capability has been a important bonus to the project. Overall, this project demonstrates that it is feasible to monitor the acoustic environment of free swimming seals, and, that such electroacoustics packages could provide important data on the seal's environment.

V. AUDITORY-EVOKED POTENTIAL FROM TRAINED DOLPHINS

Auditory-evoked potentials on dolphins and on small whales have been measured with the animals housed in a small container where there is a considerable amount of acoustic reflection from the container walls (Ridgway *et al.*, 1981; Supin and Popov, 1985). The length of these containers is generally just larger than the animal's length making it difficult to position a transducer directly on the beam axis of the animal's auditory reception system. Furthermore, brain-stem-evoked potentials have been measured with either a click or short tone-burst stimuli. With the click signal, it is impossible to specify the amplitude that would relate to a threshold obtained in a behavioral study in which the rms sound-pressure level at threshold is used. A click signal is also broadband so that it is difficult to determine what part of the spectrum the animal's auditory system is responding to. Tone bursts also have some frequency and amplitude ambiguities associated with them.

Auditory-evoked potentials recorded from the human scalp have been shown to follow the periodicity of pure-tone sinusoidal signals for frequencies up to about 1 kHz (Moushegian *et al.*, 1983) and have been termed the frequency following response. Scalp-recorded responses have also been found to follow the envelope of continuous sinusoidally amplitude modulated stimuli in humans (Richards and Clark, 1984) and in gerbils (Dolphin and Mountain, 1993) using carrier frequencies considerably higher than the pure-tone phase locking capability of the frequency following re-

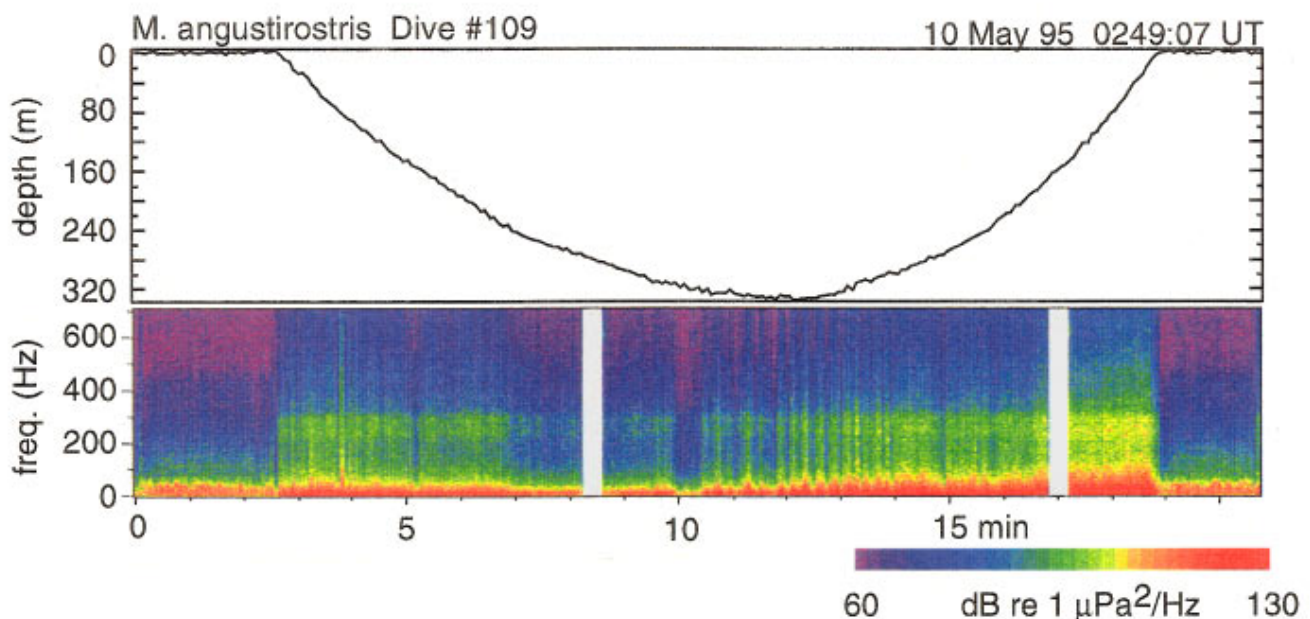


FIG. 10. Example of the acoustic signal received by the instrumentation package on a diving elephant seal (courtesy of W. Burgess).

sponse. Continuous sinusoidal amplitude-modulated (SAM) signals produced by amplitude modulating a high-frequency sinusoidal carrier with a low-frequency sine wave can be expressed as

$$s(t) = \sin(2\pi f_c t) [1 + m \sin(2\pi f_{\text{mod}} t)], \quad (1)$$

where t is time, m is the modulation depth, f_{mod} is the modulation frequency, and f_c is the carrier frequency. Such a signal will have energy only at f_c and $f_c \pm f_{\text{mod}}$. However when such a signal is passed through a nonlinear detection system, such as a mammalian auditory system, the signal is demodulated and energy at f_{mod} will appear. The evoked response to such a signal contains significant energy at the modulation, or envelope frequency of the stimulus (Dolphin and Mountain, 1993). Hence this phenomenon has been termed the envelope following response (EFR). With such a continuous signal, the rms value of the acoustic pressure can be readily measured and associated with the behavioral measure of hearing sensitivity measured at the carrier frequency.

Amplitude-modulated signals can also be produced by summing two sinusoidal or tonal signals of different frequencies so that

$$s(t) = \sin(2\pi f_1 t) + \sin(2\pi f_2 t), \quad (2)$$

where f_1 and f_2 are the frequencies of sinusoid 1 and 2, respectively. A two-tone signal (TT) contains energy only at the frequencies of the individual tones. However, the amplitude of the signal will be modulated with a period corresponding to the absolute difference in the frequency between the two tones, or $f_{2,1} = |f_2 - f_1|$.

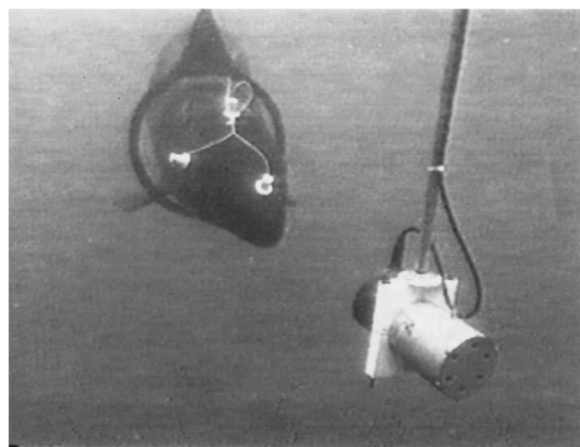


FIG. 11. A false killer whale in a hoop station with suction cup electrodes on its head, facing a J-9 sound projector.

Researchers from Boston University and the University of Hawaii have been performing auditory-evoked potential measurement, using continuous amplitude-modulated stimuli, with three species of trained dolphins (an Atlantic bottlenose dolphin, a Risso's dolphin, and a false killer whale) in an open water environment (Dolphin *et al.*, 1995). The animals are trained to wear suction cup electrodes and station themselves within a hoop which has its center at a depth of 1 m. The electrodes were silver discs (2.4-cm diameter) imbedded in custom designed latex suction cups. Evoked responses were recorded differentially from the scalp

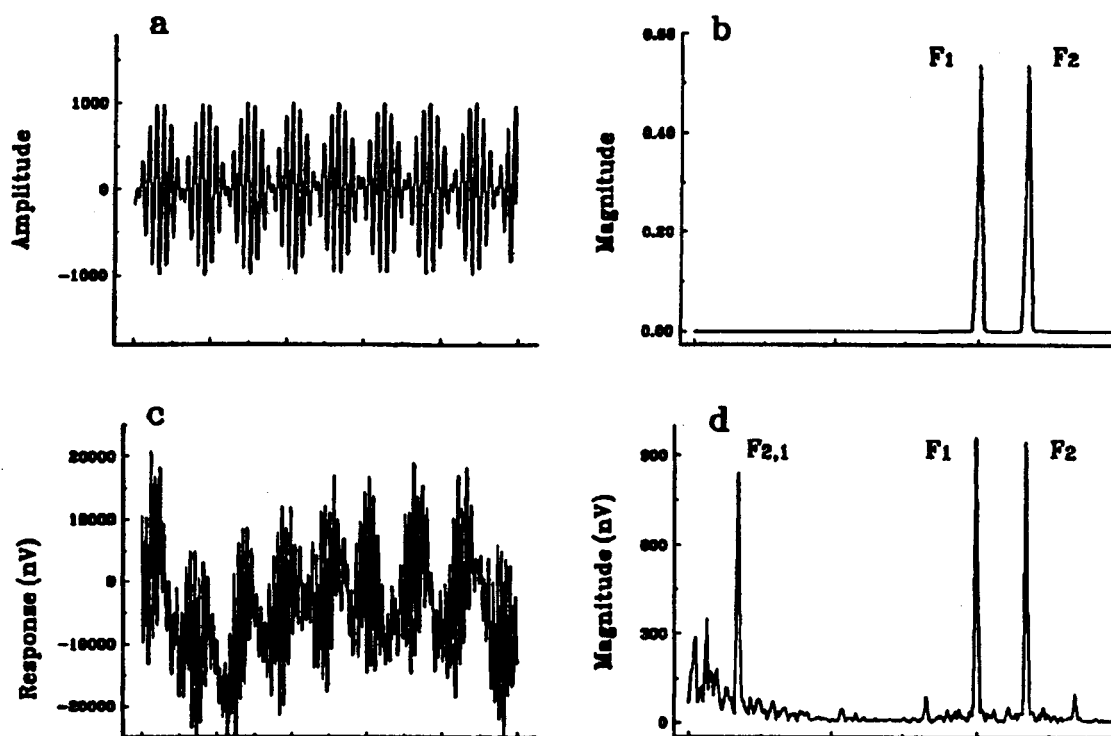


FIG. 12. (a) Example of the two-tone amplitude modulated acoustic stimulus in the time domain. (b) Frequency domain representation of the two-tone signal. (c) The evoked potential response in the time domain. (d) The evoked potential response in the frequency domain (from Dolphin *et al.*, 1995).

source between the parietal (noninverting) just posterior to the blowhole and the mastoid (inverting). A ground electrode was placed either on the melon or trunk region. A J-9 sound projector is located 1 m from the animal as shown in Fig. 11. The white suction cup electrodes on the animal's head can also be seen in the figure.

An example of the signal is depicted in the time domain [Fig. 12(a)], showing the amplitude modulation, and in the frequency domain [Fig. 12(b)], showing the energies at f_1 and f_2 . The evoked potential response averaged over 350 presentations of the stimulus in 164-ms blocks are shown in Fig. 12(c) and (d). The nonlinear property of the auditory system demodulates the amplitude modulated signal so that energy at the modulated or envelope frequency is present in the evoked response. The question of what would be the most ideal relationship between the frequencies of the carrier and modulated signals can be addressed by measuring the modulation rate transfer function. This was determined for dolphins by modulating the primary or carrier tone at frequencies between 18–4019 Hz and measuring the response magnitude and phase at f_{mod} or $f_{2,1}$. Examples of the modulation rate transfer function obtained from *Pseudorca* using both TT and SAM signals of different carrier and primary frequencies are shown in Fig. 13. From the figure one can see that there are fairly broad areas in which the magnitude of the evoked potential responses does not vary much with the modulation or difference frequency.

The hearing sensitivity of the three dolphin species to low-frequency sounds (below 1.6 kHz) has been determined behaviorally and by measuring evoked potentials. The results of both techniques agree very well, suggesting that determining hearing sensitivity with continuous amplitude modulated sound may be a good method to measure the hearing sensitivity of large whales that are beached or trapped in nets.

VI. HEARING AND ECHOLOCATION BY DOLPHINS AT DEPTH

Hearing and echolocation research on dolphins has been conducted in shallow waters with the animals usually within 2 m of the surface. However, dolphins can dive to depths greater than 300 m (Ridgway and Kanwisher, 1969) and the effects of a high-pressure environment on hearing and echolocation is not known. Sam Ridgway and colleagues at the Naval Command Control and Ocean Surveillance Center in San Diego have been studying the effects of depth on the hearing sensitivity and the echolocation signals of beluga whales, *Delphinapterus leucas* (Ridgway and Carder, 1995).

Two beluga whales were trained to dive in the open ocean and station on a test platform at depths of 6 (surface), 100, 200, and 300 m. The test platform with a whale approaching on a bite plate is shown in Fig. 14. On the test platform are video cameras to monitor the animal's behavior, a sound projecting transducer for audiometric test, a sonar target and a hydrophone to measure sonar signals. Ridgway and his team have been working in the deep waters off San Clemente Island. The whales are trained to "boat follow" an inflatable boat to travel to the test site about 1 mile off shore.

A trial consisted of a whale diving to the test platform suspended below the data collection boat and stationing on

the bite plate. Hearing thresholds were determined by presenting a sound stimulus at random intervals. Every time the whale heard a sound, it would produce a whistle, which was detected by another hydrophone and sent topside for analysis. The amplitude of the test tone was varied in a staircase fashion. Echolocation was tested by presenting a target in line with the whale's longitudinal axis, 1 m away. Each time a door hiding the target on the platform opened, the whale was trained to echolocate. The animal responded by producing a whistle if it perceived the target or remained silent when the target was absent. The echolocation signals were detected by a hydrophone and the information sent topside for recording.

Preliminary results indicate that the whales hearing sensitivity is not affected by the high pressure at depth. Their thresholds are as good or slightly better at depth than at the surface. However, preliminary results indicate that the ani-

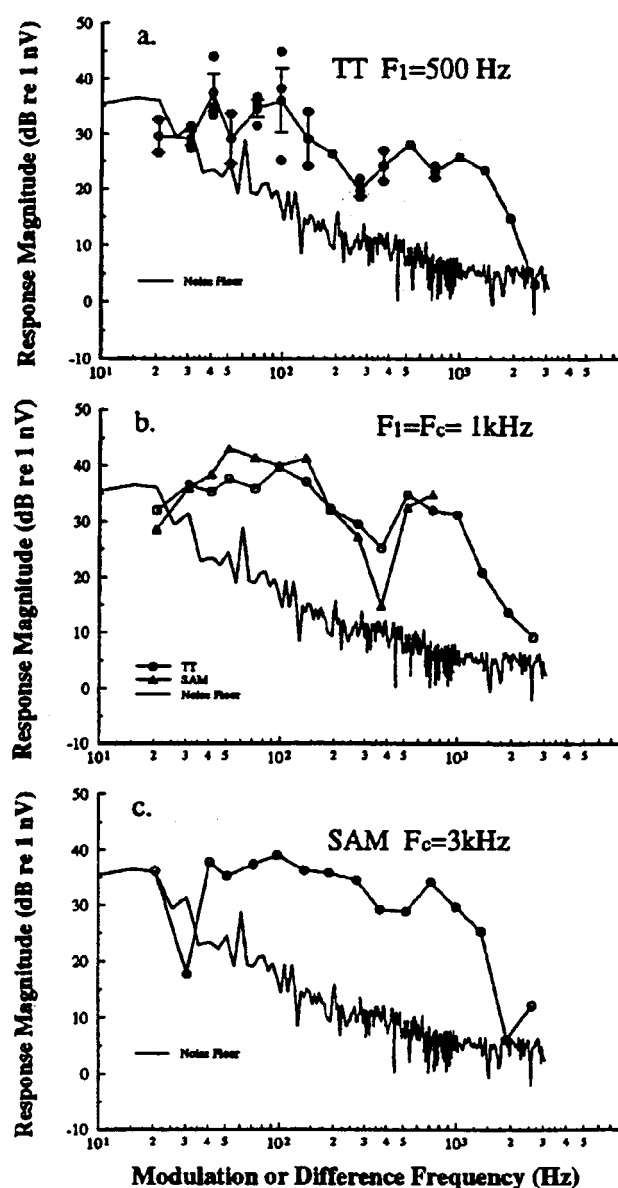


FIG. 13. Examples of the modulation rate transfer function obtained with a false killer whale (from Dolphin *et al.*, 1995).

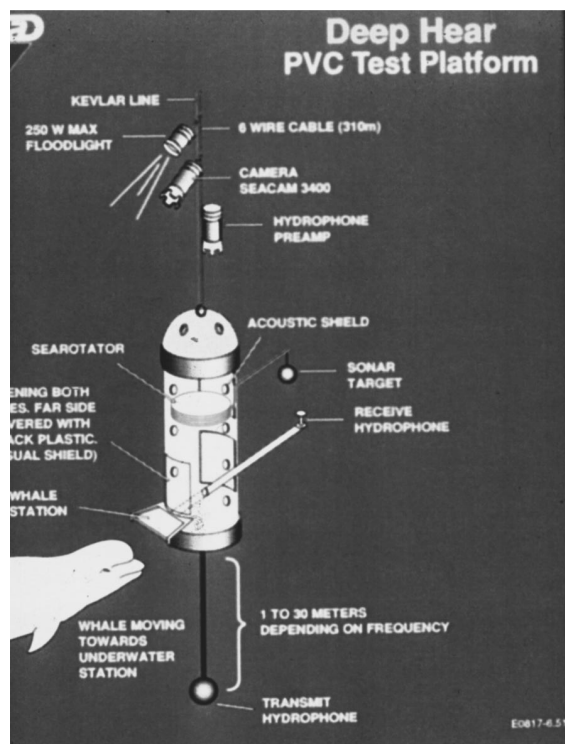


FIG. 14. The test platform used to assess the hearing sensitivity and echolocation signals of beluga whales at depth (courtesy of S. Ridgway).

mals' whistles at depth have larger high-frequency harmonics, as can be seen in Fig. 15. The lower frequency components of the whistles are present at depth and close to the surface, however, the whistles at depth have larger high frequency harmonics than at the surface. The increased pressure causes various air sacs and chambers in the nasal system to change in shape and volume, affecting their resonance frequencies. This increase of whistle frequency may have important implications concerning the mechanisms that produce whistles. Specific sound generation mechanisms in cetaceans have not yet been pin pointed. However many theories exist, based mainly on anatomical considerations.

VII. DISCUSSION

There has been an increasing interest on how man is affecting the environment in which we live, work, and recreate. Animals are an important part of this environment, and we are slowly becoming aware that our use of natural resources can have important consequences to many different life forms. Therefore, interest in animal bioacoustics has slowly been building over recent years. The U.S. military (Air Force, Army, and Navy) all have some type of bioacoustics programs to study the effects of military exercises, practices and operations on the environment of animals. The Air Force has an extensive program to determine the effects

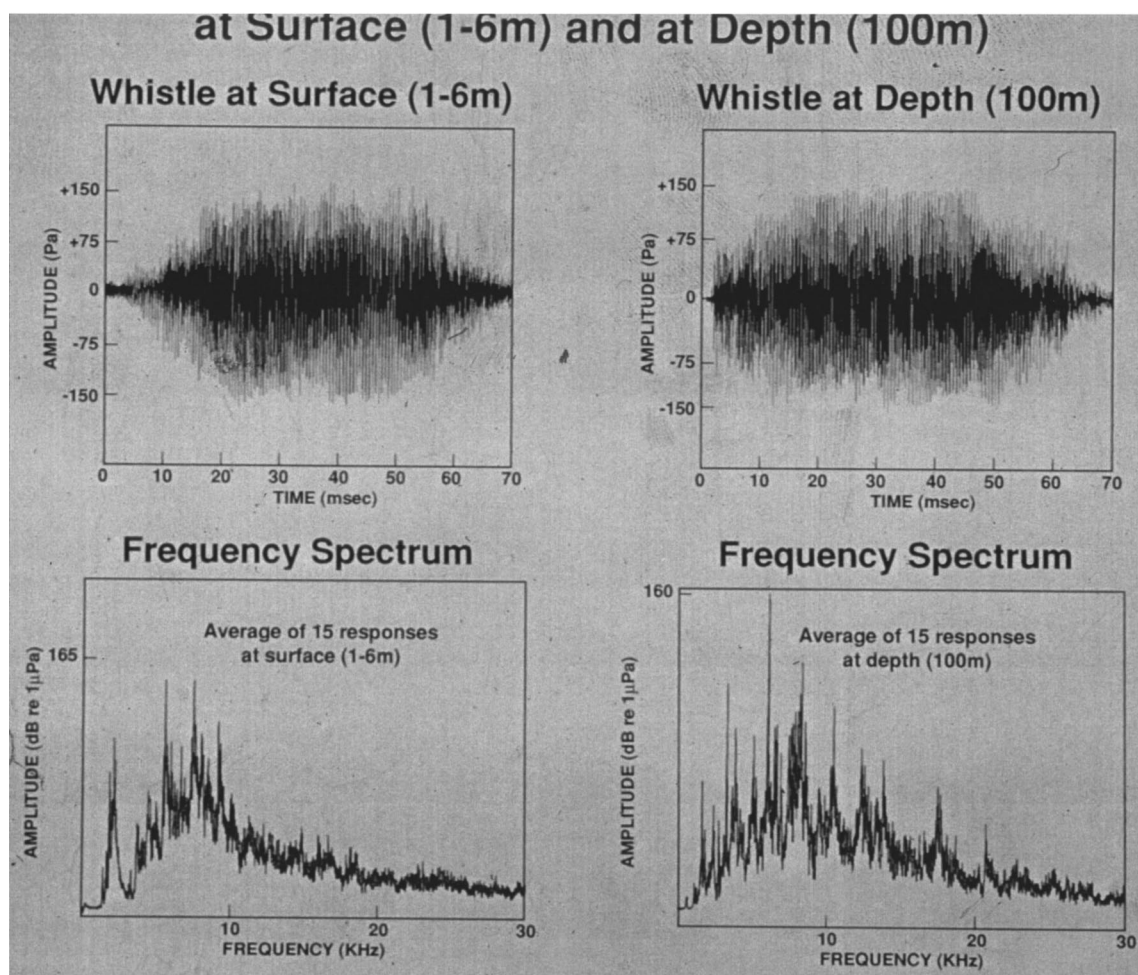


FIG. 15. Example of the beluga whale whistle at the surface and at depth (courtesy of S. Ridgway).

of aircraft noise on a wide variety of animals, from Musk Ox and Arctic foxes in the high northern latitudes to tortoise and reptiles in desert locations, and birds such as the Peregrine falcons in mountainous regions. The Army also has an extensive program studying their effects on the habitat of various endangered birds and animals. The Navy has a program to determine the effects of low-frequency underwater sounds on marine mammals. In these programs, various types of new, small, portable, and state-of-the-art devices are being developed to monitor location, physiological parameters, and environmental noise in these programs (Carter, 1996). In order to monitor and study the effects of man-made noise on different animal species, it is important to understand their auditory capabilities. Therefore, within the military's acoustic environmental assessment program, the basic hearing capabilities of different species are being determined. The military's programs are too extensive and broad to be discussed here.

Biologists are increasingly becoming aware of the usefulness of monitoring the vocalizations of animals as a good way to study them in their natural habitat with a minimum of disturbance. New electronic instrumentation that can be connected to laptop computers has been a boon to field researchers. The continual development of sophisticated signal processing and analysis software has made these tools accessible to many researchers who would not even have dreamed of using these techniques a decade ago. The maturing of satellite tracking and monitoring instrumentation has also contributed significantly in studying the life cycle and the use of acoustics by animals in the wild. The use of sophisticated laser techniques and other measurement techniques has increased the capabilities of many researchers to study the acoustic functions of different species in new and innovative ways. In short, we have entered into an exciting and interesting period for people interested in animal bioacoustics.

ACKNOWLEDGMENTS

The author acknowledges and publicly expresses his appreciation to several people who provided various written materials, 35-mm slides, and data used in this paper. These individuals are Dr. William Burgess of Stanford Research Institute, Dr. Edmund Gerstein of Florida Atlantic University, Dr. Robert Hickling of the National Center for Physical Acoustics, Dr. Burney LeBoeuf of the University of California at Santa Cruz, Dr. Sam Ridgway of the Naval Command, Control and Ocean Surveillance Center, and Dr. Daniel Robert of Cornell University. This is Hawaii Institute of Marine Biology Contribution 1021.

Burgess, W. C., Tyack, P. L., LeBoeuf, B. J., and Costa, D. P. (1996). "Acoustic measurement of cardiac function on northern elephant seals," *J. Acoust. Soc. Am.* **100**, 2709(A).

- Carter, M. (1996). "USAF monitors for recording aircraft noise received by animals," *J. Acoust. Soc. Am.* **99**, 2575(A).
- Dolphin, W. F., Au, W. W. L., Nachtigall, P. E., and Pawloski, J. (1995). "Modulation rate transfer functions to low-frequency carriers in three species of cetaceans," *J. Comp. Physiol. A* **177**, 235–245.
- Dolphin, W. F., and Mountain, D. C. (1993). "The envelope following response (EFR) in the Mongolian gerbil to sinusoidally amplitude-modulated signals in the presence of simultaneously gated pure tones," *J. Acoust. Soc. Am.* **94**, 3215–3226.
- Fletcher, S., Le Boeuf, B. J., Costa, D. P., Tyack, P. L., and Blackwell, S. B. (1996). "Onboard acoustic recording from diving northern elephant seals," *J. Acoust. Soc. Am.* **100**, 2531–2539.
- Gerstein, E. R., Gerstein, L. A., Forsythe, S. E., and Blue, J. E. (1993). "Underwater audiogram of a West Indian manatee (*Trichechus manatus*)," Proceedings, Tenth Bien. Conf. Biol. of Mar. Mamm., Galveston, Texas, p. 53(A).
- Gerstein, E. R. (1994). "Auditory assessment of the West Indian manatee *Trichechus manatus*, potential impacts of low frequency activities on manatee acoustic behavior and communication," Tech. Report DACW39-92R-0112 USAE WES, Vicksburg, MS.
- Gerstein, E. R. (1995). "Underwater audiogram of the West Indian manatee (*Trichechus manatus latirostris*)," Masters degree thesis, Florida Atlantic University, Boca Raton, Florida.
- Gerstein, E. R., Gerstein, L. A., Forsythe, S. E., and Blue, J. E. (1995). "Underwater hearing abilities of the West Indian manatee," Proceedings, Eleventh Bien. Conf. Biol. of Mar. Mamm., Orlando, Florida, p. 4(A).
- Gerstein, E. R., Gerstein, L. A., Forsythe, S. E., and Blue, J. E. (1997). "The underwater audiogram of the West Indian manatee (*Trichechus manatus*)" (in preparation).
- Hickling, R., Lee, P., Wei, W., Geyer, M., Pierce, D., Staten, R., and Henneberry, T. (1994). "Multiple Acoustic Sensor System for Detecting Pink Bollworm in Bolls," Beltwide Cotton Conference.
- Le Boeuf, B. J., Costa, D. P., Huntley, A. C., and Feldkamp, S. D. (1988). "Continuous, deep diving in female northern elephant seals, *Mirounga angustirostris*," *Can. J. Zool.* **66**, 446–458.
- Miles, R. N., Robert, D., and Hoy, R. R. (1995). "Mechanically coupled ears for directional hearing in the parasitoid fly *Ormia ochracea*," *J. Acoust. Soc. Am.* **98**, 3059–3070.
- Michelsen, A. (1992). "Hearing and sound communication in small animals: evolutionary adaptations to the laws of physics," in *The Evolutionary Biology of Hearing*, edited by D. B. Webster, R. R. Fay, and A. N. Popper (Springer-Verlag, New York), pp. 61–77.
- Moushegian, G., Rupert, A. L., and Stillman, R. D. (1983). "Scalp recorded early responses in man to frequencies in the speech range," *Electroencephalogr. Clin. Neurophysiol.* **35**, 665–667.
- Richards, R. W., and Clark, G. M. (1984). "Steady state evoked potentials to amplitude modulated tones," in *Evoked Potentials II*, edited by R. H. Nodar and C. Barber (Butterworth, Boston), pp. 163–168.
- Ridgway, S. H., Bullock, T. H., Carder, D. A., Seeley, R. L., Woods, D., and Galambos, R. (1981). "Auditory brainstem response in dolphin," *Proc. Natl. Acad. Sci. USA* **78**, 1943–1947.
- Ridgway, S. H., and Kanwisher, J. (1969). "Respiration and Deep Diving in the bottlenose porpoise," *Science* **166**, 1651–1654.
- Ridgway, S. H., and Carder, D. A. (1995). "Deep hearing and sonar studies of contintioned white whales *Delphinapterus leucas*," Proceedings, Eleventh Bien. Conf. Biol. of Mar. Mamm., Orlando, Florida, p. 4(A).
- Robert, D., Amoroso, J., and Hoy, R. R. (1992). "The evolutionary convergence of haring in a parasitoid fly and its cricket host," *Science* **258**, 1135–1137.
- Robert, D., Read, M. P., and Hoy, R. R. (1994). "The tympanal hearing organ of the parasitoid fly *ormia ochracea* (Diptera, Tachinidae, Ormiini), *Cell Tissue Res.* **275**, 63–78.
- Supin, A. Ya., and Popov, V. V. (1985). "Recovery cycles of the dolphin's brain stem responses to paired acoustic stimuli," *Dokl. Akad. Nauk SSSR* **283**, 740–743.

Radiation force on a spherical object in the field of a focused cylindrical transducer

Xucaai Chen^{a)} and Robert E. Apfel

Center for Ultrasonics and Sonics, Yale University, New Haven, Connecticut 06520-8286

(Received 9 June 1996; revised 2 November 1996; accepted 20 November 1996)

An exact solution of the radiation force on a spherical object, when positioned on the acoustic axis of a cylindrical transducer, is provided. The solution is valid for any type of sphere of any size. The radiation force function allows the calibration of high-frequency focused ultrasound fields from radiation force measurements and expands the utility of the elastic sphere radiometer developed by Dunn *et al.* [*Acustica* **38**, 58–61 (1977)]. Numerical results reveal an oscillatory behavior of the radiation force function for small spheres near the transducer surface and this behavior may present an opportunity for particle sorting based on the mechanical properties of the particle and other types of manipulation. © 1997 Acoustical Society of America. [S0001-4966(97)03405-X]

PACS numbers: 43.20.Bi, 43.20.Fn, 43.25.Qp [JEG]

INTRODUCTION

Following the work by King in 1934,¹ a rich literature on the acoustic radiation force exists.^{2–14} Of particular interest is the elastic sphere radiometer developed by Dunn *et al.*¹⁰ which is widely used for the determination of acoustic intensity from the radiation force measurement in a plane-wave field. Recently, Chen and Apfel¹⁴ presented theory and experimental results, indicating that the technique by Dunn *et al.*¹⁰ can also be used for high-frequency focused fields. For the focused cylindrical transducer, however, only the radiation force function when the sphere is located at the geometrical focus was presented. In contrast, the theory by Wu and Du¹² allowed the sphere to be located anywhere on the transducer axis. They found that the radiation force on a small sphere to be zero at the focus. However, their theory was explicitly for a small compressible sphere ($ka \ll 1$).

An exact formulation of the radiation force function for a spherical object, when positioned on the acoustic axis of a cylindrical transducer, is provided. First, the incident wave is decomposed into a summation of outgoing spherical harmonics, with the origin of the spherical coordinate system located at the center of the sphere. Then, the scattered wave by the sphere and the radiation force on the sphere are presented. Numerical results are presented and compared with existing theory.

I. THE INCIDENT WAVE

Let ρ_0 and c_0 stand for the density and sound speed of the fluid medium in which the wave travels. The wave number in the fluid medium is given by $k = \omega/c_0$ where ω is the angular frequency. The geometry of the focused transducer is shown in Fig. 1. The transducer is specified by its element radius a_0 and radius of curvature r_0 . The half aperture angle, α_0 , is determined by $\alpha_0 = \arcsin(a_0/r_0)$. When the size of the

transducer is larger than the acoustic wavelength ($ka_0 > 1$), the velocity potential in the right half space can be calculated from the Rayleigh formulation, i.e.,

$\varphi_{\text{in}}(z; r, \theta, \phi; t)$

$$= \text{Re} \left[\frac{u_0 \exp(-i\omega t)}{2\pi} \int \int_S \frac{\exp(ikR)}{R} dS \right], \quad (1)$$

where u_0 is the normal velocity amplitude on the transducer surface, R is the distance from a point source on the surface of the transducer to the point of observation, and integration is over the whole surface of the transducer. Many approximate solutions of Eq. (1) have been provided, either in polar or spherical coordinates with an origin located at the center of the transducer. In order to use the theory on radiation force by Chen and Apfel,¹⁴ it is necessary to solve Eq. (1) in spherical coordinates where the origin is located at the center of the sphere as shown in Fig. 1, such that

$$\varphi_{\text{in}}(z; r, \theta, \phi; t) = \text{Re} \left[\varphi_0 \exp(-i\omega t) \sum_{m=0}^{m=\infty} A_m (2m+1) \times i^m j_m(kr) P_m(\cos \theta) \right], \quad (2)$$

where φ_0 is the characteristic amplitude of the potential field, $j_m(x)$ is the spherical Bessel function, $P_m(x)$ is the Legendre polynomial, and A_m is the coefficient of the m th spherical wave. Once φ_0 and A_m are determined, Eq. (2) gives a complete description of the radiation field. Defining

$$\varphi_0 = (iu_0/k) \exp(ikz), \quad (3)$$

Eq. (1) can be written as

$$\varphi_{\text{in}}(z; r, \theta, \phi; t) = \text{Re} [\varphi_0 \exp(-i\omega t) D_T(z; r, \theta, \phi)], \quad (4)$$

where

$$D_T(z; r, \theta, \phi) = \frac{k^2 \exp(-ikz)}{2\pi} \int \int_S \frac{\exp(ikR)}{ikR} dS \quad (5)$$

is the nondimensional radiation pattern of the transducer. In the definition given by Eq. (3) φ_0 is actually the velocity

^{a)}Current address: Department of Medicine, University of Rochester, Rochester, NY 14642.

is written this way, B_m is independent of the incident wave. It depends on the relative mechanical properties of the sphere and the host medium. The most general spherical object is an elastic sphere, which can be described by three parameters:

$$B_m = iD_m / (1 - iD_m), \quad D_m = \frac{j_m(ka)\Phi_m - ka \cdot j'_m(ka)}{n_m(ka)\Phi_m - ka \cdot n'_m(ka)}, \quad (16)$$

$$\Phi_m = \frac{\rho_0(k_s a)^2}{2\rho^*} \frac{1}{m^2 + m - 2 - (k_s a)^2/2} \frac{[(m^2 + m)j_m(k_s a)]/[j_m(k_s a) + \Lambda_m(k_s a)] + [k_c a \cdot j'_m(k_c a)]/[\Lambda_m(k_c a)]}{[(m^2 + m)\Lambda_m(k_s a)]/[j_m(k_s a) + 2 \cdot \Lambda_m(k_s a)] - [j_m(k_c a) + 2\Lambda_m(k_c a)]/[\Lambda_m(k_c a)]},$$

$$\Lambda_m(x) = \frac{j_m(x) - x \cdot j'_m(x)}{m^2 + m - 2 - (k_s a)^2/2}.$$

Other types of spheres, such as the fluid sphere and the hard sphere, can be treated as limiting cases of the elastic sphere.¹⁴

The acoustic radiation force on the sphere due to the potential field can be calculated from

$$\mathbf{F} = \oint_{A_{\text{fix}}} d\mathbf{A} \cdot \mathbf{S}, \quad (17)$$

where $\mathbf{S} = -\langle p\mathbf{I} + \rho_0\mathbf{v}\mathbf{v} \rangle$ is the Brillouin radiation tensor, $p = \rho_0\partial(\varphi_{\text{in}} + \varphi_{\text{sc}})/\partial t$ is the acoustic pressure in the fluid, $\mathbf{v} = -\nabla(\varphi_{\text{in}} + \varphi_{\text{sc}})$ is the particle velocity in the fluid, $\langle \cdot \rangle$ stands for time average, and the integration is over a fixed surface A_{fix} which completely encloses the scatterer.

Substituting Eq. (2) into Eq. (17), the radiation force on the sphere exerted by the incident wave is found to be along the acoustic axis and is calculated from

$$F_{\parallel}(z) = E_0(\pi a^2)Y_p(z), \quad (18)$$

where $E_0 = I_0/c_0$ is the characteristic energy density, I_0 is the characteristic intensity defined as

$$I_0 = \rho_0 c_0 k^2 |\varphi_0|^2/2, \quad (19)$$

and $Y_p(z)$ is the radiation force function given by¹⁴

$$Y_p(z) = -\frac{4}{(ka)^2} \sum_{m=0}^{\infty} (m+1) [\text{Re}(A_m(z)A_{m+1}^*(z)) \times (\alpha_m + \alpha_{m+1} + 2\alpha_m\alpha_{m+1} + 2\beta_m\beta_{m+1}) + \text{Im}(A_m(z)A_{m+1}^*(z))(-\beta_m + \beta_{m+1} + 2\alpha_m\beta_{m+1} - 2\beta_m\alpha_{m+1})], \quad (20)$$

where α_m and β_m are the real and imaginary parts of B_m , i.e., $\alpha_m = \text{Re}(B_m)$, $\beta_m = \text{Im}(B_m)$. Notice that E_0 and I_0 are not necessarily the time-averaged energy density and intensity at $r=0$ when the sphere is absent. The time-averaged energy density at $r=0$ is, by definition,

its density ρ^* , compressional wave speed c_c , and shear wave speed c_s . Then $k_c = \omega/c_c$ and $k_s = \omega/c_s$ are the compressional and shear wave numbers inside the sphere. Solution for B_m is provided by Faran and others¹⁷⁻¹⁹ as

$$E(z) = \left\langle \frac{\rho_0}{2} |\mathbf{v}|^2 + \frac{1}{2\rho_0 c_0^2} |p|^2 \right\rangle \Big|_{r=0} = \frac{|A_0(z)|^2 + |A_1(z)|^2}{2} E_0. \quad (21)$$

The time-averaged intensity at $r=0$, which is along the z axis, is

$$I(z) = \langle \text{Re}(p \cdot \nu_z) \rangle \Big|_{r=0} = \text{Re}[A_0(z)A_1^*(z)]I_0. \quad (22)$$

On the other hand, the pressure amplitude on the acoustic axis is

$$P(z) = |A_0(z)|p_0. \quad (23)$$

III. RESULTS AND DISCUSSION

We now present example calculations of the radiation force function on a sphere in the field of a focused transducer. For all calculations $kr_0=500$ and $\alpha_0=0.2$ will be used. This results in a modest pressure focusing factor of $G_p \approx 10$. Figure 2 shows the pressure amplitude as well as the time-averaged energy density along the acoustic axis for such a transducer. The pressure maximum is located at $z_{\text{max}}/r_0 \approx 0.90$. While the on-axis pressure amplitude reaches zero at a limited number of locations, the time-averaged energy density never does so. This is evident from Eq. (21), since $A_0(z)$ and $A_1(z)$ cannot vanish at the same location.

The radiation force function on a small type 4310 stainless-steel sphere $Y_p(z)$ is shown in Fig. 3. For a very small sphere, such as $ka=0.01$, shown in Fig. 3(A), the radiation force is proportional to the derivative of energy density along the z axis, $\partial E/\partial z$, as reported by Wu and Du.¹² Wu and Du also pointed out that the radiation force changes sign at the pressure maximum. This is also confirmed by Fig. 3(A). Notice that there is a zero-crossing near the pressure maximum. However, these theories fail even for modest val-

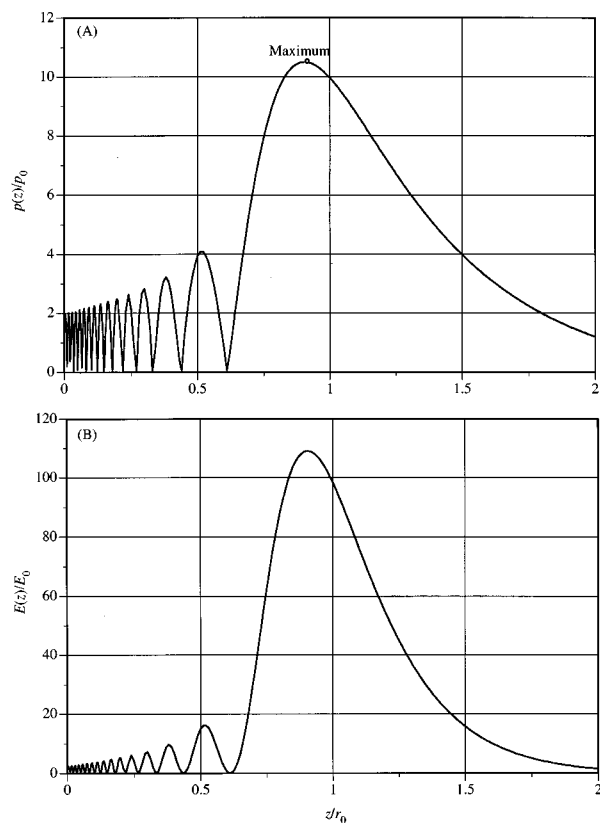


FIG. 2. The pressure amplitude (A) and the time-averaged energy density (B) along the axis of a focused transducer. Parameters used for computation: $kr_0 = 500$, $\alpha_0 = 0.02$.

ues of ka such as $ka = 0.1$, as shown in Fig. 3(B), where there is no zero-crossing near the pressure maximum.

For even larger ka values, the radiation force function behaves like the time-averaged energy density, rather than its axial derivative, as shown in Fig. 4. For $ka = 1$, the radiation force function oscillates around zero only near the transducer surface, as shown in Fig. 4(A). For $ka = 10$, the radiation force is always positive, as shown Fig. 4(B).

Figure 5 shows the radiation force function at the geometrical focus ($z = r_0$) and at pressure maximum ($z = z_{\max}$) versus the size parameter of the sphere ka . For the parameters used, $Y_p(z = r_0)$ crosses zero at $ka \approx 0.05$. The behavior of the force function at the pressure maximum is very similar to that at the geometric focus, except that the zero-crossing is too small to be observed on the scale used.

Notice that the radiation force functions shown above are normalized by the energy-density at the transducer surface. Since the energy density at the center of the sphere is related to the energy density at the transducer surface by Eq. (21) and its evaluation is straightforward, it is easy to combine the theory presented here with the elastic sphere radiometer technique developed by Dunn *et al.*¹⁰ for the calibration of focused transducer.

To demonstrate the effect of focusing on the radiation force further, Fig. 6 compares the radiation force function at the geometric focus and at the pressure maximum, normalized by the energy density at these locations, with the radiation force function on the same sphere in a plane traveling

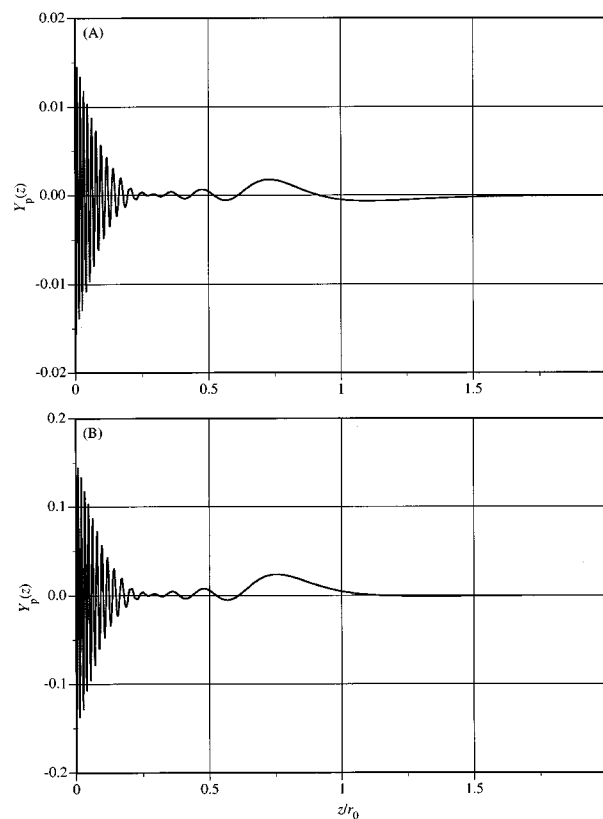


FIG. 3. Radiation force functions for a type 4310 stainless-steel sphere in water as a function of normalized position for the focused cylindrical transducer, for $ka = 0.01$ (A) and $ka = 0.1$ (B). Parameters used for computation: water $\rho_0 = 1000 \text{ kg/m}^3$, $c_0 = 1500 \text{ m/s}$; the elastic sphere $\rho^* = 7840 \text{ kg/m}^3$, $c_e = 5854 \text{ m/s}$, $c_s = 3150 \text{ m/s}$.

wave. For the particular transducer used, we have $kW_{-3 \text{ dB}} \approx 8$, where $W_{-3 \text{ dB}}$ is one half of the -3-dB beamwidth at focus. Figure 6 indicates that the focused beam theory should be used even when the size of the sphere is smaller than the -3-dB beamwidth.

Numerical calculation also indicates that when the sphere is located at the geometrical focus ($z = r_0$), the radiation force function given by Chen and Apfel¹⁴ is highly accurate, except for a very small sphere ($ka \ll 1$).

IV. CONCLUSION

An exact formulation for the radiation force on a spherical object, when positioned on the axis of a focused cylindrical transducer, is provided. This formulation utilizes the radiation pattern originally developed by Hasegawa *et al.*¹⁶ and the radiation force function for a sphere in an axisymmetric wave field presented by Chen and Apfel.¹⁴ The formulation is valid for any type of spheres (elastic, fluid, etc.) and the sphere can be of any size. For very small spheres ($ka \ll 1$), the current theory agrees with the theory presented by Wu and Du.¹² For relatively large spheres ($ka > 1$), the current theory confirms the approximation presented by Chen and Apfel.¹⁴ The radiation force function allows the calibration of high-frequency ultrasound fields from radiation force measurements and expands the utility of the elastic sphere radiometer developed by Dunn *et al.*¹⁰ Spheres as

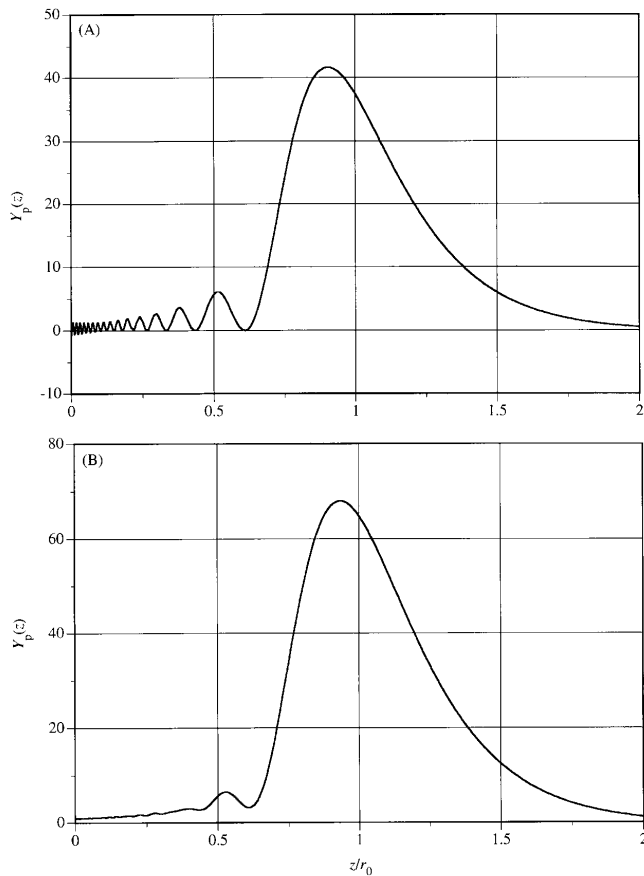


FIG. 4. Radiation force functions for a type 4310 stainless-steel sphere in water as a function of normalized position for the focused cylindrical transducer, for $ka=1$ (A) and $ka=10$ (B).

large as, or even larger than, the -3 -dB beamwidth can be used to calibrate the acoustic field produced by a high-frequency focused transducer.

Numerical results reveal an oscillatory behavior of the radiation force function for small spheres near the transducer surface, and this phenomenon may present an opportunity for particle sorting and other types of manipulation.

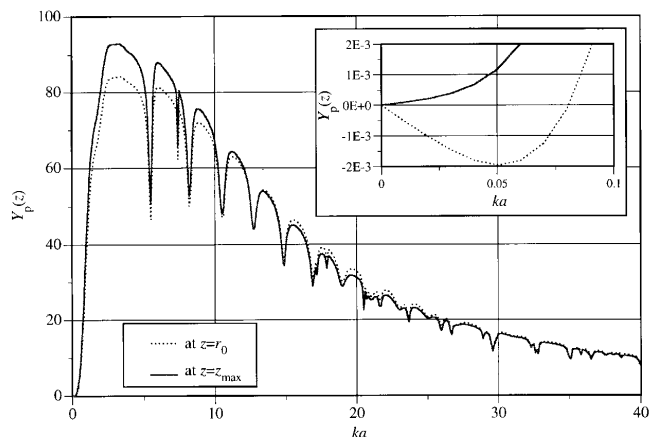


FIG. 5. Radiation force functions for a type 4310 stainless-steel sphere in water as a function of the size parameter ka at the geometrical focus and at the pressure maximum.

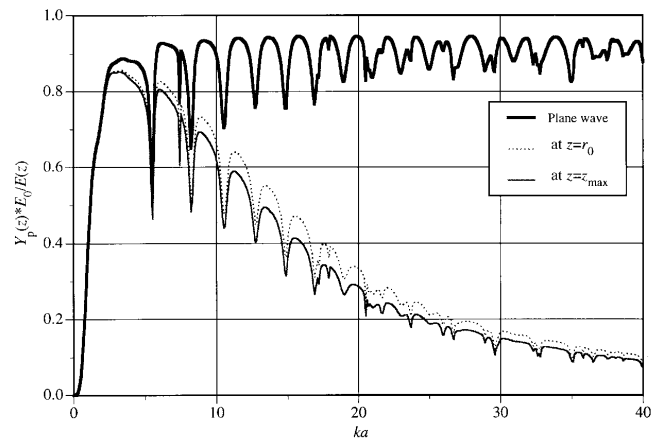


FIG. 6. Radiation force functions, normalized by the time-averaged energy density at the center of the sphere, for a type 4310 stainless-steel sphere in water, as a function of the size parameter ka at the geometrical focus and at the pressure maximum of a focused cylindrical transducer and in a plane traveling wave.

ACKNOWLEDGMENT

Partial funding by the National Institute of Health through Grant No. R04-GM30419 is acknowledged.

- ¹L. V. King, "On the acoustic radiation pressure on spheres," *Proc. R. Soc. London, Ser. A* **137**, 212 (1935).
- ²F. E. Fox, "Sound pressure of spheres," *J. Acoust. Soc. Am.* **12**, 147 (1940).
- ³T. F. W. Embleton, "Mean force on a sphere in a spherical sound field. I. (Theoretical)," *J. Acoust. Soc. Am.* **26**, 40–45 (1954).
- ⁴K. Yosioka and Y. Kawasima, "Acoustic radiation pressure on a compressible sphere," *Acustica* **5**, 167–173 (1955).
- ⁵G. Maidanik and P. J. Westervelt, "Acoustic radiation pressure due to incident plane progressive waves on spherical objects," *J. Acoust. Soc. Am.* **29**, 936–940 (1957).
- ⁶G. Maidanik, "Acoustic radiation pressure due to incident plane progressive waves on spherical objects," *J. Acoust. Soc. Am.* **29**, 738–742 (1957).
- ⁷L. P. Gor'kov, "On the forces acting on a small particle in an acoustic field in an ideal fluid," *Sov. Phys. Dokl.* **6**, 773–775 (1962).
- ⁸W. L. Nyborg, "Radiation pressure on a small rigid sphere," *J. Acoust. Soc. Am.* **42**, 947–952 (1967).
- ⁹T. Hasegawa and K. Yosioka, "Acoustic-radiation force on a solid elastic sphere," *J. Acoust. Soc. Am.* **46**, 1139–1143 (1969).
- ¹⁰F. Dunn, A. J. Averbuch, and W. D. O'Brien, Jr., "A primary method for the determination of ultrasound intensity with the elastic sphere radiometer," *Acustica* **38**, 58–61 (1977).
- ¹¹T. Hasegawa and Y. Watanabe, "Acoustic radiation pressure on an absorbing sphere," *J. Acoust. Soc. Am.* **63**, 1733–1737 (1978).
- ¹²J. Wu and G. Du, "Acoustic radiation force on a small compressible sphere in a focused beam," *J. Acoust. Soc. Am.* **87**, 997–1003 (1990).
- ¹³R. Löfstedt and S. Putterman, "Theory of long wavelength acoustic radiation pressure," *J. Acoust. Soc. Am.* **90**, 2027–2034 (1991).
- ¹⁴X. Chen and R. E. Apfel, "Radiation force on a spherical object in an axi-symmetric wave field and its application to the calibration of high frequency transducers," *J. Acoust. Soc. Am.* **99**, 713–724 (1996).
- ¹⁵M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).
- ¹⁶T. Hasegawa, N. Inoue, and K. Matsuzawa, "A new theory for the radiation from a concave piston source," *J. Acoust. Soc. Am.* **82**, 706–708 (1987).
- ¹⁷J. J. Faran, Jr., "Sound scattering by solid cylinders and spheres," *J. Acoust. Soc. Am.* **23**, 405–418 (1951).
- ¹⁸R. Hickling, "Analysis of echoes from a solid elastic sphere in water," *J. Acoust. Soc. Am.* **34**, 1582–1592 (1962).
- ¹⁹A. E. Hay and R. W. Burling, "On sound scattering and attenuation in suspensions, with marine applications," *J. Acoust. Soc. Am.* **72**, 950–959 (1982).

Double-sum technique for performing a Fourier transformation

G. Maidanik and K. J. Becker

David Taylor Model Basin (NSWC/CD), Bethesda, Maryland 20084-5000

(Received 8 July 1996; revised 9 December 1996; accepted 19 December 1996)

Structural acoustics often utilizes Fourier transformation to either reveal a phenomenon that is more directly recognized in one domain than in its Fourier conjugate domain or to employ complementarity in order to decipher a phenomenon. In this paper, a novel computational technique is introduced. This technique advantageously performs this transformation in situations in which specific spatial scales govern aliased factors in the integrand of a Fourier transformation. The technique can factorially save computational steps over corresponding computations that employ discrete Fourier transform procedures. [S0001-4966(97)00205-1]

PACS numbers: 43.20.Bi, 43.58.Ta [JEG]

INTRODUCTION

Structural acoustics deals largely with the description of wave propagations in structures and in the fluids that are in contact with their internal and external surfaces. Models of the structure are constructed so that they facilitate the derivation of these descriptions. A description may be more readily derived when the model is idealized and adapted to accommodate a given Fourier domain. Yet a phenomenon in the description may be, in some cases, more advantageously interpreted in a domain that is a Fourier conjugate to that in which the description is originally derived. Moreover, a complementary description of this kind is itself of considerable benefit in the interpretation of a phenomenon in the description. In this sense, one is often called upon to perform Fourier transformations in structural acoustics. The purpose of this paper is to propose a novel computational technique to accomplish a certain class of Fourier transformations. It transpires that a number of phenomena in structural acoustics are featured in this class; in particular, those phenomena that are related to aliasing and to pass bands and stop bands in regularly ribbed structures.¹⁻¹¹

A typical Fourier transform in this paper is expressed

$$f(x, \omega) = (2\pi)^{-1/2} \int dk \exp(-ixk) F(k, \omega), \quad (1a)$$

$$F(k, \omega) = (2\pi)^{-1/2} \int dx \exp(ixk) f(x, \omega), \quad (1b)$$

where the k domain and the x domain are Fourier conjugate domains; (k) and (x) are a Fourier variable pair; and (ω) is a silent variable in these transformations. This silent variable may be a scalar or even a vector. Moreover, (ω) may be itself a Fourier conjugate variable; e.g., $f(x, \omega)$ in Eq. (1a) may be related to the function of $\tilde{f}(x, t)$ by a Fourier transformation

$$\tilde{f}(x, t) = (2\pi)^{-1/2} \int d\omega \exp(i\omega t) f(x, \omega), \quad (1c)$$

where the ω domain and the t domain are Fourier conjugate domains. To subdue awkwardness, however, this and other more compounded variations on the theme are deferred, notwithstanding that such variabilities are, by and large, merely

natural extensions to the material presented herein. In this vein, again, (ω) is assigned to be a silent variable and only a single-dimensional Fourier transformation is considered. Equation (1a) is then conveniently abbreviated in the form

$$f(x) = (2\pi)^{-1/2} \int dk \exp(-ixk) F(k). \quad (1d)$$

In Eq. (1d) and in subsequent expressions in this paper, the presence of (ω) is implicit. For Eq. (1d) to be valid, $F(k)$ needs to be a well-behaved function of (k) so that convergence can be guaranteed. In this paper the concern is mainly focused on Eq. (1d). The integration in this equation may be cast in terms of integrations over equal intervals

$$f(x) = (2\pi)^{-1/2} \left[\dots + \int_{(\kappa_q - \kappa_1) \leq}^{(\kappa_q) >} dk + \int_{(\kappa_q) \leq}^{(\kappa_q + \kappa_1) >} dk + \int_{(\kappa_q + \kappa_1) \leq}^{(\kappa_q + 2\kappa_1) >} dk + \dots \right] \{ \exp(-ixk) F(k) \}, \quad (2)$$

where (\leq) and $(>)$ just includes and just excludes, respectively, the lower and upper limits on the integrations, (κ_1) is the wave-number extent of each interval and $\kappa_q = (q\kappa_1)$, (q) being negative and positive integers including zero. Using the self-aliased wave-number operator $S_b(k)$, Eq. (2) may be concisely stated in the form

$$f(x) = (2\pi)^{-1/2} \int_0^{\kappa_1} dk S_b(k) \{ \exp(-ixk) F(k) \}, \quad (3a)$$

$$\kappa_1 = (2\pi/b), \quad (3b)$$

where (b) is an appropriately selected spatial scale factor, and the *inclusion* and *exclusion* of the limits of integration are to be understood⁵ [cf. Eq. (2)]. The self-aliased wave-number operator $S_b(k)$ is defined

$$S_b(k) \{ N(k) \} = \sum_q N(k + \kappa_q) = M(k) = M(k + \kappa_j), \quad (4)$$

$$S_b(k) \{ K(k) G(k) \}$$

$$= K(k) S_b(k) \{ G(k) \}$$

$$\text{if } K(k) = K(k + \kappa_j), \quad \kappa_q = q\kappa_1, \quad \kappa_j = j\kappa_1,$$

and $N(k)$, $K(k)$, and $G(k)$ are arbitrary but well-behaved functions of the wave number (k). The quantities $K(k)$, by definition, and $M(k)$, by construction, are aliased; i.e., once $K(k)$ and $M(k)$ are specified in a segment that spans the width (κ_1) in the k domain, they are specified, by repetitive extrapolation, throughout that domain. In particular, if $K(k)$ and $M(k)$ are known within the segment $0 \leq k < \kappa_1$, they are known throughout the k domain. Equation (1d) is transformed into Eq. (3a) by constructing an integrand that is “aliased in (k) with respect to (κ_1)” and, therefore, the integral over a single segment (κ_1) wide suffices; all the information that is in $\{\exp(-ixk)F(k)\}$ over the entire k domain is concentrated in $S_b(k)\{\exp(-ixk)F(k)\}$ over the segment $0 \leq k < \kappa_1$ of that domain. After the aliased integrand is sampled in this segment, the integration over this segment may be evaluated adaptively or otherwise. Alternatively, the segmental span (κ_1) may be chosen to be (κ'_1) which is small enough so that a “mean value theorem” can be relied upon to evaluate the integral in Eq. (3a). At this stage, the alternative procedure is selected for the evaluation of Eq. (3a); namely

$$\begin{aligned} f(x) &= (2\pi)^{-1/2}(\kappa'_1)S_{b'}(k_0)\{\exp(-ixk_0)F(k_0)\}, \\ \kappa'_1 &= (2\pi/b'), \quad k_0 = (\alpha\kappa'_1), \\ S_{b'}(k)\{N(k)\} &= \sum_q (N + \kappa'_q), \quad \kappa'_q = q\kappa'_1, \end{aligned} \quad (3b)$$

where (α) is a positive constant less than unity [cf. Eq. (4)]. To define “small enough,” the details in the integrand in Eq. (1d) may be conveniently typified by a wave number (k_γ) and a loss factor (η_γ). The wave number (k_γ) locates the position of the (γ)th detail and (η_γ) estimates the extent of that detail in the k domain. In order to capture all relevant details in this integrand, when Eq. (1d) is converted into Eq. (3b), it is necessary to impose

$$\kappa'_1 \leq (\eta_\gamma k_\gamma), \quad (5)$$

for all relevant γ 's. The inequality stated in Eq. (5) ensures that the sampling interval κ'_1 is small enough to properly distinguish and assess the relevant details in the wave-number distribution of $F(k)$. Indeed, under the condition stated in Eq. (5), and after carrying out the operation dictated by $S_{b'}(k_0)$ in Eq. (3b), (k_0) can be safely set equal to zero by choosing (α) to be zero. In this limit, Eq. (3b) becomes a Fourier series expansion for $f(x)$.¹¹ Clearly, the infinite summation that is dictated in Eq. (3b) by the operator $S_{b'}(k_0)$ cannot be met in practice; the summation needs to be truncated. To ensure that the truncated terms can be discarded, a maximum wave number (k_M) is defined so that the convergence of $F(k)$ is substantially completed when this wave number is reached. A number (J') is then defined and it is required that

$$\begin{aligned} f(x) &\approx (2\pi)^{-1/2}(\kappa'_1)S_{b'}^{J'}(k_0)\{\exp(-ixk_0)F(k_0)\}, \\ k_M &= (J'\kappa'_1), \end{aligned} \quad (3c)$$

constitutes an adequate approximation to Eq. (3b), where

$$S_{b'}^{J'}(k)\{N(k)\} \sum_{j=-J'}^{J'} N(k + \kappa'_j), \quad \kappa'_j = (j\kappa'_1) \quad (6)$$

[cf. Eqs. (3b) and (4)]. How can an adequate approximation be defined in the context of Eq. (3c)? For this purpose one needs to specify the expected “overall spatial span” (x_M) and the “spatial sampling span” (x'_1) for the Fourier transform $f(x)$. The specification of these spatial parameters defines a number (J'_x) given by

$$x_M = (J'_x x'_1). \quad (7a)$$

The uncertainty principle and sampling theory relate $\{k_M, \kappa'_1\}$ to $\{x_M, x'_1\}$ in the form

$$(x'_1 k_M) = (x_M \kappa'_1) \geq 2\pi, \quad J'_x = J' \geq 2\pi(x'_1 \kappa'_1)^{-1}. \quad (7b)$$

Summarizing Eqs. (3) through (7) yields

$$\begin{aligned} f(x) &\approx (2\pi)^{-1/2}(\kappa'_1)S_{b'}^{J'}(k_0)\{\exp(-ixk_0)F(k_0)\}, \\ k_0 &= (\alpha\kappa'_1) \end{aligned} \quad (8)$$

and the approximation is adequate, provided

$$(\eta_\gamma)^{-1} \leq (k_\gamma / \kappa'_1) \leq J', \quad J' = (k_M / \kappa'_1) = (x_M / x'_1) \quad (9)$$

is satisfied, where (k_γ) and (η_γ) define all details in $F(k, \omega)$ in the k domain that need to be captured to render $f(x)$ a reliable Fourier transform. Usually ($x'_1 \kappa'_1$) is selected as small compared with unity and (α) is selected as zero. Choosing the equality sign in Eq. (7b) and, as such, substituting it in Eq. (8), the resulting equation can be readily manipulated to conform with the Discrete Fourier Transform (DFT). The DFT is the basis for the formulation of the Fast Fourier Transform (FFT) computational procedures.¹² In this paper, however, a different computational path is selected.

I. FORMALISM OF A DOUBLE-SUM COMPUTATIONAL PROCEDURE

The approximation that underlies Eq. (8) requires Eq. (9) to be satisfied. The latter equation implies that some knowledge of $F(k, \omega)$ is on hand in that the various vectors $\{k_\gamma, \eta_\gamma\}$ that typify details in this function can be reasonably estimated. These estimations are then used to determine the vector $\{\kappa'_1, J'\}$ that needs to be imposed on Eq. (8). This vector, in turn, determines the “wave-number sampling span” (κ'_1) and the “overall wave-number span” $k_M = (J'\kappa'_1)$. The first determines the separation between adjacent sampling positions and the second ensures convergence. Situations may arise in which the wave-number sampling span is selected, for some reason, to be (κ_1) instead of (κ'_1) and (κ_1) $>$ (κ'_1). Were this selection of (κ_1) for (κ'_1) to be made, what modification needs to be instituted in Eq. (1a) to bring in a finite summation, while ensuring that an adequate approximation is installed? Indeed, the approximation is required to match the accuracy that is enjoyed by Eq. (8) when this equation is covered by the imposition of Eq. (9).

Using the self-aliased operator $S_b(k)$, Eq. (1d) may be recast in the form

$$f(x) = (2\pi)^{-1/2} \int_0^{\kappa_1} dk S_b(k) \{\exp(-ixk) K(k) G(k)\},$$

$$\kappa_1 = (2\pi/b), \quad (10a)$$

where, for the sake of flexibility, $F(k)$ is expressed in the factorial form

$$F(k) = K(k) G(k) \quad (11)$$

[cf. Eq. (3a)]. Ensuring the convergence of $F(k)$ by an overall wave-number span $k_M = J\kappa_1$, the infinite summation implied by $S_b(k)$ is then replaced by a finite summation; namely, Eq. (10a) is approximated in the form

$$f(x) \approx (2\pi)^{-1/2} \int_0^{\kappa_1} dk S_b^J(k) \{\exp(-ixk) K(k) G(k)\},$$

$$k_M = (J\kappa_1), \quad (10b)$$

where

$$S_b^J(k) \{N(k)\} = \sum_{j=-J}^J N(k + \kappa_j),$$

$$\kappa_j = j\kappa_1, \quad \kappa_1 = (2\pi/b) \quad (12a)$$

[cf. Eqs. (3c) and (6)]. With the help of Eq. (7) it emerges that the approximation in Eq. (10b) matches that in Eq. (3c) provided the equalities and the inequality

$$x_1 = x'_1, \quad x_M = J'x_1, \quad k_M = (J'\kappa'_1) = (J\kappa_1), \quad (13)$$

$$J \geq 2\pi(x_1\kappa_1)^{-1},$$

are maintained. Since the new wave-number sampling span (κ_1) is larger than the old wave-number sampling span (κ'_1), the integral cannot be evaluated by a single segment in the manner of Eq. (3c). Each segment in the summation needs to be spanned more finely. In this vein, the integral is carried out over $(1+R)$ regular intervals, notwithstanding that an adaptive procedure may, under certain circumstances, be preferred. Executing the integral in the manner just specified, one derives from Eq. (10b) the Fourier transform in the form

$$f(x) \approx (2\pi)^{-1/2} (1+R)^{-1} (\kappa_1) \sum_{r=0}^R S_b^J(k_r) \times \{\exp(-ixk_r) K(k_r) G(k_r)\},$$

$$k_M = (J\kappa_1), \quad k_r = (1+R)^{-1} \kappa_r,$$

$$\kappa_r = (r\kappa_1), \quad \kappa_1 = (2\pi/b). \quad (14)$$

The “effective” wave-number sampling span in Eq. (14) is then given by $(1+R)^{-1} \kappa_1$. Using Eqs. (9) and (12), the approximation in Eq. (14) adequately matches that of Eq. (8) provided

$$\{\eta_\gamma(1+R)\}^{-1} \ll (k_\gamma/\kappa_1) \ll J, \quad J(1+R) \approx J', \quad (15)$$

notwithstanding that the number of terms in the double-sum in Eq. (14) is $\{(2J+1) + (1+R)\}$ vs $(2J'+1)$ in the single-sum in Eq. (8). However, the number of computational steps is $\{(2J+1)(1+R)\}$ in Eq. (14) versus $(2J'+1)$ in Eq. (8), which is in agreement with the second of Eqs. (15) when J

and $J' \gg 1$. There is then no factorial saving in computational steps between Eq. (14), as it stands, and Eq. (8); using one or the other equation is merely a matter of convenience. The crux of the paper is revealed when one or the other factors composing $F(k)$, $K(k)$, etc. is aliased in (k) with respect to (κ_1) , and (κ_1) matches the (κ_1) in Eq. (10a). Under this condition, a factorial saving of computational steps may be accrued by an appropriate manipulation of the integrand. A few examples in this regard are considered next.

II. INTEGRAND COMPOSED OF ALIASED FACTORS

The expression depicted in Eq. (11) is specialized in that the factor $K(k)$ is assumed to be aliased in (k) with respect to (κ_1) [cf. Eq. (4)]. Using the properties of the self-aliased operator $S_b(k)$ as defined in Eq. (4), the integrand in Eq. (10a) can be equivalently expressed in the form

$$S_b(k) \{\exp(-ixk) K(k) G(k)\} \\ \equiv K(k) S_b(k) \{\exp(-ixk) G(k)\} \quad \text{if } K(k) = K(k + \kappa_j). \quad (16a)$$

In terms of the approximation stated in Eq. (14), the “ R -summand” equivalence assumes the form

$$S_b^J(k_r) \{\exp(-ixk_r) K(k_r) G(k_r)\} \\ \Rightarrow K(k_r) S_b^{J_1}(k_r) \{\exp(-ixk_r) G(k_r)\}, \quad k_{M1} = J_1\kappa_1, \quad (16b)$$

where k_{M1} is the overall wave-number span with respect to the convergence of the operand under $S_b^{J_1}(k_r)$ in this equation and in general

$$S_b^J(k_r) \{N(k_r)\} = \sum_{j=-J_\beta}^{J_\beta} N(k_r + \kappa_j). \quad (12b)$$

It is noted that if the right side of Eq. (16b) is used in Eq. (14) instead of the left, a factorial saving equal to the smaller of $(2J_1+1)$ and $(2J+1)$ of computational steps is accrued. Usually, however, $J_1 = J$ and, therefore, the two factors are equal. Moreover, if the aliasing in $K(k)$ can be cast in the form

$$K(k) = S_b(k) \{K_A(k) K_B(k)\}, \quad (17a)$$

then in terms of the approximation stated in Eq. (14), the R -summand equivalence becomes

$$S_b^J(k_r) \{\exp(-ixk_r) K(k_r) G(k_r)\} \\ \Rightarrow S_b^{J_2} \{K_A(k_r) K_B(k_r)\} S_b^{J_1}(k_r) \{\exp(-ixk_r) G(k_r)\},$$

$$k_{M2} = (J_2\kappa_1), \quad (17b)$$

where k_{M2} is the overall wave-number span with respect to the convergence of the operand $\{K_A(k) K_B(k)\}$ under $S_b^{J_2}(k)$. It is noted that the factorial saving of computational steps pertaining to Eq. (17b) remains unchanged from that derived from Eq. (16b). If, in addition to the form of $K(k)$, as stated in Eq. (17a), it is assumed that $K_A(k)$ is aliased in (k) with respect to (κ_1) , then the factor $K(k)$ can be factorized further in the form

$$K(k) = K_A(k)S_b(k)\{K_B(k)\} \quad \text{if } K_A(k) = K_A(k + \kappa_j). \quad (18a)$$

In terms of the approximation stated in Eq. (14), the R -summand equivalence assumes the additional factorization, over Eq. (17b), of the form

$$\begin{aligned} S_b^J(k_r)\{\exp(-ixk_r)K(k_r)G(k_r)\} \\ \Rightarrow K_A(k_r)S_b^{J_3}(k_r)K_B(k_r)\{S_b^{J_1}(k_r)\{\exp(-ixk_r)G(k_r)\}\}, \\ k_{M3} = (J_3\kappa_1), \end{aligned} \quad (18b)$$

where k_{M3} is the overall wave-number span with respect to the convergence of the quantity $K_B(k)$ in this equation. It is noted that if in Eq. (14) the right side of Eq. (18b) is used instead of that of Eq. (16b), a factorial saving equal to the smaller of $(2J_2 + 1)$ and $(2J_3 + 1)$ of computational steps is accrued. Usually, however, $J_\beta = J$ and, therefore, the two factors are equal. Under this equality, it is noted that if the right side of Eq. (18b) is used in Eq. (14) instead of the left, a factorial saving of $(2J + 1)^2$ of computational steps is accrued. In the event that the aliasing in the factor $K_A(k_r)$, as stated in Eq. (18a), stems from its functional dependence on an aliased quantity; e.g.,

$$K_A(k) = K_A[S_b(k)\{K_{AA}(k)\}], \quad (19a)$$

the approximation stated in Eq. (14) supports the R -summand equivalence of the form

$$\begin{aligned} S_b^J(k_r)\{\exp(-ixk_r)K(k_r)G(k_r)\} \\ \Rightarrow K_A^{J_4}(k_r)S_b^{J_3}(k_r)\{K_B(k_r)\}S_b^{J_1}(k_r)\{\exp(-ixk_r)G(k_r)\}, \\ K_A^{J_4}(k_r) = K_A[S_b^{J_4}(k_r)\{K_{AA}(k_r)\}], \quad k_{M4} = (J_4\kappa_1), \end{aligned} \quad (19b)$$

where k_{M4} is the overall wave-number span with respect to the convergence of the quantity $K_{AA}(k)$ in this equation. It is noted that there is no factorial saving of computational steps between the use of Eqs. (18b) and (19b). If further factorization exists, either in $K_A(k)$, in $K_B(k)$, or in both, in the format just prescribed, one may gain more and more factorial savings. For the purpose of this paper, however, the factorization and dependence considered in Eqs. (18a) and (19a) suffice.

A special case of Eq. (11) is that in which, in addition to the aliasing of $K(k)$ in (k) with respect to (κ_1) , the function $G(k)$, in this equation, is independent of the wave number (k) ; namely

$$F(k) = G_0 K(k), \quad K(k) = K(k + \kappa_j). \quad (20)$$

Substituting Eq. (20) in Eq. (10a) derives

$$\begin{aligned} f(x) &= s_b(x)\{\delta(x)f_s(x)\}, \\ f_s(x) &= [G_0/\kappa_1](2\pi)^{-1/2} \int_0^{\kappa_1} dk \exp(-ixk)K(k), \end{aligned} \quad (21)$$

where $s_b(x)$ is a self-aliased spatial (x) operator with respect to the fundamental separation distance (b) ; $b = (2\pi/\kappa_1)$. This operator is defined

$$s_b(x)\{n(x)\} = \sum_j n(x + x_j), \quad x_j = (jb), \quad x_0 = 0 \quad (22)$$

[cf. Eq. (4)]. Executing the integral in Eq. (21) in a manner analogous to that resulting in Eq. (14), one obtains

$$f_s(x) = (2\pi)^{-1/2} [G_0/(1+R)] \sum_{r=0}^R \{\exp(-ixk_r)K(k_r)\}, \quad (23)$$

and the evaluation of $f_s(x)$ is considered, in view of Eq. (21), to be synonymous with the evaluation of $f(x)$, notwithstanding the acknowledgment that, unlike $f(x)$, the evaluation of $f_s(x)$ is meaningful only for values of $x = x_j$; $x_j = (jb)$. In this view, the recognition of the aliasing in $K(k)$, as stated in Eq. (20), is worth a factorial saving of $(2J + 1)$ of computational steps. This is similar to the factorial saving that is accrued in Eq. (16b), utilizing the same recognition. Indeed, the factorization of $K(k)$, as depicted in Eqs. (16)–(19) and beyond, are just as beneficial with respect to Eq. (23) as they are with respect to Eq. (14).

¹L. Cremer, M. Heckl, and E. Ungar, *Structure-Borne Sound, Structural Vibrations and Sound Radiation at Audio Frequencies* (Springer-Verlag, Berlin, 1988), 2nd ed.

²G. Maidanik and J. Dickey, "Velocity distributions on unloaded finitely and regularly ribbed membranes," *J. Sound Vib.* **149**, 43–70 (1991); "Response of regularly ribbed fluid loaded panels," *J. Sound Vib.* **155**, 481–495 (1992).

³G. Maidanik and J. Dickey, "Quadratic and energy estimates of the partial response of ribbed panels," *J. Acoust. Soc. Am.* **94**, 1435–1444 (1993); "Influence of variations in loss factor of a panel and line impedance of attached ribs," *J. Acoust. Soc. Am.* **94**, 1445–1452 (1993).

⁴G. Maidanik and J. Dickey, "Localization and delocalization in periodic one-dimensional dynamic systems," *Acustica* **73**, 119–128 (1991).

⁵C. B. Burroughs, "Acoustic radiation from fluid-loaded infinite circular cylinders with doubly periodic ring supports," *J. Acoust. Soc. Am.* **75**, 715–722 (1984).

⁶M. C. Junger and D. Feit, *Sound, Structures, and Their Interaction* (MIT, Cambridge, MA, 1986).

⁷F. Fahy, *Sound and Structural Vibration (Radiation, Transmission and Response)* (Academic, New York, 1985).

⁸M. P. Norton, *Fundamentals of Noise and Vibration Analysis for Engineers* (Cambridge U.P., Cambridge, England, 1989).

⁹G. Maidanik and K. J. Becker, "Computation of the modal response of regularly ribbed cylinders," *J. Acoust. Soc. Am.* **101**, 2452–2472 (1997).

¹⁰G. Maidanik and K. J. Becker, "Phenomena of aliasing and pass and stop bands in the drive in lieu of ribs on cylindrical shells," CDNSWC Report No. SIG-96/072-7030, 1996 (unpublished).

¹¹K. J. Becker, "On the response of regularly ribbed cylindrical shells," thesis (in preparation).

¹²R. J. Mayhan, *Analysis of Discrete-Time and Continuous-Time Linear Systems* (Addison-Wesley, Reading, MA, 1984).

Computation of the modal response of regularly ribbed cylinders

G. Maidanik and K. J. Becker

David Taylor Model Basin (NSWC/CD), Bethesda, Maryland 20084-5000

(Received 8 July 1996; revised 9 December 1996; accepted 19 December 1996)

Recently an efficient technique was proposed for performing a Fourier transformation on integrands that are composed of aliased factors. In this paper the proposed technique is utilized to perform Fourier transformations on quantities that pertain to the response of a regularly ribbed cylinder. The computations relate to Fourier transformations from the axial wave-number domain to the axial spatial domain. The phenomena of aliasing and pass bands and stop bands are of particular interest. It is argued that the complementarity of the data, when presented in the two Fourier conjugate domains, may be a useful analytical (and experimental) tool. [S0001-4966(97)00105-7]

PACS numbers: 43.20.Bi, 43.20.Tb, 43.40.Ey, 43.40.Rj [JEG]

INTRODUCTION

In a recent paper the authors proposed a novel computational technique to perform a certain class of Fourier transformations.¹ The class pertains to Fourier transformations in which the integrands are completely or partially composed of terms and factors that are aliased with respect to a specific scale factor in the domain in which the integrand is expressed. This class harbors a number of phenomena in structural acoustics. A few of these phenomena are employed in this paper to exemplify the use of the double-sum technique for computing Fourier transforms.¹ In particular, the circumferential modal response of a regularly ribbed cylinder is transformed from the k domain into the x domain. These Fourier conjugate domains lie in the axial direction of the cylinder. The phenomena of special concern in the present paper are those relating to aliasing and pass bands and stop bands in that modal response.²⁻⁶ Hybrid and natural cylinders are exemplified. A hybrid cylinder is defined by analytically removing the membrane-free waves from the cylinder. In this sense, the mechanical response of a hybrid cylinder is akin to that of a panel that is enjoying the same plating.² This paper assumes that the descriptions of these phenomena and the analytical models of the cylindrical structures are familiar to the reader; nonetheless a cursory analytical development is included. A more detailed analytical development can be found in Refs. 2-6. However, the analytical development of the double-sum technique that is presented in Ref. 1 is not repeated here; Ref. 1 is considered an essential companion to this paper and the reference to equations in Ref. 1 are prefaced by unity. Thus Eq. (6) in Ref. 1 is designated Eq. (1-6) in the present paper. The physical construction of the externally driven ribbed cylinder and the chosen coordinate system are sketched and depicted in Fig. 1, respectively.

In Sec. I the response and the external drive that generate it are expanded in terms of circumferential eigenfunctions.⁷ The coefficients in these expansions are the modal responses and modal external drives. The modal properties of the cylindrical shell that relate the modal response to the corresponding modal external drive are defined; this relationship defines the modal impulse response function. It

is argued that for a regularly ribbed cylinder, the expressions for the impulse response function and its terms and factors are readily stated and interpreted in the $\{k, n, \omega\}$ domain, where (ω) is the frequency variable, (n) is the mode index, and (k) is the axial wave-number variable.⁸ Indeed, in this domain the relationship between the modal response $V_n(k, \omega)$ and the modal external drive $P_{en}(k, \omega)$ may be expressed in the form

$$V_n(k, \omega) = G_{\infty n}(k, \omega) [P_{en}(k, \omega) - P_{sn}(k, \omega)],$$

where $P_{sn}(k, \omega)$ is the drive in lieu of the ribs and $G_{\infty n}(k, \omega)$ is the modal admittance of the uniform (unribbed) cylinder. The external drive is adequately defined in the simple form

$$P_{en}(k, \omega) = P_{en}^0(\omega) \exp(ikx_a),$$

where (x_a) is the position, in the x domain, to which the external drive is applied and it is assumed that the modal strength $P_{en}^0(\omega)$ is independent of (k) . The uniform cylinder modal admittance $G_{\infty n}(k, \omega)$ comprises two distinct terms; the modal mechanical surface impedance $Z_{pn}(k, \omega)$ and the modal fluid surface impedance $Z_n^f(k, \omega)$ on the outside surface of the cylinder

$$G_{\infty n}(k, \omega) = [Z_{pn}(k, \omega) + Z_n^f(k, \omega)]^{-1}.$$

On the inside of the cylinder a vacuum is assumed to prevail. The modal mechanical surface impedance $Z_{pn}(k, \omega)$, in turn, also comprises two terms

$$Z_{pn}(k, \omega) = Z_{pn}^p(k, \omega) + Z_{pn}^m(k, \omega),$$

where $Z_{pn}^p(k, \omega)$ is the modal *flexural* surface impedance and $Z_{pn}^m(k, \omega)$ is the *membrane* surface impedance.⁸⁻¹⁰ When the membrane surface impedance is analytically removed; $Z_{pn}^m(k, \omega) = 0$, the admittance $G_{\infty n}(k, \omega)$, given now by

$$G_{\infty n}(k, \omega) = [Z_{pn}^p(k, \omega) + Z_n^f(k, \omega)]^{-1},$$

describes the fluid-loaded modal admittance of a uniform (unribbed) hybrid cylinder.

The drive $P_{sn}(k, \omega)$ in lieu of the ribs is not only a functional of the properties of the uniform cylinder; defined

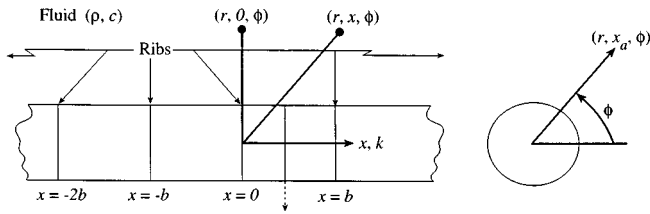


FIG. 1. A sketch of a regularly ribbed cylinder and coordinate system.

by $G_{\infty n}(k, \omega)$, but is naturally also a functional of the properties of the ribs. In a regularly ribbed cylinder the properties of the ribs are defined in terms of the modal line impedance $Z_{sn}(\omega)$ and the separation distance (b) between adjacent ribs. This line impedance can be either mass, resistance, or stiffness controlled; the change from one characteristic to another is frequency dependent.¹¹ In this connection the loss factor in the ribs is designated as (η_0) . The modal drive $P_{sn}(k, \omega)$, in addition, is a functional of the external drive. In the definition of the external drive here proposed, the dependence on the position of application (x_a) is of particular significance.

It is convenient to normalize the modal response in the form

$$\begin{aligned}\bar{V}_n(k, \omega) &= \bar{V}_{\infty n}(k, \omega) - \bar{V}_{sn}(k, \omega) \\ &= \bar{G}_{sn}(k, \omega) [\bar{P}_{en}(k, \omega) - \bar{P}_{sn}(k, \omega)], \\ \bar{V}_{\infty n}(k, \omega) &= \bar{G}_{\infty n}(k, \omega) \bar{P}_{en}(k, \omega), \\ \bar{V}_{sn}(k, \omega) &= \bar{G}_{\infty n}(k, \omega) \bar{P}_{sn}(k, \omega).\end{aligned}$$

This normalization of the modal response is achieved by the appropriate normalization of its constituent terms and factors, namely,

$$\begin{aligned}\bar{V}_n(k, \omega) &= [V_n(k, \omega) / V_{0n}(\omega)], \\ V_{0n}(\omega) &= [P_{en}^0(\omega) / (i\omega m)], \\ \bar{G}_{\infty n}(k, \omega) &= [(i\omega m) G_{\infty n}(k, \omega)], \\ \bar{P}_{sn}(k, \omega) &= [P_{sn}(k, \omega) / P_{en}^0(\omega)],\end{aligned}$$

where $(i\omega m)$ is the surface mass impedance of the plating of the cylinder and $P_{en}^0(\omega)$, as already mentioned, is the strength of the modal external drive.

The absolute values of $\bar{V}_{\infty n}(k, \omega)$, or equivalently $\bar{G}_{\infty n}(k, \omega)$, of $\bar{P}_{sn}(k, \omega)$ and of $\bar{V}_n(k, \omega)$ are computed and displayed as functions of (ak) in a frequency waterfall format, where (a) is the radius of the cylinder.¹² [If the normalized quantity $|\bar{S}(k, \omega)|$ is to be displayed as a function of (ak) in a frequency waterfall format, the plotted quantity $\bar{S}'_q(ak)$ is cast in the manner

$$\begin{aligned}\bar{S}'_q(ak) &= \log[|\bar{S}\{k, (q+1)\Delta\omega\}| 10^q] \\ &= q + \log|\bar{S}\{k, (q+1)\Delta\omega\}|,\end{aligned}$$

where $(\Delta\omega)$ is an incremental frequency and (q) is an integer bounded by $0 \leq q \leq q_f$. In a figure the identity, in $\bar{S}'_q(ak)$, of $\bar{S}(k, \omega)$ is revealed in the figure caption; e.g., in Fig. 2(a) it is $\bar{V}_{\infty n}(k, \omega)$ and in Fig. 4(b) it is $\bar{P}_{sn}(k, \omega)$. The frequency on

the stepwise frequency axis is identified by the frequency pseudonym 10^q ; e.g., in Fig. 2(a) the position marked 10^{30} locates the origin for the curve pertaining to $(\omega/\omega_c) = 31\Delta(\omega/\omega_c)$, ω_c being a normalizing frequency]. The frequency (ω) is normalized, in this paper, by the critical frequency (ω_c) of the flexural free waves with respect to the speed of sound (c) in the fluid. The discussions of these and other presentations in this paper are limited to the frequency range defined by

$$(2\omega_r/\omega_c) \leq (\omega/\omega_c) \leq 0.6, \quad \Delta(\omega/\omega_c) = 0.01,$$

$$0 \leq q \leq q_f = 59, \quad \omega_r = (c_l/a),$$

and to the wave-number range defined by

$$0 \leq (ak) \leq 75,$$

where (ω_r) is the ring frequency of the cylinder. The speed $c_l = \{(Yh/m)/(1 - \nu^2)\}^{1/2}$, where (Y) , (ν) , (m) , and (h) are the Young's modulus, the Poisson's ratio, the surface mass, and the thickness of the plating of the cylinder, respectively, is the longitudinal speed in that plating. One may find it convenient to list a set of standard parametric values to define standard displays. When variations in the parametric values are instituted, it is necessary to report only those parameters that are varied. The remaining parameters are assumed to maintain standard values.

The normalized modal response $\bar{V}_{\infty n}(k, \omega)$ of a uniform (unribbed) cylinder under standard parametric values, and a few specific variations on these parametric values, are presented in Figs. 2 and 3. The first set, i.e., Fig. 2, pertains to a hybrid cylinder and the second set, i.e., Fig. 3, pertains to a natural cylinder. The footprints of fluid loading and membrane-free waves are of particular interest in these two sets of figures.

The normalized modal drive $\bar{P}_{sn}(k, \omega)$ in lieu of the ribs under standard parametric values, and a few specific variations on these parametric values, are presented in Figs. 4 and 5. The third set, i.e., Fig. 4, pertains to a hybrid cylinder and the fourth set, i.e., Fig. 5, pertains to a natural cylinder. The footprints of aliasing and stop bands and pass bands are of particular interest in these two figures.

The normalized modal response $\bar{V}_n(k, \omega)$ of the regularly ribbed cylinder under standard parametric values, and a few specific variations on these parametric values, are presented in Figs. 6 and 7. The fifth set, i.e., Fig. 6, pertains to a hybrid cylinder and the sixth set, i.e., Fig. 7, pertains to a natural cylinder. The footprints of the normalized drive $\bar{P}_{sn}(k, \omega)$, as depicted in Figs. 4, and 5, are clearly discernible in the normalized response $\bar{V}_n(k, \omega)$. However, the strict aliasing in the drive is impaired in the response. This defect is perpetrated by the term that describes the normalized modal response $\bar{V}_{\infty n}(k, \omega)$ in the absence of ribs and by the factor $\bar{G}_{\infty n}(k, \omega)$ that multiplies the normalized aliased modal drive $\bar{P}_{sn}(k, \omega)$; neither the term nor the factor are aliased quantities. Does the impairment of the aliasing in $\bar{V}_n(k, \omega)$ imply a corresponding defect in the pass bands and stop bands of the modal response as compared with that exhibited by the modal drive in lieu of the ribs? The answer to this question is sought, in part, by using complementarity to

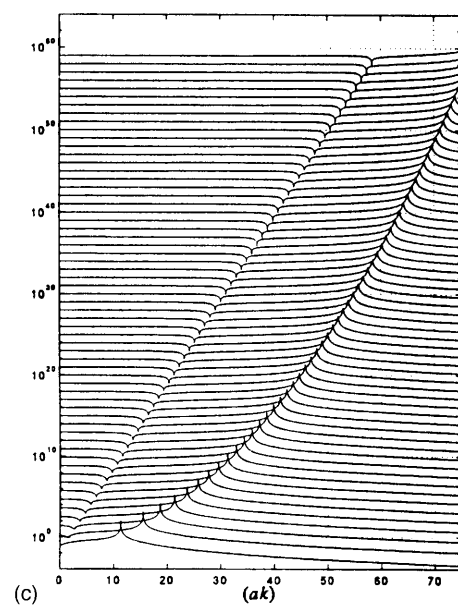
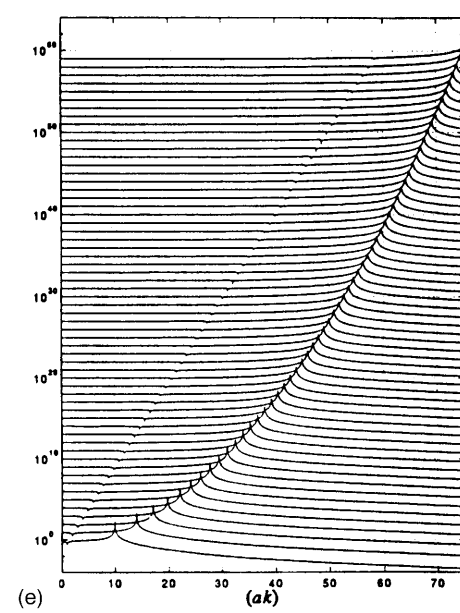
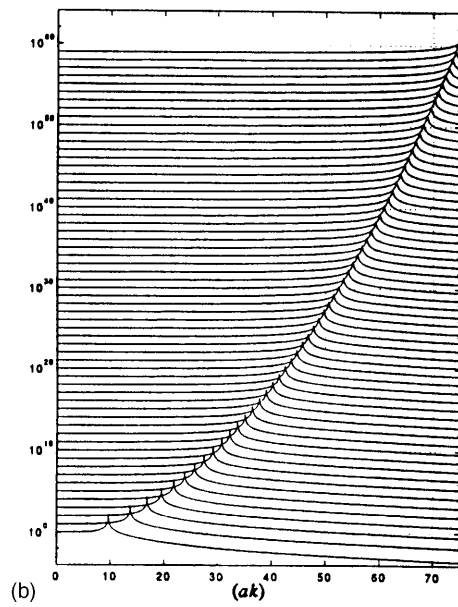
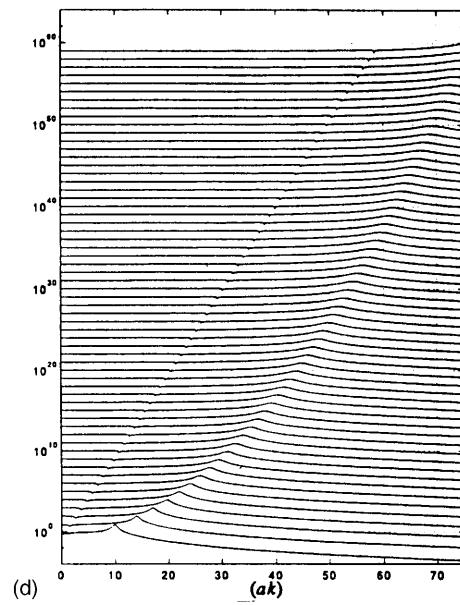
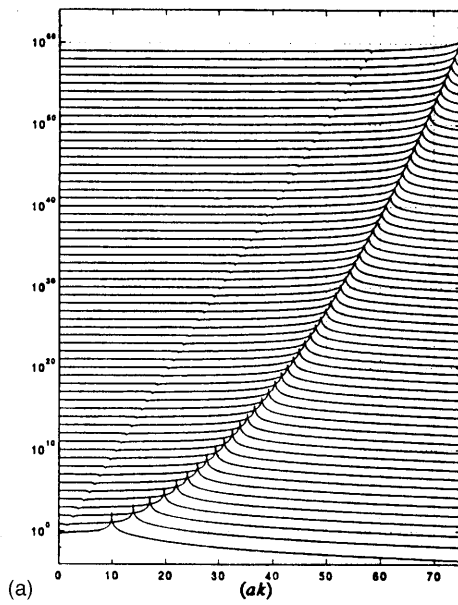


FIG. 2. Normalized modal response, $\bar{V}_{\infty n}(k, \omega)$ of a uniform *hybrid* cylinder as a function of the normalized axial wave number (ak) in a frequency waterfall format. (a) Under standard values [cf. Eq. (15)]. (b) $\epsilon_c = 10^{-4}$. (c) $\epsilon_c = 10^{-1}$. (d) $\eta_\pi = 3 \times 10^{-2}$. (e) $n = 0$.

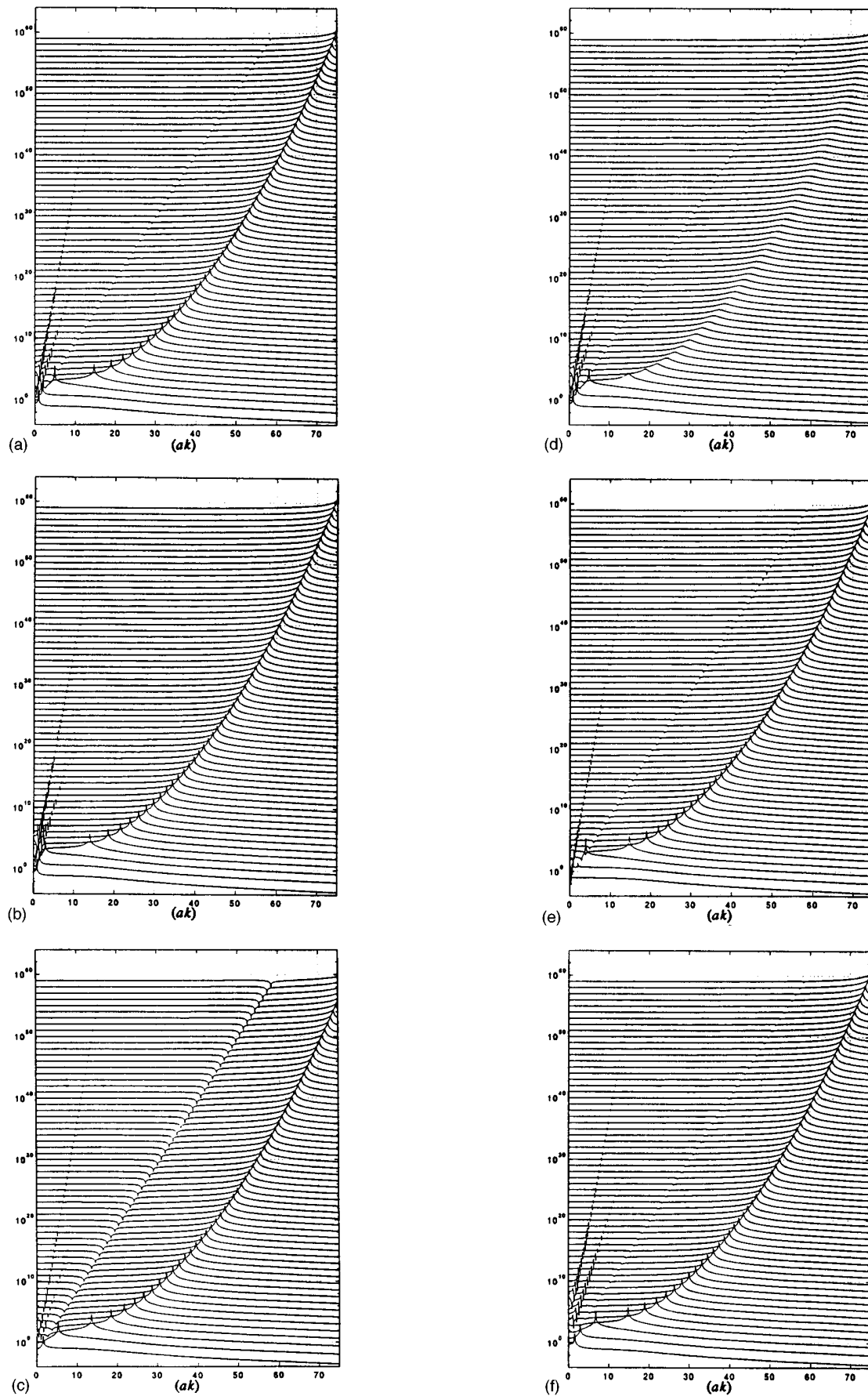


FIG. 3. Same as Fig. 2 but for a *natural* cylinder. (a) Under standard values [cf. Eq. (15)]. (b) $\epsilon_c = 10^{-4}$. (c) $\epsilon_c = 10^{-1}$. (d) $\eta_p = 3 \times 10^{-2}$. (e) $n = 0$. (f) $n = 2$. (g) $n=0$ and $\eta_p = 3 \times 10^{-2}$.

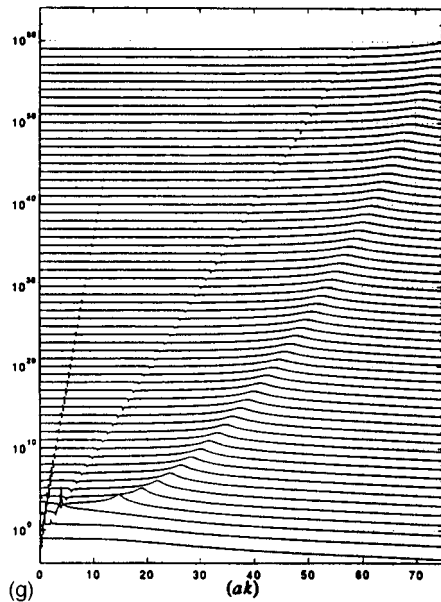


FIG. 3. (Continued.)

Figs. 2–7. This complementarity is derived by a Fourier transformation of $\bar{V}_{\infty n}(k, \omega) = [\bar{G}_{\infty n}(k, \omega) \bar{P}_{en}(k, \omega)]$, $\bar{P}_{sn}(k, \omega)$, and $\bar{V}_n(k, \omega)$ into $\bar{v}_{\infty n}(x, \omega)$, $\bar{p}_{sn}(x, \omega)$, and $\bar{v}_n(x, \omega)$, respectively. The Fourier transformation of these quantities is carried out in terms of the double-sum technique developed in Ref. 1. [Typically, the normalized quantity $\bar{s}(x, \omega)$ is displayed either as a function of (x/b) for a fixed value of (ω/ω_c) ; e.g., Fig. 8(a) or conversely, as a function of (ω/ω_c) for a fixed value (or even a few fixed values) of (x/b) ; e.g., Fig. 11(a). In the latter format, the plotted quantity $\bar{s}'_r(x/b)$ is cast in the manner

$$\bar{s}'_r(x/b) = |\bar{s}\{x, (r+1)\Delta'\omega\}|, \quad (r_f+1)\Delta'\omega = (q_f+1)\Delta\omega,$$

where $(\Delta'\omega)$ is an incremental frequency, (r) is an integer bounded by $0 \leq r \leq r_f$, $(r_f+1)(q_f+1)^{-1}$ is chosen to be an integer that equals or exceeds unity; e.g., in Figs. 11 and 12 this ratio is three (3), and $(\Delta\omega)$ and q are, respectively, the incremental frequency and the integer just adopted to serve the frequency waterfall in this paper. In a figure the identity, in $\bar{s}'_r(x/b)$, of $\bar{s}(x, \omega)$ is revealed in the figure caption; e.g., in Fig. 8(b) it is $\bar{v}_{\infty n}(x, \omega)$ and in Fig. 11 it is $\bar{p}_{sn}(x, \omega)$. Again, the frequency on the stepwise frequency axis is identified by the frequency pseudonym 10^q ; e.g., 10^{40} corresponds to $(\omega/\omega_c) = 41\Delta(\omega/\omega_c)$. This procedure renders the stepwise frequency axis, in the displays of $\bar{S}(k, \omega)$, and $\bar{s}(x, \omega)$ compatible. A compatible pair of this kind can then be complemented by matching a pair of displays along the frequency axis, e.g., Fig. 4(a).]

In Sec. II the procedure for computing $\bar{v}_{\infty n}(x, \omega)$, by a Fourier transformation of $\bar{V}_{\infty n}(k, \omega)$, is explained and the results are displayed in Figs. 8–10. Several variations in the parametric values are investigated and comparisons between Figs. 2 and 3 that feature $\bar{V}_{\infty n}(k, \omega)$ and the corresponding Figs. 8–10 that feature $\bar{v}_{\infty n}(x, \omega)$ are conducted and discussed. Interferences between flexural, fluid, and membrane-free waves at regions where their amplitudes are of compa-

table strengths are of particular interest. Such interferences are revealed when the dominant flexural free waves are subdued by selectively increasing the damping that is assigned to them. It is thus found that a complementary investigation of $\bar{V}_{\infty n}(k, \omega)$ and $\bar{v}_{\infty n}(x, \omega)$ is an analytical tool of value.

In Sec. III the procedure for computing $\bar{p}_{sn}(x, \omega)$ by a Fourier transformation of $\bar{P}_{sn}(k, \omega)$, is explained and the results are displayed in Figs. 4, 5, 11, and 12. In Figs. 4 and 5 computations of $\bar{p}_{sn}(x, \omega)$ are superimposed, frequencywise, on the corresponding figures that display $\bar{P}_{sn}(k, \omega)$. Figures 4 and 5 exhibit clearly the phenomena of aliasing and pass bands and stop bands. The complementarity between the Fourier conjugate pairs, $\bar{P}_{sn}(k, \omega)$, on the one hand, and $\bar{p}_{sn}(x, \omega)$, on the other, are more directly and simultaneously contrasted in these figures. The value of complementarity as an analytical tool is thus supported in Figs. 4 and 5. The influence of fluid loading on the phenomenon of pass bands and stop bands is investigated in Figs. 11 and 12. Moreover, the bottoming of the flexural stop bands by the fluid and membrane-free waves are discernible in these figures.

In Sec. IV the procedure for computing $\bar{v}_n(x, \omega)$ by a Fourier transformation of $\bar{V}_n(k, \omega)$ is explained and the results are displayed in Figs. 13–15. Of particular interest is the bottoming of the response in a flexural stop band by the fluid and membrane free waves. The similarity of this bottoming to the corresponding bottoming induced by damping the flexural-free waves, is striking [cf. Figs. 8–10 and 13–15]. The complementarity between Figs. 6 and 7 and the corresponding Figs. 13–15, again shows the value of its employment. Of particular interest in this final section are the comparisons between Figs. 4 and 5 and 11 and 12 on the one hand, and Figs. 6 and 7 and 13–15 on the other. These comparisons reveal that the impairment of the aliasing patterns in $\bar{V}_n(k, \omega)$, as compared with those in $\bar{P}_{sn}(k, \omega)$, is not duplicated in the pass bands and stop bands in $\bar{v}_n(x, \omega)$, as compared with those in $\bar{p}_{sn}(x, \omega)$. Indeed, if at all, the phenomenon of pass bands and stop bands in $\bar{v}_n(x, \omega)$ is sharper than it is in $\bar{p}_{sn}(x, \omega)$.

I. MODAL RESPONSE OF A RIBBED CYLINDER

The circumferential uniformity and finiteness of the cylinder, unribbed and ribbed, allow quantities to be expanded in terms of circumferential eigenfunctions $\Phi_n(\phi)$, where ϕ is the circumferential variable as shown in Fig. 1. Thus, the response $v(x, \phi, \omega)$ and the external drive $p_e(x, \phi, \omega)$ can be modalized in the x domain

$$v(x, \phi, \omega) = \sum_n \Phi_n(\phi) v_n(x, \omega), \quad (1a)$$

$$p_e(x, \phi, \omega) = \sum_n \Phi_n(\phi) p_{en}(x, \omega),$$

and in the k domain

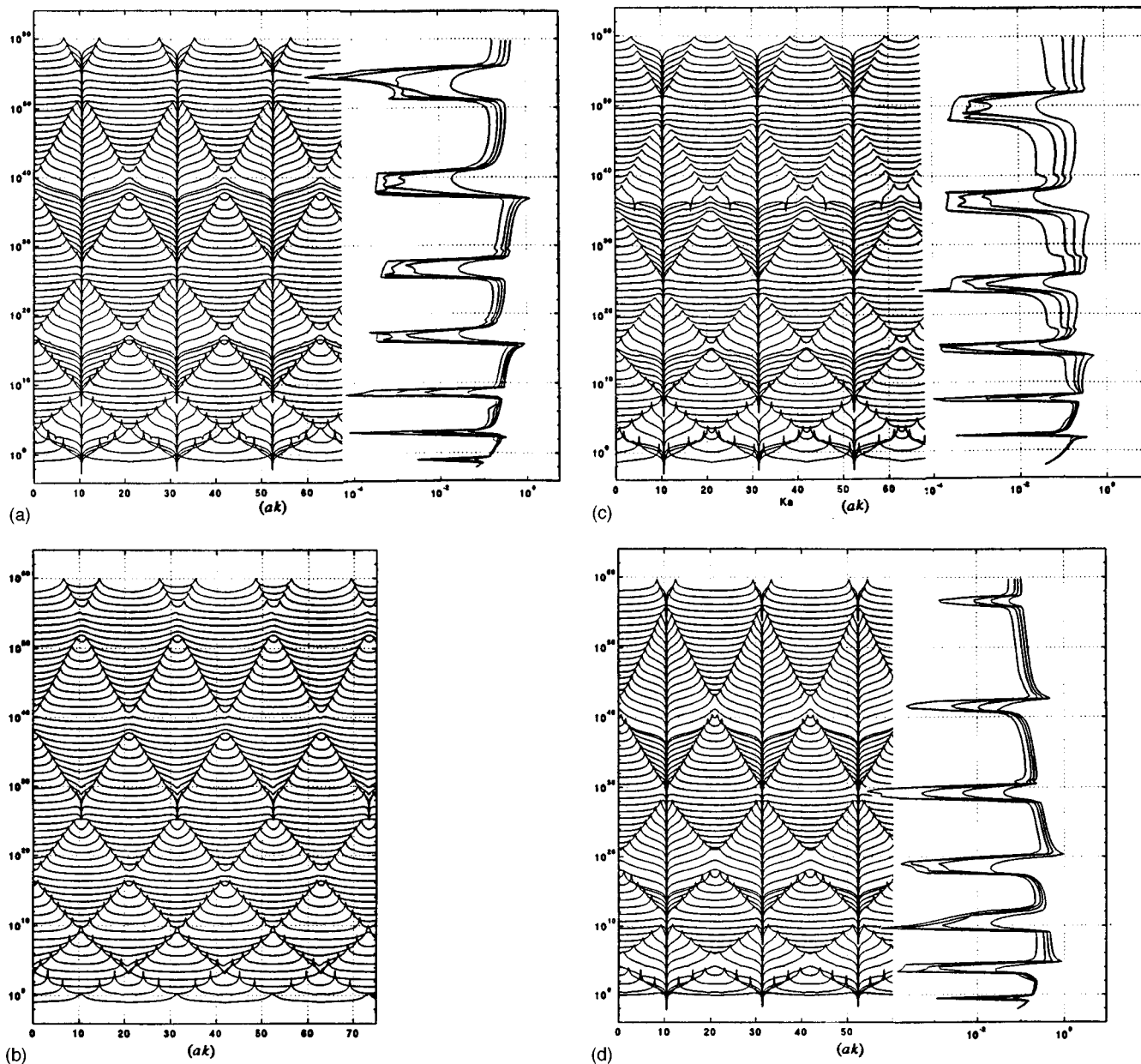


FIG. 4. Normalized modal drive $\bar{P}_{sn}(k, \omega)$ in lieu of the ribs on a *hybrid* cylinder as a function of the normalized axial wave number (ak) in a frequency waterfall format. Superposed is the normalized modal drive $\bar{p}_{sn}(x, \omega)$ in lieu of the ribs for a *hybrid* cylinder as a function of the normalized frequency (ω/ω_c) for the normalized axial distances $(x/b) = 7, 14, 21$, and 35 . (a) Under standard values [cf. Eq. (15)]. (b) $(x_a/b) = 0.3$. (c) $\epsilon_c = 10^{-1}$. (d) Rib's ring impedance stiffness controlled.

$$V(k, \phi, \omega) = \sum_n \Phi_n(\phi) V_n(k, \omega), \quad (1b)$$

$$P_e(k, \phi, \omega) = \sum_n \Phi_n(\phi) P_{en}(k, \omega).$$

Quantities in Eq. (1) are related Fourier transforms, in particular

$$v_n(x, \omega) = (2\pi)^{-1/2} \int dk \exp(-ixk) V_n(k, \omega), \quad (2a)$$

$$p_{en}(x, \omega) = (2\pi)^{-1/2} \int dk \exp(-ixk) P_{en}(k, \omega). \quad (2b)$$

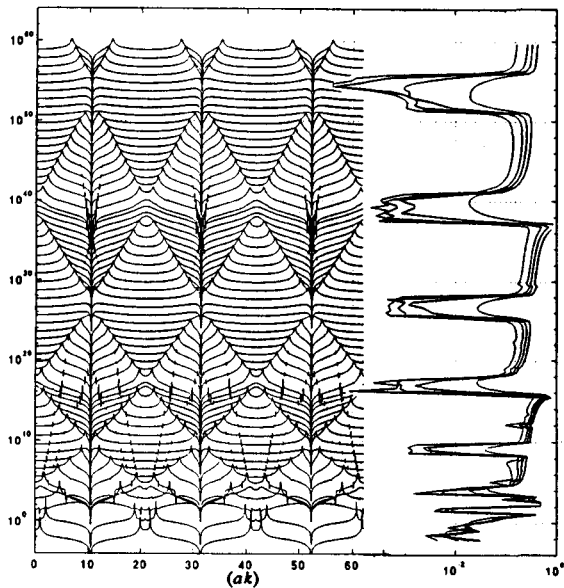
The circumferential eigenfunctions $\Phi_n(\phi)$, are assumed to be orthogonal and complete

$$\int \Phi_n(\phi) d\phi \Phi_m^*(\phi) = \delta_{nm}, \quad (3)$$

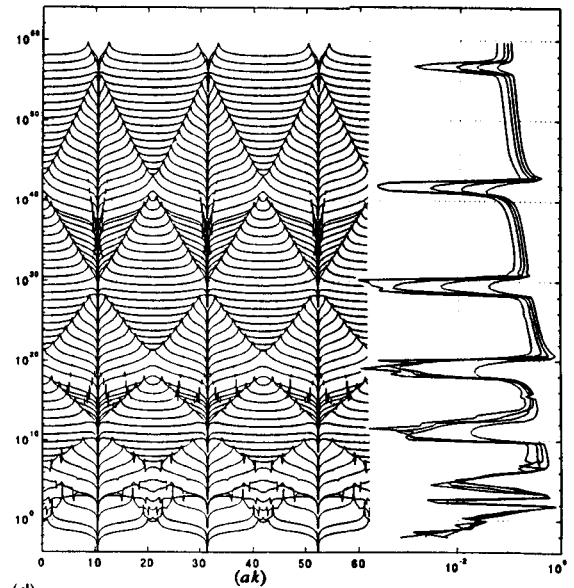
$$\sum_n \Phi_n(\phi) \Phi_n^*(\phi') = \delta(\phi - \phi'),$$

$$0 \leq (\phi - \phi') < 2\pi,$$

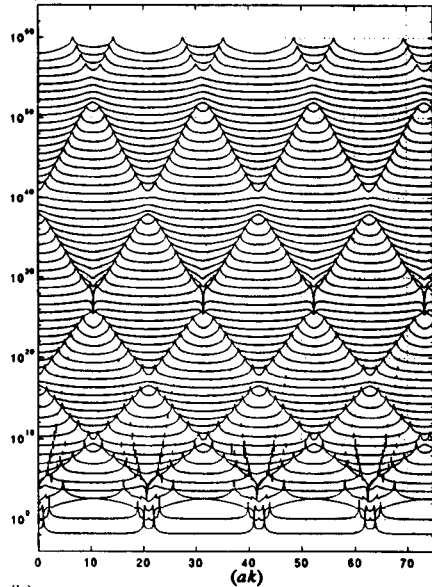
where δ_{nm} and δ are the Kronecker delta and the delta function, respectively.⁸ The modal response $v_n(x, \omega)$ of the cylinder to the modal external drive $p_{en}(x, \omega)$ in the x domain can be formally expressed in the form



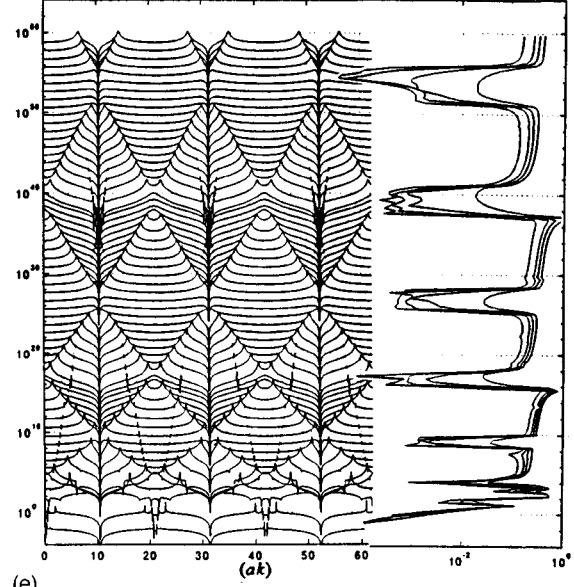
(a)



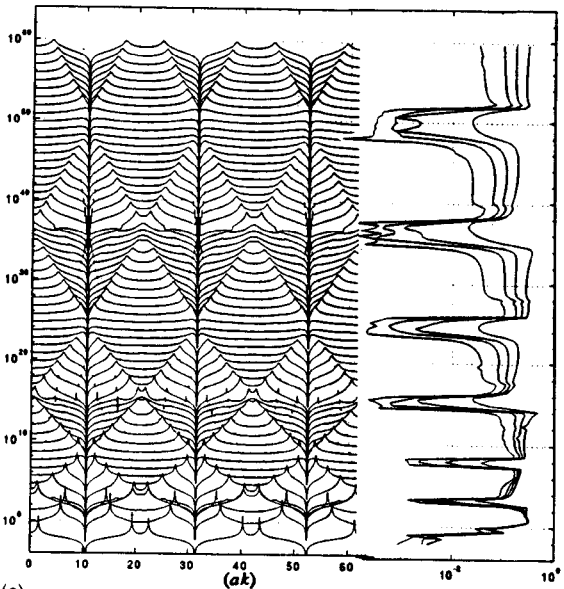
(d)



(b)



(e)



(c)

FIG. 5. Same as Fig. 4 but for a *natural* cylinder. (a) Under standard values [cf. Eq. (15)]. (b) $(x_d/b) = 0.3$. (c) $\epsilon_c = 10^{-1}$. (d) Rib's ring impedance stiffness controlled. (e) $n = 0$.

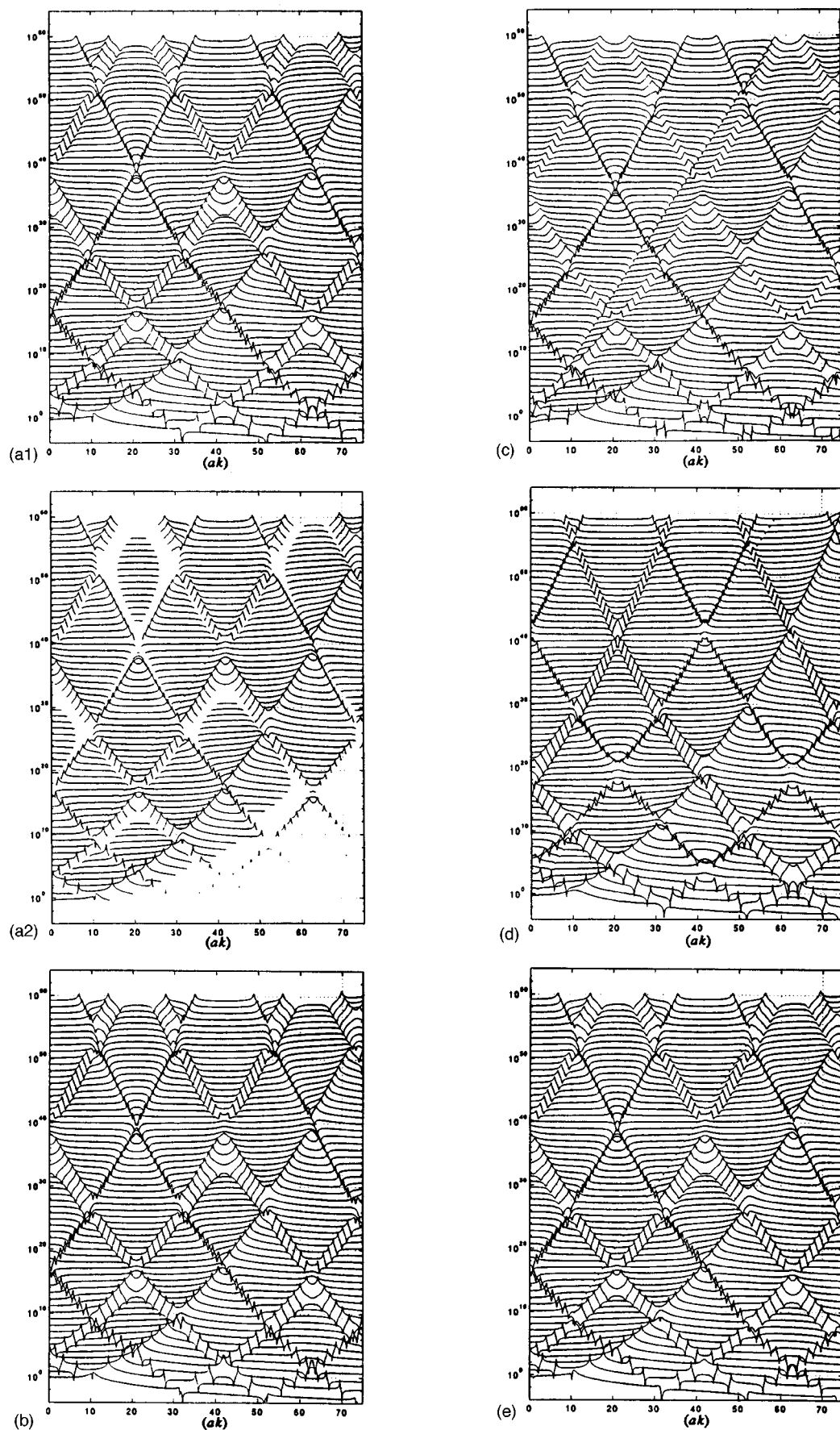


FIG. 6. Normalized response $\bar{V}_n(k, \omega)$ of a regularly ribbed *hybrid* cylinder as a function of the normalized axial wave number (ak) in a frequency waterfall format. (a) Under standard values [cf. Eq. (15)]. (a.1) Clipping of (a). (b) $\epsilon_c = 10^{-4}$. (c) $\epsilon_c = 10^{-1}$. (d) Ribs' ring impedance stiffness controlled. (e) $n = 0$.

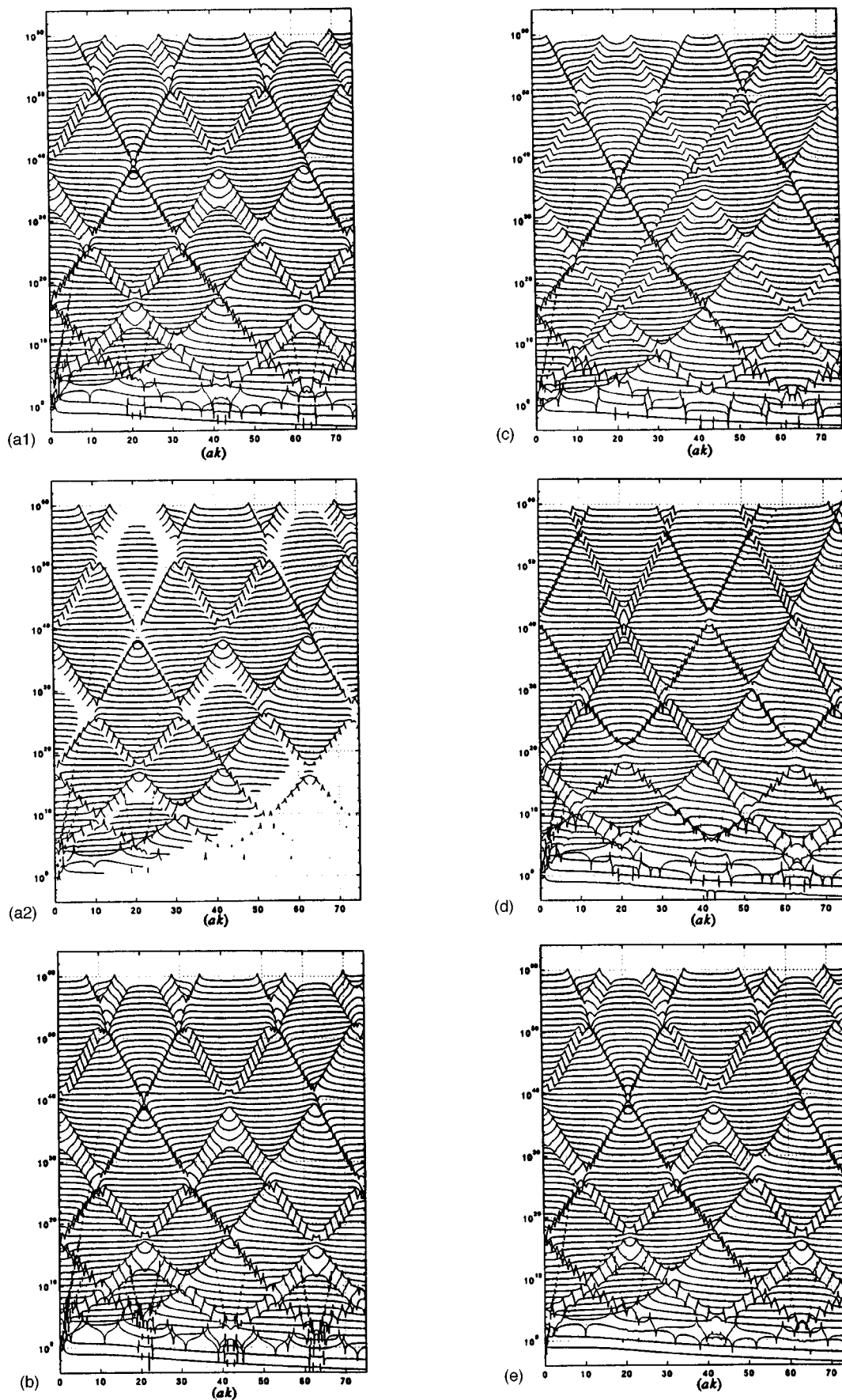


FIG. 7. Same as Fig. 6 but for a *natural* cylinder. (a) Under standard values [cf. Eq. (15)]. (a.1) Clipping of (a). (b) $\epsilon_c = 10^{-4}$. (c) $\epsilon_c = 10^{-1}$. (d) Ribs' ring impedance stiffness controlled. (e) $n = 0$.

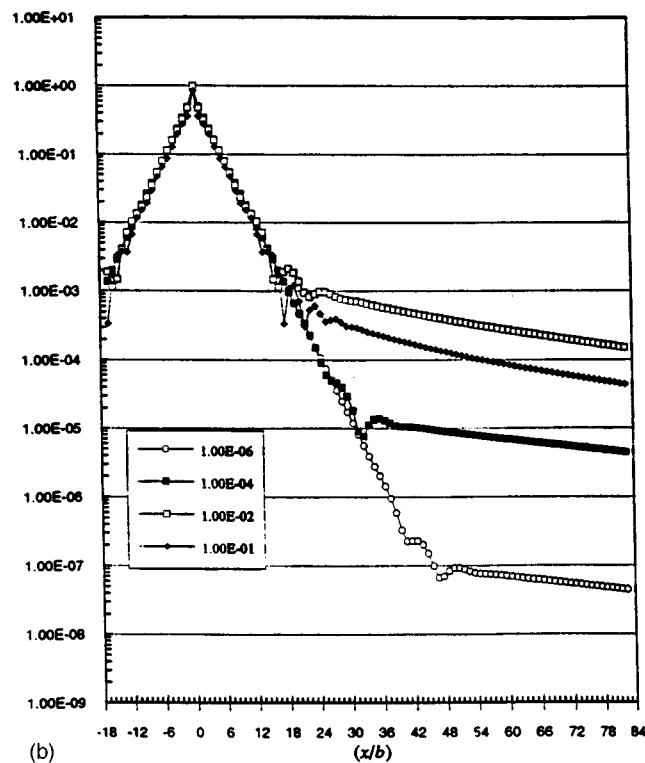
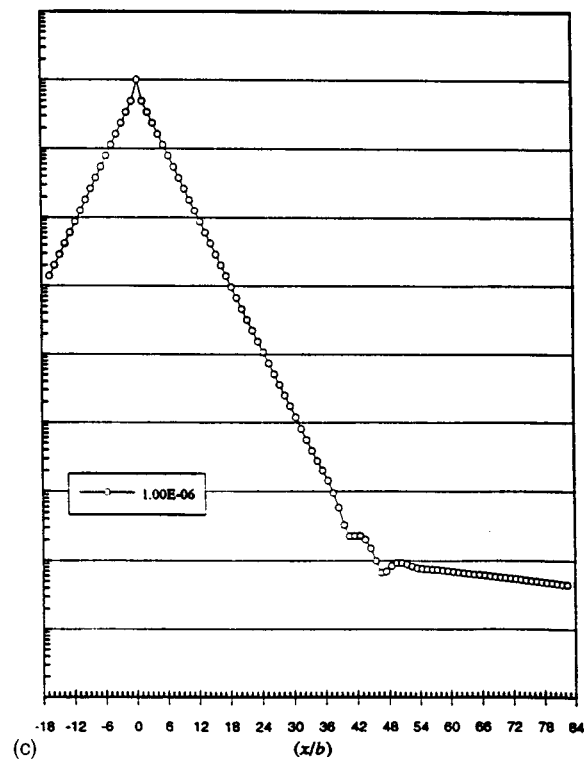
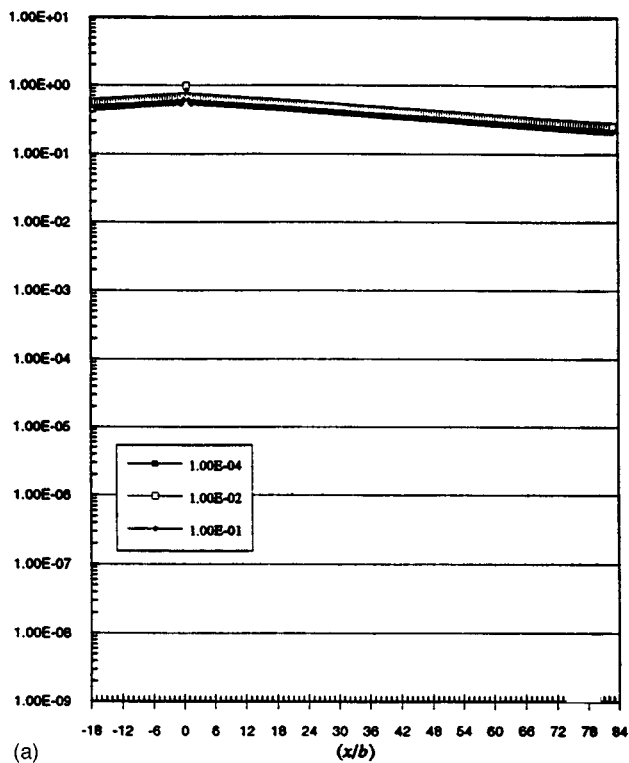


FIG. 8. Normalized modal response $\bar{v}_{\infty n}(x, \omega)$ of a uniform *hybrid* cylinder as a function of the normalized axial distance (x/b) at the normalized frequency $(\omega/\omega_c) = 0.175$. (a) Under standard values and for $\epsilon_c = 10^{-4}$, 10^{-2} , and 10^{-1} . (b) $\eta_p = 3 \times 10^{-2}$ and $\epsilon_c = 10^{-6}$, 10^{-4} , 10^{-2} , and 10^{-1} . (c) $\eta_p = 3 \times 10^{-2}$ and $\epsilon_c = 10^{-6}$.

$$v_n(x, \omega) = \int g_n(x|x', \omega) dx' p_{en}(x', \omega) \quad (4a)$$

$$V_n(k, \omega) = \int G_n(k|k', \omega) dk' P_{en}(k', \omega), \quad (4b)$$

and in the k domain in the form

where, in addition to Eq. (2), the impulse response functions in these equations are related Fourier transforms, e.g.,

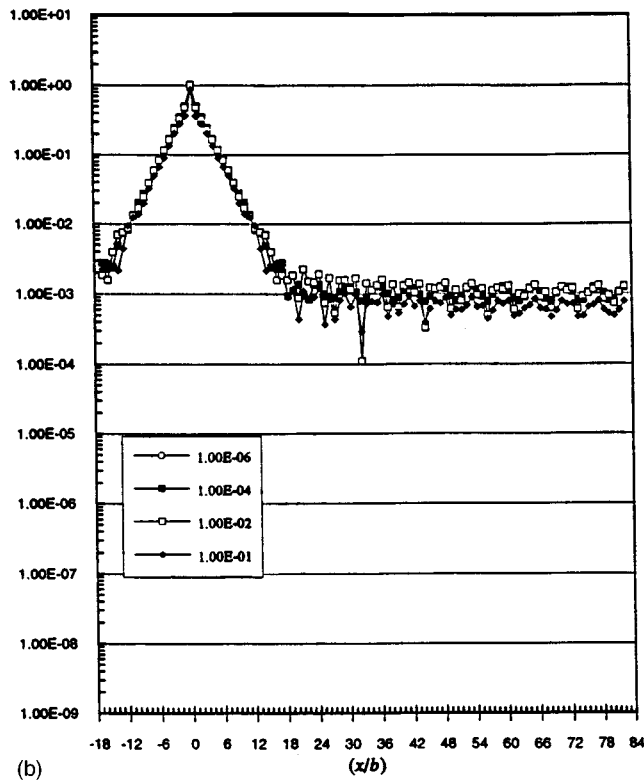
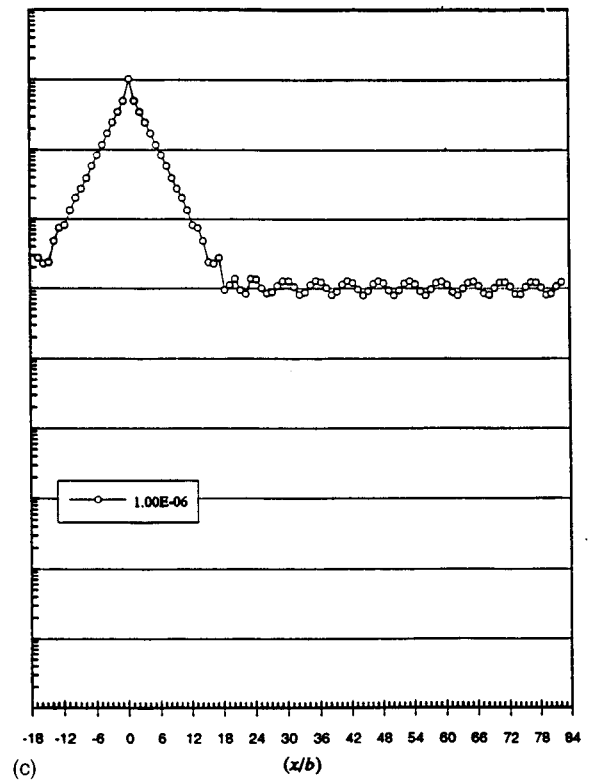
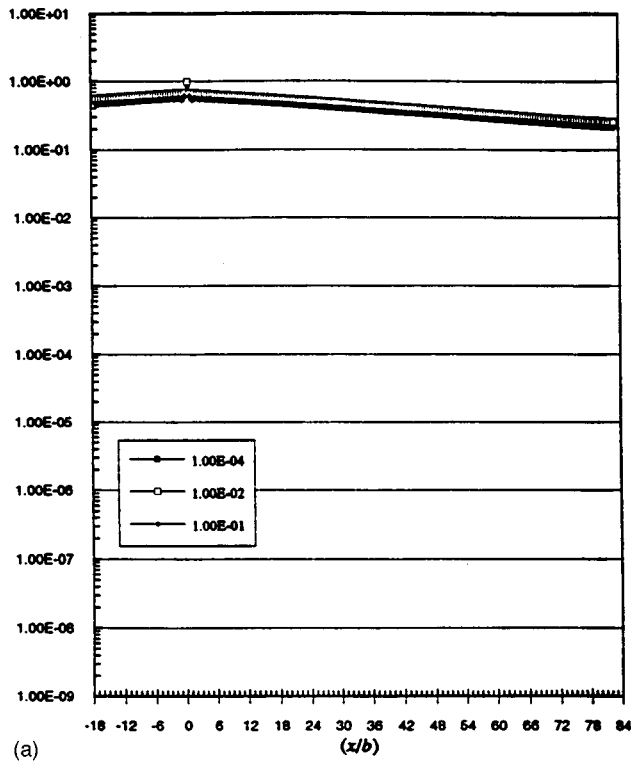


FIG. 9. Same as Fig. 8 but for a *natural* cylinder. (a) Under standard values and for $\epsilon_c = 10^{-4}$, 10^{-2} , and 10^{-1} . (b) $\eta_p = 3 \times 10^{-2}$ and $\epsilon_c = 10^{-6}$, 10^{-4} , 10^{-2} , and 10^{-1} . (c) $\eta_p = 3 \times 10^{-2}$ and $\epsilon_c = 10^{-6}$.

$$g_n(x, |x', \omega) = (2\pi)^{-1} \int dk \int dk' \times \exp[-i(xk - x'k')] G_n(k|k', \omega). \quad (5)$$

It is convenient to normalize the modal impulse response

function by the surface mass impedance ($i\omega m$) of the plating of the cylinder and the modal drive by the modal ring drive $P_{en}^0(\omega)$, namely

$$\bar{G}_n(k|k', \omega) = (i\omega m) G_n(k|k', \omega),$$

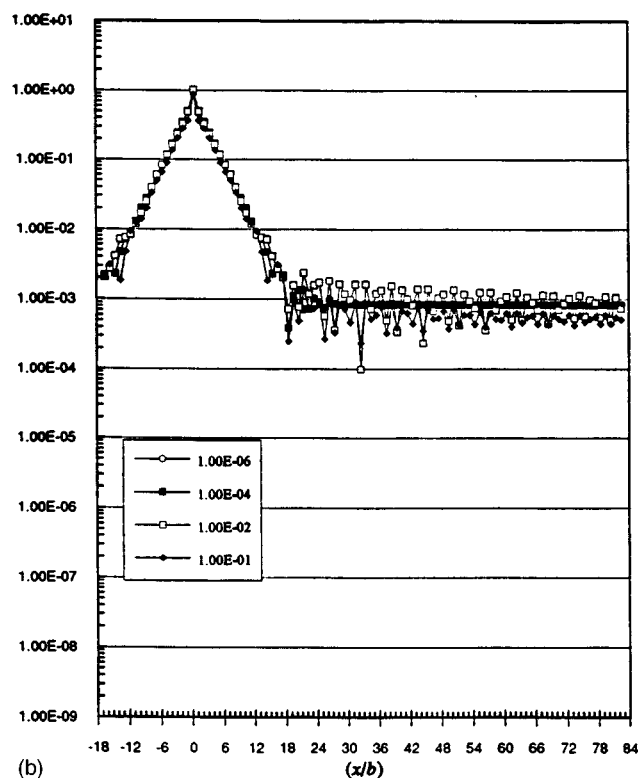
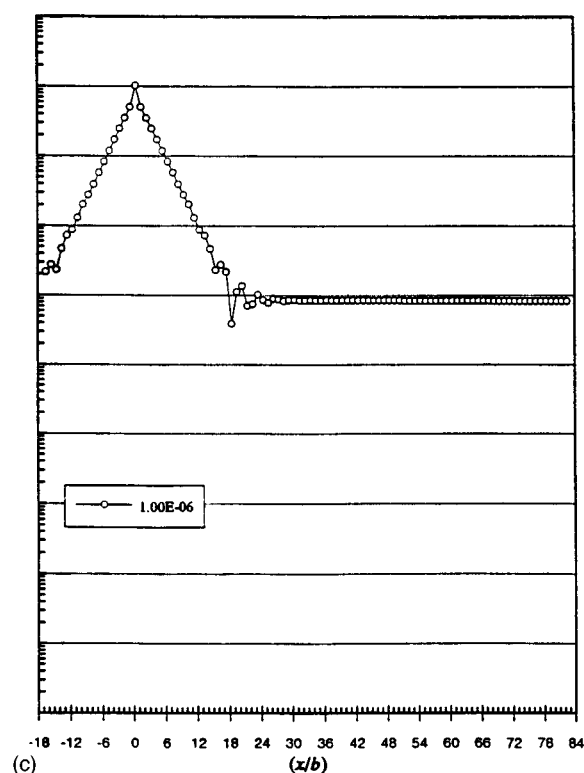
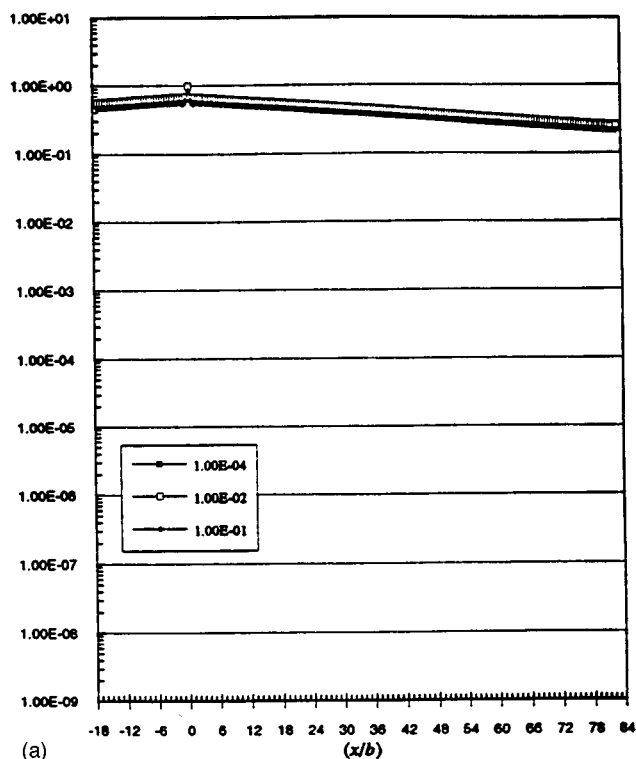


FIG. 10. Same as Fig. 9 except (n) is changed from the standard value of unity to zero. (a) $\epsilon_c = 10^{-4}$, 10^{-2} , and 10^{-1} . (b) $\eta_p = 3 \times 10^{-2}$ and $\epsilon_c = 10^{-6}$, 10^{-4} , 10^{-2} , and 10^{-1} . (c) $\eta_p = 3 \times 10^{-2}$ and $\epsilon_c = 10^{-6}$.

$$\bar{P}_{en}(k, \omega) = [P_{en}(k, \omega) / P_{en}^0(\omega)] = \exp(ix_a k), \quad (6a)$$

where (m) is the surface mass of the plating and (x_a) is the position of the ring to which the modal external drive is applied. The normalization of the modal response $V_n(k, \omega)$ follows

$$\bar{V}_n(k, \omega) = [V_n(k, \omega) / V_{n0}(\omega)],$$

$$V_{n0}(\omega) = [P_{en}^0(\omega) / (i\omega m)]. \quad (6b)$$

The modal impulse response function $\bar{G}_n(k|k', \omega)$ for a regularly ribbed cylinder is readily derived in the k domain in the simple operator form

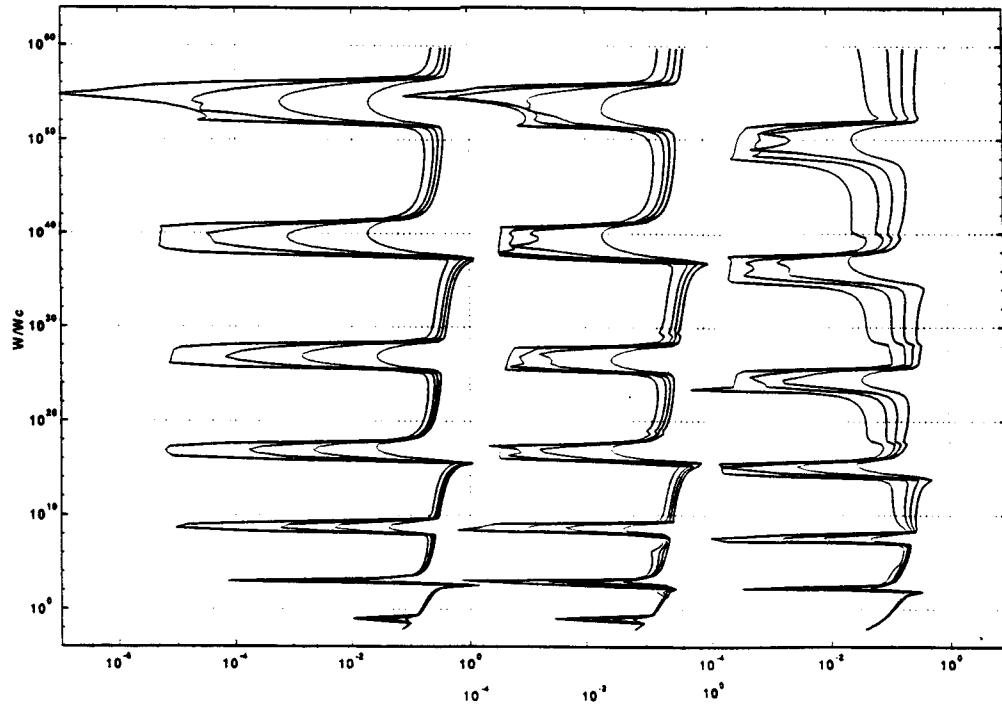


FIG. 11. Normalized modal drive $\bar{p}_{sn}(x, \omega)$ in lieu of the ribs on a *hybrid* cylinder as a function of the normalized frequency (ω/ω_c) for the normalized axial distances $(x/b) = 7, 14, 21$, and 35 and for $\epsilon_c = 10^{-4}, 10^{-2}$, and 10^{-1} .

$$\bar{G}_n(k|k', \omega) = \bar{G}_{\infty n}(k, \omega) [1 - S_b(k) \times \{I_{bn}(k, \omega) Q_{\infty n}(k, \omega)\} \delta(k - k')], \quad (7)$$

where

$$Q_{\infty n}(k, \omega) = [Z_{sn}(\omega)/b] G_{\infty n}(k, \omega), \quad (8a)$$

$$I_{bn}(k, \omega) = [1 + S_b(k) \{Q_{\infty n}(k, \omega)\}]^{-1}, \quad (8b)$$

the wave-number operator $S_b(k)$ is defined in Eq. (1-3), (b) is the separation distance between adjacent ribs, $Z_{sn}(\omega)$ is the modal ring impedance of the ribs, and $\bar{G}_{\infty n}(k, \omega)$ is the normalized modal surface admittance of the uniform (unribbed) cylinder. The incomplete curly brackets preceded by the wave-number operator $S_b(k)$ in Eq. (7) merely indicates that the wave-number operation to be per-

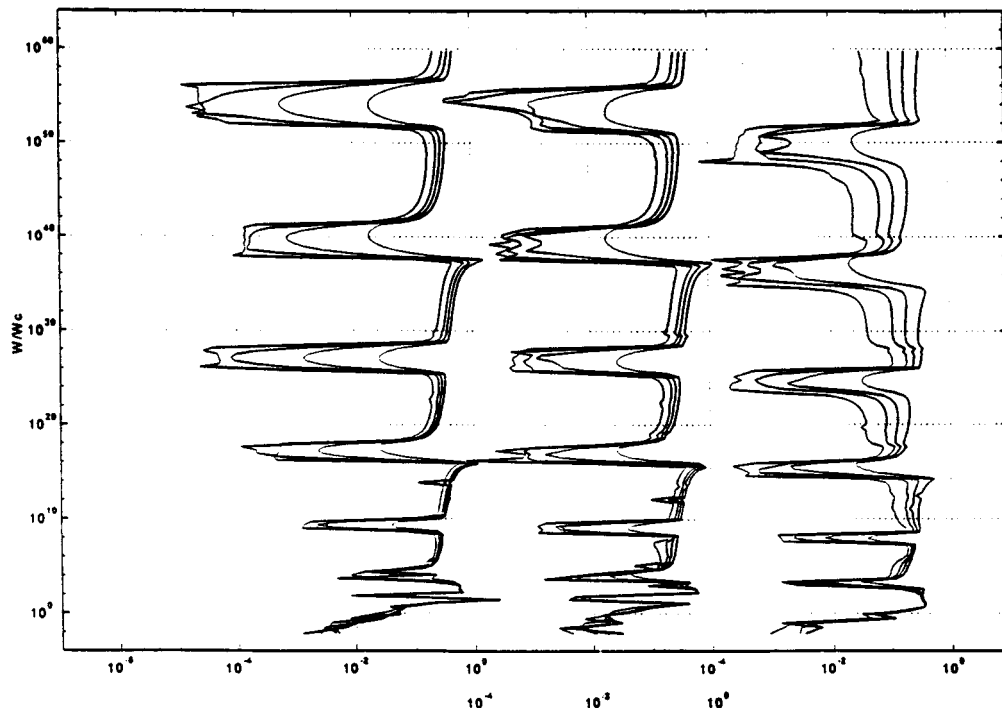


FIG. 12. Same as Fig. 11 but for a *natural* cylinder.

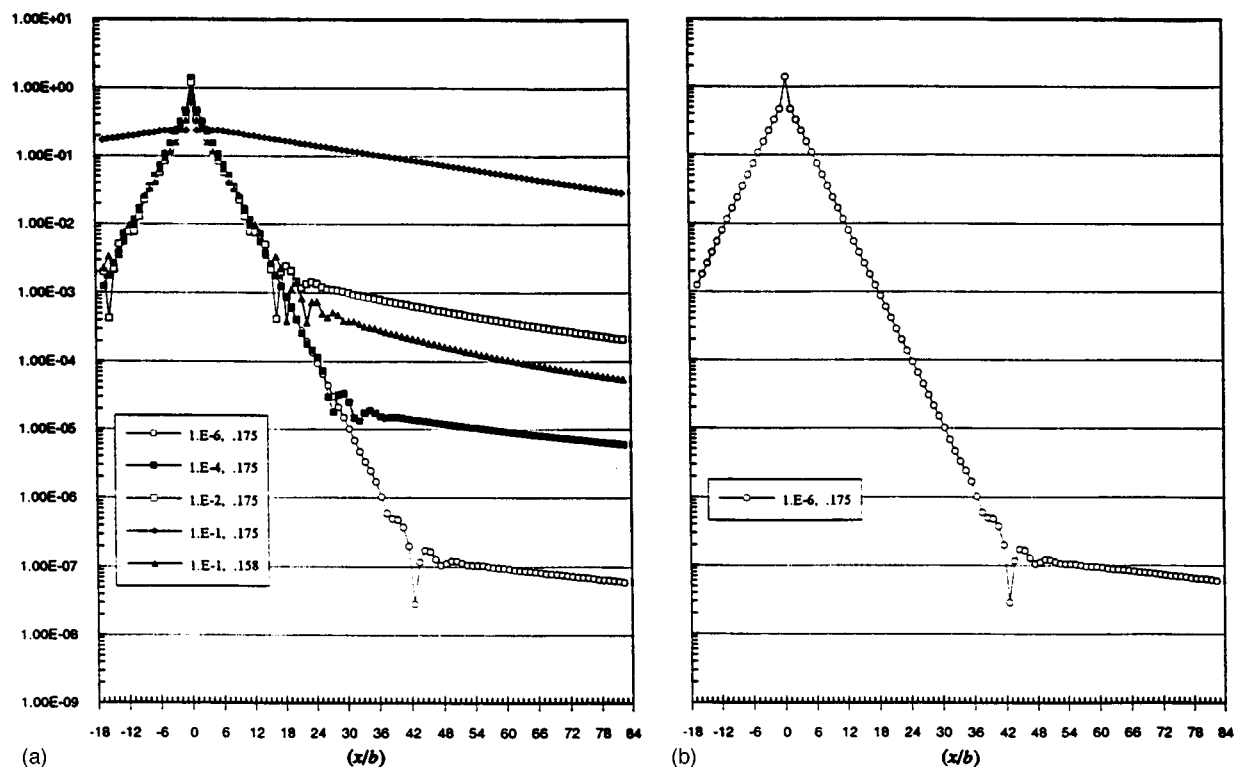


FIG. 13. Normalized modal response $\bar{v}_n(x, \omega)$ of a regularly ribbed *hybrid* cylinder as a function of the normalized axial distance (x/b) . (a) $(\omega/\omega_c) = 0.175$ with $\epsilon_c = 10^{-6}, 10^{-4}, 10^{-2}$, and 10^{-1} and $(\omega/\omega_c) = 0.158$ with $\epsilon_c = 10^{-1}$. (b) $(\omega/\omega_c) = 0.175$ with $\epsilon_c = 10^{-6}$.

formed is yet to be completed; indeed, $\bar{G}_n(k|k', \omega)$ is a wave-number operator and not just an algebraic factor. The normalized quantity $\bar{G}_{\infty n}(k, \omega)$ is the inverse of three terms

$$\begin{aligned} \bar{G}_{\infty n}(k, \omega) &= (i\omega m)G_{\infty n}(k, \omega) \\ &= [\bar{Z}_n^p(k, \omega) + \bar{Z}_n^m(k, \omega) + \bar{Z}_n^f(k, \omega)]^{-1}, \end{aligned} \quad (9a)$$

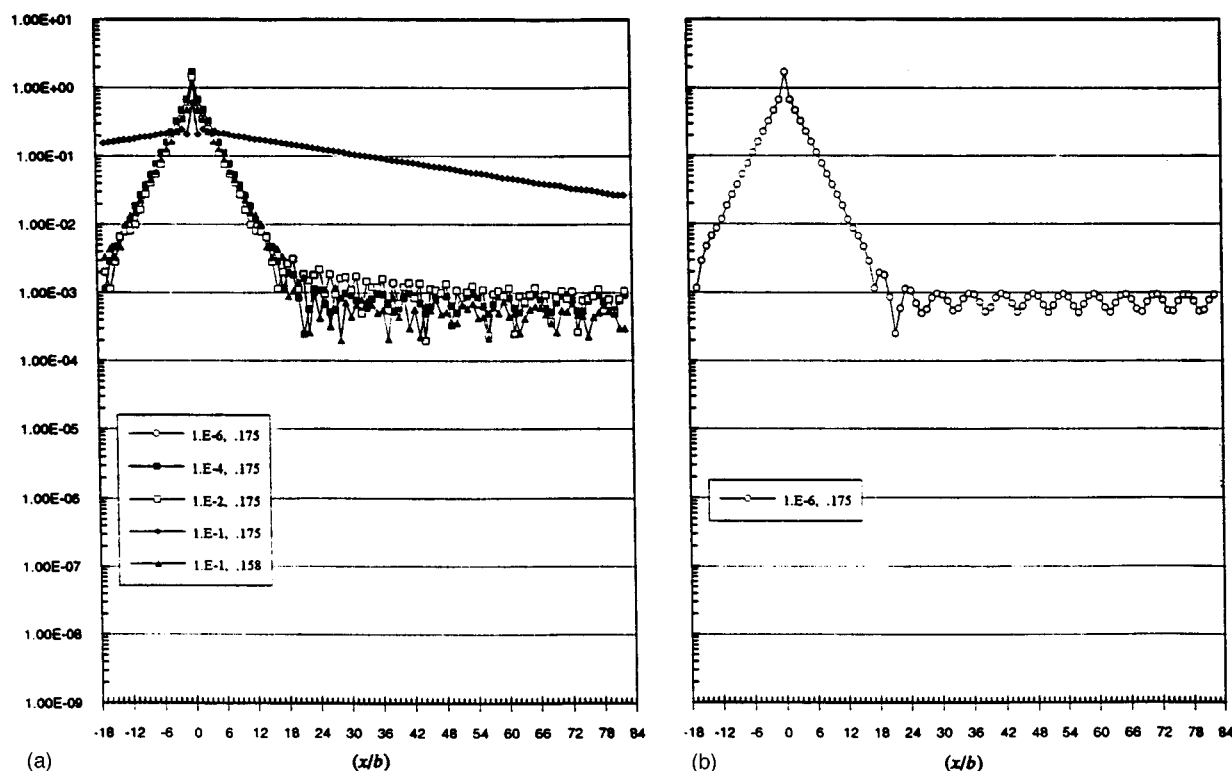


FIG. 14. Same as Fig. 13 but for a *natural* cylinder. (a) $(\omega/\omega_c) = 0.175$ with $\epsilon_c = 10^{-6}, 10^{-4}, 10^{-2}$, and 10^{-1} and $(\omega/\omega_c) = 0.158$ with $\epsilon_c = 10^{-1}$. (b) $(\omega/\omega_c) = 0.175$ with $\epsilon_c = 10^{-6}$.

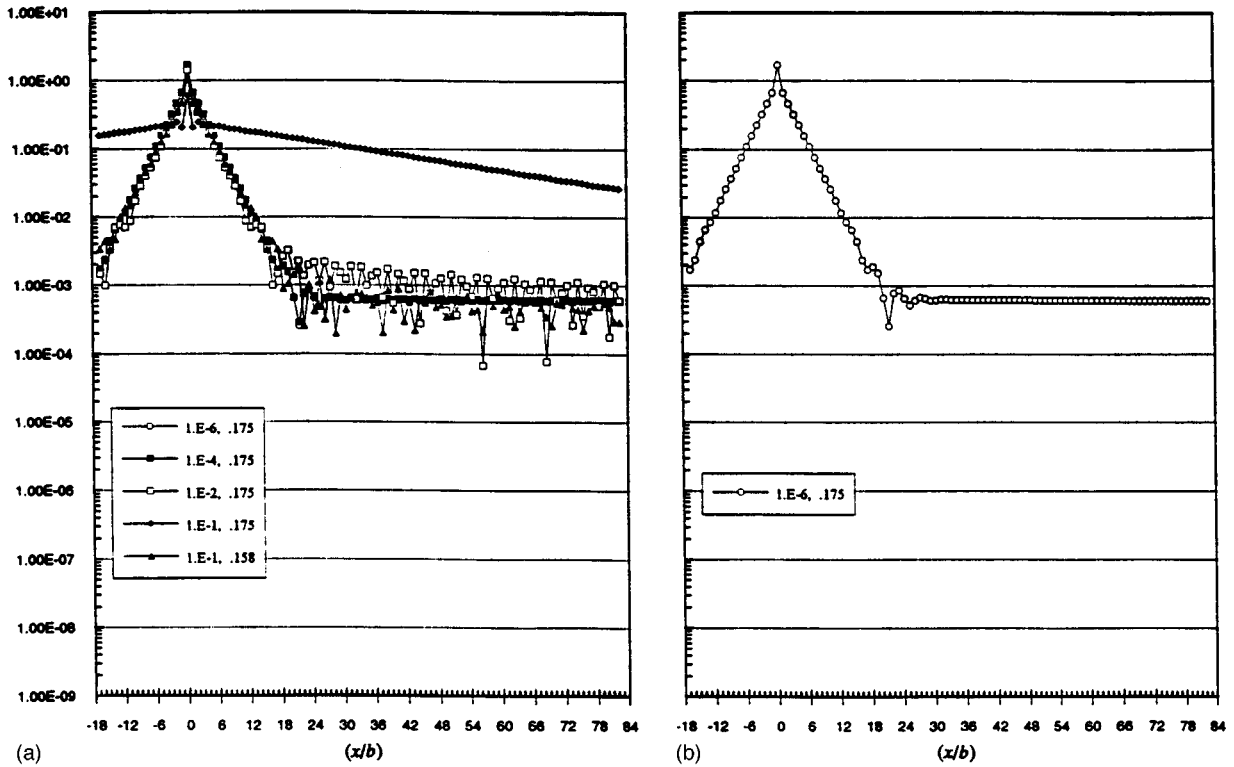


FIG. 15. Same as Fig. 14 except (n) is changed from the standard value of unity to zero. (a) $(\omega/\omega_c) = 0.175$ with $\epsilon_c = 10^{-6}, 10^{-4}, 10^{-2}$, and 10^{-1} and $(\omega/\omega_c) = 0.158$ with $\epsilon_c = 10^{-1}$. (b) $(\omega/\omega_c) = 0.175$ with $\epsilon_c = 10^{-6}$.

where $\bar{Z}_n^p(k, \omega)$, $\bar{Z}_n^m(k, \omega)$, and $\bar{Z}_n^f(k, \omega)$ are the normalized flexural, membrane, and fluid modal surface impedances, respectively, of the uniform cylinder. A hybrid cylinder is defined by a (natural) cylinder from which the membrane free waves are analytically removed by setting $\bar{Z}_n^m(k, \omega)$ equal to zero in Eq. (9a). A hybrid cylinder is defined by a normalized modal surface admittance

$$\bar{G}_{\infty n}(k, \omega) = [\bar{Z}_n^p(k, \omega) + \bar{Z}_n^f(k, \omega)]^{-1}, \quad \bar{Z}_n^m(k, \omega) = 0. \quad (9b)$$

Recognizing that the mechanical surface impedance $\bar{Z}_p(k, k_y, \omega)$ for a uniform panel relates to $\bar{Z}_n^p(k, \omega)$ in the form

$$\bar{Z}_p[k, (n/a), \omega] \equiv \bar{Z}_n^p(k, \omega), \quad (10)$$

and that the fluid surface impedance $\bar{Z}_f[k, (n/a), \omega]$ on a panel qualitatively relates to $\bar{Z}_n^f(k, \omega)$, it is inferred that a hybrid cylinder is akin to a panel, and in some cases, the results obtained on a hybrid cylinder can be inferentially translated to those on a panel.² The quantity $Q_{\infty n}(k, \omega)$ is the ratio of the equivalent modal surface impedance $[Z_{sn}(\omega)/b]$ of a rib to the modal surface impedance $[G_{\infty n}(k, \omega)]^{-1}$ of the uniform cylinder. It is convenient to express $Z_{sn}(\omega)$ in the form

$$Z_{sn}(\omega) = (i\omega M)A_n(\omega/\omega_n), \quad (11)$$

and hence from Eq. (8a) one derives

$$Q_{sn}(k, \omega) = (M/mb)A_n(\omega/\omega_n), \quad (8c)$$

where M is a line mass and $A_n(\omega/\omega_n)$ is a resonance bearing factor. In this paper $A_n(\omega/\omega_n)$ is equivalently set either to

unity or to $-[1 + (\alpha_n \omega/\omega_n)^2]^{-1}$, the ring impedance of the ribs in the former setting is mass controlled and in the latter stiffness controlled.² Substituting Eq. (7) in Eq. (4b) one derives

$$\bar{V}_n(k, \omega) = \bar{V}_{\infty n}(k, \omega) - \bar{V}_{sn}(k, \omega), \quad (12)$$

$$\bar{V}_{\infty n}(k, \omega) = \bar{G}_{\infty n}(k, \omega) \exp(ix_a k), \quad (13a)$$

$$\bar{V}_{sn}(k, \omega) = \bar{G}_{\infty n}(k, \omega) \bar{P}_{sn}(k, \omega), \quad (13b)$$

where

$$\begin{aligned} \bar{P}_{sn}(k, \omega) &= [P_{sn}(k, \omega)/P_{en}^0(\omega)] \\ &= S_b(k) \{I_{bn}(k, \omega) Q_{\infty n}(k, \omega) \exp(ix_a k)\}, \end{aligned} \quad (14a)$$

is the normalized modal drive in lieu of the ribs. Clearly, $\bar{P}_{sn}(k, \omega)$ is an aliased quantity in (k) with respect to (κ_1) ; $\kappa_1 = (2\pi/b)$ [cf. Eq. (1-17a)]. Moreover, one recognizes that $I_{bn}(k, \omega)$ as stated in Eq. (8b) is similarly aliased in (k) with respect (κ_1) [cf. Eq. (1-19a)]. In consequence $\bar{P}_{sn}(k, \omega)$ can be factorized in the form

$$\bar{P}_{sn}(k, \omega) = I_{bn}(k, \omega) S_b(k) \{Q_{\infty n}(k, \omega) \exp(ix_a k)\} \quad (14b)$$

[cf. Eq. (1-18a)]. It is apparent from Eqs. (12)–(14) that the behavior of the modal response $\bar{V}_n(k, \omega)$ and a number of terms and factors in its composition; e.g., the modal response $\bar{V}_{\infty n}(k, \omega)$ in the absence of ribs and the drive $\bar{P}_{sn}(k, \omega)$ in lieu of the ribs, are merely functional of the properties of the fluid loaded uniform cylinder, the ribs, and the external drive. These properties are described by the modal surface

admittance $G_{\infty n}(k, \omega)$ of the uniform cylinder, the modal line impedance $Z_{sn}(\omega)$, and the separation distance (b) of the ribs and the ring position (x_a) of the external drive. The absolute values of the quantities $\bar{V}_{\infty n}(k, \omega)$ or equivalently $\bar{G}_{\infty n}(k, \omega)$, $\bar{P}_{sn}(k, \omega)$, and $\bar{V}_n(k, \omega)$, as functions of (ak) , are computed and presented in typical frequency waterfall displays in Figs. 2 and 3, 4 and 5, and 6 and 7, respectively.¹² The even numbered figures in each set, e.g., Fig. 2, relate to a hybrid cylinder and the odd numbered, e.g., Fig. 3, relate to a natural cylinder. A typical case is defined by standard parametric values. These standard values may also serve to define other cases by merely selecting changes in specific standard values. The standard parametric values and their definitions follow: the flexural-free waves are defined by a wave number k_p ,

$$k_p = k_{p0}(1 - i\eta_p), \quad k_{p0}^2 = (\omega\omega_c/c^2), \quad (15a)$$

$$(h/a) = 10^{-2}, \quad \eta_p = 10^{-3},$$

the longitudinal free waves are defined by a wave number k_l ,

$$k_l = k_{l0}(1 - i\eta_l), \quad k_{l0} = (\omega/c_l), \quad c_l^2 = (Y/\rho_s)(1 - \nu^2)^{-1}, \quad (15b)$$

$$m = h\rho_s, \quad (c_l/c) = 3.5, \quad \eta_l = 5 \times 10^{-4},$$

the shear free waves are defined by a wave number k_s ,

$$k_s = k_{s0}(1 - i\eta_s), \quad k_{s0} = (\omega/c_s), \quad (c_s/c_l)^2 = [(1 - \nu)/2], \quad (15c)$$

$$\nu = (1/3), \quad \eta_s = 5 \times 10^{-4},$$

the fluid loading is defined by a fluid loading parameter ϵ_c ,

$$\epsilon_c = (\rho c/\omega_c m), \quad (a\omega_c/c) = 97.53, \quad \epsilon_c = 10^{-2}, \quad (15d)$$

the ring frequency ω_r and the critical frequency ω_c are defined

$$\omega_r = (c_l/a), \quad \omega_c = \sqrt{12}(c^2/hc_l), \quad (15e)$$

$$(\omega_r/\omega_c) = 3.6 \times 10^{-2},$$

and the definition of the ribs and mode index (n) are specified

$$(b/a) = 0.3, \quad (M/mb) = 0.2, \quad (x_a/b) = 0.5, \quad (15f)$$

$$A_n(\omega/\omega_n) = 1, \quad \eta_0 = 0.3, \quad n = 1,$$

where (a) is the radius, Y is Young's modulus, (ν) is the Poisson's ratio, and (ρ_s) is the density of the plating of the cylinder, the loss factors η_p , η_l , and η_s describe propagation losses, (ρ) and (c) are the density and sound speed in the fluid. The interest in this paper is confined to the spectral range

$$\{0, (2\omega_r/\omega_c)\} \lesssim \{(ak), (\omega/\omega_c)\} \lesssim \{75, 0.6\}, \quad (15g)$$

notwithstanding that some displays are extended, but not discussed, into the frequency range $(\omega/\omega_c) \lesssim (2\omega_r/\omega_c)$.¹³ Computations and their representations in this low frequency are discussed under a separate cover.^{2,14} The first subfigures in the first and second sets; i.e., Figs. 2(a) and 3(a), are computed under standard parametric values. The flexural free

waves and the fluid free waves are clearly discernible in Fig. 2(a). In Fig. 3(a), in addition to the flexural and fluid free waves, membrane free waves are clearly discernible. The membrane free waves are composed of longitudinal and shear free waves as well as of "curvature free waves," in the low-frequency range and the low-mode indices, the curvature free waves supersede the flexural free waves. The curvature free waves and the higher surface impedances and origins of the longitudinal and shear free waves are all found largely to reside in Fig. 3(a), in the low-frequency range; $(\omega/\omega_c) \lesssim (2\omega_r/\omega_c)$.¹⁴ Again, discussion of these features are not conducted in this paper.² Figures 2(a) and 3(a) are repeated in Figs. 2(b) and (c) and 3(b) and (c), respectively, except that the fluid loading parameter (ϵ_c) is changed from the standard value of 10^{-2} to 10^{-4} and 10^{-1} , respectively. In this connection it is observed that a fluid loading with $\epsilon_c = 10^{-4}$ is light, with $\epsilon_c = 10^{-2}$ is moderate, and with $\epsilon_c = 10^{-1}$ is heavy. The change of (ϵ_c) from 10^{-2} to 10^{-4} renders the fluid free waves hardly discernible in Figs. 2(b) and 3(b), little else is different between these figures and Figs. 2(a) and 3(a), respectively. The change of (ϵ_c) from 10^{-2} to 10^{-1} intensifies the fluid free waves in Fig. 2(c) and 3(c). Moreover, this change in fluid loading shifts the free waves, especially the subsonic flexural waves, to a higher wave-number range and weakens the strengths of the supersonic membrane free waves. The shifts are due to the "added surface mass" and the weakening is due to the "radiation damping" effects associated with the fluid loading. Figures 2(d) and 3(d) repeat Figs. 2(a) and 3(a), respectively, except that the flexural loss factor (η_p) is changed from the standard value of 10^{-3} to 3×10^{-2} ; a substantial increase in the damping of the flexural free waves. Indeed, the flexural free waves in Figs. 2(d) and 3(d) are substantially subdued compared with those in Figs. 2(a) and 3(a). Figures 2(a) and 3(a) are repeated in Figs. 2(e) and 3(e), respectively, except that the mode index (n) is changed from the standard value of unity to zero. This change hardly induces any change in the flexural free waves; however, a drastic change is induced in the membrane free waves. The membrane free waves in Fig. 3(a) are comprised of longitudinal and shear free waves, in Fig. 3(e) of longitudinal free waves only. The shear free waves join the longitudinal when the mode index (n) equals or exceeds unity. Thus, in Fig. 3(f), which repeats Fig. 3(a) except that the mode index (n) is set equal to (2), the membrane free waves are composed of both the longitudinal and shear free waves. Finally, the first set is completed by Fig. 3(g). This figure repeats Fig. 3(e), except that the flexural loss factor (η_p) is changed from the standard value of 10^{-3} to 3×10^{-2} . Again, the flexural free waves are substantially subdued by this change [cf. Figs. 2(d) and 3(d)].

The first subfigures in the third and fourth sets, i.e., Figs. 4(a) and 5(a), are computed under standard parametric values. The aliasing of the normalized modal drive $\bar{P}_{sn}(k, \omega)$ in lieu of the ribs is apparent in these figures; the dominant patterns are those associated with the flexural free waves. Nonetheless, aliasing in the membrane free waves are discernible in Fig. 5(a). Aliasing is basically a scattering phenomenon and although the aliasing patterns of the membrane free waves are scant, the scattering of the membrane free

waves by “simple” ribs are clearly established in Fig. 5(a). In part, the visibility of the membrane free waves in Fig. 5(a) is aided by the locus minima in $\bar{P}_{sn}(k, \omega)$. These locus minima occur in $\bar{P}_{sn}(k, \omega)$, in both Figs. 4(a) and 5(a), whenever $(ak) = \pm (a\kappa_1)[(2j+1)/2]$ and is due to the symmetry in this quantity. The symmetry results from the standard position of the external drive at $(x_a/b) = 0.5$. This symmetry is removed in Figs. 4(b) and 5(b). These figures repeat Figs. 4(a) and 5(a), respectively, except that (x_a/b) is changed from 0.5 to 0.3. The locus of minima are absent in the patterns of Figs. 4(b) and 5(b). Moreover, the patterns relating to the membrane free waves in Fig. 5(b) are weakened. The fluid free waves are hardly discernible in Figs. 4(a) and (b) and 5(a) and (b). A major feature in the patterns of these figures is the cancellations between the peaks of a pair of flexural aliasing orders in the region in which they cross. Regions of these kind lie regularly in certain frequency bands. The cancellations are made possible by the interactions of the ribs among themselves via the cylindrical shells; in the absence of such interactions the cancellations are also absent.^{2,15} Figures 4(a) and 5(a) are repeated in Figs. 4(c) and 5(c), respectively, except that the fluid loading parameter is increased from the standard value of 10^{-2} to 10^{-1} . The influence of this increase is largely discernible in the shifts in the frequency bands at which pairs of various flexural aliasing orders cross; the shifts are to lower frequencies and are consistent with the wave-number shifts observed between Figs. 2(a) and 3(a) and Figs. 2(c) and 3(c). In addition, the flexural free waves in both figures and the membrane free waves in Fig. 5(c) are subdued by this increase in the fluid loading. Remembering that under the influence of ribs the subsonic flexural free waves can become reasonable radiators, the subdued flexural and, again, the membrane free waves can be attributed to the increase in radiation damping with the increase in fluid loading. However, the fluid free waves, as such, are still hardly discernible in Figs. 4(c) and 5(c) [cf. Figs. 4(a) and (b) and 5(a) and (b), respectively]. Figures 4(d) and 5(d) repeat Figs. 4(a) and 5(a), except that the standard mass controlled ring impedance of the ribs, i.e., with $A_n(\omega/\omega_n) = 1$, is changed to that of stiffness controlled, i.e., with $A_n(\omega/\omega_n) = -[1 + (5\omega/\omega_c)^2]^{-1}$. The change is largely discernible in the shifts in the frequency bands at which pairs of various flexural aliasing orders cross; the shifts are to higher frequencies. Otherwise the influence of the change on patterns of the aliased modal drive $\bar{P}_{sn}(k, \omega)$ is insignificant, notwithstanding that the cancellations are weaker and narrower at the crossings of pairs of various flexural aliasing orders as the ring impedance of the ribs is diminished, by definition, at the higher frequency ranges. Figure 5(a) is repeated in Fig. 5(e), except that the mode index (n) is changed from the standard value of unity to zero. The absence of the shear free waves in Fig. 5(e) is consistent with Fig. 3(e).

The first subfigures in the fifth and sixth sets; i.e., Figs. 6(a) and 7(a), are computed under standard parametric values. The patterns in the normalized modal response $\bar{V}_n(k, \omega)$ manifest footprints of the aliasing in the normalized drive $\bar{P}_{sn}(k, \omega)$. However, since neither the term

$\bar{P}_{en}(k, \omega) [= \exp(ix_a k)]$ nor the factor $\bar{G}_{\infty n}(k, \omega)$, in Eqs. (13) and (14), are aliased in (k) with respect to (κ_1) , the aliasing in $\bar{P}_{sn}(k, \omega)$ is spoiled in $\bar{V}_n(k, \omega)$; the quantity $\bar{V}_n(k, \omega)$ is not strictly aliased. The “ribs” free term” $\bar{V}_{\infty n}(k, \omega)$ in this quantity accentuates and reemphasizes the zeroth-order patterns and the factor $\bar{G}_{\infty n}(k, \omega)$ in the “ribs” term” $\bar{V}_{sn}(k, \omega)$ subdues the aliased patterns in $\bar{P}_{sn}(k, \omega)$ once $[(ak)^2 + (n/a)^2] > (k_{po})^2$ is satisfied. The subduing increases the more this inequality is satisfied. The first of these effects is manifested in that the patterns in Figs. 2(a) and 3(a), are, respectively, superimposed in Figs. 4(a) and 5(a). The second is manifested in that the patterns in Figs. 4(a) and 5(a) are suppressed in Figs. 6(a) and 7(a) in spectral ranges that lie beyond the flexural free waves. The second effect is made clearer in Figs. 6(a.1) and 7(a.1). These figures are the clipped version of Figs. 6(a) and 7(a), respectively; values that lie below a given threshold in Figs. 6(a) and 7(a) are removed in Figs. 6(a.1) and 7(a.1).^{2,3} For the sake of completeness Figs. 6(a) and 7(a) are repeated in Figs. 6(b)–(d) and 7(b)–(e), with changes that correspond to those made in Figs. 4(b)–(d) and 5(b)–(e) with respect to Figs. 4(a) and 5(a), respectively. The features in the patterns of Figs. 6(a) and 7(a) drawn from the patterns in Figs. 2(a) and 4(a) and 3(a) and 5(a), respectively, can be generalized to the other subfigures in these categories; e.g., the patterns in Fig. 7(c) are superimposed of patterns in Fig. 3(c) and of patterns in Fig. 5(c). The patterns in the latter figure are, however, subdued in spectral ranges that lie beyond the flexural free waves. Thus, changes in the parametric values and accompanied influences in Figs. 2–5, if correspondingly selected, impose similar influences in Figs. 6 and 7.

The derivation and manipulation of the modal response $\bar{V}_n(k, \omega)$ and its composition are found to be relatively simple in the k domain. The algebraic nature of the modal response $\bar{V}_{\infty n}(k, \omega)$ induced by the normalized modal external drive $\bar{P}_{en}(k, \omega)$ in the absence of the ribs, the harmonic nature of the drive $\bar{P}_{sn}(k, \omega)$ in lieu of the ribs and, again, the algebraic nature of the modal response $\bar{V}_{sn}(k, \omega)$ induced by this drive, render efficacious the k domain for formulating the modal response $\bar{V}_n(k, \omega)$. The descriptions of this quantity and terms and factors in its composition are presented in Figs. 2–7. These descriptions demonstrate that the flexural free waves are the dominant response contributors, that the fluid free waves are weak contributors even when fluid loading is heavy and, finally, although the membrane free waves are discernible and they are evidently scattered by the ribs, the membrane free waves are nonetheless weak contributors to the response, especially at the higher frequency range, where $(2\omega_r/\omega_c) \leq (\omega/\omega_c)$. In Figs. 2(f) and 3(f) and (g) the flexural free waves are substantially weakened by increasing the flexural loss factor (η_p) by more than an order of magnitude. The other free waves remain substantially unaltered so that they are relatively accentuated. The quantitative assessments of the weakening of the flexural free waves and the accentuation, thereby, of the other free waves are not readily afforded by Figs. 2 and 3 which include Figs. 2(f) and 3(f) and (g). In this light, can a Fourier transformation of $\bar{V}_{\infty n}(k, \omega)$ into the x domain assist in these assessments?

II. MODAL RESPONSE OF A UNIFORM CYLINDER IN THE SPATIAL DOMAIN

The expression for the normalized modal response $\bar{V}_{\infty n}(k, \omega)$ of a uniform cylinder is stated in Eq. (13a). As already observed, neither the term $\exp(ix_a k)$ nor the factor $\bar{G}_{\infty n}(k, \omega)$ are aliased and, therefore, the modal response $\bar{V}_{\infty n}(k, \omega)$ in the k domain is not aliased. The Fourier transformation of $\bar{V}_{\infty n}(k, \omega)$ into the spatial x domain affords no factorial saving performing the transformation either via Eq. (1-10b) or via Eq. (1-14); indeed, computations of this kind are relatively simple. Nonetheless, to obtain compatibility and in preparation for performing computations that are more compounded, the technique described in Eq. (1-14) is utilized to Fourier transform $\bar{V}_{\infty n}(k, \omega)$ to $\bar{v}_{\infty n}(x, \omega)$. From Eqs. (1-14) and (13a), and after an appropriate normalization, one obtains

$$\bar{v}_{\infty n}(x, \omega) \approx e_n(\omega) [\kappa_1 (1 + R)^{-1}] \sum_{r=0}^R S_b^J(k_r) \times \{\bar{G}_{\infty n}(k_r, \omega) \exp[-i(x - x_a)k_r]\},$$

$$e_n(\omega) = [(2/\pi^2 k_{p0}^2) \{(n/ak_{p0})^4 - 1\}]^{1/2}, \quad (-1)^{1/2} = (-i). \quad (16)$$

The normalization adopted in Eq. (16) yields for a hybrid cylinder in the absence of fluid loading, at the position of the external ring drive; i.e., at $(|x - x_a|/b) = 0$, a unit amplitude. In Fig. 8(a) the normalized modal response $\bar{v}_{\infty n}(x, \omega)$ is displayed as a function of (x/b) for the normalized frequency $(\omega/\omega_c) = 0.175$. The computations in this figure are performed under standard parametric values for a hybrid cylinder. Computations pertaining to changes in the fluid loading parameter (ϵ_c) from the standard value of 10^{-2} to 10^{-4} and 10^{-1} are superimposed on the original Fig. 8(a) [cf. Fig. 2(a)–(c)]. The influence of the fluid surface mass loading is revealed in this figure; the curves pertaining to $\epsilon_c = 10^{-2}$ and 10^{-4} substantially overlap, whereas the curve for $\epsilon_c = 10^{-1}$ lies just below, but consistently adjacent. There is, however, no evidence of fluid free waves in Fig. 8(a); the dominance of the flexural free waves appear overwhelming. This figure is repeated in Figs. 9(a) and 10(a) except that a natural cylinder is substituted for the hybrid cylinder and the mode index (n) is selected to be unity and zero, respectively. Figures 9(a) and 10(a) are identical to Fig. 8(a) and, therefore, it appears that the dominance of the flexural free waves overwhelms not only the fluid free waves, but also the membrane free waves. To emphasize this identity, Fig. 8(a) is jointly designated Figs. 9(a) and 10(a). Can the flexural free waves be rendered weak enough so that those other types of free waves, which are faintly, but clearly discernible in Figs. 2 and 3, may be boldly exposed and investigated? Figures 2(f) and 3(f) and (g) suggest that damping the flexural free waves may induce the desired weakening in these free waves. Figures 8(a), 9(a), and 10(a) are repeated in Figs. 8(b), 9(b), and 10(b), respectively, except that the flexural loss factor (η_p) is changed from the standard value of 10^{-3} to 3×10^{-2} , a substantial increase in damping. This increase in damping of the flexural free waves appears more dramatic and revealing in Figs. 8(b), 9(b), and 10(b) than in the corresponding Figs.

2(d), 3(d), and 3(g) in the k domain. In Figs. 8(b), 9(b), and 10(b) the increase spatial decay of the flexural free waves is bottomed; this bottoming is not due to lack of computational accuracy, but, rather, due to the existence and presence of other types of free waves. In part, to help substantiate this statement Figs. 8(c), 9(c), and 10(c) are added. In these figures the fluid loading parameter (ϵ_c) is set equal to 10^{-6} and, in conformity, the flexural loss factor (η_p) is retained equal to 3×10^{-2} . Figures 8(c), 9(c), and 10(c) are also superimposed on Figs. 8(b), 9(b), and 10(b). The presence of fluid free waves emerges in Fig. 8(b) and the presence of fluid and membrane free waves emerge in Figs. 9(b) and 10(b). Moreover, manifestation of interferences among the various types of free waves are discernible in these figures.¹³ Interferences occur in “regions” in which more than one type of free waves are present with strengths within an order of magnitude. Thus, in Fig. 8(b) and (c) interference occurs in the region of the bottoming knee; this interference is between the flexural and fluid free waves. In this region the strengths of these two types of free waves are within an order of magnitude. Increasing the spatial distance $(|x - x_a|/b)$ beyond the knee, leaves the fluid free waves alone; the flexural free waves are weakened out by the high damping. Changes in the strength of the fluid free waves are then exhibited in Fig. 8(b); the knees appropriately adjust to these changes over several orders of magnitude. In passing it is recalled that in Fig. 8(a) the fluid loading significantly influences the surface impedance of the cylinder only when $\epsilon_c > 10^{-2}$. Indeed, this influence causes the saturation and reversal in the strength of fluid free waves as the fluid loading parameter (ϵ_c) is increased through 10^{-2} on to 10^{-1} . The apparent anomaly is associated with Le Chartier’s principle.² In Figs. 9(b) and (c) and 10(b) and (c), the membrane free waves join in; in Fig. 9(b) and (c) the membrane free waves are comprised of two components; longitudinal and shear free waves and in Fig. 10(b) and (c) of one component; longitudinal free waves only. These differences in the two sets of figures help decipher the free waves that are involved in the interferences. For example, it is clear from Figs. 9(c) and 10(c) that the strengths of the longitudinal and shear free waves, at $(\omega/\omega_c) = 0.175$, are within an order of magnitude and, therefore they significantly interfere in Fig. 9(c), whereas this interference is absent in Fig. 10(c). In these two figures the fluid free waves do not participate in the interference; the fluid free waves are rendered weak by selecting the fluid loading parameter (ϵ_c) equal to 10^{-6} in these figures. Clearly the fluid free waves participate in the interferences depicted in Figs. 9(b) and 10(b) for a fluid loading parameter $\epsilon_c \geq 10^{-2}$; in particular with the membrane free waves at distances $(|x - x_a|/b)$ beyond the bottoming knees. The details that are captured in the interferences in these figures attest to the accuracy attained in the computations via Eq. (1-14). The simplicity of the summand and, therefore, the simplicity of the computations performed to generate Figs. 8–10, may not convincingly constitute a critical proof of the viability of Eq. (1-14). Nonetheless, the kind of figures just presented are of intrinsic value, especially as they relate and assist in deciphering figures displaying more compounded

computations; those that account for the attachment of ribs on the cylinder.

III. MODAL DRIVE IN LIEU OF THE RIBS IN THE SPATIAL DOMAIN

The normalized modal drive $\bar{P}_{sn}(k, \omega)$ in lieu of the ribs is typically depicted in the k domain for a hybrid and a natural cylinder in Figs. 4 and 5, respectively. These figures are computed under standard parametric values and for a selected variety in these values. These figures demonstrate that this quantity is aliased in (k) with respect to (κ_1); the expression for this quantity in Eq. (14) confirms this aliasing. Moreover, the expression for $\bar{P}_{sn}(k, \omega)$ indicates that Eq. (1-21) suitability befits the formula for performing the Fourier transformation for this quantity. Designating $\bar{p}_{sn}(x, \omega)$ the Fourier transform of $\bar{P}_{sn}(k, \omega)$, Eqs. (1-14), (1-21)–(1-23) and (14) are used to derive

$$\bar{p}_{sn}(x, \omega) = s_b(x) \{ \delta(x) \bar{p}_{sn}^0(x, \omega) \}, \quad (17)$$

$$\bar{p}_{sn}(x, \omega) = [p_s(x, \omega) / p_{en}^0(\omega)],$$

where

$$\bar{p}_{sn}^0(x, \omega) \approx (1 + R)^{-1} \sum_{r=0}^R \exp(-ixk_r) I_{bn}^J(k_r, \omega) S_b^J(k_r) \times \{ Q_{\infty n}(k_r, \omega) \exp(ix_a k_r) \}, \quad (18a)$$

$$I_{bn}^J(k, \omega) = [1 + S_b^J(k) Q_{\infty n}(k, \omega)]^{-1}, \quad (18b)$$

and it is recognized that in Eqs. (17) and (18) there exist a factorial saving of $(2J+1)^2$ in computational steps as compared with estimating $\bar{p}_{sn}(x, \omega)$ from $\bar{P}_{sn}(k, \omega)$ employing the DFT procedure, as stated in Eq. (1-2c). The absolute values of $\bar{p}_{sn}^0(x, \omega)$ as a function of the normalized frequency (ω/ω_c) for a hybrid and a natural cylinder are presented in Figs. 11 and 12, respectively. The computations are performed for $(x/b) \Rightarrow (x_j/b) = 7, 14, 21$, and 35 and for a mode index (n) equal to unity. Each of these figures is composited of three values of the fluid loading parameter; namely, $\epsilon_c = 10^{-4}$, 10^{-2} , and 10^{-1} . The presence of the phenomenon of pass band and stop bands with respect to the flexural free waves emerges in these figures. The influence of fluid loading on this phenomenon is also distinguished in these figures. In particular there is a shift to lower frequencies of the stop bands with increase in fluid loading; the shift is more pronounced the higher the fluid loading is. This shift in frequencies is related to the added surface mass component in the fluid surface impedance as already discussed with respect to Figs. 2–5. This and other complementary features in $\bar{P}_{sn}(k, \omega)$ and $\bar{p}_{sn}^0(x, \omega)$ are facilitated by superimposing the corresponding computations performed on each quantity in the same figure. Thus, Figs. 4 and 5 simultaneously exhibit both quantities. It is observed in these figures that the high spatial decay in the modal drive $\bar{p}_{sn}^0(x, \omega)$ in a (flexural) stop band in the x domain, is appropriately commensurate with the uniformity in the drive $\bar{P}_{sn}(k, \omega)$ in a stop band in the k domain. This uniformity is derived from cancellations between pairs of flexural dispersive aliasing orders in those spectral regions in which they cross. In the absence of these

cancellations, as is the case in a first-order model in which interactions among the ribs via the cylindrical shell are neglected, pass bands and stop bands are also absent.^{2,15} The pass bands and stop bands are largely dominated by the flexural free waves; the other types of free waves are poorly manifested by a few minor kinks in the curves depicting $\bar{p}_{sn}^0(x, \omega)$ and a few corresponding kinks in the curves depicting $\bar{P}_{sn}(k, \omega)$. Other features in these quantities can be similarly related from the k domain to the x domain and vice versa; e.g., the shifts in frequencies of the stop bands due to a change in the character of the ring impedance of the ribs from mass to stiffness controlled correctly track in the complementary representations in Figs. 4(d) and 5(d). There are, of course, features that are more indigenous to one domain than another. The aliasing in $\bar{P}_{sn}(k, \omega)$ are clearly a k -domain phenomenon. On the other hand, in a stop band, when the distance ($|x_j - x_a|/b$) exceeds a specific value, the decay in $\bar{p}_{sn}^0(x, \omega)$ is bottomed. The bottomings in the (flexural) stop bands are clearly a x -domain phenomenon. The bottoming phenomenon in a stop band is reminiscent of the same named phenomenon induced by the increase in damping of the flexural free waves in the absence of ribs. Is there a commonality in these two phenomena?

IV. RESPONSE OF A REGULARLY RIBBED CYLINDER IN THE SPATIAL DOMAIN

The normalized modal response $\bar{V}_n(k, \omega)$ is composed of two terms, $\bar{V}_{\infty n}(k, \omega)$ and $\bar{V}_{sn}(k, \omega)$, as stated in Eqs. (12) and (13). Formally, from Eqs. (1-1a), (12), and (13) one obtains

$$\bar{v}_n(x, \omega) = \bar{v}_{\infty n}(x, \omega) - \bar{v}_{sn}(x, \omega), \quad (19)$$

$$\bar{v}_{\infty n}(x, \omega) = e_n(\omega) (2\pi)^{-1/2} \int dk \exp(-ixk) \bar{V}_{\infty n}(k, \omega), \quad (20a)$$

$$\bar{v}_{sn}(x, \omega) = e_n(\omega) (2\pi)^{-1/2} \int dk \exp(-ixk) \bar{V}_{sn}(k, \omega), \quad (20b)$$

where $e_n(\omega)$ is the normalization factor defined in Eq. (16). The Fourier transform of the first term is dealt with in Sec. II; in this section the second term is added to the computational task. As already discussed in the k domain, neither $\bar{V}_{\infty n}(k, \omega)$ nor $\bar{V}_{sn}(k, \omega)$ are strictly aliased. In the first term neither of its two factors are aliased and, therefore, no factorial saving can be accrued from the form of its evaluation. In the second term the factor $\bar{G}_{\infty n}(k, \omega)$ is not aliased; however, the second factor, being the modal drive $\bar{P}_{sn}(k, \omega)$ in lieu of the ribs, is aliased. Identifying $\bar{G}(k, \omega)$ as $\bar{G}_{\infty n}(k, \omega)$ and $K(k, \omega)$ as $\bar{P}_{sn}(k, \omega)$ in Eq. (1-14), and with Eqs. (1-16b) and (14a) in tow, a factorial saving of $(2J+1)$, and with Eqs. (1-19b) and (14b) in tow, a factorial saving of $(2J+1)^2$ of computational steps may be accrued in the evaluation of $\bar{v}_{sn}(x, \omega)$ in Eq. (20b). [This kind of saving is not afforded using Eq. (1-2c).] From Eqs. (1-14), (1-19b), (14b), and (20b) one obtains

$$\begin{aligned}\bar{v}_{sn}(x, \omega) &\approx e_n(\omega) [\kappa_1(1+R)^{-1}] \sum_{r=0}^R I_n^J(k_r, \omega) S_b^J(k_r) \\ &\quad \times \{Q_{\infty n}(k_r, \omega) \exp(ix_a k_r)\} \\ &\approx S_b^J(k_r) \{\exp(-ix k_r) \bar{G}_{\infty n}(k_r, \omega)\}, \\ e_n(\omega) &= [(2/\pi^2 k_{p0}^2) \{(n/ak_{p0})^4 - 1\}]^{1/2}, \quad (-1)^{1/2} = (-i). \quad (21)\end{aligned}$$

The normalization and evaluation of the second term in Eq. (19) are rendered compatible, in Eq. (21), with those of the first term in Eq. (16). A question may be raised: Since the aliasings phenomenon in $\bar{P}_{sn}(k, \omega)$ is impaired in $\bar{V}_n(k, \omega)$ and the pass bands and stop bands phenomenon is revealed in the modal drive $\bar{P}_{sn}(x, \omega)$, is the pass bands and stop bands phenomenon impaired in the modal response $\bar{v}_n(x, \omega)$? In other words, how interrelated are these two phenomena? To the extent that both phenomena require the regularity of the ribs, they are akin. However, as argued in Ref. 2, the phenomenon of pass bands and stop bands requires, in addition, interactions among the ribs. When interactions are allowed, the normalized drive $\bar{P}_{sn}(k, \omega)$ in lieu of these ribs remains, nonetheless, strictly aliased (cf. Figs. 4 and 5).² Moreover, as already discussed, $\bar{P}_{sn}(k, \omega)$ and its Fourier transform $\bar{p}_{sn}(x, \omega)$ show patterns and curves that indicate the presence of the phenomenon of pass bands and stop bands (cf. Figs. 4, 5, 11, and 12). On the other hand, the aliasing in the normalized modal response $\bar{V}_n(k, \omega)$ is impaired compared with the strict aliasing in $\bar{P}_{sn}(k, \omega)$ (compare Figs. 6 and 7 with Figs. 4 and 5). Is then the phenomenon of pass bands and stop bands in the Fourier transform of $\bar{V}_n(k, \omega)$; namely, in $\bar{v}_n(x, \omega)$, also impaired? The absolute value of the normalized modal response $\bar{v}_n(x, \omega)$ is displayed as a function of (x/b) , for a hybrid cylinder in Fig. 13 and for a natural cylinder in Figs. 14 and 15 the mode index is set equal to unity and zero, respectively. In Figs. 13(a), 14(a), and 15(a) the frequency is set at $(\omega/\omega_c) = 0.175$ with the fluid loading parameter $\epsilon_c = 10^{-6}$, 10^{-4} , 10^{-2} , and 10^{-1} and at $(\omega/\omega_c) = 0.158$ with $\epsilon_c = 10^{-1}$. It is observed that data pertaining to the frequency $(\omega/\omega_c) = 0.175$ with $\epsilon_c = 10^{-6}$, 10^{-4} , and 10^{-2} and the frequency $(\omega/\omega_c) = 0.158$ with $\epsilon_c = 10^{-1}$ lie in a (flexural) stop band. The data pertaining to a frequency $(\omega/\omega_c) = 0.175$ with $\epsilon_c = 10^{-1}$ substantially lie in pass band; the shift in the frequencies of the stop bands under the influence of heavy fluid loading accounts for this feature in the data with respect to $\epsilon_c = 10^{-1}$ (cf. Figs. 11 and 12). Figures 13(a), 14(a), and 15(a) convincingly indicate that the pass bands and stop bands phenomenon is not impaired in $\bar{v}_n(x, \omega)$, there may even be an improvement compared with this phenomenon in $\bar{p}_{sn}(x, \omega)$.

It further emerges that data in Figs. 13(a), 14(a), and 15(a) that lie in the stop band at $(\omega/\omega_c) = 0.175$ with $\epsilon_c = 10^{-6}$, 10^{-4} , and 10^{-2} and at $(\omega/\omega_c) = 0.158$ with $\epsilon_c = 10^{-1}$, are reminiscent of the corresponding data in Figs. 8(b), 9(b), and 10(b), respectively. This reminiscence holds even with respect to the reversal of strength between the fluid free waves with $\epsilon_c = 10^{-2}$ and 10^{-1} ; the anomaly is particularly discernible in Figs. 8(b) and 13(a). In the vein of Figs.

8(c), 9(c), and 10(c), the curves in Figs. 13(a), 14(a), and 15(a) that pertain to a fluid loading parameter $\epsilon_c = 10^{-6}$ are, in isolation, displayed in Figs. 13(b), 14(b), and 15(b), respectively. Notwithstanding that there exist minor and accountable differences in levels, the resemblance between in Figs. 13(a) and (b), 14(a) and (b), and 15(a) and (b) and Figs. 8(b) and (c), 9(b) and (c), and 10(b) and (c), respectively, is striking. The resemblance even matches the bottoming and interference phenomena described earlier. Indeed, comparison among the corresponding figures in the two sets leads one to conclude that the regularly attached ribs in a stop band play a role that is commensurate with the role played by the increase in flexural damping in the absence of ribs. This conclusion does not take account of the undulations in level within a bay that exist in the former set, but are absent in the latter set of figures.¹⁶ The undulations are avoided in the former set of figures by an appropriate selection of the position (x_a/b) of the external drive and the regularity of the sampling positions (x/b) along the axis of the cylinder.¹⁶ These undulations are not significant enough, however, to altercate the striking resemblance between these two sets of figures.

Figures 13(a) and (b), 14(a) and (b), and 15(a) and (b) are made possible only with an accuracy and reproductive details that transcend several orders of magnitude. In the computations that underlie these figures the overall wave-number span (k_M) and the sampling wave-number span (κ_1) are related largely by $J = 50$, yielding a factorial savings of 10^2 and, in some cases, of 10^4 of computational steps in computations performed with the double-sum technique versus computations performed with techniques that are directly akin to the DFT procedures. Whether this kind of factorial savings is crucial to the generation of the figures presented in this paper is a matter of memory and patience. The former can be bought and the latter can be acquired by practice, notwithstanding that shortages on both accounts are thinning out. Nonetheless, that the double-sum technique has been successfully employed to furnish the figures just presented is by now a fact.

¹ G. Maidanik and K. J. Becker, "A double-sum technique for performing a Fourier transformation" *J. Acoust. Soc. Am.* **101**, 2448–2451 (1997).

² G. Maidanik and K. J. Becker, "Phenomena of aliasing and pass and stop bands in the drive in lieu of ribs on cylindrical shells," CDNSWC-SIG-96/072-7030, 1996 (unpublished).

³ L. Cremer, M. Heckl, and E. Ungar, *Structure-Borne Sound, Structural Vibrations and Sound Radiation at Audio Frequencies*, 2nd ed. (Springer-Verlag, Berlin, 1988).

⁴ M. C. Junger and D. Feit, *Sound Structures, and Their Interaction* (MIT, Cambridge, MA, 1986).

⁵ F. Fahy, *Sound and Structural Vibration (Radiation, Transmission and Response)* (Academic, New York, 1985).

⁶ M. P. Norton, *Fundamentals of Noise and Vibration Analysis for Engineers* (Cambridge U.P., Cambridge, England, 1989).

⁷ P. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).

⁸ C. B. Burroughs, "Acoustic radiation from fluid-loaded infinite circular cylinders with doubly periodic ring supports," *J. Acoust. Soc. Am.* **75**, 715–722 (1984).

⁹ W. Vogel and D. Feit, "Response of a point excited infinitely long cylindrical shell immersed in an acoustic medium," DTNSRDC Report No. 80/061, 1980 (unpublished).

¹⁰ Y. Liu and A. Tucker, "The distribution of radiated and vibratory powers of a point-driven infinite cylindrical shell, Part I-Uncoated structure,"

- DTNSRDC Report No. 84/043, 1984 and "Part II-Structures with a compliant coating," DTNSRDC Report no. 84/044, 1984 (unpublished).
- ¹¹G. Maidanik and J. Dickey, "Acoustic behavior of ribbed panels," *J. Sound Vib.* **123**, 293–307 (1988).
- ¹²G. Maidanik and J. Dickey, "Velocity distributions on unloaded finitely and regularly ribbed membranes," *J. Sound Vib.* **149**, 43–70 (1991); "Response of regularly ribbed fluid loaded panels," *J. Sound Vib.* **155**, 481–495 (1992).
- ¹³D. M. Photiadis, "The propagation of axisymmetric waves on a fluid-loaded cylindrical shell," *J. Acoust. Soc. Am.* **88**, 239–250 (1990).
- ¹⁴G. Maidanik and K. J. Becker, "Phenomenon of leaky free waves in the modal response of a uniform cylinder," NSWCCD-SIG-96/120-7030, 1996 (unpublished).
- ¹⁵G. Maidanik and A. J. Tucker, "Proper and first order solutions of regularly ribbed panels," *J. Sound Vib.* **55**, 267–274 (1976).
- ¹⁶K. J. Becker, "On the response of regularly ribbed cylindrical shell," thesis (in preparation).

Transmission and reflection of transient elastodynamic waves at a linear slip interface

Martin D. Verweij^{a)} and Christopher H. Chapman

Schlumberger Cambridge Research, High Cross, Madingley Road, Cambridge CB3 0EL, England

(Received 8 May 1996; accepted 5 December 1996)

A method is presented for the analysis of the transmitted and reflected transient elastodynamic wave field at a fracture that may be modeled as a linear slip interface. With this method, the space-time domain problem is first carried over to the transform domain by means of a temporal Laplace transformation and a two-dimensional, horizontal spatial Fourier transformation. Subsequently, for both the *SH* and the *P/SV* wave systems, expressions are determined for the transmitted and reflected wave fields at the interface. Afterward, a transformation back to the space-time domain is carried out. The method is capable of providing the exact transmitted and reflected waves (including the body wave, interface wave, and head wave contributions) that are caused by a spatially curved, incident wave field due to a point source. Moreover, it is especially designed to yield the time domain waveforms. Various exact numerical results are given. For the *SH* case, the directionally dependent filter behavior of the fracture is shown. For the *P/SV* case, first the reflection, transmission, and conversion of body waves is presented. The next set of results concerns the interface wave contribution and includes pictures of the particle movement on both sides of the fracture when either the transversal or the normal specific compliance of the fracture is equal to zero. Finally, the presence of head waves in case of *SV*→*P* and *SV*→*SV* reflection is shown. © 1997 Acoustical Society of America. [S0001-4966(97)04505-0]

PACS numbers: 43.20.Fn, 43.40.Ph, 68.35.Ja, 91.60.Ba [ANN]

INTRODUCTION

Solid objects—from rocks within the Earth's crust to ceramic artefacts—may locally be weakened by material damage in the form of fractures. In this paper we will use the word fracture to indicate any feature that forms an elongated domain with a compliance that is significantly higher than that of the surrounding material, such as a single fault, a non-welded contact, or a band of multiple (micro)cracks. Fractures will affect the propagation of elastic waves to an extent that depends upon their width and their relative compliance (Schoenberg and Haugen¹). In the literature two alternative methods are frequently employed to evaluate the influence of fractures on the elastodynamic wave field. The first method, which consists of replacing the fractured solid by a homogeneous, anisotropic, equivalent medium, is particularly suitable for the analysis of the collective effect of many fractures (O'Connell and Budiansky,² Hudson,³ Schoenberg,⁴ White,⁵ Schoenberg and Sayers⁶). With the second method, a fracture is modeled by means of boundary conditions (Kendall and Tabor,⁷ Schoenberg,⁸ Kitsunezaki⁹). This is most useful when a detailed picture of the behavior of the elastodynamic wave field at an individual fracture is needed (Pyrak-Nolte *et al.*¹⁰). In this paper we will focus on the second method.

As long as the elastic waves only involve sufficiently small changes of the stress and the strain, as is generally the

case in seismics, it suffices to model the fracture behavior by means of linearized boundary conditions for the stress and the strain. Here we will consider the so-called linear slip interface model of a fracture (Schoenberg⁸), which is also known as the displacement discontinuity model, the incomplete interface model, or the imperfect interface model. This model prescribes that at an interface with a normal vector ν_j , the traction $\tau_{ij}\nu_j$ has to be continuous while the displacement u_i may show a jump discontinuity with a size that is proportional to the traction, i.e.,

$$\Delta(\tau_{ij}\nu_j) = 0, \quad (1)$$

$$\Delta u_i = Z_{ij}\tau_{jk}\nu_k, \quad (2)$$

over the interface. The values of the elements of the specific compliance matrix Z_{ij} , which are in fact the gradients of the displacement versus traction curves of the fracture, in general strongly depend on the values of the static tractions at the fracture (Pyrak-Nolte *et al.*,¹⁰ Hsu and Schoenberg¹¹). From energy considerations it follows that the specific compliance matrix is real, symmetric and non-negative definite (Schoenberg⁸). The precise form of the matrix Z_{ij} for several types of anisotropic fracture behavior has been discussed by Schoenberg⁸ and Schoenberg and Haugen.¹ Equations (1) and (2) describe a lossless fracture; for a lossy fracture (e.g., a fracture that is filled with a viscous fluid) the linear slip interface conditions may be written as

$$\Delta(\tau_{ij}\nu_j) = 0, \quad (3)$$

$$\Delta u_i = Z_{ij}^* \tau_{jk}\nu_k, \quad (4)$$

^{a)}Work performed as a Visiting Scientist from the Laboratory of Electromagnetic Research, Department of Electrical Engineering, Centre for Technical Geoscience, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands.

in which the elements of the specific compliance matrix are relaxation functions. The symbol $*$ denotes a temporal convolution.

In a number of papers the experimental verification of the linear slip interface condition has been addressed (Pyrak-Nolte *et al.*,^{10,12} Hsu and Schoenberg¹¹), and in several other papers the effect of a linear slip interface on the elastodynamic wave field has been investigated numerically by means of the finite-difference technique (Savić,¹³ Coates and Schoenberg¹⁴) or the boundary integral equation method (Gu *et al.*¹⁵). In a third class of papers, the linear slip interface condition has been employed to theoretically investigate the influence of a fracture on the elastodynamic wave field (Schoenberg,⁸ Pyrak-Nolte and Cook,¹⁶ Schoenberg and Haugen¹). Virtually all of the latter theoretical investigations deal with plane waves in single-frequency steady state. For these types of waves, the reflection, transmission, and conversion of body waves at a linear slip interface has been analyzed; further it has been shown that under certain circumstances interface waves and head waves may propagate along a linear slip interface.

However, these studies do not reveal to what extent the body waves, interface waves and head waves are present in the transmitted and reflected wave fields that are due to a physical (finite-size) source. Moreover, they do not directly aim at the generation of the time domain synthetic waveforms. The fact that knowledge of these aspects is of interest for, e.g., seismic prospecting has motivated us to conduct the investigation that is reported in this paper. Our key problem is the exact theoretical and numerical analysis of the transmitted and reflected transient elastodynamic waves (including body waves, interface waves, and head waves) that are excited when a wave, generated by a point source, encounters a plane and infinitely large linear slip interface in between two half-spaces that are composed of the same isotropic medium. In this analysis, all three spatial directions will be taken into account.

For our theoretical investigation we will employ the integral transformation approach: First we will apply integral transformations with respect to those coordinates in which the configuration is invariant (the time coordinate and the spatial coordinates in the plane of the interface), then the resulting transform domain problem is solved, and finally we will apply an inverse transformation procedure to return to the space-time domain. This approach may be recognized in the outline of this paper. Section I contains a description of the configuration and a formulation of the basic elastodynamic equations. In Sec. II we describe how the space-time domain quantities are carried over to the transform domain by means of a temporal Laplace transformation and a two-dimensional, horizontal spatial Fourier transformation. The transform domain wave propagation problem in the half-spaces surrounding the interface will be solved in Sec. III, while the transmission and the reflection problem at the interface is treated in Sec. IV. In Sec. V the transformation back to the space-time domain is presented. In this inverse transformation process we will employ elements of the Cagniard-De Hoop method (De Hoop,¹⁷⁻¹⁹ Van der Hijden²⁰) and the WKBJ seismogram method (Chapman²¹).

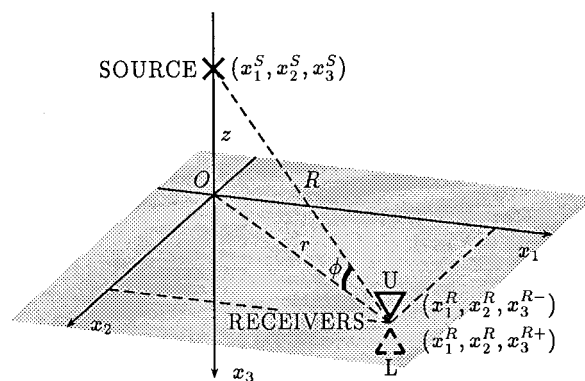


FIG. 1. The basic configuration with the linear slip interface (shaded), the source, and the receiver U at an infinitesimal distance above the interface and the receiver L at an infinitesimal distance below the interface.

The results of many of our numerical experiments are shown in Sec. VI. This paper ends with Sec. VII, which contains some conclusions that follow from the presented analysis.

I. CONFIGURATION AND BASIC EQUATIONS

The configuration that forms the basis of our analysis is depicted in Fig. 1. The spatial coordinates $\{x_1, x_2, x_3\}$ are used to indicate position with respect to the given right-handed, orthonormal, Cartesian reference frame. The coordinate t is used to indicate time. At $x_3=0$ an infinite, plane, linear slip interface is present that forms the separation between two homogeneous, isotropic and lossless elastic half-spaces. The elastodynamic wave field is generated by a point source located above the interface. The incident and reflected wave fields are detected by a receiver that is placed at an infinitesimal distance above the interface ($x_3^R = x_3^{R-} = 0^-$), while the transmitted wave field is detected by a receiver that is placed at an infinitesimal distance below the interface ($x_3^R = x_3^{R+} = 0^+$).

We assume that in both half-spaces the wave propagation is described by the basic space-time domain elastodynamic equations

$$\partial_j \tau_{ij}(x_k, t) - \rho \partial_t v_i(x_k, t) = -f_i(x_k, t), \quad (5)$$

$$\frac{1}{2} [\partial_j v_i(x_k, t) + \partial_i v_j(x_k, t)] - s_{ijpq} \partial_t \tau_{pq}(x_k, t) = \delta_{ij} q(x_k, t). \quad (6)$$

Here the subscript notation and the summation convention have been employed, where the lower-case Latin subscripts (except t) range from 1 to 3. The elastodynamic wave field is represented by the stress τ_{ij} and the particle velocity v_i . The symbols ∂_j and ∂_t denote a differentiation with respect to the spatial coordinate x_j and the time coordinate t , respectively. The action of the point source is contained in the volume density of volume force f_i and the volume density of volume injection rate q ; the symbol δ_{ij} indicates the Kronecker delta. Due to the horizontal and temporal shift invariance of

the configuration, without loss of generality we may assume that the point source is located on the vertical axis of the reference frame and that it starts to act at $t=0$, i.e.,

$$f_i(x_k, t) = F_i^S(t) \delta(x_1, x_2, x_3 - x_3^S), \quad (7)$$

$$q(x_k, t) = Q^S(t) \delta(x_1, x_2, x_3 - x_3^S), \quad (8)$$

with the source signatures $F_i^S(t)=0$ and $Q^S(t)=0$ for $t<0$. The medium in both half-spaces is the same and is characterized by the scalar mass density ρ and the tensorial stiffness s_{ijpq} . These are chosen in such a way that the medium possesses a compressional wave speed c^P and a shear wave speed c^S and, consequently, the Lamé coefficients $\lambda = \rho(c^P)^2 - 2\rho(c^S)^2$ and $\mu = \rho(c^S)^2$.

Unless otherwise noted, we assume that the fracture is lossless and, consequently, that the elements of the specific compliance matrix Z_{ij} are constants. Thus in the actual case the linear slip interface conditions are given by

$$\Delta \tau_{i3} = 0, \quad (9)$$

$$\Delta u_i = Z_{ij} \tau_{j3}. \quad (10)$$

Moreover, we will only consider interfaces with up-down and rotational symmetry. This implies that the specific compliance matrix has the form

$$Z_{ij} = z^t \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + z^n \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (11)$$

where z^t and z^n represent the transversal and the normal specific compliances of the fracture (cf. Schoenberg⁸ and Schoenberg and Haugen¹).

II. FORWARD TRANSFORMATION

We exploit the time invariance of the configuration by subjecting the space-time domain quantities to a one-sided, temporal Laplace transformation. This transformation replaces, for example, the space-time domain quantity $v_i(x_k, t)$ by its Laplace-transformed counterpart

$$\hat{v}_i(x_k, s) = \int_0^\infty v_i(x_k, t) \exp(-st) dt. \quad (12)$$

Reversely, an inverse Laplace transformation in fact performs the solution of Eq. (12) considered as an integral equation with a known function $\hat{v}_i(x_k, s)$ (Carson²²). According to Lerch's theorem (Widder²³), the causal solution $v_i(x_k, t)$ of this integral equation is unique if $\hat{v}_i(x_k, s)$ is bounded for the infinite sequence $\{s_n\}$, where $s_n = s_0 + n\ell$ with s_0 real, positive and sufficiently large; $n=0, 1, 2, \dots$; and ℓ positive and real. This implies that it is perfectly valid to apply inverse Laplace transformation recipes in which only real and sufficiently large values of s show up, i.e., in which complex values of s (like those on the contour of the well-known Laplace inversion integral) do not occur. Following Cagniard,²⁴ in our analysis we keep s real and positive.

Subsequently, we employ the horizontal shift invariance of the configuration by performing a two-dimensional, two-sided, Radon-type spatial Fourier transformation with respect to x_1 and x_2 . In this way the Laplace-transformed quantity $\hat{v}_i(x_k, s)$ is replaced by its transformed counterpart

$$\begin{aligned} \tilde{v}_i(\alpha_1, \alpha_2, x_3, s) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{v}_i(x_k, s) \\ &\times \exp[is(\alpha_1 x_1 + \alpha_2 x_2)] dx_1 dx_2. \end{aligned} \quad (13)$$

The inverse Fourier transformation is

$$\begin{aligned} \hat{v}_i(x_k, s) &= \left(\frac{s}{2\pi}\right)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{v}_i(\alpha_1, \alpha_2, x_3, s) \\ &\times \exp[-is(\alpha_1 x_1 + \alpha_2 x_2)] d\alpha_1 d\alpha_2. \end{aligned} \quad (14)$$

Provided that the dependence may be deduced uniquely, from now on some or all of the arguments of the applied functions will be omitted for brevity.

After elimination of the stress components that may show a jump discontinuity across the interface ($\tilde{\tau}_{i1}$ and $\tilde{\tau}_{i2}$), six transform domain state quantities remain for representing the wave field. These are conveniently arranged into the state vector

$$\tilde{b}_I = (\tilde{v}_1, \tilde{v}_2, \tilde{v}_3, -\tilde{\tau}_{13}, -\tilde{\tau}_{23}, -\tilde{\tau}_{33})^T, \quad (15)$$

where the upper case Latin subscript ranges from 1 to 6. The analysis of the state vector in each of the half-spaces will be dealt with in Sec. III. The investigation of its behavior at the interface forms the subject of Sec. IV.

III. PROPAGATION IN THE HALF-SPACES

In both half-spaces, the state vector satisfies the differential equation

$$\partial_3 \tilde{b}_I + s A_{IJ} \tilde{b}_J = \tilde{F}_I. \quad (16)$$

For the derivation of the system matrix A_{IJ} and the notional source vector \tilde{F}_I we refer to Van der Hijden²⁰ (pp. 45–47 and pp. 57–58). Instead of directly solving this equation, it is more convenient to first introduce the wavevector \tilde{w}_J through the linear transformation

$$\tilde{b}_I = N_{IJ} \tilde{w}_J. \quad (17)$$

The so-called composition matrix N_{IJ} consists of columns that are normalized eigenvectors of the system matrix A_{IJ} . Using the column ordering and normalization as given on p. 52 of Van der Hijden,²⁰ we obtain

$$N_{IJ} = \begin{pmatrix} i\alpha_1 & i\alpha_1\gamma^s & -i\alpha_2 & i\alpha_1 & -i\alpha_1\gamma^s & -i\alpha_2 \\ i\alpha_2 & i\alpha_2\gamma^s & i\alpha_1 & i\alpha_2 & -i\alpha_2\gamma^s & i\alpha_1 \\ -\gamma^p & -(\alpha_1^2 + \alpha_2^2) & 0 & \gamma^p & -(\alpha_1^2 + \alpha_2^2) & 0 \\ -2i\mu\alpha_1\gamma^p & -2i\mu\alpha_1\chi & i\mu\alpha_2\gamma^s & 2i\mu\alpha_1\gamma^p & -2i\mu\alpha_1\chi & -i\mu\alpha_2\gamma^s \\ -2i\mu\alpha_2\gamma^p & -2i\mu\alpha_2\chi & -i\mu\alpha_1\gamma^s & 2i\mu\alpha_2\gamma^p & -2i\mu\alpha_2\chi & i\mu\alpha_1\gamma^s \\ 2\mu\chi & 2\mu(\alpha_1^2 + \alpha_2^2)\gamma^s & 0 & 2\mu\chi & -2\mu(\alpha_1^2 + \alpha_2^2)\gamma^s & 0 \end{pmatrix}, \quad (18)$$

in which

$$\chi = \frac{1}{2}(c^s)^{-2} + \alpha_1^2 + \alpha_2^2, \quad (19)$$

$$\gamma^p = [(c^p)^{-2} + \alpha_1^2 + \alpha_2^2]^{1/2}, \quad (20)$$

$$\gamma^s = [(c^s)^{-2} + \alpha_1^2 + \alpha_2^2]^{1/2}. \quad (21)$$

The latter two quantities are the vertical compressional slowness and the vertical shear slowness. Substitution of Eq. (17) into Eq. (16) yields the wave vector differential equation

$$\partial_3 \tilde{w}_I + s \Lambda_{IJ} \tilde{w}_J = \tilde{k}_I \delta(x_3 - x_3^S). \quad (22)$$

In this equation there appears the eigenvalue matrix

$$\Lambda_{IJ} = \text{diag}(-\gamma^p, -\gamma^s, -\gamma^s, \gamma^p, \gamma^s, \gamma^s). \quad (23)$$

As a consequence of the choice in Eq. (18), the wave vector components \tilde{w}_1 , \tilde{w}_2 and \tilde{w}_3 are exponential functions that represent upward propagating *P* (compressional), *SV* (vertically polarized shear) and *SH* (horizontally polarized shear) waves, and the components \tilde{w}_4 , \tilde{w}_5 , and \tilde{w}_6 are exponential functions that represent downward propagating *P*, *SV* and *SH* waves. The nonzero wave vector components of the incident wave field at the interface are given by

$$\tilde{w}_4^i = \tilde{k}_4 \exp[-s\gamma^p(x_3^R - x_3^S)], \quad (24)$$

$$\tilde{w}_5^i = \tilde{k}_5 \exp[-s\gamma^s(x_3^R - x_3^S)], \quad (25)$$

$$\tilde{w}_6^i = \tilde{k}_6 \exp[-s\gamma^s(x_3^R - x_3^S)]. \quad (26)$$

The amplitudes of these waves depend on the specific source type. For a volume source we find

$$\tilde{k}_4 = \frac{3\lambda + 2\mu}{2\gamma^p(\lambda + 2\mu)} \hat{Q}^S, \quad (27)$$

$$\tilde{k}_5 = 0, \quad (28)$$

$$\tilde{k}_6 = 0, \quad (29)$$

and in case of a force source we get

$$\tilde{k}_4 = \frac{(c^s)^2}{2\gamma^p\mu} (i\alpha_1 \hat{F}_1^S + i\alpha_2 \hat{F}_2^S + \gamma^p \hat{F}_3^S), \quad (30)$$

$$\tilde{k}_5 = \frac{(c^s)^2}{2(\alpha_1^2 + \alpha_2^2)\mu} \left(i\alpha_1 \hat{F}_1^S + i\alpha_2 \hat{F}_2^S + \frac{\alpha_1^2 + \alpha_2^2}{\gamma^s} \hat{F}_3^S \right), \quad (31)$$

$$\tilde{k}_6 = \frac{1}{2\gamma^s(\alpha_1^2 + \alpha_2^2)\mu} (i\alpha_2 \hat{F}_1^S - i\alpha_1 \hat{F}_2^S). \quad (32)$$

Apart from the amplitude factors, the expressions for the nonzero wave vector components \tilde{w}_4^t , \tilde{w}_5^t and \tilde{w}_6^t of the transmitted wave field at the interface and \tilde{w}_1^r , \tilde{w}_2^r and \tilde{w}_3^r of the reflected wave field at the interface are of the same form as the wave vector components \tilde{w}_1^i , \tilde{w}_2^i and \tilde{w}_3^i . The amplitude factors for these waves follow upon multiplication of the amplitude factors of the incident wave with the relevant transmission and reflection factors of the interface. The latter will be derived in the next section.

IV. TRANSMISSION AND REFLECTION AT THE INTERFACE

The linear slip interface model of Eqs. (9) and (10) gives rise to the transform domain boundary conditions

$$\Delta \tilde{\tau}_{i3} = 0, \quad (33)$$

$$\Delta \tilde{v}_i = s Z_{ij} \tilde{\tau}_{j3}, \quad (34)$$

at $x_3 = 0$. As a consequence, we find that the limiting values $\tilde{b}_I^- = \lim_{x_3 \uparrow 0} \tilde{b}_I$ and $\tilde{b}_I^+ = \lim_{x_3 \downarrow 0} \tilde{b}_I$ of the state vector are related through

$$\tilde{b}_I^- = \begin{pmatrix} I & sZ \\ O & I \end{pmatrix}_{IJ} \tilde{b}_J^+ = S_{IJ} \tilde{b}_J^+. \quad (35)$$

Here S_{IJ} is a shorthand notation of the 6×6 matrix that is composed of the 3×3 submatrices I_{ij} (unit matrix), Z_{ij} (specific compliance matrix) and O_{ij} (zero matrix). In view of Eq. (17), the corresponding relation for the wave vector is

$$\tilde{w}_I^- = N_{IK}^{-1} S_{KL} N_{LJ} \tilde{w}_L^+ = Q_{IJ} \tilde{w}_J^+. \quad (36)$$

For later use, it is convenient to subdivide the 6×6 matrix Q_{IJ} into four 3×3 submatrices $Q_{ij}^{\delta\epsilon}$ according to

$$Q_{IJ} = \begin{pmatrix} Q^{11} & Q^{12} \\ Q^{21} & Q^{22} \end{pmatrix}_{IJ}. \quad (37)$$

Next we define the transmission matrix T_{ij} as the matrix that connects the transmitted $\{\tilde{w}_4, \tilde{w}_5, \tilde{w}_6\}$ part of \tilde{w}_I^+ with the incident $\{\tilde{w}_4, \tilde{w}_5, \tilde{w}_6\}$ part of \tilde{w}_I^- . Analogously, we introduce the reflection matrix R_{ij} as the matrix that relates the reflected $\{\tilde{w}_1, \tilde{w}_2, \tilde{w}_3\}$ part of \tilde{w}_I^- and the incident part of \tilde{w}_I^- . The role of these matrices is depicted schematically in Fig. 2. In view of Eq. (36), the transmission and reflection matrices have to satisfy the matrix equation

$$\begin{pmatrix} R \\ I \end{pmatrix}_{ij} = \begin{pmatrix} Q^{11} & Q^{12} \\ Q^{21} & Q^{22} \end{pmatrix}_{IK} \begin{pmatrix} O \\ T \end{pmatrix}_{Kj}. \quad (38)$$

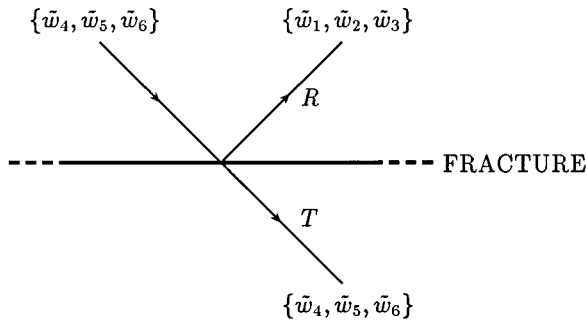


FIG. 2. The role of the transmission and reflection matrices of the interface.

From this equation it may be deduced that

$$T_{ij} = (Q^{22})_{ij}^{-1}, \quad (39)$$

$$R_{ij} = Q_{ik}^{12} (Q^{22})_{kj}^{-1}. \quad (40)$$

Determination of the expressions for the elements of Q_{IJ} shows us that its submatrices $Q_{ij}^{\delta\epsilon}$ are of the form

$$Q_{ij}^{\delta\epsilon} = \begin{pmatrix} Q_{11}^{\delta\epsilon} & Q_{12}^{\delta\epsilon} & 0 \\ Q_{21}^{\delta\epsilon} & Q_{22}^{\delta\epsilon} & 0 \\ 0 & 0 & Q_{33}^{\delta\epsilon} \end{pmatrix}. \quad (41)$$

From Eqs. (39)–(41) it follows that the transmission and reflection matrices are of the same form as the matrices $Q_{ij}^{\delta\epsilon}$. This implies that at the interface the P and the SV waves become coupled (i.e., mode conversion may occur between P and SV waves, see Schoenberg and Haugen¹), while the SH waves remain uncoupled from the P and SV waves (i.e., there occurs neither mode conversion between SH and P waves nor between SH and SV waves). For this reason we will evaluate the scalar reflection and transmission factors for the SH wave system separately from the 2×2 transmission and reflection matrices for the P/SV wave system.

A. Transmission and reflection factors for the SH wave system

The actual transmission factor for SH waves turns out to be [cf. Eq. (39)]

$$T^{SH} = (Q_{33}^{22})^{-1} = \frac{1}{1 + s(\frac{1}{2}z^t \mu \gamma^s)}. \quad (42)$$

From this expression we see that in case of transmission of SH waves, the interface acts as a (directionally dependent) low-pass filter. The reflection factor for the SH waves is [cf. Eq. (40)]

$$R^{SH} = Q_{33}^{12} (Q_{33}^{22})^{-1} = \frac{s(\frac{1}{2}z^t \mu \gamma^s)}{1 + s(\frac{1}{2}z^t \mu \gamma^s)}. \quad (43)$$

This expression shows us that in case of reflection of SH waves the interface forms a (directionally dependent) high-pass filter.

B. Transmission and reflection matrices for the P/SV wave system

In the P/SV case it makes sense to first introduce the 2×2 matrices

$$Q_{\alpha\beta}^{\delta\epsilon} = \begin{pmatrix} Q_{11}^{\delta\epsilon} & Q_{12}^{\delta\epsilon} \\ Q_{21}^{\delta\epsilon} & Q_{22}^{\delta\epsilon} \end{pmatrix}, \quad (44)$$

where the lower case Greek subscripts, like the lower case Greek superscripts introduced earlier, range from 1 to 2. Further we need to evaluate [cf. Eqs. (39) and (40)]

$$(Q^{22})_{\alpha\beta}^{-1} = \frac{1}{\det(Q_{\alpha\beta}^{22})} \begin{pmatrix} Q_{22}^{22} & -Q_{12}^{22} \\ -Q_{21}^{22} & Q_{11}^{22} \end{pmatrix}. \quad (45)$$

Using the expressions for the elements of $Q_{\alpha\beta}^{22}$, we find that the determinant is of the form

$$\det(Q_{\alpha\beta}^{22}) = (1 + sac)(1 + sbc), \quad (46)$$

with

$$a = 2\mu z^t \gamma^p, \quad (47)$$

$$b = 2\mu z^n \gamma^s, \quad (48)$$

$$c = (c^s)^2 \left(\frac{\chi^2}{\gamma^p \gamma^s} - \alpha_1^2 - \alpha_2^2 \right). \quad (49)$$

For all real values of α_1 and α_2 , the quantities a , b , and c are real valued and we may always find a $s_0 > 0$ such that the determinant is nonzero for all $s \geq s_0$. This implies that for all values s , α_1 , and α_2 that play a role in the formal inverse transformations of Sec. II, the inverse matrix $(Q^{22})_{\alpha\beta}^{-1}$ exists. (During the actual inverse transformation process in Sec. V, however, we will see that for certain complex values of α_1 and/or α_2 the quantity c may become zero. This fact will be associated with the occurrence of interface waves.) The transmission matrix for the P/SV wave system turns out to be [cf. Eqs. (39), (45), and (46)]

$$T_{\alpha\beta}^{P/SV} = (Q^{22})_{\alpha\beta}^{-1} = \frac{1}{c(1 + sac)} T_{\alpha\beta}^t + \frac{1}{c(1 + sbc)} T_{\alpha\beta}^n, \quad (50)$$

in which

$$T_{\alpha\beta}^t = (c^s)^2 \begin{pmatrix} -(\alpha_1^2 + \alpha_2^2) & \frac{\chi(\alpha_1^2 + \alpha_2^2)}{\gamma^p} \\ \frac{-\chi}{\gamma^s} & \frac{\chi^2}{\gamma^p \gamma^s} \end{pmatrix}, \quad (51)$$

$$T_{\alpha\beta}^n = (c^s)^2 \begin{pmatrix} \frac{\chi^2}{\gamma^p \gamma^s} & \frac{-\chi(\alpha_1^2 + \alpha_2^2)}{\gamma^p} \\ \frac{\chi}{\gamma^s} & -(\alpha_1^2 + \alpha_2^2) \end{pmatrix}. \quad (52)$$

From these expressions we see that the interface acts as a (directionally dependent) low-pass filter for transmitted P/SV waves. The reflection matrix for the P/SV wave system is [cf. Eqs. (40), (45), and (46)]

$$R_{\alpha\beta}^{P/SV} = Q_{\alpha\eta}^{12} (Q_{\eta\beta}^{22})^{-1} = \frac{sac}{c(1+sac)} R_{\alpha\beta}^t + \frac{sbc}{c(1+sbc)} R_{\alpha\beta}^n, \quad (53)$$

in which

$$R_{\alpha\beta}^t = (c^s)^2 \begin{pmatrix} -(\alpha_1^2 + \alpha_2^2) & \frac{\chi(\alpha_1^2 + \alpha_2^2)}{\gamma^p} \\ \frac{\chi}{\gamma^s} & \frac{-\chi^2}{\gamma^p \gamma^s} \end{pmatrix}, \quad (54)$$

$$R_{\alpha\beta}^n = (c^s)^2 \begin{pmatrix} \frac{-\chi^2}{\gamma^p \gamma^s} & \frac{\chi(\alpha_1^2 + \alpha_2^2)}{\gamma^p} \\ \frac{\chi}{\gamma^s} & -(\alpha_1^2 + \alpha_2^2) \end{pmatrix}. \quad (55)$$

These expressions show us that the interface forms a (directionally dependent) high-pass filter for reflected P/SV waves. Note that both the transmission matrix and the reflection matrix consist of a part that only depends on the transversal fracture behavior and a part that only depends on the normal fracture behavior.

V. INVERSE TRANSFORMATION

At this stage, all the ingredients are available for finding the transform domain state vector of the incident, transmitted, or reflected wave field at the interface: Eqs. (27)–(32) give the amplitudes of the P , SV , and SH waves that are emitted by the source, Eqs. (42), (43), (50) and (53) show to what extent the incident waves are transmitted and reflected, and Eq. (17) describes how to obtain the state vector corresponding to the wave vector. The only step that remains is the transformation of the state vector components back to the space-time domain. In this section we will present an exact inverse transformation process that uses the the slowness contour deformation that characterizes the Cagniard–De Hoop method (De Hoop,^{17–19} Van der Hijden²⁰), and that employs the Laplace inversion of the integrand in the same way as the WKBJ seismogram method (Chapman²¹). Our approach has two important features: It enables a systematic identification of all relevant wave phenomena (body waves, interface waves and head waves) and it remains valid in case of a lossy interface.

A. Body waves

The inverse transformation process is most easily explained by starting with an example that only involves body waves. Here we will consider the particle velocity v_i of a transmitted²⁵ SH wave due to a force source f_j . The general form of the expression of the transform domain particle velocity \tilde{v}_i is

$$\tilde{v}_i = \Pi_{ij} \hat{F}_j^s \frac{\exp(-sz\gamma^s)}{1 + s(\frac{1}{2}z^t \mu \gamma^s)}, \quad (56)$$

in which $z = x_3^R - x_3^S$, and $\Pi_{ij} = \Pi_{ij}(\alpha_1, \alpha_2)$ describes the coupling of the SH wave with the source and the receiver

according to N_{IJ} and N_{IJ}^{-1} . As a first step in the inversion process, it is customary to introduce the transform domain Green's function \tilde{g}_{ij} through

$$\tilde{v}_i = s^2 \tilde{g}_{ij} \hat{F}_j^s. \quad (57)$$

Combination of Eqs. (56), (57) and (14) yields the Laplace-domain equation

$$\hat{g}_{ij} = \left(\frac{1}{2\pi} \right)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Pi_{ij} \times \frac{\exp[-s(i\alpha_1 x_1^R + i\alpha_2 x_2^R + z\gamma^s)]}{1 + s(\frac{1}{2}z^t \mu \gamma^s)} d\alpha_1 d\alpha_2. \quad (58)$$

Further, the horizontal slowness p and the new variable q (not to be confused with the volume density of volume injection rate q) are introduced through the rotation

$$\alpha_1 = -ip \cos \theta + q \sin \theta, \quad (59)$$

$$\alpha_2 = -ip \sin \theta - q \cos \theta, \quad (60)$$

in which $\theta = \arctan(x_2^R/x_1^R)$ is the polar angle of the receiver with respect to the source. Substitution of Eqs. (59) and (60) into Eq. (58) results in

$$\hat{g}_{ij} = \left(\frac{1}{2\pi} \right)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i \bar{\Pi}_{ij} \frac{\exp[-s(pr + z\bar{\gamma}^s)]}{1 + s(\frac{1}{2}z^t \mu \bar{\gamma}^s)} dp dq. \quad (61)$$

Here, $r = [(x_1^R)^2 + (x_2^R)^2]^{1/2}$ forms the horizontal offset between the source and the receiver. An overbar indicates those quantities in which α_1 and α_2 have been transformed into p and q . For example, the vertical shear slowness has become

$$\bar{\gamma}^s = [(c^s)^{-2} - p^2 + q^2]^{1/2}. \quad (62)$$

Next, the integrand is analytically continued into the complex p plane. We define $\text{Re}\{\bar{\gamma}^s\} \geq 0$ to keep the square root in Eq. (62) single valued. Branch points occur at $\pm p^s = \pm [(c^s)^{-2} + q^2]^{1/2}$ due to the occurrence of $\bar{\gamma}^s$ in the integrand of Eq. (61); these will be supplemented with appropriate branch cuts as indicated in Fig. 3. In the right half of the complex p plane, the complex solutions of

$$\tau = pr + z\bar{\gamma}^s = \text{real} \quad (63)$$

form the so-called Cagniard contour. In Fig. 3 the upper and lower parts of this contour have been labeled C^+ and C^- . In the present case, the Cagniard contour is given by

$$p = \frac{r\tau}{R^2} \pm i \frac{z}{R^2} [\tau^2 - T^2(q)]^{1/2}, \quad (64)$$

where $R = (r^2 + z^2)^{1/2}$ is the total source–receiver distance and

$$T(q) = R[(c^s)^{-2} + q^2]^{1/2} \quad (65)$$

is the value of the parameter τ in the point p_0^s where the contour crosses the real p axis ($p_0^s < p^s$). Jordan's lemma and Cauchy's theorem form the tools that enable the deformation of the imaginary path of the slowness integral into the Cagniard contour; note that upon deformation no interference with p^s and the right branch cut occurs.²⁶ In view of

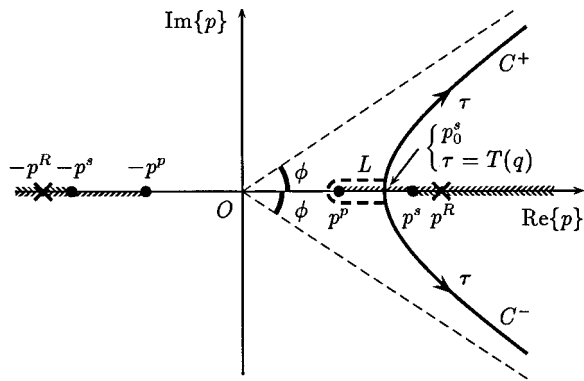


FIG. 3. The complex p plane with the shear branch points $\pm p^s$, the associated branch cuts, and the Cagniard contour $C^+ + C^-$ for the shear wave speed. In the P/SV case, the poles $\pm p^R$ show up, which are associated with the occurrence of interface waves. In case of $SV \rightarrow P$ and $SV \rightarrow SV$ transmission, the compressional branch points $\pm p^p$ need be added. This may necessitate an additional loop L , which is associated with the occurrence of a head wave.

the symmetries of the integrand and the Cagniard contour with respect to $q=0$ and the real p axis, it is allowed to write

$$\hat{g}_{ij} = \frac{1}{\pi^2} \int_0^\infty \text{Im} \left\{ \int_{C^+} \bar{\Pi}_{ij} \frac{\exp(-s\tau)}{1 + s(\frac{1}{2}z^t \mu \bar{\gamma})} dp \right\} dq. \quad (66)$$

At this stage we perform the inverse Laplace transformation of the integrand by inspection. This results in the space-time domain equation

$$g_{ij} = \frac{1}{\pi^2} \int_0^\infty \text{Im} \left\{ \int_{C^+} \bar{\Pi}_{ij} \frac{1}{\frac{1}{2}z^t \mu \bar{\gamma}} H(t-\tau) \times \exp \left(\frac{\tau-t}{\frac{1}{2}z^t \mu \bar{\gamma}} \right) dp \right\} dq. \quad (67)$$

Since τ increases monotonically when going from p_0^s to infinity along C^+ , we may replace the integration over complex p by an integration over real τ . This gives

$$g_{ij} = \frac{1}{\pi^2} \int_0^\infty \int_{T(q)}^\infty \text{Im} \left\{ \bar{\Pi}_{ij} \frac{\partial p}{\frac{1}{2}z^t \mu \bar{\gamma}} \exp \left(\frac{\tau-t}{\frac{1}{2}z^t \mu \bar{\gamma}} \right) \right\} \times H(t-\tau) d\tau dq. \quad (68)$$

By interchanging the integration sequence we arrive at the final form

$$g_{ij} = \frac{H(t-T_{\text{arr}}^{BW})}{\pi^2} \int_{T_{\text{arr}}^{BW}}^t \int_0^{Q(\tau)} \text{Im} \left\{ \bar{\Pi}_{ij} \frac{\partial p}{\frac{1}{2}z^t \mu \bar{\gamma}} \times \exp \left(\frac{\tau-t}{\frac{1}{2}z^t \mu \bar{\gamma}} \right) \right\} dq d\tau, \quad (69)$$

in which $T_{\text{arr}}^{BW} = T(0) = R/c^s$ denotes the arrival time of the SH body wave and $Q(\tau)$ is the inverse of $T(q)$ for $q \geq 0$. The fact that both integrals in this expression are bounded eases its numerical evaluation in comparison with Eq. (68). Knowing the space-time domain Green's function g_{ij} , the space-time domain state vector component v_i is found as [cf. Eq. (57)]

$$v_i = \partial_t^2 (g_{ij} * F_j^S). \quad (70)$$

Note that we may also apply the above inverse transformation process in case of a lossy fracture, provided that for the relevant function $\hat{z}^t(s)$ the inverse Laplace transformation of the integrand of Eq. (66) is known. This is the case, for example, when a lossy fracture may be represented by a rational function $\hat{z}^t(s) = \sum_{m=0}^M a_m s^m / \sum_{n=0}^{N \geq M} b_n s^n$. The main problem is then to perform the inverse Laplace transformation of the factor

$$\frac{1}{1 + s[\frac{1}{2}\hat{z}^t(s)\mu\bar{\gamma}]} = \frac{\sum_{n=0}^{N \geq M} b_n s^n}{\sum_{n=0}^{N \geq M} b_n s^n + \frac{1}{2}\mu\bar{\gamma} \sum_{m=0}^M a_m s^{m+1}} \quad (71)$$

that occurs in the integrand of Eq. (66). This may always be achieved by first applying the partial fraction representation of the latter rational function (Henrici²⁷) and subsequently recognizing the time domain equivalents of the terms thus obtained.

B. Interface waves

For the discussion of interface waves we consider a state vector quantity that is associated with a transmitted²⁸ P or SV wave that is caused by an incident P wave (although interface waves may also be generated by incident SV waves). The analysis that has been presented in the previous subsection may easily be adapted to deal with this case. A second set of branch points $\pm p^p = \pm[(c^p)^{-2} - p^2 + q^2]^{1/2}$ needs to be introduced in view of the occurrence of γ^p in Eqs. (51) and (52), and the Cagniard contour now has to satisfy $\tau = pr + z\bar{\gamma}^p = \text{real}$ because the propagation occurs with the compressional wave speed. An important point of difference with the analysis above is that the factor c in the denominators of Eq. (50) always gives rise to two poles $\pm p^R$ on the real p axis.²⁹ The location of these poles is the same as for the classical Rayleigh wave poles (Aki and Richards,³⁰ De Hoop¹⁸). Since the point p_0^p where the Cagniard contour crosses the real p axis is located such that $p_0^p < p^p < p^s < p^R$, no special provisions are necessary to enable the deformation of the imaginary integration path into the Cagniard contour. However, the right pole will exert its influence on the result of the integration along the contour. This effect is especially important when the angle $\phi = \arctan(z/r)$ is small, i.e., when the contour is close to the real p axis. The proximity of the pole will manifest itself as a contribution to the time domain result that may be associated with the occurrence of interface waves. In the limiting case $\phi=0$, the Cagniard contour forms a loop around the right branch cut, and the pole will give rise to an explicit interface wave. The occurrence of s in the denominators of Eq. (50) indicates that, as opposed to the classical Rayleigh waves, the interface waves are dispersive (hence the alternative notion generalized Rayleigh waves). The existence of these dispersive interface waves has first been derived by plane wave, frequency-domain analysis (Pyrak-Nolte and Cook¹⁶) and has later been confirmed by direct observation (Pyrak-Nolte *et al.*,^{31,32} Pyrak-Nolte and Nolte³³). In the lat-

ter references it has been noted that these waves may be viewed as two mutually coupled Rayleigh waves that propagate along the surfaces of both half-spaces.

C. Head waves

To complete the discussion of the different wave phenomena, we consider a state vector quantity that is associated with a transmitted²⁸ P or SV wave that is caused by an incident SV wave. In this situation, the procedure in Subsection V A remains applicable, except that the occurrence of γ^p in Eqs. (51) and (52) gives rise to two additional branch points $\pm p^p$ on the real p axis. Especially for small angles $\phi = \arctan(z/r)$ it may occur that $p_0^s > p^p$. If this happens it is necessary to supplement the Cagniard contour with an additional loop L in order to perform the contour deformation, see Fig. 3. This will result in a contribution to the time domain result that is associated with the occurrence of a head wave. Substitution of $p = 1/c^p$ in Eq. (63) shows that the arrival time of the head wave equals

$$T_{\text{arr}}^{HW} = r/c_p + z[(c^s)^{-2} - (c^p)^{-2}]^{1/2}. \quad (72)$$

For a detailed analysis of the head wave case we refer to Van der Hijden²⁰ and De Hoop.¹⁹ We remark that the head wave contribution may be present in both the transmitted P wave and the transmitted SV wave that are caused by an incident SV wave. The same effect shows up in case of reflection.

VI. NUMERICAL RESULTS

We have implemented the above theory in a computer program. The core task of this program is the numerical evaluation of integrals like those in Eq. (69). To reduce the required computing time, we have removed the singularity of the integrand at $q = Q(\tau)$ by the substitution

$$q = Q(\tau) \sin(\psi), \quad (73)$$

and we have made the integrand less localized near $\tau = t$ by the substitution

$$\tau = t + \kappa \ln(\tau'). \quad (74)$$

In the SH case, the real constant κ has to be assigned such a value that $1/\kappa$ has the same order of magnitude as $\text{Re}\{1/\frac{1}{2}z^t\mu\bar{\gamma}^s\}$. In the P/SV case, the order of magnitude of $1/\kappa$ must be the same as $\text{Re}\{1/ac\}$ or $\text{Re}\{1/bc\}$, whichever is relevant. We have subsequently invoked the NAG Fortran Library³⁴ routines D01AJF and D01AHF for the numerical integrations with respect to ψ and τ' .

With the aid of this computer program we have obtained exact numerical results for the incident, transmitted and reflected space-time domain wave fields in configurations with a layout as depicted in Fig. 1. Unless noted otherwise, in our numerical experiments we have employed the wave speeds $c^p = 5200$ m/s and $c^s = 3000$ m/s, the mass density $\rho = 2650$ kg/m³, the specific compliances $z^t = z^n = 10^{-9}$ m/Pa, the source–receiver distance $R = 500$ m, and $x_2^R = 0$ M. In most experiments we have used the four-point optimum Blackmann pulse of unit amplitude and 10 ms duration for the source signature.

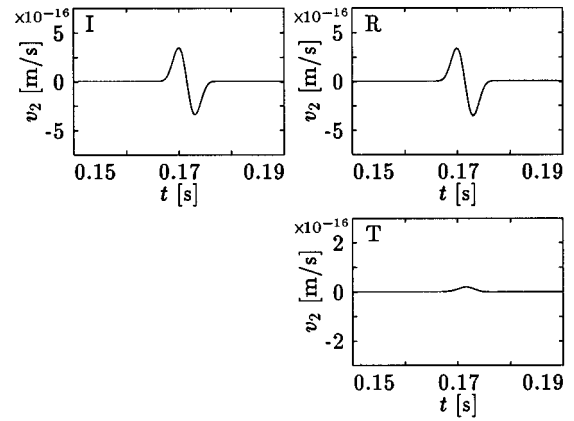


FIG. 4. The temporal behavior of the particle velocity component v_2 of the incident (I), the reflected (R), and the transmitted (T) SH wave field due to the force source component f_2 . The graphs apply to the angle $\phi = 90^\circ$.

Our first set of numerical results concerns the temporal behavior of the particle velocity component v_2 of the transmitted and reflected SH wave fields that are caused by an incident SH wave field. The latter is due to a force source with a nonzero component f_2 . Figures 4 and 5 show the results for the angles $\phi = 90^\circ$ (source vertically above the receiver) and $\phi = 5^\circ$ (source relatively close to the fracture surface). To obtain these results, in this case we have used the specific compliances $z^t = z^n = 10^{-8}$ m/Pa. For both angles the reflected wave field is almost of the same magnitude as the incident wave field, while the transmitted wave field is much smaller. This is caused by the relatively weak coupling between the half-spaces, i.e., the relatively high fracture compliance. For $\phi = 90^\circ$ the reflected wave field is of the same form as the incident wave field, and the transmitted wave field has the shape of the integrated incident waveform. This may be explained from the fact that for this angle the vertical slowness is that large that the term 1 in the denominators of Eqs. (42) and (43) may be neglected, in which case T^{SH} acts as an integrator and R^{SH} shows a proportional behavior. On the other hand, for $\phi = 5^\circ$ the shape of the reflected wave field tends to the differentiated incident waveform, and the transmitted wave field has a shape that tends to the incident waveform. Now the vertical slowness is that

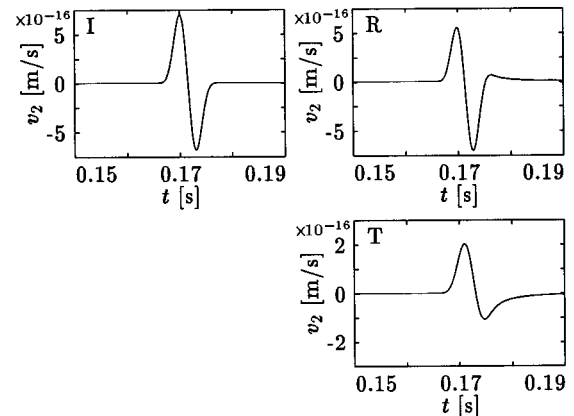


FIG. 5. The temporal behavior of the particle velocity component v_2 of the incident (I), the reflected (R), and the transmitted (T) SH wave field due to the force source component f_2 . The graphs apply to the angle $\phi = 5^\circ$.

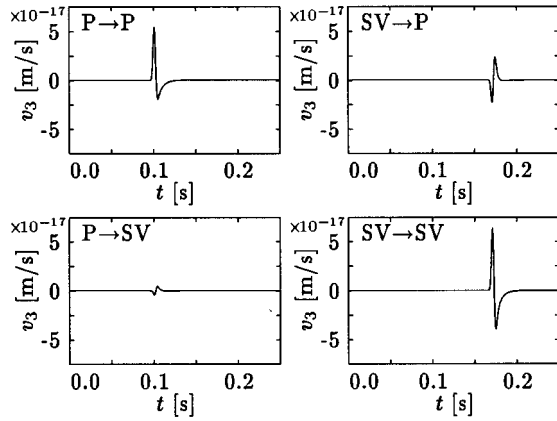


FIG. 6. The temporal behavior of the particle velocity component v_3 of the P and SV wave fields that are caused by transmission of the incident P and SV wave fields due to the force source component f_3 . The graphs apply to the angle $\phi = 60^\circ$.

small that the term $s(\frac{1}{2}z^t\mu\gamma^s)$ in the denominators of Eqs. (42) and (43) is relatively unimportant, and thus T^{SH} shows mainly a proportional behavior and R^{SH} mainly acts as a differentiator. These results clearly demonstrate that the linear slip interface forms a directionally dependent filter.

Next we turn our attention to the P/SV case. In analogy with the SH case above, we have determined the particle velocity component v_3 of the wave fields that are associated with the individual transmission and reflection factors in Eqs. (50) and (53). Figures 6 and 7 show the results for the transmitted and reflected wave fields, respectively, that are caused by the incident P and SV wave fields due to the force source component f_3 . For the chosen value of $\phi = 60^\circ$ no head waves occur and no interface wave contributions are visible. Note that both at transmission and at reflection a significant conversion of P waves into SV waves, and vice versa, may occur (cf. the numerical results obtained by Coates and Schoenberg¹⁴ with the aid of the finite-difference technique).

In Fig. 8 we have plotted the particle velocity component v_3 of the $P \rightarrow P$ reflected wave field due to a

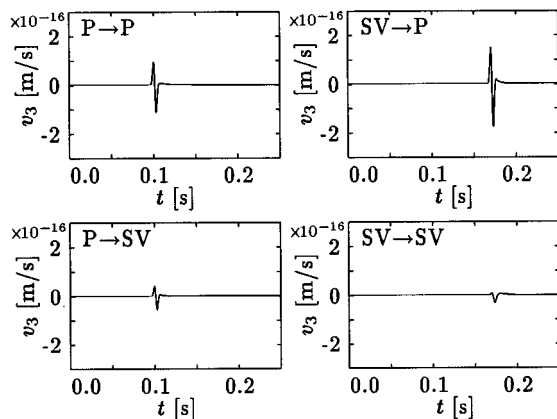


FIG. 7. The temporal behavior of the particle velocity component v_3 of the P and SV wave fields that are caused by reflection of the incident P and SV wave fields due to the force source component f_3 . The graphs apply to the angle $\phi = 60^\circ$.

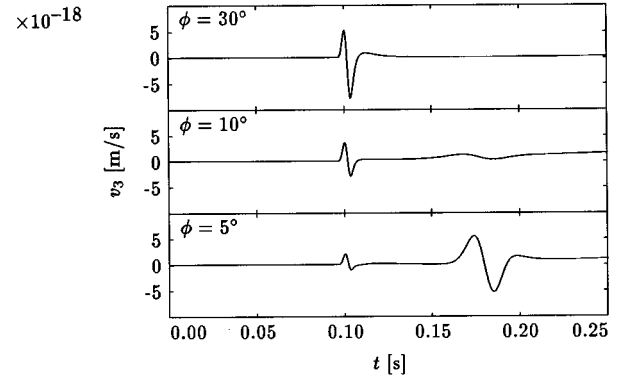


FIG. 8. The temporal behavior of the particle velocity component v_3 of the P wave field that is caused by reflection of an incident P wave field due to the force source component f_3 . The angles $\phi = 30^\circ$, 10° , and 5° are considered in order to show the relative importance of the interface wave contribution (between $t = 0.15$ s and $t = 0.20$ s) and the body wave (around $t = 0.1$ s) in relation to the relative source position.

force source component f_3 for different values of ϕ . For $\phi = 30^\circ$, the only event that may be observed is the arrival of the body wave around $t = 0.1$ s. When ϕ is decreased to 10° , the body wave becomes smaller, while a small interface wave contribution shows up between $t = 0.15$ s and $t = 0.20$ s. For $\phi = 5^\circ$, the interface contribution has become dominant over the body wave. The dominance of the body wave for large values of ϕ and the interface wave for small values of ϕ is also manifest in the $SV \rightarrow SV$ reflected wave field, as may be concluded from the experimental results in Pyrak-Nolte *et al.*³¹

If, instead, we take a fixed angle of $\phi = 10^\circ$ and employ different specific compliances $z^t = z^n$, the behavior of the particle velocity component v_3 of the $P \rightarrow P$ transmitted wave field associated with the force source component f_3 is as given in Fig. 9. Figure 10 contains the corresponding results for the $P \rightarrow P$ reflected wave field. As expected, a decreased fracture compliance results in an increased transmission of the body wave (cf. the experimental results in Pyrak-

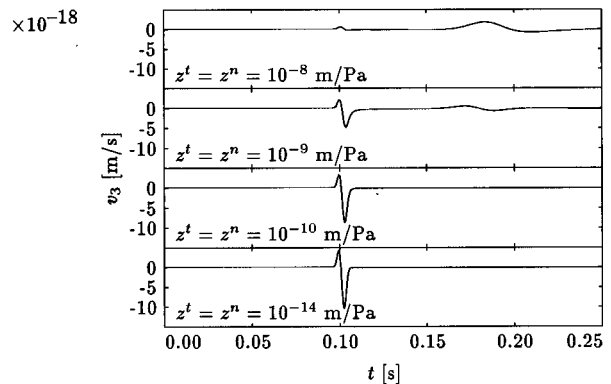


FIG. 9. The particle velocity component v_3 of the P wave field that is caused by transmission of an incident P wave field due to the force source component f_3 . Four different values of $z^t = z^n$ are considered in order to show the relative importance of the body wave and the interface wave contribution in relation to the fracture compliance. The graphs apply to the angle $\phi = 10^\circ$.

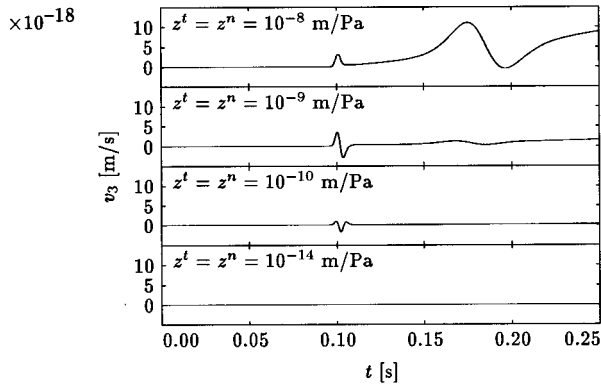


FIG. 10. The particle velocity component v_3 of the P wave field that is caused by reflection of an incident P wave field due to the force source component f_3 . Four different values of $z^t = z^n$ are considered in order to show the relative importance of the body wave and the interface wave contribution in relation to the fracture compliance. The graphs apply to the angle $\phi = 10^\circ$.

Nolte *et al.*,¹⁰ Pyrak-Nolte and Nolte³³) and a decreased reflection of the body wave. Below $z^t = z^n = 10^{-9}$ m/Pa, the interface wave contribution becomes invisibly small. Further we observe that for the reflected wave field the interface wave contribution vanishes at a higher specific compliance than the body wave.

In Fig. 11 we have depicted the displacement of two corresponding points on opposite sides of the fracture during the time interval in which the interface wave contribution is present. The applied source is of the volume injection type, the source signature is a Dirac delta pulse, and $\phi = 5^\circ$. The movement of the point at the upper fracture surface (U) is the result of the combined effects of the incident P wave field and the reflected P and SV wave fields, while the movement of the point at the lower fracture surface (L) is the result of the combined effects of the transmitted P and SV wave fields (note that the curves start with a nonzero offset due to the

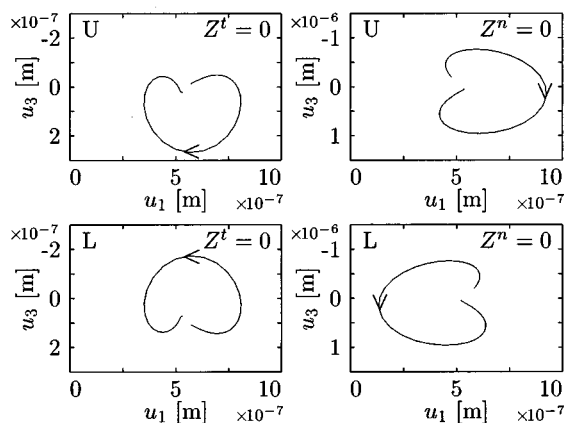


FIG. 11. The displacement u_3 vs u_1 for a particle at the upper fracture surface (U) and the lower fracture surface (L). The observation time ranges from $t = 0.15$ s to $t = 0.25$ s. During this period, the significant contribution is associated with the occurrence of interface waves. The two extreme cases $z^t = 0$ m/Pa, $z^n = 10^{-9}$ m/Pa (left) and $z^t = 10^{-9}$ m/Pa, $z^n = 0$ m/Pa (right) have been considered. The source has a delta pulse signature and is of the volume injection type. In both cases the angle $\phi = 5^\circ$ has been applied.

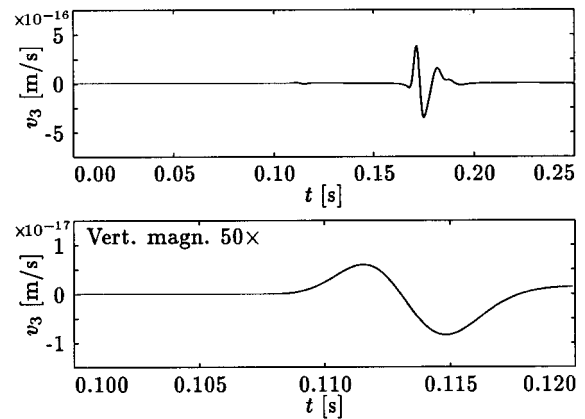


FIG. 12. The particle velocity component v_3 of the SV wave field that is caused by reflection of an incident P wave field due to the force source component f_3 . Here the angle $\phi = 5^\circ$ has been applied. The arrival around $t = 0.11$ s is the head wave, and between $t = 0.16$ s and $t = 0.20$ s the body wave and the partly overlapping interface wave contribution arrive. In the lower graph a close-up of the head wave is shown.

movement before $t = 0.15$ s). In the left two graphs, which apply to the case $z^t = 0$ m/Pa and $z^n = 10^{-9}$ m/Pa, both points have equal transversal displacements and opposite normal displacements. In the right two graphs, which apply to the dual case $z^t = 10^{-9}$ m/Pa and $z^n = 0$ m/Pa, both points show opposite normal displacements and equal transversal displacements. In both cases the particle movement is retrograde, just like a conventional Rayleigh wave (cf. the numerical results obtained by Gu *et al.*¹⁵ with the aid of the boundary integral equation method). The displacements in the graphs on the left side and the right side of Fig. 11 correspond to those of the fast and slow interface waves, respectively, that have been described by Pyrak-Nolte and Cook.¹⁶

In Fig. 12 we have plotted the particle velocity component v_3 of the $SV \rightarrow P$ reflected wave field due to a force source component f_3 for an angle $\phi = 5^\circ$. In the upper graph, at $t = 0.11$ s the arrival of a head wave is visible. The dominant part of the response (between $t = 0.16$ s and $t = 0.20$ s) is due to the body wave and the partially overlapping interface wave contribution. In the lower graph we have presented a close-up of the head wave. In the same way as in Fig. 10, in Fig. 13 we have plotted the head wave for different values of $z^t = z^n$. Obviously, just like the reflected body wave and the interface wave contribution, the head wave becomes invisibly small below a certain value of the fracture compliance.

Finally, in Fig. 14 we have presented the counterpart of Fig. 12 for the $SV \rightarrow SV$ reflected wave field. The figure shows the arrival of a head wave at $t = 0.11$ s and the occurrence of a body wave and a partially overlapping interface wave contribution between $t = 0.16$ s and $t = 0.20$ s. The presence of a head wave in case of $SV \rightarrow SV$ reflection is in close agreement with our theory.

VII. CONCLUSIONS

In this paper we have presented a method for the exact analysis of the transmitted and reflected elastodynamic wave

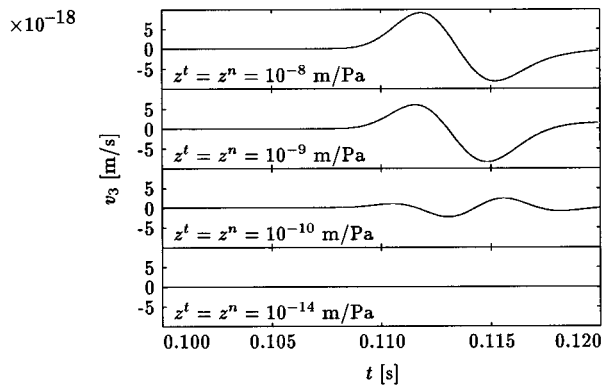


FIG. 13. The particle velocity component v_3 of the SV wave field that is caused by reflection of an incident P wave field due to the force source component f_3 . Four different values of $z^t = z^n$ are considered in order to show the size of the head wave in relation to the specific fracture compliance. In all four graphs the angle $\phi = 5^\circ$ has been applied.

fields at a fracture in a homogeneous and isotropic material. As a model of this fracture, we have employed a plane, infinitely large, linear slip interface in between two half-spaces containing the same medium. The interface behavior is assumed to show up-down and rotational symmetry. Our method distinguishes itself through the capability to deal with the transmission and the reflection of a spatially curved incident wave field due to a point source, and the ability to generate time domain waveforms.

In the theoretical part of the paper we have first subjected the space-time domain problem to a temporal Laplace transformation and a two-dimensional, horizontal spatial Fourier transformation. The subsequent transform domain analysis has provided us with the transmission and reflection factors for the decoupled SH and P/SV wave systems. To complete the method, we have presented a process for finding the space-time domain counterparts of the transform domain results. The application of the complex slowness plane

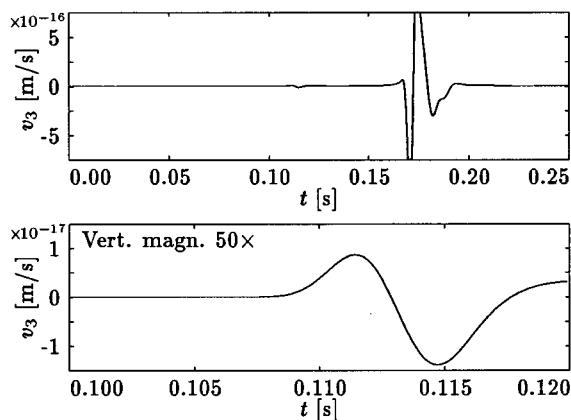


FIG. 14. The particle velocity component v_3 of the SV wave field that is caused by reflection of an incident SV wave field due to the force source component f_3 . Here the angle $\phi = 5^\circ$ has been applied. The arrival around $t = 0.11$ s is the head wave, and between $t = 0.16$ s and $t = 0.20$ s the body wave and the partly overlapping interface wave contribution arrive. In the lower graph a close-up of the head wave is shown.

in the inverse transformation process has enabled us to clearly identify the different wave phenomena (body waves, interface waves, and head waves). Although the analysis in this paper has been carried out for lossless fractures, we have discussed that it requires only a minor modification of the inverse transformation process to deal with lossy fractures.

In the numerical part of the paper, we have presented various exact results that have been generated by means of a computer implementation of the theory. Using SH waves, we have demonstrated the directionally dependent filter behavior of the fracture. For the P/SV case, besides the $P \rightarrow P$ and $SV \rightarrow SV$ reflected and transmitted wave fields, we have shown all reflected and transmitted wave fields that result from conversions of P waves into SV waves and vice versa. Further we have seen that in the P/SV case there exist contributions that are associated with the occurrence of interface waves propagating along the fracture. This interface wave contribution becomes significant when the source is relatively close to the fracture and the transversal and normal specific compliances of the fracture are relatively large. By taking either the transversal specific compliance or the normal specific compliance equal to zero, we have obtained particle movements that correspond to those of the fast and slow interface waves described in the literature. Moreover, as predicted by our theory, for both the $SV \rightarrow P$ and $SV \rightarrow SV$ reflected waves we have found that head waves show up when the source is relatively close to the fracture and when the specific compliances are relatively large. Since the numerical results clearly show the dependence of the generated waveforms and wave phenomena on the specific compliances of the fracture, we expect that for an unknown fracture these parameters may be obtained from the measured transmitted or reflected wave fields by way of inversion.

ACKNOWLEDGMENTS

The research of the first author has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences. The research reported in this paper has been financially supported through research grants from the Stichting Fund for Science, Technology and Research (a companion organization to the Schlumberger Foundation in the U. S. A.).

- ¹M. Schoenberg and G. U. Haugen, "Fracture characterization," in *Proceedings of the 3rd SEGJ/SEG International Symposium — Geotomography: Fracture Imaging* (The Society of Exploration Geophysicists of Japan, Tokyo, 1995), pp. 31–38.
- ²R. J. O'Connell and B. Budiansky, "Seismic velocities in dry and saturated cracked solids," *J. Geophys. Res.* **79**, 5412–5426 (1974).
- ³J. A. Hudson, "Wave speeds and attenuation of elastic waves in material containing cracks," *Geophys. J. R. Astron. Soc.* **64**, 133–150 (1981).
- ⁴M. Schoenberg, "Reflection of elastic waves from periodically stratified media with interfacial slip," *Geophys. Prosp.* **31**, 265–292 (1983).
- ⁵J. E. White, *Underground Sound* (Elsevier Science, New York, 1983), Chap. 3, pp. 72–81.
- ⁶M. Schoenberg and C. M. Sayers, "Seismic anisotropy of fractured rock," *Geophysics* **60**, 204–211 (1995).
- ⁷K. Kendall and D. Tabor, "An ultrasonic study of the area of contact between stationary and sliding surfaces," *Proc. R. Soc. London, Ser. A* **323**, 321–340 (1971).
- ⁸M. Schoenberg, "Elastic wave behavior across linear slip interfaces," *J. Acoust. Soc. Am.* **68**, 1516–1521 (1980).

- ⁹C. Kitsunezaki, "Behavior of plane elastic waves across a plane crack," J. Mining Coll. Akita Univ. Ser. A **6**, 173–187 (1983).
- ¹⁰L. J. Pyrak-Nolte, L. R. Myer, and N. G. W. Cook, "Transmission of seismic waves across single natural fractures," J. Geophys. Res. **95**, 8617–8638 (1990).
- ¹¹C.-J. Hsu and M. Schoenberg, "Elastic waves through a simulated fractured medium," Geophysics **58**, 964–977 (1993).
- ¹²L. J. Pyrak-Nolte, L. R. Myer, and N. G. W. Cook, "Anisotropy in seismic velocities and amplitudes from multiple parallel fractures," J. Geophys. Res. **95**, 11 345–11 358 (1990).
- ¹³M. Savić, "Ultrasonic scattering from a hydraulic fracture: Theory, computation and experiment," Thesis, Delft University of Technology, Delft, The Netherlands, 1995, Chap. 4.
- ¹⁴R. T. Coates and M. Schoenberg, "Finite-difference modeling of faults and fractures," Geophysics **60**, 1514–1526 (1995).
- ¹⁵B. Gu, N. G. W. Cook, K. T. Nihei, and L. R. Myer, "Modeling elastic wave propagation in a medium with non-welded interfaces using the boundary integral equation method," in *Computer Methods and Advances in Geomechanics*, edited by H. J. Siriwardane and M. M. Zaman (Balkema, Rotterdam, 1994), pp. 387–392.
- ¹⁶L. J. Pyrak-Nolte and N. G. W. Cook, "Elastic interface waves along a fracture," Geophys. Res. Lett. **14**, 1107–1110 (1987).
- ¹⁷A. T. de Hoop, "A modification of Cagniard's method for solving seismic pulse problems," Appl. Sci. Res. Sec. B **8**, 349–356 (1960).
- ¹⁸A. T. de Hoop, "Theoretical determination of the surface motion of a uniform elastic half-space produced by a dilatational, impulsive, point source," in *Proceedings Colloque International du C.N.R.S.* (Marseille, 1961), No. 111, pp. 21–31.
- ¹⁹A. T. de Hoop, "Acoustic radiation from impulsive sources in a layered fluid," Nieuw Archief voor Wiskunde **6**, 111–129 (1988).
- ²⁰J. H. M. T. van der Hijden, *Propagation of Transient Elastic Waves in Stratified Anisotropic Media* (North-Holland, Amsterdam, The Netherlands, 1987), Chaps. 4, 5.
- ²¹C. H. Chapman, "A new method for computing synthetic seismograms," Geophys. J. R. Astron. Soc. **54**, 481–518 (1978).
- ²²J. R. Carson, *Electric Circuit Theory and the Operational Calculus* (McGraw-Hill, New York, 1926), p. 19.
- ²³D. V. Widder, *The Laplace Transform* (Princeton U. P., Princeton, 1946), Chap. II, pp. 59–63.
- ²⁴L. Cagniard, *Réflexion et Réfraction des Ondes Sismiques Progressives* (Gauthier-Villars, Paris, 1939) [in French; translated and revised version: E. A. Flinn and C. H. Dix, *Reflection and Refraction of Progressive Seismic Waves* (McGraw-Hill, New York, 1962)].
- ²⁵In the expressions for the reflected wave field, a factor $s(\frac{1}{2}z'\mu\gamma^s)/[1+s(\frac{1}{2}z'\mu\gamma^s)]$ occurs, which may be written as $1-1/[1+s(\frac{1}{2}z'\mu\gamma^s)]$. The first term gives rise to an expression that may be inverted using the conventional Cagniard–De Hoop method [De Hoop (Ref. 17), Van der Hijden (Ref. 20)] and the second term leads to an expression that may be inverted in the same way as the expression for the transmitted wave field.
- ²⁶In some circumstances a factor p^2-q^2 may show up in the denominator of $\bar{\Pi}_{ij}$, in which case a pole at $p=q$ must be encircled. Since the residue of this pole is always odd in q and therefore does not contribute to the result of the integral with respect to q , we will pay no attention to the possible occurrence of this pole.
- ²⁷P. Henrici, *Applied and Computational Complex Analysis* (Wiley, New York, 1974), Vol. 1, Chap. 7, pp. 553–562.
- ²⁸In the expressions for the reflected wave field, the factors $sac/(1+sac)$ and $sbc/(1+sbc)$ occur, which may be written as $1-1/(1+sac)$ and $1-1/(1+sbc)$, respectively. The first terms give rise to expressions that may be inverted using the conventional Cagniard–De Hoop method [De Hoop (Ref. 17), Van der Hijden (Ref. 20)] and the second terms lead to expressions that may be inverted in the same way as the expressions for the transmitted wave field.
- ²⁹For real and positive s the factors $1+sac$ and $1+sbc$ in the denominators do not cause poles anywhere in the complex p plane.
- ³⁰K. Aki and P. G. Richards, *Quantitative Seismology: Theory and Methods* (Freeman, San Francisco, 1980), Vol. I, Chap. 5, pp. 155–163, Chap. 6, pp. 214–224, 239–240.
- ³¹L. J. Pyrak-Nolte, J. Xu, and G. M. Haley, "Elastic interface waves propagating in a fracture," Phys. Rev. Lett. **68**, 3650–3653 (1992).
- ³²L. J. Pyrak-Nolte, S. Roy, and B. Mullenbach, "Interface waves propagated along a fracture," in *Proceedings of International Symposium on Recent Advances in Exploration Geophysics* (Faculty of Engineering, Kyoto University, Kyoto, Japan, 1995), pp. 51–58.
- ³³L. J. Pyrak-Nolte and D. D. Nolte, "Wavelet analysis of velocity dispersion of elastic interface waves propagating along a fracture," Geophys. Res. Lett. **22**, 1329–1332 (1995).
- ³⁴*The NAG Fortran Library Manual, Mark 16* (NAG Ltd., Oxford, 1993), Vol. 1.

Solving ray acoustic problems with a DNA computer

Yang Yang

SFA, Inc., 1401 McCormick Drive, Landover, Maryland 20785

Yuling Li

Human Genome Sciences, Inc., 9620 Medical Center Drive, Rockville, Maryland 20850

Luise S. Couchman

Code 7131, Naval Research Laboratory, Washington, DC 20375-5000

(Received 25 February 1996; revised 20 September 1996; accepted 26 September 1996)

A computational scheme that integrates the advantages of both the electronic computer and DNA computer (DNA experiment) is proposed to solve the ray acoustic problems at high frequencies. More explicitly, a conventional electronic computer is used to perform most of the sequential operations, such as solving differential equations, designing the DNA experiment, and verifying the final experimental solutions; whereas DNA annealing reaction is used to provide huge parallel operations to assemble the solutions to the scattered field for all the specified incident directions. To illustrate this idea, an example is given where the DNA reaction, rather than carrying out experimentally, is tested and simulated on an electronic computer. © 1997 Acoustical Society of America. [S0001-4966(97)04705-X]

PACS numbers: 43.20.Dk, 43.30.Cq [JEG]

INTRODUCTION

Building molecular computers has been of a great interest within the scientific community for a few decades.^{1,2} Considerable efforts have been focused on building biochips and molecular devices that mimic some components of electronic computers.²⁻⁵ However, no actual computations had even been attempted at the molecular level until, recently, Adleman⁶ realized that specific computation could be carried out on DNA molecules long before a universal molecular computer would appear. Using standard protocols of molecular biology, Adleman showed experimentally how to solve an instance of the Hamiltonian path problem (HPP) with DNA molecules. In fact, he built the first molecular computer that was loaded with not only the hardware but also the software for execution of the specific computation (his computer is referred as a DNA computer throughout this paper). More recently, Lipton⁷ proposed a scheme to solve the "satisfaction" problem (SAT) on a DNA computer. Both HPP and SAT are combinatorial problems and have been proven to be NP complete for which no efficient algorithm exists. It seems that Adleman's method is inherently suitable to solve combinatorial problems, but the impact of Adleman's pioneer work goes far beyond this specific field; it may fundamentally change the way that we do computations in physics and engineering. We will show here that a class of interesting physics problems, such as ray acoustic problems, can also be solved efficiently on a DNA computer.

Computation of many fundamental wave problems in physics and engineering is known to be very time consuming at high frequencies. The ray method, originated in geometric optics, is often used to overcome this kind of difficulty. It replaces the problem of solving linear partial differential equations; i.e., wave equations, with that of solving a set of nonlinear first-order ordinary differential equations, i.e., ray equations.^{8,9} This approximation is particularly appropriate

at high frequencies where energy flows predominantly along rays. Since it is much easier and more efficient to solve ordinary differential equations than to solve partial differential equations, this method is of practical importance and, therefore, has been extensively used to solve the wave problems in electromagnetism, acoustics, and seismology.^{10,11} It may turn out to be a powerful technique for the analysis of wave scattering from large complex structures, such as aircraft¹² or submarines.^{13,14}

In general, ray tracing is a nonlinear initial-value problem and may be cast in the forms of first-order differential equations i.e., ray equations,^{10,11,16-18}

$$\frac{d\mathbf{v}}{ds} = \mathbf{v}, \quad \frac{d\mathbf{v}}{ds} = \mathbf{G}(\mathbf{y}, \mathbf{v}), \quad (1)$$

with a general boundary condition:

$$B(s, \mathbf{y}, \mathbf{v})|_{\partial D} = 0, \quad (2)$$

where \mathbf{y} is the position vector of a point along a ray which is parameterized by its arc length s , the vector \mathbf{v} specifies the tangent direction along the ray, \mathbf{G} is a nonlinear vector function which is given and depends on the local properties of the propagation medium and the propagation direction, and the symbol ∂D denotes the exterior boundary as well as the interior interfaces. A ray path described by Eq. (1) is, in general, a curve in a Euclidean two- or three-dimensional space or in a non-Euclidean two-dimensional space, i.e., a curved surface. Efficient numerical schemes have been developed on electronic computers for tracing a ray in a 2D smooth inhomogeneous medium¹⁵ and on a smooth arbitrarily shaped surface.¹³ However, in the presence of material or geometric discontinuities in a propagation medium, determination of the scattered field via tracing ray with the existing schemes becomes very difficult if not completely impossible, because a ray may split into a number of rays whenever it hits an

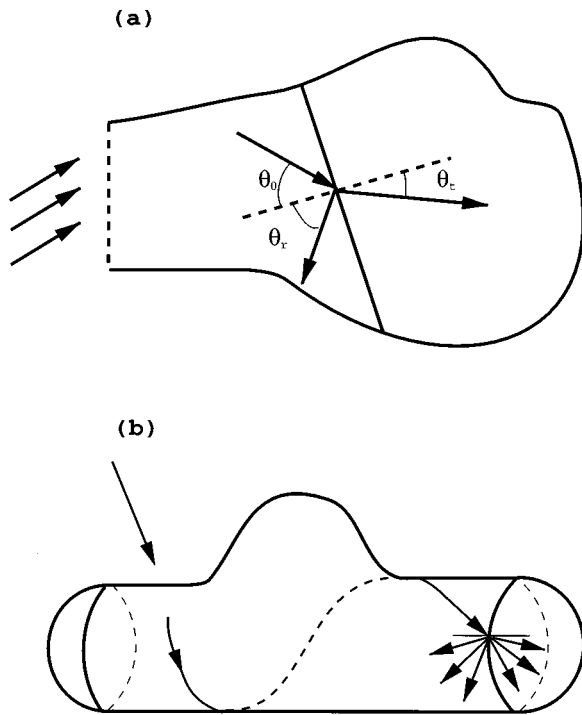


FIG. 1. A schematic picture of rays incident on a two-dimensional arbitrarily shaped rigid cavity filled with two fluid media (a), and an idealized submarine (b). On the interface in (a), a ray splits into two new rays, a reflected and a transmitted ray; while on an interface in (b), a ray splits into six new ones of three distinct types of elastic waves, longitudinal, transverse, and flexural waves.

interface of discontinuity, and, consequently, the total number of rays grows exponentially during the ray tracing. For example, in the scattering from an open-ended arbitrarily shaped rigid cavity filled with two different kinds of fluid as shown in Fig. 1(a), a ray splits into a reflected and a transmitted ray on the interface. The situation becomes worse for the rays (elastic waves) traveling over a submarine's surface as shown in Fig. 1(b), where a ray of any type of wave: A longitudinal or a transverse or a flexural wave, may split into six new rays, i.e., the transmitted and the reflected rays of three distinct wave types, at the interfaces of curvature discontinuity.

According to ray theory,^{10,11,14,16} the scattered field in a given observation direction is the sum of the contributions from those rays that pass through the receiver in the same observation direction. Searching for those rays from the source to the receiver via all the possible rays is a formidable task even on a state-of-the-art supercomputer. It is true that ray tracing is a deterministic problem, but it requires a great number of parallel operations to trace all the rays independently. This difficulty may be overcome by the molecular method similar to that of Adleman⁶ where the DNA annealing reaction is used to provide enormous parallel operations. We will show that, with some modifications on Adleman's method, this class of ray-tracing problems can be efficiently solved on a DNA computer. The main modification that needs to be introduced is that of implementing the physics laws described by Eqs. (1) and (2) into the DNA experiment.

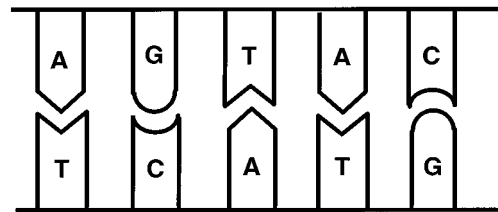


FIG. 2. A diagram of a DNA complementary structure.

This paper is outlined as follows. In Sec. I, we briefly introduce the Watson–Crick model of DNA which is the core of the molecular computer, the basic reactions that perform the parallel computation on DNA, and some standard protocols in molecular biology that may be used to extract the solutions. Then, in Sec. II, we propose a computing scheme that integrates the use of both the electronic computer and a DNA computer (DNA experiments). To illustrate the basic ideal of the scheme, each step of the scheme is tested and simulated on an electronic computer for a relative simple example in Sec. III. Finally, the numerical and experimental errors that probably occurred in the computing are discussed.

I. DNA MODEL AND ITS BASIC REACTIONS

The simplest and the most important model of DNA was proposed by Watson and Crick in 1953.¹⁹ According to their model, the DNA consists of two polymer strands. Attached to each strand are four different types of base, small ring molecules often labeled by letters of the alphabet $\{A, G, T, C\}$. The sequence of these bases along one strand may be completely arbitrary, but the sequences in both strands of a DNA must be structurally complementary, like a key fitted into a lock, that is, $A \leftrightarrow T$ and $G \leftrightarrow C$ (see the illustration in Fig. 2).

The intermolecular bonds that keep the two complementary strands of a DNA together are much weaker than those connecting nucleotides into each strand, the backbone of DNA. For example, consider a test tube containing many double-stranded DNAs of distinct sequences. When the tube is heated to a certain temperature, the intermolecular bonds are ruptured while the backbones of a DNA remain intact; consequently, all the double-stranded DNAs are separated into single-stranded DNAs. This process is called DNA melting. On the contrary, as the temperature decreases, a single-stranded DNA can find its complementary partner from billions of possible choices to form a double-stranded DNA in a matter of seconds. This reaction is called DNA annealing, which is the main operation adopted in DNA computing.⁶

One of important techniques recently developed in molecular biology is polymerase chain reaction method (PCR).²⁰ Given the sequences of two short segments, called primers, consisting of a few base pairs and located at two, sequentially distant sites along a DNA, the PCR technique can be used to multiply the very part of the DNA between the two primers by millions. The state-of-the-art automated machines for performing the PCR technique are now commercially available in the market. Adleman has

demonstrated in his experiment that the PCR instrument is an ideal instrument for amplifying the real solutions resulted from DNA computing.⁶

Nowadays, synthesizing and sequencing of DNA became routine operations and have been completely automated. All of these new technologies, however, were not developed for the general purpose of computing: In fact, they cannot perform any meaningful computation unless an effective scheme is given. Therefore, the focus of the rest of this paper is on the construction of such a scheme with which a class of ray acoustic problems can be effectively solved. The DNA computer mentioned in this paper is referred to as a set of experimental procedures that perform the following routine operations with use of the techniques described in the above: (1) synthesis of a single-stranded DNA of distinct sequence; (2) conducting DNA annealing reaction; (3) use of PCR technique to amplify the part of a DNA between two primers of given sequences; (4) sequence of a long DNA.

II. DESCRIPTION OF THE COMPUTING SCHEME

An electronic computer (EC) is conceivably much superior to a DNA computer in terms of solving nonlinear differential equations; on the other hand, a DNA computer has shown a tremendous advantage over an electronic computer in parallel computation, such as the operations of searching and matching.⁶ Integrating the advantages of both electronic and DNA computers may provide powerful leverage in computation of difficult problems. Therefore, it is of practical importance to appropriately distribute a complex task among these two different types of computers before carrying out any computation. Here, for instance, a complete ray path may be viewed as an ensemble of many elementary rays which only travel within each smooth domain and never cross any interface of discontinuity. We may first trace all the possible elementary rays on an electronic computer by solving Eq. (1), use synthetic single-stranded DNAs, i.e., oligonucleotides, to represent these rays, then implement the physics laws (the boundary condition) by building a number of joints among related rays in terms of another set of single-stranded DNAs, and finally let a DNA computer perform an annealing reaction to assemble the complete ray paths from all the elementary rays. To further illustrate this idea, let us consider the scattering by a two-dimensional open-ended rigid cavity filled with two media, as shown in Fig. 1(a). This problem has important applications in reduction of radar scattering cross sections from aircraft.²¹ The following scheme is proposed to obtain the complete ray paths for all the possible incident directions by using both an electronic and a DNA computer:

(1) Discretize the aperture and the interface of discontinuity into a number of elements and nodal points as shown in Fig. 3, and assign each nodal point an integer, $i_m = \text{ind}(m, i)$, where $m = 1, 2$ denote the planes of the aperture and the interface respectively, and i means that the node is the i th point in the plane of $m = 1$ or 2. The position vector of a nodal point is store in an array, denoted by $\text{cord}(i_m, j)$ with $j = 1, 2$.

(2) Generate all the elementary rays within each domain of the same medium: On each nodal point, say i_m , sample

the angular space evenly to create a number of initial vectors for the elementary rays radiated out from i_m , and then trace these rays on an electronic computer by solving Eq. (1) until they hit the other nodal points on the interface or on the aperture. An elementary ray, no matter how complicated it may be, is labeled by an integer as a name tag, that is,

$$e_{imn} = \text{nod_ray}(i_m, n), \quad (3)$$

where integer n refers to the n th path radiated out from the point i_m . On a DNA computer, each of these integers, e_{imn} , corresponds to a distinct 20-mer oligonucleotide, i.e., a single-stranded DNA of 20 units, except for those rays incident at the aperture. They are labeled by e_k^{in} with $k = 1, 2, 3, \dots$, and correspond to only 10-mer oligonucleotides. The choice of the length or the number of nucleotides for each oligonucleotide is based on the consideration of maximizing the number of possible distinct oligonucleotides while minimizing the cost of synthesizing these oligonucleotides. The detailed shape of each elementary ray is stored on an electronic computer (EC). For example, two elementary rays displayed in Fig. 3, one in the open-ended rigid cavity and the other on the surface of an idealized submarine, are determined numerically on an EC through the use of the ray-tracing scheme developed by Yang *et al.*;¹³ their coordinates are stored on an EC but can be retrieved by their name tags, i.e., the integers e_{imn} . The sequence information will be output from the electronic computer to the DNA computer for synthesis of the oligonucleotides. Note that due to the discretization of the interface and the aperture, the end point of an elementary path may not be exactly at a nodal point. We may move it to its nearest nodal point as an approximation. The numbers of two nodal points connected by an elementary path, say e_{imn} , are stored in an array: $i_1 = \text{nop}(e_{imn}, 1)$ and $i_2 = \text{nop}(e_{imn}, 2)$; the path, e_{imn} , starts from point i_1 and ends at point i_2 .

(3) Build the joint between two elementary rays which are connected by the physics law such as Snell's law:

$$c_0^{-1} \sin \theta_0 = c_1^{-1} \sin \theta_1, \quad (4)$$

where c_0 and c_1 are the wave speeds along the two rays, θ_0 is the angle between the normal to the interface and the direction of the elementary ray incident on the interface, and θ_1 is the angle between the normal and the direction of the reflected or transmitted elementary ray [Fig. 1(a)]. When both rays are in the same medium, then $c_1 = c_0$ and Eq. (4) is just the specular reflection condition: $\theta_1 = \theta_0 = \theta_r$ where θ_r is the reflection angle defined in Fig. 1(a); whereas, if they are in two different media, then $c_1 \neq c_0$, and Eq. (4) gives the transmission condition: $\theta_1 = \theta_t \neq \theta_0$ where θ_t is the transmission angle shown in Fig. 1(a). For an elementary ray, say e_0 , that hits the interface at the nodal point $i_2 = \text{nop}(e_0, 2)$, we search for its reflected and transmitted ray, say e_r and e_t , according to the condition (4), on an electronic computer from all the rays radiated out from the node i_2 . The joints between the ray, e_0 , and its reflected and transmitted ray, e_r and e_t , are labeled by an array of integers, $\text{noe}(e_0, j)$ with $j = 1, 2$, respectively. Each of these joints then corresponds to a distinct 20-mer oligonucleotide in such a way that its first ten nucleotides are Watson-Crick complemen-

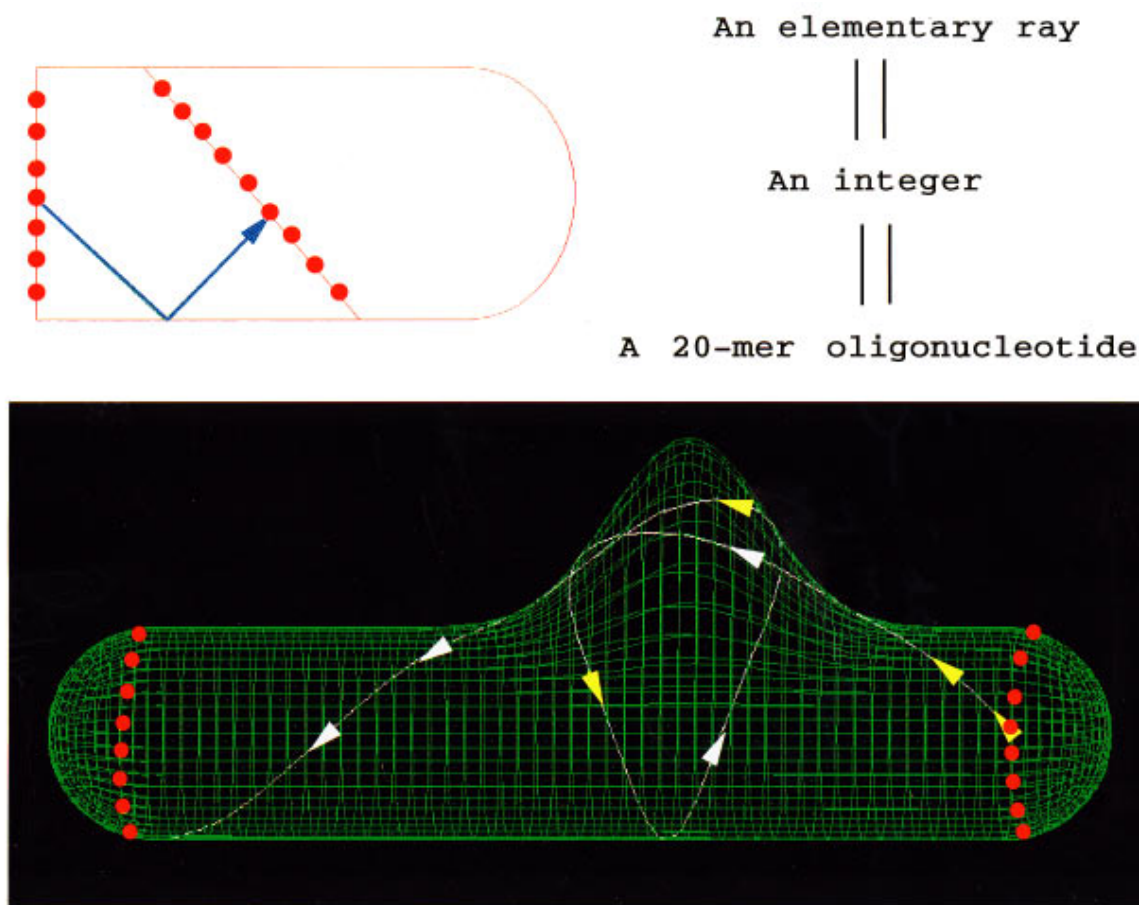


FIG. 3. The interfaces and the aperture are discretized into a number of elements and nodes. An elementary ray is launched from one node and traced numerically by using the scheme of Yang *et al.*¹³ to the other node. Each elementary ray is labeled by a name tag, i.e., an integer which corresponds to a distinct 20-mer oligonucleotide in the DNA experiment. The white or the yellow arrow is used to indicate whether the particular part of the ray is on the front or the back side of the submarine.

tary to the last ten nucleotides of the oligonucleotide associated with the elementary ray e_0 , while its last ten nucleotides are then complementary to the first ten nucleotides of the oligonucleotide associated the elementary ray e_r or e_t (see the illustration in Fig. 4). This step demonstrates the main difference between the present method and the method in Ref. 6: Here, Snell's law must be enforced at each node and thus requires that a ray pointing to a node can be connected only to its reflected and transmitted ray. As a result, there are a number of distinct joints on each node; whereas, for the HPP problem considered in Ref. 6, there is only one unique joint associated with each node because no physics laws or constraints are required to be enforced at each node so that an edge pointing to a node is connected to all the edges leaving that node.

(4) A number of vectors, so-called observation vectors, are introduced to specify different observation directions and are labeled by integers: O_m , $m=1,2,\dots$. Each of these vectors corresponds to a distinct 10-mer oligonucleotide, and is then connected by the joints to those elementary rays, denoted by e_j^{out} with $j=1,2,\dots$, that exit the cavity in the same direction of the observation vector O_m . The oligonucleotide associated with the joint between O_m and e_j^{out} is built in a way similar to those shown in the previous step: its first ten nucleotides are complementary to the last ten nucle-

otides of the oligonucleotide associated with the ray e_j^{out} while its last ten nucleotides are complementary to all the ten nucleotides of O_m .

(5) Mix together all the oligonucleotides representing the elementary rays, the joints, and the observation vectors in the experiment of DNA annealing reaction. Following the specific base pairing: A with T and G with C in the process of DNA annealing, the oligonucleotide associated with each joint is most likely to form a connection between those two whose first or second half part is complementary, respectively, to the second or the first half of the joint oligonucleotide (Fig. 4). Consequently, many double-stranded DNA molecules are generated during the annealing reaction; they encode all the solutions for all the specified incident directions.

(6) Amplify the real solution by the PCR technique:²⁰ Use the oligonucleotides corresponding to \vec{e}_k^{in} and O_m as primers to run PCR. Here, the sequence of \vec{e}_k^{in} is Watson-Crick complementary to that of e_k^{in} . Thus, only those paths which begin with the specified incident direction and end with the specified common observation vector are amplified. The resulting solutions may contain none or several different paths that exit the cavity at different nodal points but in the same observation direction. Furthermore, one may use the

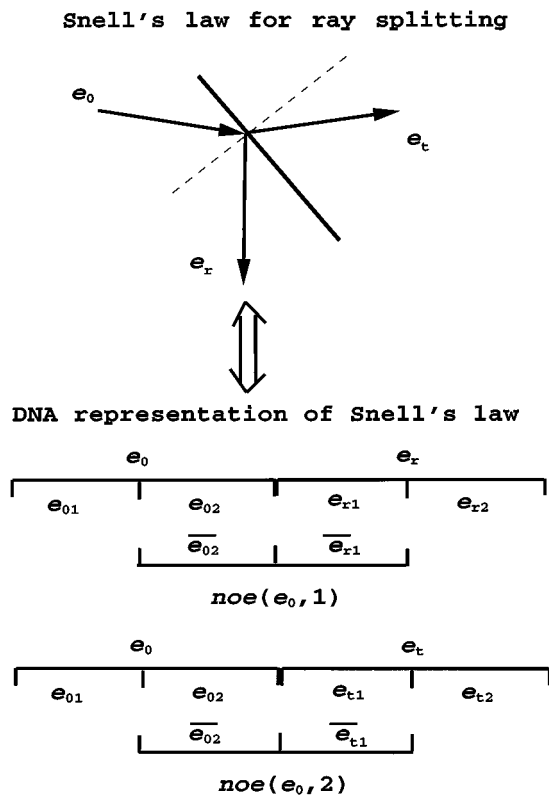


FIG. 4. A schematic picture to show how Snell's law is enforced on the construction of the joints, labeled by $noe(e_0, 1)$ and $noe(e_0, 2)$, between the ray, e_0 , incident on the interface and its reflected and transmitted ray, e_r and e_t , respectively. Both the joints and the elementary rays correspond to distinct 20-mer oligonucleotides. Each of these oligonucleotides is divided into two parts of equal length, i.e., ten nucleotides. The first and the last half of the oligonucleotide associated with the joint, $noe(e_0, m)$ with $m=1$ or 2 , are complementary, respectively, to the last half of the oligonucleotide associated with the elementary ray e_0 and to the first half of the oligonucleotide associated with the elementary ray e_r or e_t .

oligonucleotides corresponding to \bar{e}_k^{in} and e_j^{out} as primers to further differentiate these different paths.

(7) Sequence the amplified DNAs to obtain the real solutions.

(8) Examine the experiment results on an electronic computer to ensure that the physics law is satisfied at the joint between any two connected elementary rays.

In summary, steps 1–4 can be executed efficiently on an electronic computer, which then outputs to a DNA computer the sequence information of the oligonucleotides corresponding to all the elementary rays, their joints, and the observation vectors. Based on this information, the oligonucleotides are prepared; steps 5–7 are then executed experimentally. Finally, the experiment results are fed back to the electronic computer for verification (step 8).

III. EXAMPLE AND DISCUSSION

The experimental procedures involved in steps 5–7 of the proposed scheme are all routine protocols in molecular biology. Note that the kinds of operations provided by these protocols, such as annealing, PCR, and sequencing, can be simulated, at least approximately, on an electronic computer. Thus, instead of doing the actual experiment by synthesizing

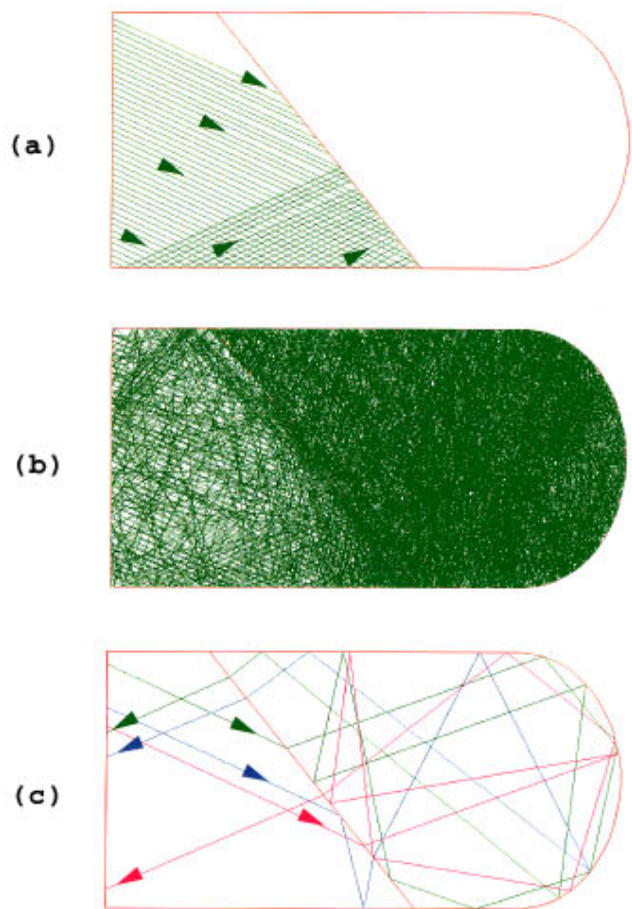


FIG. 5. (a) The 40 rays are incident in the direction of -20° on a rigid cavity composed of two parallel straight lines and a half-circle, and hit the interface of two acoustic media. (b) These rays are traced up to six number of times of transmission across the interface. This figure contains all the solutions for any given incident direction. (c) The ray paths in (b) are then rendered with the simulation of steps 6 and 7 and only three of them are left in the observation direction of 200° .

a large number of oligonucleotides, we test and simulate every step of the present scheme on an conventional electronic computer (EC). The objective of carrying out this kind of simulation on EC is to elucidate the basic ideas of the present scheme and the class of the problems that the scheme can efficiently solve. A conceivable difficulty is to fully realize on an EC the huge parallel operations provided by the DNA annealing reaction. Here we simulate the reaction sequentially by tracing rays up to a certain number of times of reflection and transmission as an approximation. For instance, based on a given incident direction, we determine the labels (the name tags) of all the beginning elementary rays, then elongate each path according to the connectivity information, $noe(e_0, j)$. Whenever a path hits the interface, it splits into two new paths and elongation continues along these new paths. This process is repeated many times until the paths exit the cavity or after an assigned number of times of reflection and transmission. Finally, the paths leaving the cavity in all different observation directions may be obtained.

A relatively simple example is presented in Fig. 5. The 2-D rigid cavity is composed of two parallel straight lines

and a half-circle as an endcap. Inside the cavity there are two different fluid media separated by an interface inclined at the angle of 135° with respect to the axis of the cavity. The ratio of the sound speed of fluid 1 to that of fluid 2, i.e., c_1/c_2 , is 1.5. The aperture and the interface are equally divided into 40 and 60 elements, respectively; each node launches an elementary ray at every 6° . Figure 5(a) shows 40 rays of plane acoustic wave incident upon the cavity in the direction of -20° with respect to the central axis of the cavity. These rays are traced until they leave the cavity or up to six transmissions across the interface; the resulting ray paths are displayed in Fig. 5(b). One can see that the domain with the endcap is densely and uniformly covered by the elementary rays. Figure 5(b) results from the simulation of step 5. After rendering the result as shown in Fig. 5(b) with the simulation of steps 6 and 7; we obtain the complete ray paths leaving the cavity in the direction of 200° and plot them in Fig. 5(c). Of course, there would be more than those three complete ray paths in Fig. 5(c) if one traces each ray for a longer distance or includes more initial rays.^{12,21}

Once the ray paths are determined, it is trivial to compute the ray-tube area and the amplitude associated with each ray path, and, furthermore, the scattered field in all possible observation directions. For a ray traveling in a 2-D or 3-D Euclidean space, the detailed expressions of the ray-tube areas and the associated amplitude are given by Eqs. (8-7.8), (8-8.9), and (8-8.13) in the text of Pierce.¹¹ For a ray traveling on a smooth curved surface, the ray-tube area may be obtained by numerically integrating a set of ordinary differential equations along the ray path using the scheme developed by Yang *et al.*¹³ The difficulty of dealing with caustics may be overcome by extending the real ray-tube area into complex space through the use of Gaussian beam methods.¹⁴

Conceivably, the present scheme is directly applicable to the class of the ray-tracing problems shown in Fig. 1. These problems may share most of the oligonucleotides in the experiments; thus the experimental cost may be reduced significantly if all these experiments are carried out in a parallel fashion.

If the present scheme is executed as it is proposed, the errors may come from two primary sources: the numerical and the experimental errors. The former arises from the discretization of the geometric configuration or the numerical integration of Eq. (1) on an electronic computer; the latter is from the possible binding between two unrelated elementary rays during the annealing. The numerical errors have been very well understood and may be reduced either by increasing the number of the nodal points and of the elementary rays radiated out from each node, or by decreasing the step size of the integration.¹³ The experimental errors are possible to be minimized and identified. For a given length of oligonucleotides, say 20 units (nucleotides), there are 4^{20} possible choices of distinct sequences. One may develop an algorithm to generate 20-mer oligonucleotides with little similarity in their sequences. It is also possible to adjust the temperature range of the cooling and heating cycles in the PCR process in order to reduce the possibility of the mismatching between two unrelated elementary rays. At the bottom line, one may identify the errors of the mismatching by checking the valid-

ity of Snell's law at each node when the experiment results are sent back to an electronic computer.

In the near future, one may not be able to build a molecular computer that outperforms a traditional electronic computer in every aspect, but one can certainly build a molecular computer for certain intrinsically complex problems which otherwise cannot be solved on an electronic computer. Integrated with electronic computers, molecular computers may further extend its applicability to large scale problems and, therefore, greatly strengthen our current computational ability.

ACKNOWLEDGMENTS

The authors thank Dr. Geoffrey Main for his valuable suggestions. This work was supported by Office of Naval Research.

- ¹R. P. Feynman, in *Miniaturization*, edited by D. H. Gilbert (Reinhold, New York, 1961), pp. 282–296.
- ²F. L. Carter, Ed., *Molecular Electronic Devices* (Dekker, New York, 1982).
- ³R. C. Haddon and A. A. Lamola, "The molecular electronic device and the biochip computer: Present status," *Proc. Natl. Acad. Sci. USA* **82**, 1874–1878 (1985).
- ⁴B. H. Robinson and N. C. Seeman, "The design of a biochip: a self-assembling molecular-scale memory device," *Protein Eng.* **1**, 295–300 (1987).
- ⁵U. P. Wild, S. Bernet, B. Kohler, and A. Renn, "From supramolecular photochemistry to the molecular computer," *Pure Appl. Chem.* **64**, 1335–1342 (1992).
- ⁶L. M. Adleman, "Molecular computation of solutions to combinatorial problems," *Science* **266**, 1201–1204 (1994).
- ⁷R. J. Lipton, "DNA solution of hard computational problems," *Science* **268**, 542–545 (1995).
- ⁸M. Born and E. Wolf, *Principles of Optics* (Pergamon, New York, 1959).
- ⁹S. Weinberg, "Eikonal method in magnetohydrodynamics," *Phys. Rev.* **126**, 1899–1909 (1962).
- ¹⁰L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves* (Prentice-Hall, Englewood Cliffs, NJ, 1973).
- ¹¹A. D. Pierce, *Acoustics* (Acoustical Society of America, Woodbury, NY, 1989), pp. 317–419.
- ¹²J. Baldauf, S. Lee, L. Lin, S. Jeng, S. Scarborough, and C. L. Yu, "High frequency scattering from trihedral corner reflectors and other benchmark targets: SBR versus experiment," *IEEE Trans. Antennas Propag.* **39**, 1345–1351 (1991).
- ¹³Y. Yang, A. N. Norris, and L. S. Couchman, "Ray tracing over smooth elastic shells of arbitrary shape," *J. Acoust. Soc. Am.* **99**, 55–64 (1996).
- ¹⁴Y. Yang, A. N. Norris, and L. S. Couchman, "Acoustic scattering from fluid-loaded elastic shells: A Gaussian beam approach," *J. Acoust. Soc. Am.* **98**, 611–622 (1995).
- ¹⁵R. Nowack and K. Aki, "The two-dimensional Gaussian beam synthetic method: testing and application," *J. Geophys. Res.* **89**, 7797–7819 (1984).
- ¹⁶A. N. Norris, "Rays, beams and quasimodes on thin shell structures," *Wave Motion* **21**, 127–147 (1995).
- ¹⁷A. N. Norris and D. A. Rebinsky, "Membrane and flexural waves on thin shells," *ASME J. Vib. Acoust.* **116**, 457–467 (1994).
- ¹⁸P. L. Marston, "Geometrical and catastrophe optics methods in scattering," in *Physical Acoustics*, Vol. 21, edited by A. D. Pierce and R. N. Thurston (Academic, New York, 1992), pp. 1–234.
- ¹⁹J. D. Watson and F. H. C. Crick, "Molecular structure of nucleic acids," *Nature* **171**, 737–738 (1953).
- ²⁰K. B. Mullis and F. A. Faloona, "Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction," *Methods Enzymol.* **155**, 335 (1987).
- ²¹H. Ling, R. Chou, and S. Lee, "Shooting and bouncing rays: calculating the RCS of an arbitrarily shaped cavity," *IEEE Trans. Antennas Propag.* **37**, 194–205 (1989).

New formulation of the resonance scattering theory

Myoung-Seon Choi

Nondestructive Evaluation Team, Korea Atomic Energy Research Institute, P.O. Box 105, Yusong, Taejeon 305-600, Korea

(Received 12 September 1996; accepted for publication 9 January 1997)

According to the resonance scattering theory, a process of background subtraction is always required to obtain the resonance spectrum of a target. A new formulation of the theory is proposed here in which the background subtraction is not necessary and the resonance spectrum can be obtained directly. Contrary to the previous theory, this formulation suggests that, within each partial wave, there are three contributions: the resonance, the background, and their correlation. Also, it is shown that the resonance contribution (at least for empty elastic shells) can be synthesized by considering the resonances of Breit–Wigner form and their pair correlation. © 1997 Acoustical Society of America. [S0001-4966(97)05005-4]

PACS numbers: 43.20.Fn, 43.20.Ks, 43.40.Ey [ANN]

INTRODUCTION

The scattering problems of acoustic, elastic, or electromagnetic waves from an object in a medium have been extensively analyzed during the past several decades. One of the most important problems was to model the scattering process between a target and a monochromatic plane wave. Most targets in the scattering problems were of canonical geometry such as cylindrical or spherical geometry, because the closed form solution could be obtained in the case of simple geometry.¹

The Rayleigh normal mode series (also called “partial waves”) solution, which is the classical solution for the scattering problems, is useful only for low frequencies because of its slow convergence for high frequencies. This limitation can be considerably improved by applying the Watson–Sommerfeld transformation to the series solution.² When target size and wavelength are of considerable dimensions, however, the scattered field is so complicated that the above approximate approaches are not valid any more; thus, more general approaches, such as the resonance scattering theory (RST) and the singularity expansion method,³ are required. This is particularly true when the targets are penetrable and admit interior fields, since the size of the resonance region for penetrable targets is considerably larger than that for impenetrable targets. The RST, established originally for nuclear scattering and applied to acoustic, elastic, and electromagnetic scattering,^{4–8} postulates that, within each partial wave, there are two interacting contributions: a nonresonant background term which is a smooth function of frequency, and a resonance term which is significant at and near the resonance frequencies. The former is mainly due to reflection and diffraction from the target surface and the latter is due to transmission into the target and reradiation.

A long-standing difficulty in the application of RST is that it always requires a process of background subtraction to obtain the resonance spectrum of a target; nevertheless, an exact expression for the background of the field scattered by any penetrable target had been elusive over the past two decades. Fortunately, a general approach to a simple expression for the exact background, named the *inherent* back-

ground coefficient, has been recently proposed, and its effectiveness has been successfully demonstrated for submerged, empty, elastic shells of spherical and cylindrical geometry.^{9–11} Although the inherent background coefficients are utilized, however, the background subtraction is still required according to RST.

In this paper, a new formulation of RST is proposed in which the background subtraction is not necessary and the resonance spectrum of a target can be obtained directly. This new formulation may be applicable to any target of canonical geometry. However, it is impossible to deal with all geometries at once, since the mathematical formalism is different for each geometry. Therefore, first of all, we deal with fluid-loaded, empty, elastic, spherical shells, which have received much attention over the years. However, we use a methodology easily applicable to other geometries. Contrary to the previous RST, the new theory suggests that, within each modal scattering coefficient, there are three interacting contributions: the resonance coefficient, the background coefficient, and their correlation. It is analytically and numerically shown that the residual coefficient, often used in the previous RST and given by subtracting the background coefficient from the scattering coefficient, is equal to the resonance coefficient in their moduli but there is a difference in their phases. In the total scattered field, the residual response may be thus considerably different from the resonance feature of a target. Also, it is shown that the resonance coefficient can be synthesized by considering the resonances of Breit–Wigner form¹² and their pair correlation. Finally, the application to a cylindrical geometry is discussed.

I. THEORY

A. Scattering coefficient and the inherent background coefficient

For an empty spherical elastic shell of outer radius a and inner radius b (thus, thickness $d=a-b$) insonified normally by a plane acoustic wave of unit amplitude, the scattering form function of the n th normal mode is expressed by

$$f_n(x_1, \theta) = \frac{2}{x_1 i} (2n+1) R_n P_n(\cos \theta), \quad (1)$$

where $i \approx \sqrt{-1}$, R_n is a coefficient from the five boundary conditions (three at the outer surface and two at the inner surface of the shell), and P_n is a Legendre polynomial.¹³ The coefficient R_n is expressed as a ratio of the 5×5 determinants: $R_n = B_n/D_n$. The elements of these determinants B_n and D_n , listed in Ref. 13, contain the spherical Bessel and outgoing Hankel functions (j_n , y_n , and $h_n^{(1)}$), and their derivatives (j'_n , y'_n , and $h_n^{(1)'}(x)$) with respect to the argument x_i ($\equiv k_i a$, $i=1, L, T$) or y_i ($\equiv k_i b$, $i=L, T$), where k_1 is the wave number of the incident wave in the outer fluid and $k_{L,T}$ are those of longitudinal and transverse waves in the shell. The dimensionless frequencies of the elastic bulk waves are linearly related to that of the incident wave by $x_{L,T} = x_1(C_1/C_{L,T})$ and $y_{L,T} = x_{L,T}(1-h)$, where C_1 and $C_{L,T}$ are the velocities of the incident wave and the elastic bulk waves and h is the ratio of thickness to outer radius of shell, i.e., $h=d/a$. For the sake of simplicity, we put $x_1 = x$ hereinafter.

By expanding the secular determinants, B_n and D_n , with respect to their first column, the scattering coefficient can be written as follows:

$$R_n = - \frac{x j'_n(x) - F_n j_n(x)}{x h_n^{(1)'}(x) - F_n h_n^{(1)}(x)}, \quad (2)$$

where $F_n = -(\rho_1/\rho_2)x_1^2 D_n^{11}/D_n^{21}$, ρ_1 and ρ_2 are material densities of the outer fluid and the shell, and D_n^{11} and D_n^{21} are the 4×4 minor determinants of the first two elements in the first column of D_n . The function F_n , called the modal surface admittance, depends on the density ratio of the outer fluid to the shell explicitly and the characteristics of the elastic bulk waves in the shell implicitly.

When the material damping of the elastic bulk waves in the shell is not taken into account, the admittance function F_n has real value. Then, it has been recently shown by the author and his colleague that, within each modal admittance, there are two interacting contributions: one is a nonresonant component which is constant (independent of frequency) and corresponds to the inherent background coefficients, the other is a resonant component which is particularly significant at and near the resonances and is small off resonances.¹⁰ These two components interact over almost all frequencies; thus, the constant component seldom manifests itself. However, when the shear wave in the shell is neglected, like in the liquid-shell background by Veksler,¹⁴ the constant component reveals itself furtively near zero frequency. As a result, the constant component, Ω_n , is expressed as

$$\Omega_n = \frac{\rho_1}{\rho_2} \frac{n + (n+1)(1-h)^{2n+1}}{1 - (1-h)^{2n+1}}, \quad (3)$$

and the inherent background coefficient is given as

$$R_n^{(b)} = - \frac{x j'_n(x) - \Omega_n j_n(x)}{x h_n^{(1)'}(x) - \Omega_n h_n^{(1)}(x)}. \quad (4)$$

From Eq. (4), it is easily shown that the inherent background approaches the soft background in the low-frequency regime

of $x \ll \Omega_n$ and the rigid background in the high-frequency regime of $x \gg \Omega_n$.

B. Resonance coefficient and its synthesis

Now we introduce the partial-wave scattering functions, S_n and $S_n^{(b)}$, and scattering phase shifts, δ_n and $\delta_n^{(b)}$, defined by

$$S_n = 1 + 2R_n \equiv \exp(2i\delta_n) \quad (5a)$$

and

$$S_n^{(b)} = 1 + 2R_n^{(b)} \equiv \exp(2i\delta_n^{(b)}). \quad (5b)$$

By substituting Eqs. (2) and (4) into Eq. (5), the scattering functions are expressed as follows:

$$S_n = S_n^{(s)} \frac{z_n^{(2)} - F_n}{z_n^{(1)} - F_n} \quad (6)$$

and

$$S_n^{(b)} = S_n^{(s)} \frac{z_n^{(2)} - \Omega_n}{z_n^{(1)} - \Omega_n}, \quad (7)$$

where $S_n^{(s)} = -h_n^{(2)}(x)/h_n^{(1)}(x)$ and $z_n^{(i)} = x h_n^{(i)'}(x)/h_n^{(i)}(x)$, $i=1,2$. The function $S_n^{(s)}$ is the partial-wave scattering function for an acoustically perfect soft sphere. It is easy to prove that all the scattering functions satisfy the unitary condition $|S_n| = |S_n^{(b)}| = |S_n^{(s)}| = 1$. By combining Eq. (6) with Eq. (7), the scattering function for the elastic shell is given as

$$S_n = S_n^{(b)} \frac{1/\bar{z}_n^{(2)} - 1/\bar{F}_n}{1/\bar{z}_n^{(1)} - 1/\bar{F}_n}, \quad (8)$$

where $\bar{F}_n = F_n - \Omega_n$ and $\bar{z}_n^{(i)} = z_n^{(i)} - \Omega_n$, $i=1,2$. In Eq. (8), it should be noted that the numerator and the denominator of the fractional expression are complex conjugates of each other. This implies that we can introduce another scattering function and phase shift, $S_n^{(\text{reso})}$ and $\delta_n^{(\text{reso})}$, defined by

$$S_n^{(\text{reso})} = 1 + 2R_n^{(\text{reso})} \equiv \exp(2i\delta_n^{(\text{reso})}), \quad (9)$$

where

$$R_n^{(\text{reso})} = \frac{-i \operatorname{Im}[1/\bar{z}_n^{(1)}]}{1/\bar{z}_n^{(1)} - 1/\bar{F}_n}. \quad (10)$$

Then Eq. (8) can be expressed by

$$S_n = S_n^{(b)} S_n^{(\text{reso})}. \quad (11)$$

From Eqs. (5), (9), and (11) one can obtain

$$R_n = R_n^{(b)} + R_n^{(\text{reso})} + 2R_n^{(\text{reso})} R_n^{(b)} \quad (12)$$

and, since $|S_n^{(b)}| = 1$,

$$|R_n - R_n^{(b)}| = |R_n^{(\text{reso})}|. \quad (13)$$

Equation (13) says that, in moduli, the residual coefficient obtained by subtracting the background coefficient from the scattering coefficient is equal to the coefficient $R_n^{(\text{reso})}$ given by Eq. (10). This implies that the coefficient $R_n^{(\text{reso})}$ is the resonance coefficient corresponding to the resonances of an elastic shell. Therefore, we can say that, in

obtaining the resonance feature of a shell, a process of background subtraction is not necessary. The direct use of Eq. (10) is much more effective even in the viewpoint of reducing the computing time. Equation (11) indicates that the scattering function for the elastic shell is represented as the product of the unimodular background and resonance functions. This kind of the product expansion has been suggested,^{15,16} even though the exact expression for the resonance function and the proof were not given. As a result of the product expansion, Eq. (12) postulates that, within each partial wave, there is another contribution as well as the resonance and background coefficients. The third term represents a correlation between the resonance coefficient and the background coefficient. It will cause the phase of the modal residual coefficient to deviate from that of the modal resonance coefficient. Thus, in the total scattered field, the residual response may be considerably different from the resonance feature of a shell, even in their moduli.

The resonance frequencies x_{nl} are found from the condition that the real part of the denominator in Eq. (10) should be equal to zero;

$$\text{Re}[1/\bar{z}_n^{(1)}(x_{nl})] - 1/\bar{F}_n(x_{nl}) = 0. \quad (14)$$

In the vicinity of a resonance frequency, it may be expanded in a Taylor series as

$$\text{Re}[1/\bar{z}_n^{(1)}(x)] - 1/\bar{F}_n(x) \approx (x - x_{nl})\beta_{nl}, \quad (15)$$

where β_{nl} is a constant. The function $\text{Im}[1/\bar{z}_n^{(1)}(x)]$, which appears both in the numerator and in the denominator, can be approximated by itself at $x = x_{nl}$ because it is a slowly varying function.^{17,18} If now the width of resonance Γ_{nl} is introduced according to the definition

$$\Gamma_{nl} = 2 \text{Im}[1/\bar{z}_n^{(1)}(x_{nl})]/\beta_{nl}, \quad (16)$$

then the resonance coefficient $R_n^{(\text{reso})}$ can be written in the Breit-Wigner form¹²

$$R_n^{(\text{reso})} \approx W_{nl} = \frac{-i\Gamma_{nl}/2}{x - x_{nl} + i\Gamma_{nl}/2}. \quad (17)$$

The following numerical examination will show that, over the entire frequencies, the resonance coefficient $R_n^{(\text{reso})}$ can be approximately expressed as

$$R_n^{(\text{reso})} \approx \sum_{l=0}^{\infty} W_{nl} + 2 \sum_{l \neq q} W_{nl} W_{nq}, \quad (18)$$

where the second term on the right side indicates a pair correlation among resonances.

II. NUMERICAL ANALYSIS

The scattering, background, and resonance coefficients of Eqs. (2), (4), and (10) are now examined. The corresponding backscattering ($\theta = \pi$) form functions, $f_n(x, \pi)$, $f_n^{(b)}(x, \pi)$, and $f_n^{(\text{reso})}(x, \pi)$, are also calculated. All calculations are performed for a 99% thick ($h = 0.99$), empty, steel shell in water. The material properties of the pertinent substances are as follows: $\rho_1 = 1.0 \text{ g cm}^{-3}$, $\rho_2 = 7.9 \text{ g cm}^{-3}$, $C_1 = 1480 \text{ m s}^{-1}$, $C_L = 5780 \text{ m s}^{-1}$, and $C_T = 3090 \text{ m s}^{-1}$. The calculation step size of $\Delta x = 0.01$ is used.

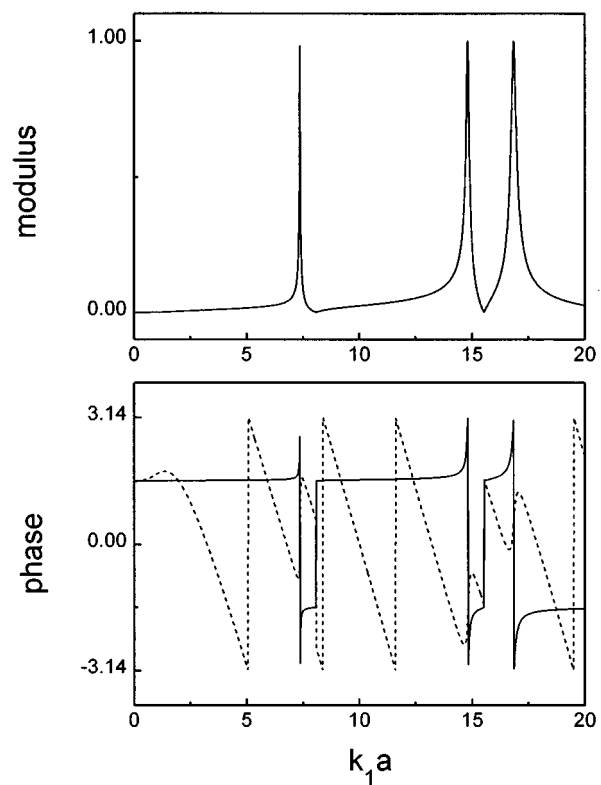


FIG. 1. Comparison of the resonance coefficient (solid line) and the residual coefficient (dashed line) of the $n = 1$ partial wave for a 99% thick, empty, spherical, steel shell in water.

Figure 1 shows a comparison of the resonance coefficient, $R_n^{(\text{reso})}$, and the residual coefficient, $R_n - R_n^{(b)}$, for the $n = 1$ partial wave. There is no distinguishable difference in the moduli, so they look like one curve. Three resonances are shown at $x = 7.37$, 14.82 , and 16.86 . Their widths, which are measured by fitting a Lorentzian function to square of the modulus data of each resonance, are 0.03 , 0.12 , and 0.21 , respectively. However, there is a considerable difference in the phases. The phase of the resonance coefficient is so simple that it can be characterized by only two kinds of the phase change of π : one (from $\pi/2$ to $-\pi/2$) takes place in the vicinity of the resonance frequencies gradually, the other (from $-\pi/2$ to $\pi/2$) at the positions of the vanishing modulus (shown like a dip) between two successive resonances abruptly. The narrower the resonance width, the faster the phase change is. The abrupt phase jump appears near the narrower resonance. In the next paragraph, it will be shown that the abrupt phase jump and the vanishing modulus are attributed to an interference and a pair correlation of resonances. Figure 2 illustrates the total scattered fields, which are obtained by summing up to the $n = 25$ partial wave. The resonance response $|\Sigma f_n^{(\text{reso})}|$ is no longer equal to the residual response $|\Sigma (f_n - f_n^{(b)})|$. As mentioned above, this is due to the fact that the correlation between the resonance coefficients and the background coefficients still remains in the residual response.

Figure 3 illustrates a synthesis of the dipole ($n = 1$) resonance coefficient using Eq. (18). When the pair-correlation term is taken into account, the approximation (solid line) is in excellent agreement with the exact reso-

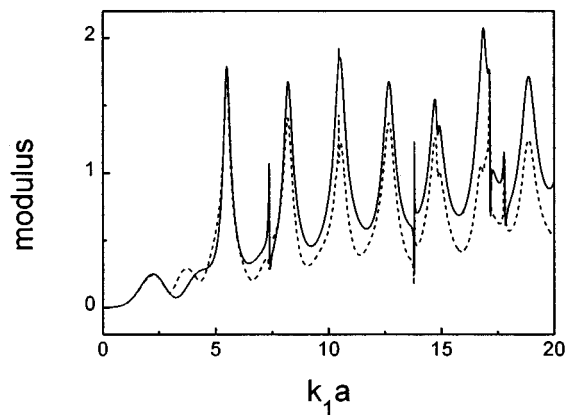


FIG. 2. Moduli of the total backscattering form function of the resonance response (solid line) and of the residual response (dashed line) for a 99% thick, empty, spherical, steel shell in water.

nance coefficient shown in Fig. 1. However, when the correlation term is not taken into account, the result (dashed line) fails to reproduce the abrupt phase jump between the resonances. Although it is not so clear with this example, the correlation term is more significant for close resonances.¹⁵ Then, if the term is omitted from Eq. (18), the result may fail even to describe the unimodal resonance peaks and the vanishing modulus of the exact resonance coefficient.

III. APPLICATION TO A CYLINDRICAL GEOMETRY

For an empty cylindrical shell submerged in a fluid,^{11,19} the modal scattering form function and the modal scattering coefficient are given by

$$f_n(x, \theta) = \frac{2}{\sqrt{\pi i x}} \varepsilon_n R_n \cos(n\theta) \quad (19)$$

and

$$R_n = - \frac{x J'_n(x) - F_n J_n(x)}{x H_n^{(1)'}(x) - F_n H_n^{(1)}(x)}, \quad (20)$$

where ε_n is the Neumann factor ($\varepsilon_0 = 1$, $\varepsilon_n = 2$ for $n \geq 1$), J_n and $H_n^{(1)}$ are the cylindrical Bessel and outgoing Hankel functions, and $F_n = -(\rho_1/\rho_2)x_T^2 D_n^{11}/D_n^{21}$. The elements of the secular determinant D_n were listed in Ref. 19. In the previous work,¹¹ it was shown that the inherent background coefficient for this case is given by

$$R_n^{(b)} = - \frac{x J'_n(x) - \Omega_n J_n(x)}{x H_n^{(1)'}(x) - \Omega_n H_n^{(1)}(x)}, \quad (21)$$

where for $n = 0$

$$\Omega_n = \frac{\rho_1}{\rho_2} \frac{-1}{\ln(1-h)}, \quad (22a)$$

and for $n \geq 1$

$$\Omega_n = \frac{\rho_1}{\rho_2} n \frac{1 + (1-h)^{2n}}{1 - (1-h)^{2n}}. \quad (22b)$$

Therefore, the formulation from Eq. (5) to Eq. (18) is identically applicable to the cylindrical geometry, with the exception of replacing the spherical functions by the corresponding cylindrical functions.

IV. CONCLUSIONS

In this study, it has been shown that the resonance spectrum of a target can be obtained directly without using a process of background subtraction. Within the modal scattering coefficient, there are three interacting contributions: the resonance coefficient, the background coefficient, and their correlation. The resonance coefficient is equal to the residual coefficient in their moduli, but there is a difference in their phases. In the total scattered field, therefore, the residual response may be considerably different from the resonance feature of a target. The resonance coefficient (at least for empty elastic shells) can be synthesized by considering the resonances of Breit-Wigner form and their pair-correlation.

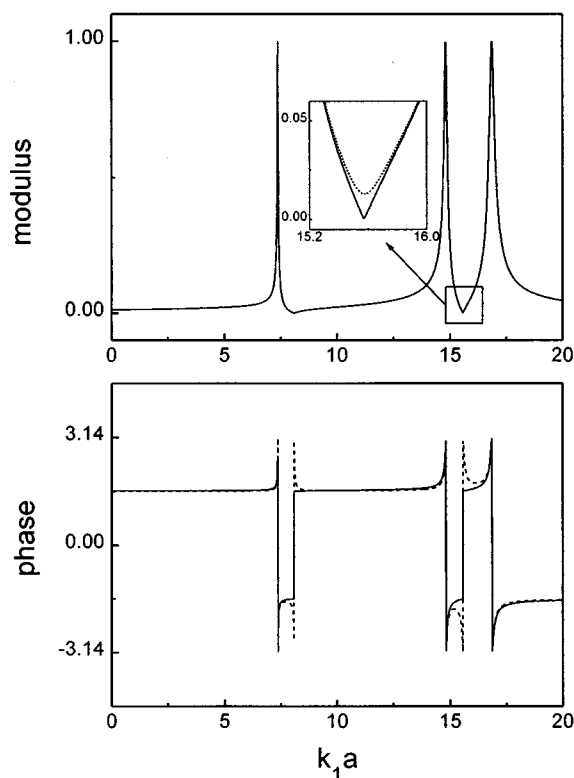


FIG. 3. Synthesis of the dipole resonance coefficient (solid line in Fig. 1) by Eq. (18) with (solid line) and without (dashed line) the pair correlation term.

¹J. Bowman, T. Senior, and P. Uslenghi, *Electromagnetic and Acoustic Scattering by Simple Shapes* (Wiley-Interscience, North-Holland, Amsterdam, 1969).

²G. N. Watson, "The diffraction of elastic waves by the earth," *Proc. R. Soc. London, Ser. A* **95**, 83-99 (1918); A. Sommerfeld, *Partial Differential Equations in Physics* (Academic, New York, 1949).

³H. Überall, P. P. Delsanto, J. D. Alemar, E. Rosario, and A. Nagl, "Application of the singularity expansion method to elastic wave scattering," *Appl. Mech. Rev.* **43**, 235-249 (1990).

⁴H. Feshbach, D. C. Peaslee, and V. F. Weisskopf, "On the scattering and absorption of particles by atomic nuclei," *Phys. Rev.* **71**, 145-158 (1947).

⁵L. Flax, L. R. Dragonette, and H. Überall, "Theory of elastic resonance excitation by sound scattering," *J. Acoust. Soc. Am.* **63**, 723-731 (1978).

⁶D. Brill and G. Gaunard, "Resonance theory of elastic wave ultrasonically scattered from an elastic sphere," *J. Acoust. Soc. Am.* **81**, 1-21 (1987).

⁷G. C. Gaunard and M. F. Werby, "Acoustic resonance scattering by submerged elastic shells," *Appl. Mech. Rev.* **43**, 171-208 (1990); and

- also, G. C. Gaunard, "Elastic and acoustic resonance wave scattering," *ibid.* **42**, 143–193 (1989).
- ⁸H. Überall, P. J. Moser, J. P. Murphy, A. Nagl, G. Igiri, J. V. Subrahmanyam, G. Gaunard, D. Brill, P. P. Delsanto, J. D. Alemar, and E. Rosario, "Electromagnetic and acoustic resonance scattering theory," *Wave Motion* **5**, 307–329 (1983).
- ⁹M.-S. Choi, Y.-S. Joo, and J.-P. Lee, "The inherent background in acoustic wave scattering by a submerged target," *J. Acoust. Soc. Am.* **99**, 2594 (1996).
- ¹⁰M.-S. Choi and Y.-S. Joo, "Theory of the background amplitudes in acoustic resonance scattering," *J. Acoust. Soc. Am.* **101**, 2083–2087 (1997).
- ¹¹M.-S. Choi, Y.-S. Joo, and J.-P. Lee, "Inherent background coefficients for submerged cylindrical shells," *J. Acoust. Soc. Am.* **101**, 1743–1745 (1997).
- ¹²G. Breit and E. P. Wigner, "Capture of slow neutrons," *Phys. Rev.* **49**, 519–531 (1936).
- ¹³G. C. Gaunard and M. F. Werby, "Lamb and creeping waves around submerged spherical shells resonantly excited by sound scattering," *J. Acoust. Soc. Am.* **82**, 2021–2033 (1987); "Sound scattering by resonantly excited, fluid-loaded, elastic spherical shell," *J. Acoust. Soc. Am.* **90**, 2536–2550 (1991).
- ¹⁴N. D. Veksler, "Intermediate background in problems of sound waves scattering by elastic shells," *Acustica* **76**, 1–9 (1992).
- ¹⁵P. L. Marston, S. G. Kargl, and N. H. Sun, "Elastic resonance amplitudes described by generalized GTD and by product expansion of the *S*-matrix," in *Acoustic Resonance Scattering*, edited by H. Überall (Gordon and Breach, Philadelphia, 1992), pp. 305–333.
- ¹⁶N. D. Veksler and J.-L. Izbicki, "Modal resonances of the Franz waves," *Acustica* **82**, 18–26 (1996); "Modal resonances of peripheral waves," *ibid.* 401–410 (1996); N. D. Veksler, J.-M. Conoir, J.-L. Izbicki, and P. Rembert, "Methods of isolation of modal resonances (a review)," *J. Acoust. Soc. Am.* **99**, 2599 (1996).
- ¹⁷J. D. Murphy, E. D. Breitenbach, and H. Überall, "Resonance scattering of acoustic waves from cylindrical shells," *J. Acoust. Soc. Am.* **64**, 677–683 (1978).
- ¹⁸J. D. Murphy, J. George, A. Nagl, and H. Überall, "Isolation of the resonant component in acoustic scattering from fluid-loaded elastic spherical shells," *J. Acoust. Soc. Am.* **65**, 368–373 (1979).
- ¹⁹P. L. Marston and N. H. Sun, "Backscattering near the coincidence frequency of a thin cylindrical shell: surface wave properties from elastic theory and an approximate ray synthesis," *J. Acoust. Soc. Am.* **97**, 777–783 (1995).

Circumferential creeping waves around a fluid-filled cylindrical cavity in an elastic medium

Waled Hassan and Peter B. Nagy

Department of Aerospace Engineering and Engineering Mechanics, University of Cincinnati, Cincinnati, Ohio 45221

(Received 22 July 1996; accepted for publication 7 January 1997)

The dispersion behavior of circumferential creeping waves around a fluid-filled cylindrical cavity in an infinite elastic medium is analyzed by computational methods. Phase and group velocity as well as attenuation curves are constructed by numerically solving the dispersion equation. A comparison of the corresponding modes for elastic and rigid hosts is presented. The modes in both cases exhibit essentially the same series of cutoff frequencies corresponding to radial resonances at which the phase velocity of the associated modes becomes infinite and the group velocity assumes a limiting value of $\pi c_f/2$, where c_f is the compressional wave velocity in the fluid. Attenuation of the circumferential creeping modes in a cylindrical cavity is caused solely by losing energy to the surrounding elastic bulk. Therefore, for all modes, the attenuation diminishes at high frequencies as leakage into the surrounding solid becomes negligible. This is in sharp contrast with the case of leaky Rayleigh wave propagation along the plane surface of a solid-fluid interface where attenuation is caused solely by radiation of energy into the fluid, which causes the frequency to have an opposite effect on the degree of leakage in these situations. © 1997 Acoustical Society of America. [S0001-4966(97)03905-2]

PACS numbers: 43.20.Ks, 43.20.Mv [ANN]

INTRODUCTION

Recently, the development of a novel ultrasonic inspection technique that detects radial fatigue cracks on the far side of so-called “weep” holes in thin airframe stiffeners was reported.¹ These cracks tend to be located on the upper part of the weep hole (at the 12 o'clock position) and therefore are not readily detectable by conventional ultrasonic inspection techniques from the lower skin of the wing. The new technique utilizes circumferential creeping waves propagating around the inner surface of the hole to perform the inspection. However, the wet wing has to be emptied and dried out before inspection because even a small amount of fluid fuel trapped in these rather small (approximately 6–7 mm in diameter) holes would strongly affect the propagation of circumferential creeping waves.² We have searched the literature for published results on circumferential creeping wave propagation around fluid-filled cylindrical cavities in elastic media. Surprisingly, although the analytical solution of this canonical problem can be readily constructed from existing building blocks, very little was found in terms of numerical results that could be used to gain better understanding of the phenomenon. This motivated us to attack the problem by numerically solving the dispersion equation and constructing the corresponding dispersion and attenuation curves for a specific case of interest, namely, for that of a water-filled cylindrical hole in aluminum.

The problem of acoustic wave propagation along the interface of a solid half-space in contact with a fluid half-space is well understood (see, for example, the paper of Bertoni and Tamir³ and the references in it). If the velocity of the Rayleigh wave propagating on the free surface of the solid half-space is higher than the sound velocity in the fluid the fluid-loaded Rayleigh mode becomes leaky into the fluid,

i.e., it radiates some of its energy into the fluid as it propagates. For curved surfaces, Keller and Karal^{4,5} introduced the geometrical theory of surface waves, which provides a high-frequency asymptotic approximation for surface waves under fairly general conditions. Grimshaw⁶ and Gregory⁷ also investigated the propagation of Rayleigh waves on free curved surfaces at high frequencies. The problem of dispersive Rayleigh wave propagation on both convex and concave cylindrical surfaces was solved by Rulf.⁸ As for fluid loading, extensive literature is available on the exterior problem, i.e., on elastic waves propagating around solid cylinders and spheres immersed in fluid.^{9–11} In these cases, the coupling between the fluid and the solid is relatively weak because of the usually large difference between their acoustic impedance. Of course much stronger acoustic coupling occurs when the solid rod or sphere is substituted by a thin shell. The problem of elastic waves running around air-filled elastic shells submerged in fluid has been extensively investigated by Gaunaurd, Werby, Überall and others.^{10–17} Kaduchak and Marston evaluated the surface displacement of a cylindrical shell insonified by plane waves perpendicular to the shell's axis of symmetry.¹⁸ They mathematically decomposed these displacements into circumferential traveling waves propagating in opposite directions and presented phase and group velocity curves for the resulting Lamb-type circumferential waves [Fig. 5(a) in their paper]. Optical visualization of the diffraction of a longitudinal pulse by a cylindrical cavity in an optically transparent solid was discussed and presented by Ying.¹⁹

From the point of view of energy partition, there are two types of circumferential waves that can propagate around a fluid-filled cylindrical hole in an elastic host, namely whispering gallery modes that carry most of their energy in the

fluid and a Rayleigh-type mode that propagates essentially in the solid. Whispering gallery or halo modes in air were first discussed by Lord Rayleigh.²⁰ According to the classical physical explanation based on ray theory, such modes are produced by the repeated reflection of the compressional bulk wave in the fluid at the cylindrical wall. With respect to air the solid host can be assumed to be an ideally rigid perfect reflector so that these modes are not attenuated at all and carry their energy entirely within the fluid. For an elastic host, the circumferential modes are slightly attenuated as some of their energy leaks into the solid. The Rayleigh-type circumferential mode is the interface wave propagating around the fluid-filled surface of the cavity, which carries most of its energy on the solid side. The Rayleigh mode becomes dispersive on any curved surface. In addition, on concave surfaces, it also becomes attenuated as some of its energy is leaked into the solid host.^{8,17} With fluid loading on the inside of the cylindrical hole, the dispersive Rayleigh mode becomes leaky into the fluid as well, which changes its propagation properties very significantly. To the best of our knowledge, the resulting mode coupling between the Rayleigh-type surface mode and the whispering gallery type halo modes has not been investigated before. Longitudinal wave scattering from fluid-filled cylindrical holes was studied experimentally in great detail by Sachse *et al.*^{21–23} However, in spite of the fairly strong acoustical coupling between the solid host and the fluid within the hole, their results did not show the presence of a Rayleigh-type surface mode propagating around the hole as the longitudinal mode in the solid is but very weakly coupled to this mode. We have found that vertically polarized (normal to the axis of the hole) shear waves are much more strongly coupled to the Rayleigh-type modes and generate relatively weaker whispering gallery modes in the fluid.²

It should be mentioned that Solomon *et al.* studied resonance scattering of elastic waves from fluid-filled cylindrical cavities.²⁴ They first mathematically formulated the scattered wave field for both empty and fluid-filled cylindrical cavities, then subtracted the real and imaginary parts of the scattered field amplitudes for the empty cavity from the corresponding quantities for the fluid-filled case. Finally, the results were converted to dispersion curves displayed in Fig. 11 in their paper. Of course these curves do not display the effect of the elastic host on the circumferential elastic waves, since the subtraction process essentially eliminates this effect and produces the results for the case of a fluid-filled cylindrical cavity in an infinitely stiff host in spite of starting out with an elastic host (aluminum).

The complete analytical solution for elastic waves propagating around free convex and concave surfaces was first obtained by Rulf who presented the exact analytical form of the dispersion equation describing the propagation of the Rayleigh-type circumferential wave around a cylindrical cavity in an infinite elastic medium.⁸ In this paper, the dispersion equation of circumferential waves around a concave fluid-loaded cylindrical surface will be first presented based on Rulf's approach. This general equation will then be specialized for the case of a fluid-filled cylindrical cavity in an infinitely stiff, i.e., rigid host which will provide a check on

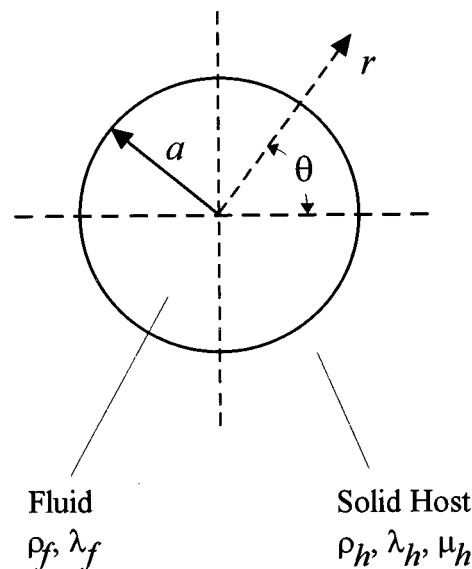


FIG. 1. Geometry of a fluid-filled cylindrical cavity in a solid host and the coordinate system.

our results as we compare them to published data.^{20,24} Finally, dispersion and attenuation curves for a fluid-filled cylindrical cavity in an elastic host will be calculated and discussed with particular attention paid to the cutoff frequencies and the high-frequency asymptotic behavior of the different modes.

I. THEORETICAL BACKGROUND

Figure 1 shows the schematic diagram of a fluid-filled cylindrical cavity in a solid host and the coordinate system. The physical properties of the fluid and the elastic host along with the field variables in them will be distinguished by subscripts *f* and *h* respectively. The boundary conditions require the continuity of the normal stress and displacement components and the vanishing of the tangential stress component, which gives⁸

$$\begin{bmatrix} -\lambda_f p_f^2 & -2\mu_h L_1 + \lambda_h p_h^2 & -2\mu_h L_2 \\ 0 & -2\mu_h L_2 & \mu_h(2L_1 + s_h^2) \\ L_3 & -L_3 & -L_4 \end{bmatrix} \begin{bmatrix} \phi_f \\ \phi_h \\ \psi_h \end{bmatrix}_{r=a} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad (1)$$

where $L_1 = \partial^2/\partial r^2$, $L_2 = (1/r)(\partial^2/\partial r \partial \theta) - (1/r^2)(\partial/\partial \theta)$, $L_3 = \partial/\partial r$, and $L_4 = (1/r)(\partial/\partial \theta)$. μ and λ denote the Lamé constants of the material and $s = \omega/c_s$ and $p = \omega/c_d$ are the shear and longitudinal wave numbers, respectively. Here, ω denotes the angular frequency, $c_s = \sqrt{\mu/\rho}$ is the shear wave velocity, $c_d = \sqrt{(\lambda + 2\mu)/\rho}$ is the longitudinal wave velocity, and ρ is the density of the material. The scalar and vector potentials satisfy the $(\nabla^2 + p^2)\phi = 0$ and $(\nabla^2 + s^2)\psi = 0$ uncoupled wave equations. The scalar potential in the fluid ϕ_f must be regular at $r=0$, therefore it can be expressed as $\phi_f(r, \theta) = A_1 J_\nu(r p_f) e^{i\nu\theta}$, where J_ν denotes the first kind

Bessel function of order ν . Furthermore, ϕ_h and ψ_h in the host must satisfy the radiation condition as $r \rightarrow \infty$, therefore $\phi_h(r, \theta) = A_2 H_\nu^{(1)}(r p_h) e^{i\nu\theta}$ and $\psi_h(r, \theta) = A_3 H_\nu^{(1)}(r s_h) e^{i\nu\theta}$, where $H_\nu^{(1)}$ denotes the first kind Hankel function of order ν . Substituting these formal solutions of the cylindrical wave equation into Eq. (1), the boundary conditions can be expressed in the matrix form of $QA = 0$, where the elements of the secular matrix Q are listed in Ref. 8. For a fluid-filled cylindrical cavity in an elastic host, the characteristic equation is given as $|Q| = 0$. For the case of a fluid-filled cylindrical cavity in a rigid host the characteristic equation reduces to $Q_{31} = 0$ which was originally derived by Lord Rayleigh.²⁰

II. NUMERICAL RESULTS AND DISCUSSION

The dispersion equations for the cases of a cylindrical cavity filled with fluid in elastic and infinitely stiff hosts were introduced in the previous section. These two equations were solved numerically using Mathematica Version 2.2.3. Our calculations have been made for the case of a water-filled cylindrical cavity in aluminum. The density of aluminum was taken as $\rho_h = 2.7 \times 10^3 \text{ kg/m}^3$ and its Lamé constants as $\lambda_h = 5.52 \times 10^{10} \text{ N/m}^2$ and $\mu_h = 2.6 \times 10^{10} \text{ N/m}^2$. These parameters give the speeds of the longitudinal and shear waves as $c_d = 6301 \text{ m/s}$ and $c_s = 3103 \text{ m/s}$, respectively. The density and Lamé constant of water were assumed to be $\rho_f = 10^3 \text{ kg/m}^3$ and $\lambda_f = 2.25 \times 10^9 \text{ N/m}^2$, respectively, which give the compressional wave speed as $c_f = 1500 \text{ m/s}$. Of course no shear waves can propagate in the inviscid fluid ($\mu_f = 0$).

The dispersion equations were solved for the angular wave number ν , which is a complex quantity in general, as a function of the normalized frequency $\Omega = s_h a = \omega a / c_s$. First, the dispersive phase velocity was calculated as $c_p(\omega) = a \omega / \text{Re}\{\nu\}$, then the group velocity is obtained by numerical differentiation from

$$c_g = a \frac{\partial \omega}{\partial \text{Re}\{\nu\}} = \frac{c_p}{1 - (\omega / c_p)(\partial c_p / \partial \omega)}. \quad (2)$$

The attenuation coefficient of the circumferential mode is $\alpha = \text{Im}\{\nu\} / a$. The normalized attenuation was calculated as the total attenuation over the semi-circumference of the cylindrical hole, i.e., $\alpha_n = 8.686 \pi \text{Im}\{\nu\}$ in dB.

A. Phase velocity

Figure 2 shows the dispersion curves of the first 12 circumferential modes propagating around a cylindrical cavity in an elastic host (solid line). The phase velocity is normalized with respect to the compressional wave velocity in water (1500 m/s) and is plotted as a function of the normalized frequency. In order to point out the effects of the acoustic loading of the elastic host on the free vibrations of the fluid column, we also plotted in dashed lines the corresponding modes of a fluid-filled cylindrical cavity in a rigid medium. Generally, the modes for the elastic host very closely follow the corresponding modes for the rigid host. The acoustic loading of the elastic host causes strong perturbation of the mode structure only in a narrow band where the phase ve-

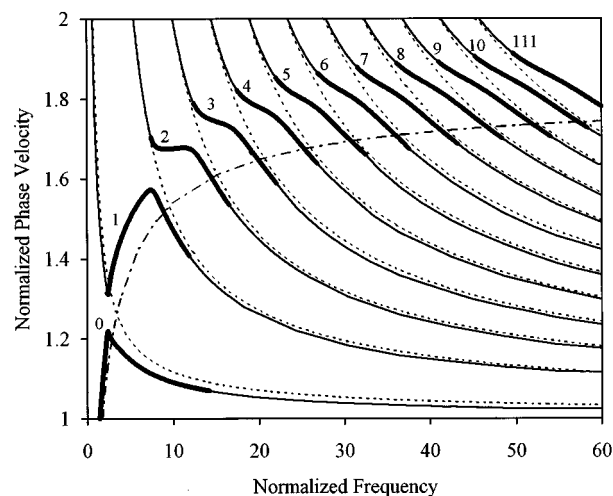


FIG. 2. Dispersion curves for circumferential waves propagating around the inner surface of a water-filled cylindrical cavity in aluminum (—), in an ideal rigid host (·····) and on the free surface of a cylindrical cavity in aluminum (— · — · —).

locity of the whispering gallery modes coincides with the phase velocity of the Rayleigh-type surface wave running around the inside surface of the cylindrical hole. The dash-dot-dash line represents this dispersive Rayleigh mode for the case of a free cylindrical hole in an elastic medium. An obvious feature of the curves displayed in Figure 2 is the repulsion of the modes of the fluid-filled cylindrical cavity in an elastic medium (the highlighted parts of the solid curves). These highlighted portions of the modes seem to form a discontinuous leaky Rayleigh-type surface mode, that follows the well-known dispersive leaky Rayleigh mode that would propagate around the free surface of the cylindrical hole in the same elastic medium. Mode repulsion is a common feature of guided wave propagation in acoustically coupled systems. In our case it means that in the region where the modes of the unloaded fluid column (in a rigid host) would cross the unloaded Rayleigh mode (on the free surface of the cylindrical hole in the elastic medium) the modes of the loaded case (fluid-filled hole in an elastic host) jump to the next higher order mode of the unloaded fluid column, thereby forming the branch that follows the Rayleigh mode instead of directly crossing it. Resonances within the fluid column and the attenuation mechanism associated with the circumferential waves propagating around the cavity will be discussed in more detail later in this section after the presentation of the attenuation and group velocity curves.

The high-frequency behavior of the four lowest order modes is shown in Fig. 3. As the frequency increases, all modes but the lowest order (zeroth) are approaching the same asymptotic value of unity for both elastic and rigid hosts. In other words, in the high-frequency limit these modes asymptotically approach the compressional wave in the fluid. In contrast, the zeroth order mode for the elastic host asymptotically approaches the Scholte–Stoneley mode, which has a slightly lower velocity than the compressional mode in the fluid depending on the degree of the acoustic coupling between the fluid and the solid. This phenomenon

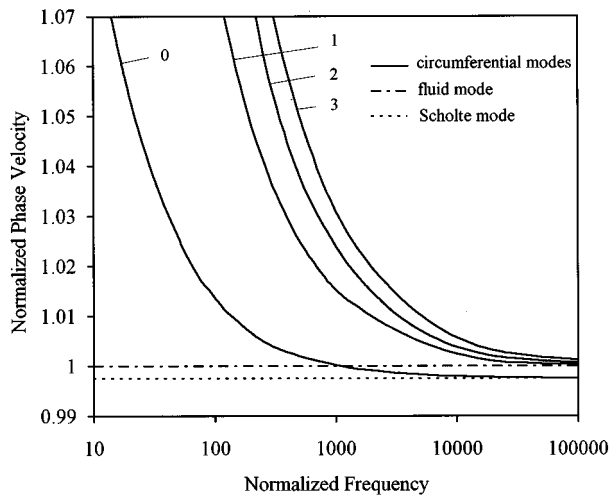


FIG. 3. High-frequency behavior of the four lowest order modes (—). The velocities of the compressional mode in the fluid (— · — · —) and the Scholte–Stoneley mode on a flat fluid–solid interface (·····) are also shown to indicate the different asymptotic behavior of the lowest order mode from that of all higher order modes.

is similar to the high-frequency asymptotic behavior of Lamb waves propagating in plates, where the lowest order symmetric and asymmetric modes approach the Rayleigh surface wave velocity while all the higher order modes approach the somewhat higher bulk shear wave velocity. It should be mentioned that the difference between the compressional fluid velocity and the Scholte–Stoneley interface wave velocity for an aluminum–water interface is less than 0.3%, more than one order of magnitude smaller than the approximately 6% difference between the shear bulk wave velocity and the Rayleigh surface wave velocity for aluminum.

The eigenvalues of the secular determinant for $\nu=0$ define a series of cutoff frequencies at which the phase velocity of the corresponding modes becomes infinite. This observation was first reported by Rayleigh²⁰ in connection with the free vibrations of a cylindrical fluid column in a rigid host. Solomon *et al.*²⁴ also reported similar results using resonance scattering techniques. The phase velocity of the first mode for a fluid-filled hole in a rigid host becomes infinite at very low frequencies (zero represents the cutoff frequency for this mode). As the frequency increases, the phase velocity starts to decrease until a normalized frequency of about 2.4 is reached, where it is still in a very good agreement with the corresponding first mode for the case of the rigid host. Increasing the frequency any further results in a local increase of the phase velocity until it reaches a maximum close to the next higher order mode associated with the rigid host. Above this local peak any additional increase in the frequency will cause the phase velocity to decrease again as the mode follows the next higher order mode of the rigid host. The agreement between the two curves becomes better and better as the frequency increases until they both reach the same high-frequency asymptotic limit. The lowest order mode which follows the Rayleigh mode at very low frequencies reaches a local maximum at about 2.4, where it starts to decrease with increasing frequency. This mode essentially

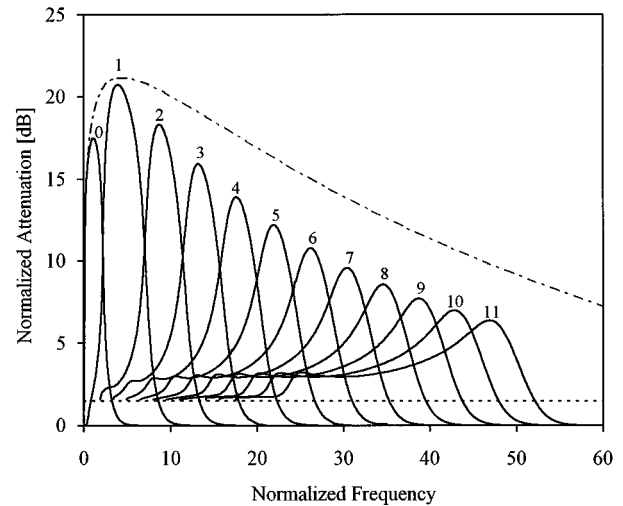


FIG. 4. Normalized attenuation curves for circumferential waves propagating around the inner surface of a water-filled cylindrical cavity in aluminum (—) and on the free surface of a cylindrical cavity in aluminum (— · — · —). The limiting value of the normalized attenuation at the cutoff frequencies is also shown (·····).

follows the first mode associated with the rigid host but stays below it and eventually approaches the Scholte–Stoneley velocity at very high frequencies.

B. Attenuation

The normalized attenuation coefficient α_n is shown in Fig. 4 as a function of the normalized frequency. For all modes we notice that the attenuation is very small at low frequencies. As the frequency increases, the attenuation increases until it reaches a maximum, after which any further increase in the frequency results in a reduction in the normalized attenuation and the attenuation coefficient eventually asymptotically goes to zero. All the higher order modes start at their respective cutoff frequencies with a normalized attenuation of approximately 1.6 dB. In the next section, we shall demonstrate that at the cutoff frequencies the fluid column undergoes diagonal vibrations. These resonances are obviously damped by transmission into the solid host at each interaction. It is interesting to point out that the 1.6 dB normalized attenuation at the cutoff frequencies is in good agreement with the 1.5 dB reflection loss predicted using the simple formula for calculating the reflection loss for the case of normal incidence of plane waves on a plane water–aluminum interface.

Except for the lowest order mode, which follows the Rayleigh mode at low frequencies, the peaks of the attenuation curves decrease as the mode number increases. The dash-dot-dash line in Fig. 4 represents the attenuation curve for the case of a free cylindrical cavity in an elastic medium. This curve constitutes an upper limit for the attenuation coefficient in the case of the fluid-filled cylindrical cavity. The peaks for the fluid-filled case approximately follow the pattern of the smooth attenuation curve for the free case. Of course the modes of the fluid-filled cylindrical cavity in a rigid host are not attenuated.

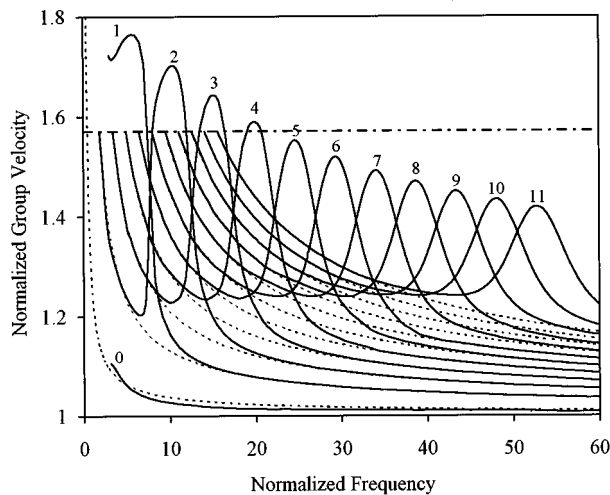


FIG. 5. Normalized group velocity curves for circumferential waves propagating around the inner surface of a water-filled cylindrical cavity in aluminum (—) and in an ideal rigid host (·····). The $\pi/2$ limiting value of the normalized group velocity at the cutoff frequencies is also shown (— · — · — ·).

C. Group velocity

The group velocity represents the speed at which energy propagates from one point on the circumference of the cavity to another. Figure 5 shows the normalized group velocity curves for a water-filled cylindrical cavity in an ideally rigid medium and in aluminum. Equation (2) was used to generate this figure from the phase velocity dispersion curves previously shown in Fig. 2. Several features associated with Fig. 5 are worth noting. First, for the case of an infinitely stiff medium, we notice that the group velocity curves start at certain cutoff frequencies that are given as the roots of the corresponding dispersion equation in the limiting case of $\nu \rightarrow 0$. Except for the first mode, the group velocity of each mode has the same magnitude at their respective cutoff frequencies. This constant limit of the group velocity is $\pi c_f/2$, which is also shown in Fig. 5 by the horizontal dash-dot-dash line. In the case of a rigid host the characteristic equation^{20,24} simplifies to $J'_\nu(ap_f)$ and it can be proven mathematically that $\pi c_f/2$ is the limiting value for the group velocity. In general, the group velocity can be calculated from $c_g = d\omega/dk$. For the n th mode, the cutoff frequency is $\omega_n = j'_{\nu,n} c_f/a$, where $j'_{\nu,n}$ is the n th positive zero of J'_ν . Also the wave number k is given as $k = \nu/a$, therefore the group velocity can be expressed as follows

$$c_g = c_f \frac{dj'_{\nu,n}}{d\nu}. \quad (3)$$

For small values of ν and large orders of n it can be shown²⁵ that $\partial j'_{\nu,n}/\partial \nu \rightarrow \pi/2$, i.e., the normalized group velocity indeed approaches a limiting value of $\pi/2$ at the cutoff frequencies. The physical explanation is even simpler. Infinite circumferential phase velocity corresponds to a pure radial resonance, i.e., the energy actually oscillates diagonally over a distance of $2a$ with the compressional wave velocity of c_f . However, the group velocity is calculated from the circumferential wave number ν by using Eqs. (2) and (3), i.e.,

by measuring the propagation distance along the perimeter of the cylindrical hole, which is $\pi/2$ times higher than the propagation distance within the fluid. Accordingly the group velocity along the longer semi-circumferential distance is $\pi/2$ higher. The radial resonance interpretation of the infinite circumferential phase velocity provides an easy and direct approach to evaluate the limiting value of the group velocity. Generally, infinite phase velocity in a given direction corresponds to energy propagation normal to that direction. Therefore, in the more familiar case of a plane interface, the group velocity is zero along the boundary at the cutoff frequencies.

The solid lines in Fig. 5 represent the group velocity curves for the case of a water-filled cylindrical cavity in aluminum. It is clear that all modes except for the lowest order (zeroth) start from the same cutoff frequency as the corresponding mode in the case of the rigid host. As the frequency increases the group velocity for the fluid-filled cylindrical cavity in an elastic medium starts to deviate from that of the corresponding mode in an ideally rigid host. The former starts to increase until it reaches a local maximum, after which it starts to decrease again approaching the next higher order mode associated with the rigid host case. For high frequencies that correspond to short wavelengths, we see that the normalized group velocities of the modes for both cases being investigated approach the same limiting value of unity, i.e., the energy at high frequencies propagates around the hole at a group velocity equal to the compressional wave velocity in the fluid.

Special attention has to be paid to the behavior of the first order mode in the case of the cylindrical cavity in a rigid host. We notice that, unlike all other modes, the first order mode does not attain a finite group velocity as its phase velocity approaches an infinite value, i.e., at $\omega = 0$. Instead, the group velocity continues to increase as the frequency decreases. This anomalous behavior can be justified based on the fact that at zero frequency there is no energy propagation and the same state exists at all points around the inner surface of the cylindrical cavity at all times. This manifests itself as an infinite group velocity, which explains the continuing rapid growth of the curve as the frequency decreases.

At very low frequencies, the group velocity curves for the zeroth and first order modes in the case of an elastic host display a different behavior than that of the higher order modes around their corresponding cutoff frequencies. To avoid any confusion, these two modes were not shown below $\Omega = 3$ in Fig. 5. Their somewhat complicated mode structure at low frequencies is displayed separately in Fig. 6. Initially, the first order mode exhibits a behavior similar to that of the corresponding rigid host mode described earlier, but switches over to the Rayleigh-type surface mode above $\Omega \approx 2.4$. The zeroth order mode exhibits a complementary behavior. Initially, it follows the Rayleigh-type surface mode, but switches over to the first order fluid mode in a rigid host above the same normalized frequency of $\Omega \approx 2.4$.

D. Mode coupling

The apparently rather complicated mode structure observed in Figs. 2–6 is the result of highly localized mode

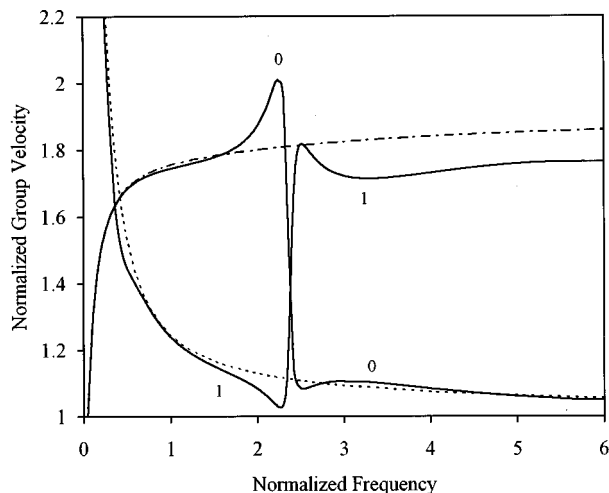


FIG. 6. Low-frequency normalized group velocity curves for the two lowest order circumferential waves propagating around the inner surface of a water-filled cylindrical cavity in aluminum (—) and in an ideal rigid host (·····). The group velocity of the Rayleigh wave propagating around the inner surface of a free cylindrical cavity in aluminum is also shown for comparison (— · — · — · —).

coupling between the two principal types of waves that can propagate around the fluid-filled hole. Figure 7 shows the schematic diagram of leaky Rayleigh wave propagation along the circumference of a fluid-filled cylindrical cavity illustrating both reradiated bulk and fluid waves. The reradiated bulk wave can be considered as a weak perturbation caused by the curvature of the surface. The resulting attenuation decreases with frequency as previously shown by the dash-dot-dash line in Fig. 4. The reradiated fluid wave however is not lost as it will mode-convert back to the Rayleigh mode after crossing the hole and hitting the fluid-solid interface at exactly the Rayleigh angle. Figure 8 illustrates the schematic diagram of shear wave scattering from a fluid-

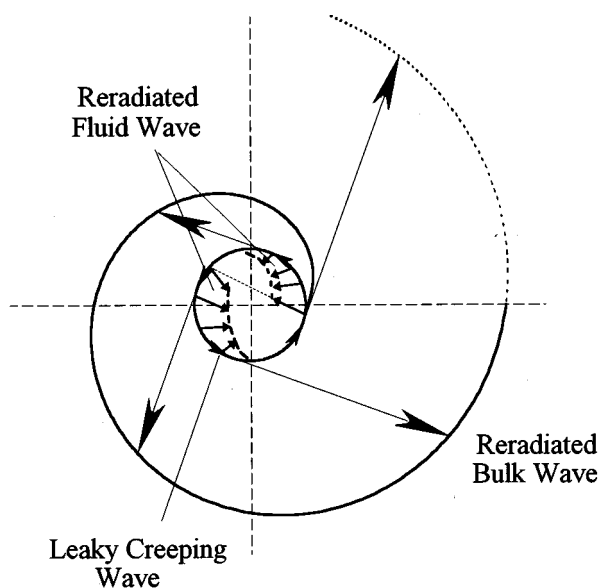


FIG. 7. Schematic diagram of leaky Rayleigh wave propagation around a fluid-filled cylindrical cavity.

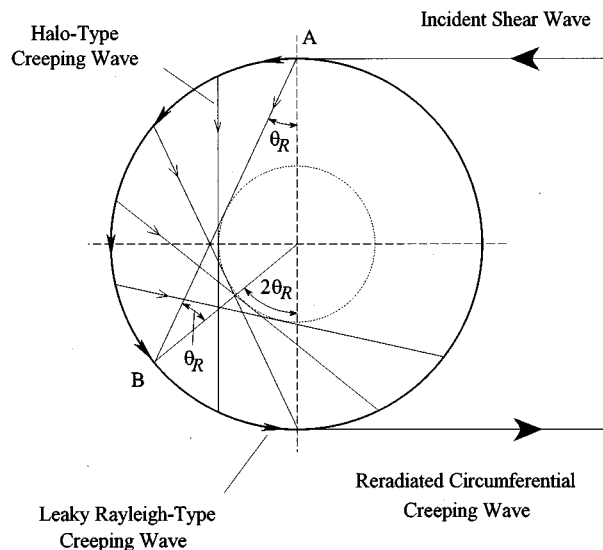


FIG. 8. Schematic diagram of leaky Rayleigh surface wave propagating around the surface of a water-filled cylindrical cavity in an elastic solid.

filled cylindrical hole in an elastic solid, which is of particular interest in nondestructive testing applications.^{1,2} By virtue of reciprocity, only those circumferential modes that are strongly attenuated by leaking into the elastic solid can be excited from within the solid. As it was demonstrated by the highlighted segments in Fig. 2, these branches form a more or less continuous Rayleigh-type surface mode which radiates into the fluid at approximately the so-called Rayleigh angle. In comparison, creeping modes of very high phase velocity would radiate towards the center of the cylindrical cavity. Only one of the peripheral incident rays is shown, which excites this Rayleigh-type surface wave at point A. As this mode starts to propagate around the surface of the cavity, it leaks energy into the fluid at an angle equal to the Rayleigh angle θ_R (for the case of aluminum-water $\theta_R \approx 30^\circ$). Let us consider the ray in the fluid which starts very close to point A and hits the opposite surface of the cylindrical cavity at point B after crossing the fluid-filled hole along the AB chord. This ray makes an angle of θ_R with the normal to the tangent at point A and hits the surface of the cavity at an angle of incidence equal to θ_R , therefore it will excite a Rayleigh surface wave that will propagate along the inner concave surface of the hole from point B. This mode is again a leaky Rayleigh wave that will reradiate its energy both into the bulk of the solid and into the enclosed fluid. The above process repeats itself for every ray reradiated into the fluid at subsequent points along the circumference of the cavity. The concentric circle of radius $a \sin \theta_R = a c_f / c_R$ shown by the dotted line in Fig. 8 represents the approximate inner limit of the whispering gallery, i.e., the maximum depth of penetration for strong halo modes towards the center of the cylindrical cavity in the case of shear wave excitation from the outside. A natural similarity exists between this process and the phenomenon of whispering galleries discussed by Rayleigh.²⁰ One important distinction between the two phenomena is that in a classical whispering gallery the source is within the fluid column therefore

the rays can hit the wall at any angle and not necessarily at the Rayleigh angle. When no surface waves are excited, essentially perfect specular reflection occurs. In the simple case of a rigid wall considered by Rayleigh, no acoustic coupling into the host can occur. Even in the case of an elastic host, the coupling is negligible except when the halo waves hit the surface at the Rayleigh angle when all of the incident energy mode-converts into the leaky Rayleigh mode. On a plane fluid–solid interface total internal reflection would occur, i.e., all the incident energy would be ultimately reradiated into the fluid without any loss to the bulk of the solid. In the case of a concave surface, some of the energy is lost to the bulk of the solid before it is reradiated into the fluid, therefore the circumferential modes become strongly attenuated by leakage into the bulk of the solid whenever phase matching occurs between the velocity of the halo modes and the Rayleigh-type surface mode.

The highlighted portions of the dispersion curves presented in Fig. 2 represent regions where strong leaky Rayleigh waves are generated by the halo modes. This explains the obvious agreement between these portions of high attenuation and the leaky Rayleigh mode that would propagate around the free surface of the cylindrical hole. It takes a propagation time of $a\pi/c_R$ for the Rayleigh mode to travel the semi-circumferential distance half-way around the cylindrical cavity. Of course, this wave is much more attenuated than the somewhat lagging second arrival which skipped once as a halo wave and propagated only the remaining $2\theta_{Ra}$ distance as a Rayleigh mode. The delay between these two principal waves can be calculated as follows

$$\Delta t \approx \frac{2a \cos \theta_R}{c_f} - \frac{a(\pi 2\theta_R)}{c_R}. \quad (4)$$

Similar calculations predicting the total time of travel of the backscattered rays from a longitudinal wave incident on a circular cylindrical cavity in an elastic solid were presented by Pao and Sachse.²³ Interference between these two components gives rise to periodically occurring strong attenuation at frequencies separated by $\Delta\omega = 2\pi/\Delta t$, which can be expressed in a normalized form as

$$\Delta\Omega = \frac{2\pi a}{\Delta t c_{sh}} = \frac{\pi}{\cot \theta_R - \frac{1}{2}\pi + \theta_R}, \quad (5)$$

where we approximated the Rayleigh velocity simply with the shear velocity. For a water-filled cylindrical cavity in aluminum, from Eq. (5) we get $\Delta\Omega \approx 4.2$, which is in very good agreement with the periodicity of the attenuation peaks shown in Fig. 4 at high frequencies. The above approximations are very simple and consequently rather rough, but the resulting simplicity of the results more than compensates for the loss of rigor. Equations (4) and (5) are based on ray theory and relate to the very high frequencies where the attenuation peaks appear fairly periodically.

Part of the energy of the leaky Rayleigh mode gets reradiated back into the fluid, thereby contributing again to the halo modes in the fluid. The rest of the energy (assuming ideal conditions and neglecting any losses associated with

viscosity, friction and thermal conductivity) is lost to the surrounding infinite elastic medium, thereby causing these modes to be attenuated. This effect manifests itself in the attenuation curves as an increase in the normalized attenuation coefficient until it reaches a maximum value, after which it starts to decrease again. The falling part of the curves can be justified on the basis that as the frequency increases the wavelength becomes shorter and shorter compared to the radius of the cavity, thus approaching the case of a Rayleigh mode propagating along a plane surface that has no curvature. This allows more and more energy to be reradiated into the fluid within the cavity and less and less energy to the surrounding elastic solid. Consequently, the normalized attenuation coefficient decreases until the frequency becomes so high that plane surface conditions prevail and the modes become practically unattenuated by leakage into the solid host. It is important to note that, although the two cases of a non-dispersive Rayleigh mode propagating along a plane solid–fluid interface and a dispersive Rayleigh wave propagating around the surface of a fluid-filled cylindrical cavity in an elastic solid have some features in common, the attenuation mechanisms associated with them are very different. While the Rayleigh mode in the first case is attenuated by losing energy to the fluid, the attenuation in the second case is solely caused by reradiating energy to the surrounding infinite solid. This energy will be lost and never recovered, while the energy reradiated into the fluid-filled hole is continuously recovered. This behavior causes the frequency to have an opposite effect on the degree of leakage in the two cases.

Beside the decreasing leakage caused by the surface curvature at increasing frequencies, the reduction in the peaks of the normalized attenuation curves shown in Fig. 4 is also due to the decreasing actual propagation length along the interface. The effective propagation length before reradiation into the fluid occurs, which is related to the so-called Schoch displacement, becomes smaller and smaller as the frequency increases thereby reducing the possibility of losing energy to the elastic solid. This translates into a direct reduction in the maximum normalized attenuation coefficient of each successive halo resonance within the cavity.

III. CONCLUSIONS

In this paper, the problem of circumferential creeping waves propagating along the surface of a fluid-filled cylindrical cavity was numerically solved to obtain dispersion and attenuation curves for both rigid and elastic hosts. The characteristic equations for both cases were obtained by the previously published method of Rulf.⁸ A better understanding of the physical nature of these modes was achieved by analyzing the obtained numerical results. It was found that the circumferential creeping modes around a fluid-filled cylindrical cavity exhibit a rather strange behavior. First, they follow the whispering gallery modes of the unloaded fluid column in a rigid host at low frequencies, then they are attracted by the Rayleigh mode propagating on the free surface of the cylindrical hole in the elastic host, and finally they tend towards

the whispering gallery modes of the unloaded fluid column at high frequencies, but with a jump to the next higher order mode.

It has been confirmed that all but the lowest order two modes will have cutoff frequencies at which the phase velocity approaches infinity, the group velocity reaches a limiting value of $\pi c_f/2$ and the normalized attenuation is approximately equal to the reflection loss of plane waves at normal incidence. These cutoff frequencies correspond to pure radial compressional resonances in the fluid column and are similar to the well-known pure longitudinal and shear through-thickness resonances in the Lamb mode structure of elastic plates except for the non-zero group velocity brought about by the cylindrical configuration. Acoustic ray theory along with the concept of non-specular reflection at the Rayleigh angle were utilized to explain the dispersion and attenuation curves. The energy leaked into the fluid at the Rayleigh angle from the Rayleigh mode propagating around the surface of the cavity will hit the opposite surface in a series of reflections exciting leaky Rayleigh waves upon each encounter, which again reradiate most of their energy back into the fluid. The process is repeated until all the energy is lost to the bulk of the surrounding solid, which is the only attenuation mechanism in the case of the fluid-filled cylindrical cavity in an elastic medium. This is markedly different from what takes place in the case of leaky Rayleigh wave propagation along the plane surface of a solid half-space in contact with a fluid half-space, where the attenuation is caused solely by energy leakage into the fluid.

The presented conventional solution of the dispersion equation describes the behavior of the circumferential waves in the frequency domain. This form of the solution has great practical importance in narrow-band experiments when waves lasting for a very long time are utilized (e.g., tone bursts). Unfortunately, this form of the solution is not as useful in broadband experiments where pulses lasting for a very short time are used, such as those experiments done by Nagy *et al.* to detect fatigue cracks in weep holes.¹ To be more useful in the interpretation of this kind of experiments the solution of the problem will have to be obtained in the time domain using the Green's function method. In spite of their limited use in broadband experiments, the results obtained are of great importance as the conventional solution of a basic canonical problem and serve the purpose of deeper understanding some aspects of circumferential creeping wave propagation around fluid-filled cylindrical cavities.

ACKNOWLEDGMENT

This work was partially supported by the NDE Branch of the Metals and Ceramics Division, Wright Laboratory, Wright-Patterson AFB.

- ¹P. B. Nagy, M. Blodgett, and M. Golis, "Weep hole inspection by circumferential creeping waves," *NDT&E Int.* **27**, 131 (1994).
- ²W. Hassan and P. B. Nagy, "Feasibility of fatigue crack detection in fluid-filled cylindrical holes using circumferential creeping waves," in *Review of Progress in Quantitative Nondestructive Evaluation* (Plenum, New York, 1997), Vol. 16, pp. 43–50.
- ³H. L. Bertoni and T. Tamir, "Unified theory of Rayleigh-angle phenomena for acoustic beams at liquid-solid interfaces," *Appl. Phys.* **2**, 157 (1973).
- ⁴J. B. Keller and F. C. Karal, "Surface wave excitation and propagation," *J. Appl. Phys.* **31**, 1039 (1960).
- ⁵J. B. Keller and F. C. Karal, "Geometrical theory of elastic surface wave excitation and propagation," *J. Acoust. Soc. Am.* **36**, 32 (1964).
- ⁶R. Grimshaw, "Propagation of surface waves at high frequencies," *J. Inst. Math. Appl.* **4**, 174 (1968).
- ⁷R. D. Gregory, "The propagation of Rayleigh waves over curved surfaces at high frequencies," *Proc. Camb. Philos. Soc.* **70**, 103 (1971).
- ⁸B. Rulf, "Rayleigh waves on curved surfaces," *J. Acoust. Soc. Am.* **45**, 493 (1969).
- ⁹R. D. Doolittle, H. Überall, and P. Uginčius, "Sound scattering by elastic cylinders," *J. Acoust. Soc. Am.* **43**, 1 (1968).
- ¹⁰G. C. Gaunaurd, "Elastic and acoustic resonance wave scattering," *Am. Soc. Mech. Eng.* **42**, 143 (1989).
- ¹¹L. R. Dragonette, "Evaluation of the relative importance of circumferential or creeping waves in acoustic scattering from rigid and elastic solid cylinders and cylindrical shell," *NRL Report 8216*, Naval Research Laboratory, Washington, DC (1978).
- ¹²G. C. Gaunaurd and M. F. Werby, "Acoustic resonance scattering by submerged elastic shells," *Am. Soc. Mech. Eng.* **43**, 171 (1990).
- ¹³G. C. Gaunaurd and H. Überall, "Relationship between creeping-wave acoustic transients and the complex-frequency poles of the singularity expansion method," *J. Acoust. Soc. Am.* **78**, 234 (1985).
- ¹⁴C. W. Horton and M. V. Mechler, "Circumferential waves in a thin-walled air filled cylinder in a water medium," *J. Acoust. Soc. Am.* **51**, 295 (1972).
- ¹⁵P. Smith, "Phase velocities and displacement characteristics of free waves in a thin cylindrical shell," *J. Acoust. Soc. Am.* **27**, 1065 (1955).
- ¹⁶P. Uginčius, "Creeping-wave analysis of acoustic scattering by elastic cylindrical shells," *Naval Weapons Lab, Tech. Report TR-2128* (1968).
- ¹⁷I. Viktorov, *Rayleigh and Lamb Waves* (Plenum, New York, 1967).
- ¹⁸G. Kaduchak and P. L. Marston, "Traveling-wave decomposition of surface displacements associated with scattering by a cylindrical shell: Numerical evaluation displaying guided forward and backward wave properties," *J. Acoust. Soc. Am.* **98**, 3501 (1995).
- ¹⁹C. F. Ying, "Photoelastic visualization and theoretical analyses of scatterings of ultrasonic pulses in solids," in *Physical Acoustics* (Academic, San Diego, 1990), Vol. XIX, pp. 291–343.
- ²⁰J. W. S. Rayleigh, *The Theory of Sound* (Dover, New York, 1945), 2nd ed.
- ²¹W. Sachse, "Ultrasonic spectroscopy of a fluid-filled cavity in an elastic solid," *J. Acoust. Soc. Am.* **56**, 891 (1974).
- ²²W. Sachse and C. T. Chain, "Determination of the size and mechanical properties of a cylindrical fluid inclusion in an elastic solid," *Mater. Eval.* **33**, 81 (1975).
- ²³Y. Pao and W. Sachse, "Interpretation of the time records and power spectra of scattered ultrasonic pulses in solids," *J. Acoust. Soc. Am.* **56**, 1478 (1974).
- ²⁴S. G. Solomon, H. Überall, and K. B. Yoo, "Mode conversion and resonance scattering of elastic waves from a cylindrical fluid-filled cavity," *Acustica* **55**, 147 (1984).
- ²⁵M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).

A study of wave propagation in varying cross-section waveguides by modal decomposition. Part II. Results

N. Amir, V. Pagneux, and J. Kergomard

Laboratoire d'Acoustique de l' Université du Maine, URA 1101 CNRS, BP 535, Avenue Olivier Messiaen, 72017 Le Mans Cedex, France

(Received 25 March 1996; revised 13 December 1996; accepted 19 December 1996)

The full derivation of the equations governing the generalized impedance matrix Z , the pressure, and the velocity were presented in Part I of this series [Pagneux *et al.*, J. Acoust. Soc. Am. **100**, 2034–2048 (1996)]. Here only the results of that paper, i.e., the final set of equations which needed to be solved are repeated. Other factors influencing the solution are the boundary conditions at the end of the waveguide: Source and radiation conditions are also presented. Finally, the details of the numerical implementation are also relevant, and will be discussed in some detail. © 1997 Acoustical Society of America. [S0001-4966(97)02605-2]

PACS numbers: 43.20.Mv, 43.20.Ks, 43.75.Fg [JEG]

LIST OF SYMBOLS

z	axial coordinate
r	radial cylindrical coordinate
p	acoustic pressure
\mathbf{P}	column vector of coefficients of p for modal decomposition
\mathbf{v}	particle velocity
v_z	axial component of particle velocity
\mathbf{U}	column vector of coefficients of axial velocity for modal decomposition
c	speed of sound in free space
ω	angular frequency
k	wave number in free space
k_i	wave number of the i th mode in the waveguide
K	diagonal matrix of modal wave numbers; $K_{ij} = k_i^2 \delta_{ij}$

Z	generalized impedance matrix
Z_c	characteristic matrix impedance of a cylindrical pipe
Z_r	value of Z at the free end of the waveguide
Y	generalized admittance matrix
Y_r	value of Y at the free end of the waveguide
ψ	column vector of eigenfunctions in the waveguide
S	cross-sectional area of waveguide
R	local cross-sectional radius of waveguide
γ_i	i th zero of Bessel function J_1
A	diagonal matrix of γ_i/R
f	frequency
$\epsilon, \epsilon_r, \epsilon_v$	values used to calculate viscothermal losses
$D_i(\tau)$	function used to calculate radiation impedance
J_m	Bessel function of order m .

INTRODUCTION

This paper is the sequel to Part I,¹ where the theoretical background was presented. Here in Part II we will study in depth two typical waveguides. The first is a trombone bell, in which the flare increases very rapidly toward the end. We will calculate the input impedance and compare it to the measured input impedance, in order to show how the multimodal analysis improves on the results of plane-wave analysis. We will also examine the calculated pressure and velocity field along the horn. The second case is a waveguide with circular cross section, and an elliptic longitudinal section, in which we will study focalization effects and reflection properties. In both cases we show the effects of varying the parameters at our disposal on the solution, e.g., the number of modes taken into account, error thresholds, etc. As a matter of fact, the main goal of this part is the analysis of the field inside the guide with varying cross section, taking advantage of the specificity of the method. Incidentally, we compare the results with finite element results and experimental results, but this aspect is of secondary importance.

We will begin by reviewing in Sec. II the equations to

be solved numerically, followed by a description of the boundary conditions used for various types of radiation, and the numerical procedures used. Sections II and III then deal with the trombone horn and elliptic waveguide, respectively.

I. THEORETICAL BASIS

A. Equations

We wish to solve the matricial Riccati equation for the generalized impedance matrix Z :²

$$Z' = -j \frac{k\rho c}{S} + j \frac{S}{k\rho c} ZKZ + \frac{S'}{S} (QZ + ZQ). \quad (1)$$

Since Z is a matrix of infinite order, we must truncate it to order $n \times n$, where n must somehow be determined.

Once Z is known along the waveguide, we can find the pressure or axial velocity along the waveguide, integrating one of the two following equations:³

$$\mathbf{P}' = \frac{S'}{S} Q\mathbf{P} - j \frac{k\rho c}{S} \mathbf{U},$$

$$\mathbf{U}' = -j \frac{S}{k\rho c} \mathbf{K} \mathbf{P} - \frac{S'}{S} {}^t \mathbf{Q} \mathbf{U}. \quad (2)$$

Once P or U are found, the pressure or axial velocity are given by

$$p = \sum \psi_i P_i, \quad v_z = \frac{1}{S} \sum \psi_i U_i. \quad (3)$$

In certain cases it may be inconvenient to use the Riccati equation (1) for Z . If we define a generalized admittance matrix Y so that $U = YP$, we can obtain from (2) a Riccati equation for Y in the same way as (1) was obtained in Part I.¹ This equation will be

$$Y' = \frac{-jS}{k\rho c} \mathbf{K} - \frac{S'}{S} ({}^t \mathbf{Q} Y + Y \mathbf{Q}) + \frac{jk\rho c}{S} Y Y. \quad (4)$$

A further refinement to these equations is the introduction of viscothermal losses created at the boundaries. Keeping in mind that the expression K is simply a diagonal matrix of propagation constants for the various modes, we can simply replace it in Eqs. (1) and (4) by an expression for the lossy propagation factors. This can be found in Ref. 4. For the axisymmetric modes it is as follows:

$$k_i^2 \approx k^2 - \left(\frac{\gamma_i}{R} \right)^2 + 2k[\text{Im}(\epsilon_i) - j \text{Re}(\epsilon_i)]. \quad (5)$$

At atmospheric pressure and room temperature, ϵ_i can be approximated as

$$\epsilon_i = \left(1 - \frac{\gamma_i^2}{R^2 k^2} \right) \epsilon_v + \epsilon_t, \quad (6)$$

where $\text{Re}(\epsilon_v) = \text{Im}(\epsilon_v) = 2.03 \times 10^{-5} f^{1/2}$, and $\text{Re}(\epsilon_t) = \text{Im}(\epsilon_t) = 0.95 \times 10^{-5} f^{1/2}$ (f in s^{-1}).

B. Boundary conditions

The general problem we aim to solve in this paper is as follows: Given a rigid walled waveguide, we have a source at one end, and a radiating termination at the other. The source can be a pressure or velocity source, thus fixing either P or U at one end. The radiation condition must fix the value of the matrix Z at the other end, which we will call Z_r .

One simple case is a termination consisting of an infinite cylindrical tube. In this case, Z_r is simply a diagonal matrix of characteristic impedances for each mode and $Z_r = Z_c$. Another simple case is the "ideal reflection" of an open end; i.e., each mode has a pressure node at the open end, and thus Z_r is uniformly zero. Such a boundary condition for the impedance matrix is a very rough approximation of an open end. Even in the plane-wave approximation, a length correction has to be taken into account. On the other hand, if we wish to model a closed end, the axial velocity is zero, thus elements of Z_r need to be infinite. In this case it is easier to solve the equation for Y , Eq. (4), with a uniformly zero boundary condition for Y_r . The value of Z at the input can be obtained by inverting Y .

The most complicated case is the most realistic one, of radiation into free space. This is quite difficult to solve, therefore we chose to adopt instead the radiation into a half-

space; the open end of the waveguide is thus placed in an infinite baffle. This case has been treated before in Refs. 5 and 6. Initially, the analytical expression for Z_r is

$$Z_r = \frac{j\omega\rho}{S^2} \int_S \int_S \psi(\mathbf{w}) G(\mathbf{w}, \mathbf{w}') {}^t \psi(\mathbf{w}') ds dS', \quad (7)$$

where $G(\mathbf{w}, \mathbf{w}')$ is the Green's function for radiation into the half-space, and \mathbf{w} is a vector on S . This integral is singular and thus difficult to evaluate directly. Therefore we chose to use the formulation presented by Zorumski,⁵ which seems to be posed most efficiently for numerical evaluation. The expression for Z_r presented by the latter are

$$Z_{ij} = -j \int_0^\infty \tau (\tau^2 - 1)^{-1/2} D_i(\tau) D_j(\tau) d\tau. \quad (8)$$

Here, Z_{ij} is a nondimensional form of Z_r and D_i is defined by

$$D_i(\tau) = k^2 \int_0^r r J_0(\tau k r) \psi_i(r) dr. \quad (9)$$

Zorumski provides an exact analytical solution for (9), and another form of (8) which can be numerically evaluated more readily:

$$Z_{ij} = \int_0^{1/2\pi} \sin \phi D_i(\sin \phi) D_j(\sin \phi) d\phi - j \int_0^\infty \cosh \xi D_i(\cosh \xi) D_j(\cosh \xi) d\xi. \quad (10)$$

These results were compared to another calculation based on Ref. 6, giving very good agreement.

Since this expression for the radiation impedance is the closest available to us, we compared the theoretical results to experiments both with and without an "infinite" baffle, in order to have an idea of how far this radiation condition is from radiation with no baffle. An extension of the calculation of the case with infinite baffle to the case of no baffle could be made by iterative procedure (see Williams⁷).

C. Numerical procedures

The basic "engine" used to solve the Riccati equations for Z or Y was a third- or fourth-order Runge-Kutta procedure. This proved to be relatively robust in most cases. Divergence did occur when using simplified radiating conditions where no energy was allowed to radiate from the waveguides (the closed or "ideal open" terminating conditions), since the impedance or admittance peaks became infinite. In these cases it was imperative to incorporate boundary layer thermal and viscous losses into the calculation, as described above.

It is also imperative to use an adaptive stepsize, even with radiation or viscous losses, since the peaks in Z or Y can be pronouncedly sharp, and there is no way to predict the necessary minimum stepsize. This type of solution requires selection of an error threshold for determining when the stepsize must be reduced; this threshold must be selected by trial and error. If it is not small enough, the computation will generally diverge. It may noticed that the effect of increasing

the number of modes selected reduces the Runge–Kutta stepsize, because with more modes we have more peaks for the impedance.

Once Z or Y have been integrated from the radiating end to the source, it remains to calculate the pressure or velocity field, integrating one of (2) in the other direction, i.e., from the source to the radiating end. This requires knowledge of Z or Y at each point of integration. This is a touchy point: As described in Part I, an attempt to reintegrate either of these values back up the waveguide always fails. This means that Z or Y must be stored at each point during the numerical solution of the Riccati equations, to be used later when integrating for \mathbf{P} or \mathbf{U} . This creates two problems: It can take up much memory, and it fixes the integration points for their computation. The first problem has no easy remedy—the RAM memory at the program’s disposal must be sufficiently large, or disk space should be used (slowing down the calculation considerably). The second problem means that an adaptive stepsize cannot be used for the second integration.

In practice, Z was stored during the first integration on a grid of regularly spaced points along the axis, by “forcing” the adaptive stepsize algorithm to “visit” these points in addition to any intermediate points required by the algorithm. The second calculation was then performed with a fixed stepsize. As a coherence check, the energy flux was calculated along the waveguide, using lossless propagation. When it came out nearly constant (relative error of less than 0.1%) the calculation was judged to be sufficiently accurate. This procedure necessitates a very fine grid for storing Z , although apparently this is the only method possible.

We will demonstrate in the later sections how various choices of parameters influence the results of the computation.

II. FIRST WAVEGUIDE: THE TROMBONE BELL

The first case to be studied in detail is the trombone bell, which comprises approximately the last 55 cm of the instrument. It is characterized by a rapid flare at the end, which should render one-dimensional analysis methods relatively inaccurate. In this section we will study the theoretically computed input impedance and compare it to results obtained experimentally. We will also examine the pressure field inside the horn.

The objective of this study is not only to verify the theoretical results by comparison to experiments, but also to study the influence of various parameters on the theoretical results. We wish to see the influence of various radiation conditions, the effect of the number of modes used on the accuracy, and the interaction between these parameters. Another point to be examined is how the assumption of an infinite baffle can change the experimentally measured input impedance.

We will begin with a description of the geometry of the horn, and its relation to the frequency range considered. This will be followed by a short description of the experimental apparatus, and then by comparisons of theoretically and experimentally obtained input impedances. The final part of the section will present results concerning the pressure field.

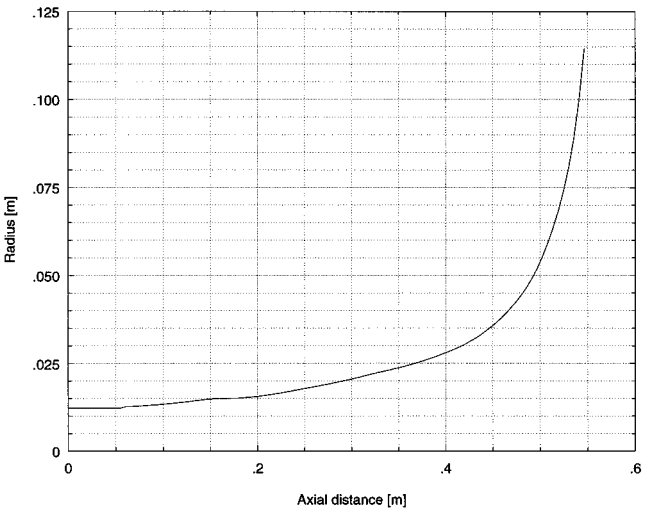


FIG. 1. Radius of the trombone horn versus axial distance.

A. Geometry and frequency

The part of trombone studied begins with a cylindrical section of 25.4-mm diameter. It then begins to flare modestly, terminating in an abrupt flare to $R=115.4$ mm The radius as a function of distance appears in Fig. 1. Table I summarizes the points that were measured. Intermediate points were obtained by third order polynomial interpolation; a polynomial was calculated for four points at a time, and used for calculating values between the two middle points only.

The frequency range to be studied was between 0 and 1200 Hz, this being approximately the musically useful range of the trombone. It is interesting to note that for this range of frequencies, all of the higher order modes are in cutoff, even at the mouth of the horn. At the throat of the horn this is 7.2% of the first cutoff frequency, and at the mouth it is 65% of the first cutoff frequency. The excitation is assumed to be a plane piston, i.e., only the plane-wave mode is excited at the throat, and any influence of higher order modes is due to mode coupling induced by the flare. These higher-order modes are assumed to be so deeply in cutoff at the throat that they can be ignored there, therefore enabling us to treat the input impedance as a simple scalar value, which is the corner value of the matrix Z . The validity of this assumption will be examined later.

B. Experimental apparatus and results

The experimental apparatus available enabled us to measure the input impedance of the horn, but not the pressure field. This apparatus consisted of a velocity source and a

TABLE I. Radius of the trombone horn versus axial distance.

distance [cm]	0	5.5	5.6	5.7	5.8	5.9	8.3	12.4
radius [cm]	12.25	12.25	12.25	12.25	12.65	12.65	13	14
distance [cm]	16.2	20.3	25.3	31.2	36	40	43.2	45.7
radius [cm]	15	16.5	18.5	21.5	24.5	28	32	37
distance [cm]	48.2	50	51.5	52.7	53.6	54.4		
radius [cm]	44.5	52	62	74.5	89.5	114.5		

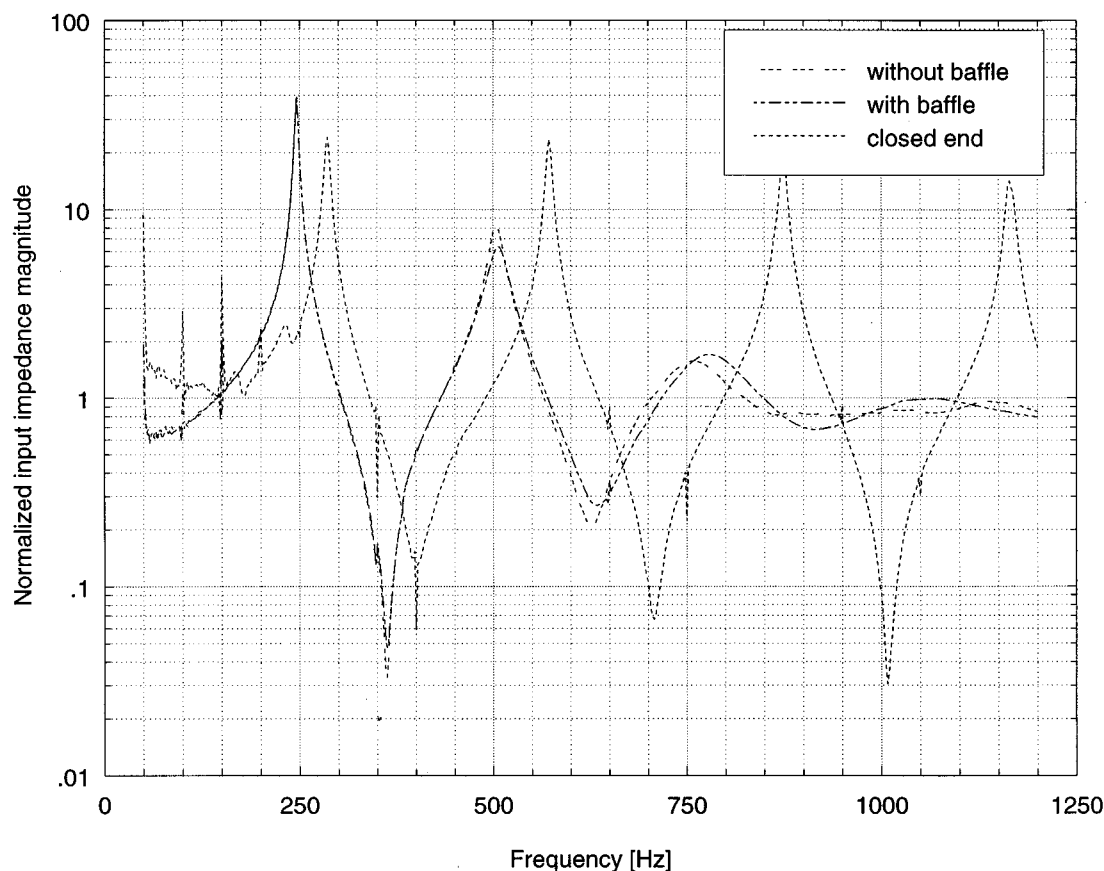


FIG. 2. Measured input impedance with baffle, without baffle, and with rigid sealed end. Note the difference between the first two, at high frequencies.

pressure sensitive microphone at the throat of the horn, both connected to a lock-in amplifier. The latter is connected to a PC, which controls the frequency sweep. This sweep is performed adaptively, so that smaller increments are taken near the impedance peaks, ensuring good accuracy in these regions. The entire apparatus was placed in an anechoic chamber, and performed at a controlled temperature. Calibration was carried out using a cylindrical pipe of approximately 1 m length. For further details of this apparatus, refer to Ref. 8.

The measurements were carried out for three cases: one where the horn radiated freely into the room, one where the mouth was placed in a large baffle (1.5×1.5 m), and one where the end of the horn was shut hermetically with a flat Plexiglass plate.

Figure 2 shows the results of these three measurements. Note that there is a subtle but noticeable difference between measurements made with and without the baffle, at higher frequencies. This is important, since the theoretical calculations use the baffled value of the radiation impedance. This is also used often in the literature for the one-dimensional models, although without always specifying whether the experiments used a baffle also. Fortunately, at low frequencies the two measurements give practically identical results; at such frequencies a good approximation of an infinite baffle would have to be very large. At frequencies where the two measurements differ, the baffle used was in fact a good approximation of an infinite one. The reasons for this behavior are explained in Ref. 9, where it is shown that musical horns are

quite insensitive to the radiation impedance at low frequencies, and that even plane-wave calculations can give good results at these frequencies.

Figure 2 also shows a big difference between measurements with an open and closed horn. The closed horn has no radiation losses, and being quite short the viscothermal losses are quite small also, therefore the impedance peaks are similar in height, and at different frequencies than those found for the open horn.

C. Theoretical results for input impedance

Before comparing the theoretical and experimental results, we examine the purely theoretical results from several points of view.

As stated above, the parameter we will be dealing with will simply be the corner value Z_{00} of the matrix Z . Before continuing, the assumption that this parameter is the main one of importance must be justified. In Fig. 3 we show the diagonal of Z as a function of frequency, and the first row of Z as a function of frequency. These represent, respectively, the “self” impedances of each mode, and the “cross” impedance between the higher-order modes and the plane-wave mode. The very low values of all the other elements in relation to the corner value shows that their importance is minimal. They are also generally quite smooth, as opposed to the peaks and dips found in Z_{00} , thus giving no information concerning resonances or antiresonances.

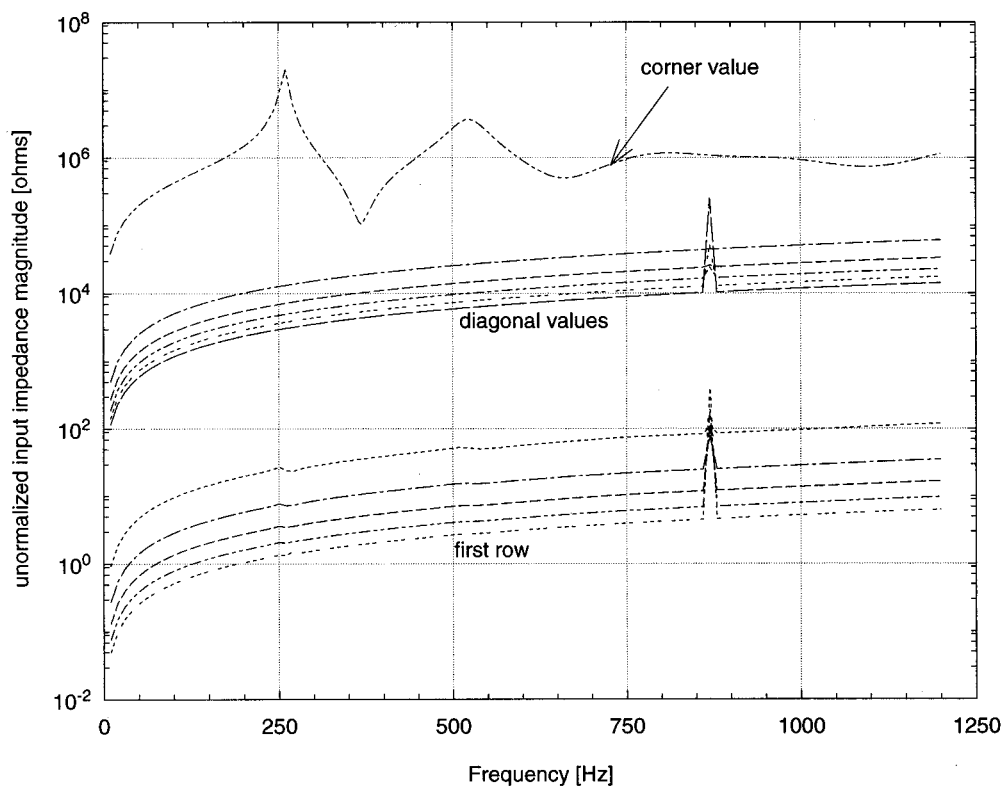


FIG. 3. The diagonal and first row of the input impedance matrix Z (arbitrary units).

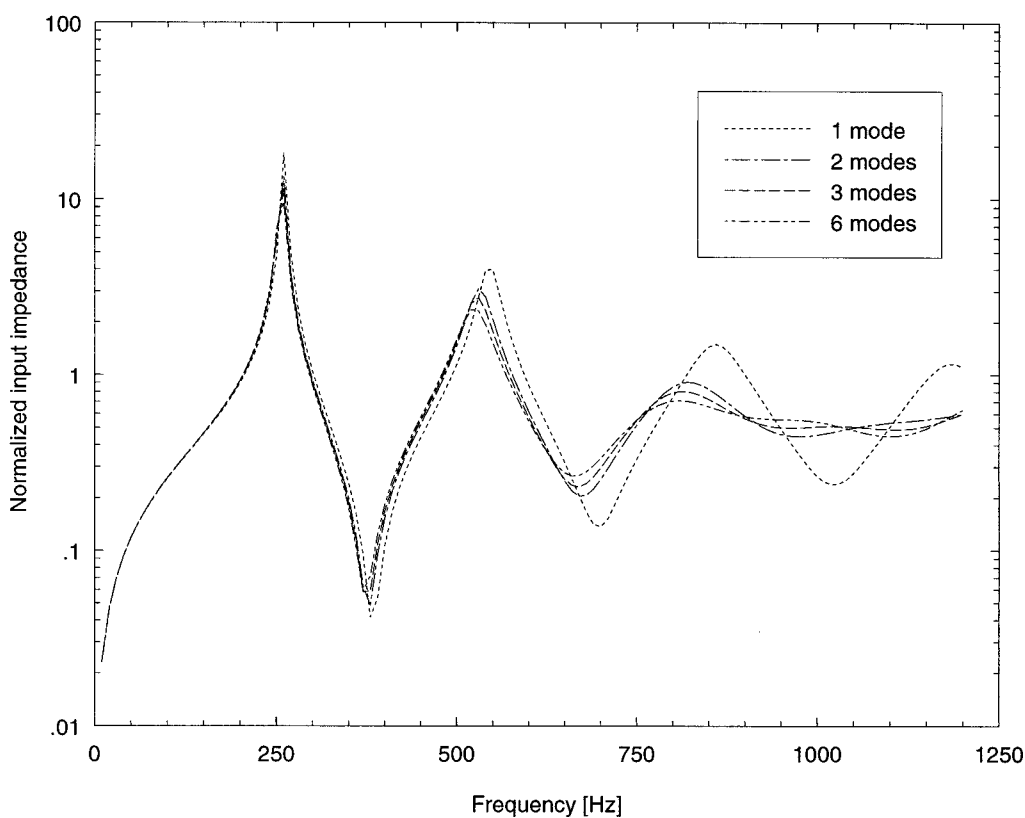


FIG. 4. The corner value of Z calculated with progressively more modes. One mode=plane-wave mode only.

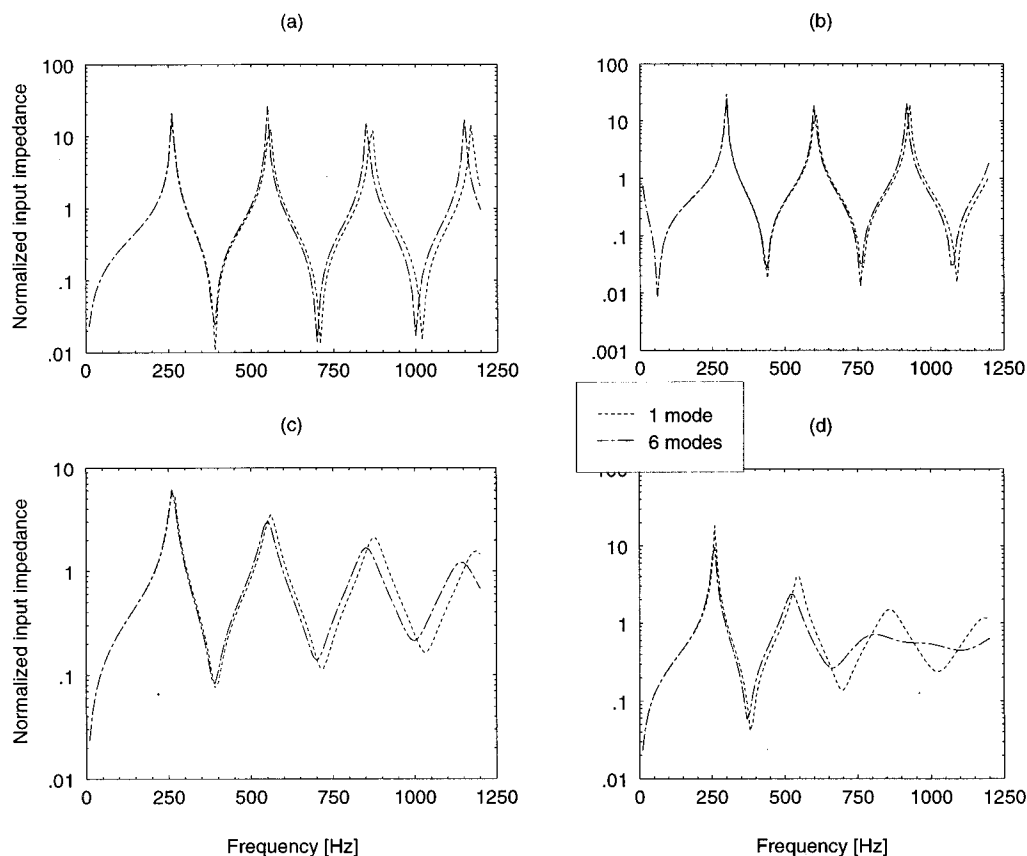


FIG. 5. Comparison of input impedance calculated with plane-wave mode and with six modes for four different end conditions: (a) closed end; (b) ideal open end; (c) infinite pipe termination; and (d) baffled radiation model. Note that the largest difference between the two calculations is found in (d).

An interesting artifact found in the other elements of Z is the appearance of spurious peaks. These are simply artifacts of the numerical calculation, and can be made to disappear if a smaller stepsize is constrained, by using a smaller error threshold in the Runge–Kutta algorithm. This lengthens the calculation time, with very little effect on the resulting value of Z_{00} . Therefore, to a certain extent these artifacts were allowed. It is the adaptive stepsize which “returns” the diverging curves to their normal values; using a fixed stepsize allowed these artifacts to cause the whole algorithm to diverge.

The most important parameter of the calculation is the number of modes used, i.e., the order to which the matrix Z is truncated. In Fig. 4 we show the corner value computed with different sizes of Z . The most radical change occurs when adding just one mode to the basic plane-wave model. Afterward the differences are small, becoming progressively smaller, and more obvious at higher frequencies. For our calculations we generally used six modes, even though all the higher order modes were evanescent. In general, this parameter strongly depends on the frequency range being studied, and on the irregularity of the horn walls.

It is of primary interest to view how much the introduction of the higher-order modes changes the results, as compared to plane-wave analysis, or spherical wave analysis. In this context we studied a number of possible radiation impedances: the hole-in-infinite-baffle, zero radiation impedance (“ideal open end”), infinite impedance (closed end),

and termination by an infinite cylindrical pipe. Not all of these, of course, could be implemented experimentally.

In Fig. 5 we see for each of these a comparison of the input impedance calculated with one and six modes. In all cases there is good agreement at low frequency, less so at higher frequency. The difference is most marked for the most practical case of radiation into a baffled space. For a closed horn, on the other hand, we can expect good results even with a plane-wave model, as has already been shown in Ref. 9.

Another interesting point is to compare the input impedances calculated with different boundary conditions. This is carried out in Fig. 6 for 1 and 6 mode calculations. As also shown in Ref. 9, the resonance frequencies for all cases except for the closed end are very near each other for the plane-wave calculation. This quality of musical horns is actually an artifact of the plane wave model, as shown by experiments in Ref. 9, and borne out by the six mode calculation shown here; at low frequencies it remains true, but at higher frequencies the peaks for the baffled radiation case are very low, and at very different frequencies than the others.

D. Comparing theory and experiments

The most important comparison carried out here is that of the multimode calculation with the experimental results. This appears for theoretical and experimental results in Fig. 7. A good agreement is observed, better than the result ob-

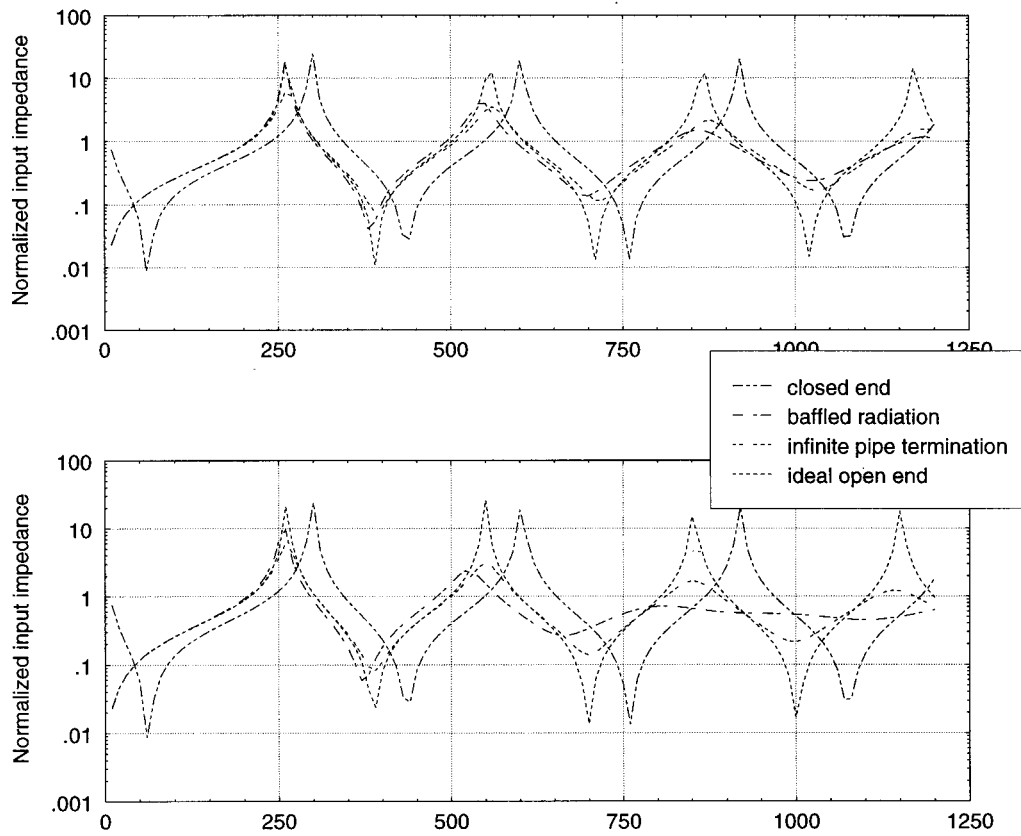


FIG. 6. Comparison of input impedances for four end conditions: (a) with plane-wave mode only; (b) with six modes.

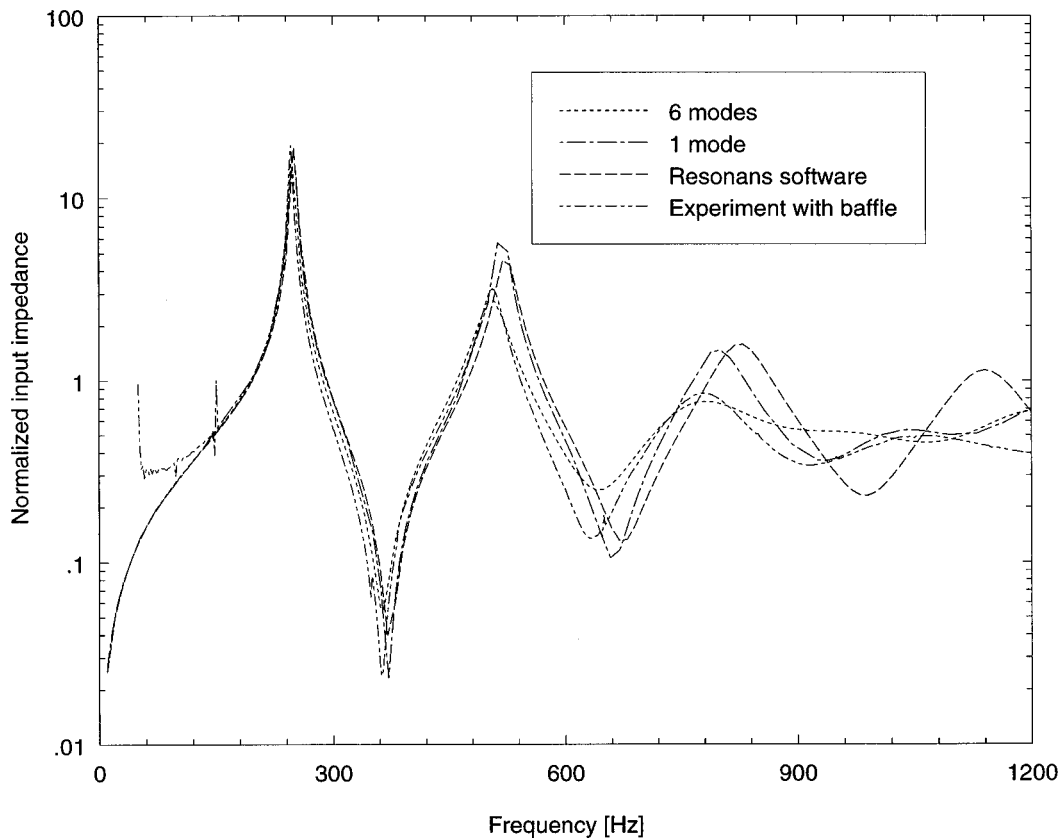


FIG. 7. Comparison of experimental input impedance with various theoretical curves. Best agreement is achieved with the multimodal calculation.

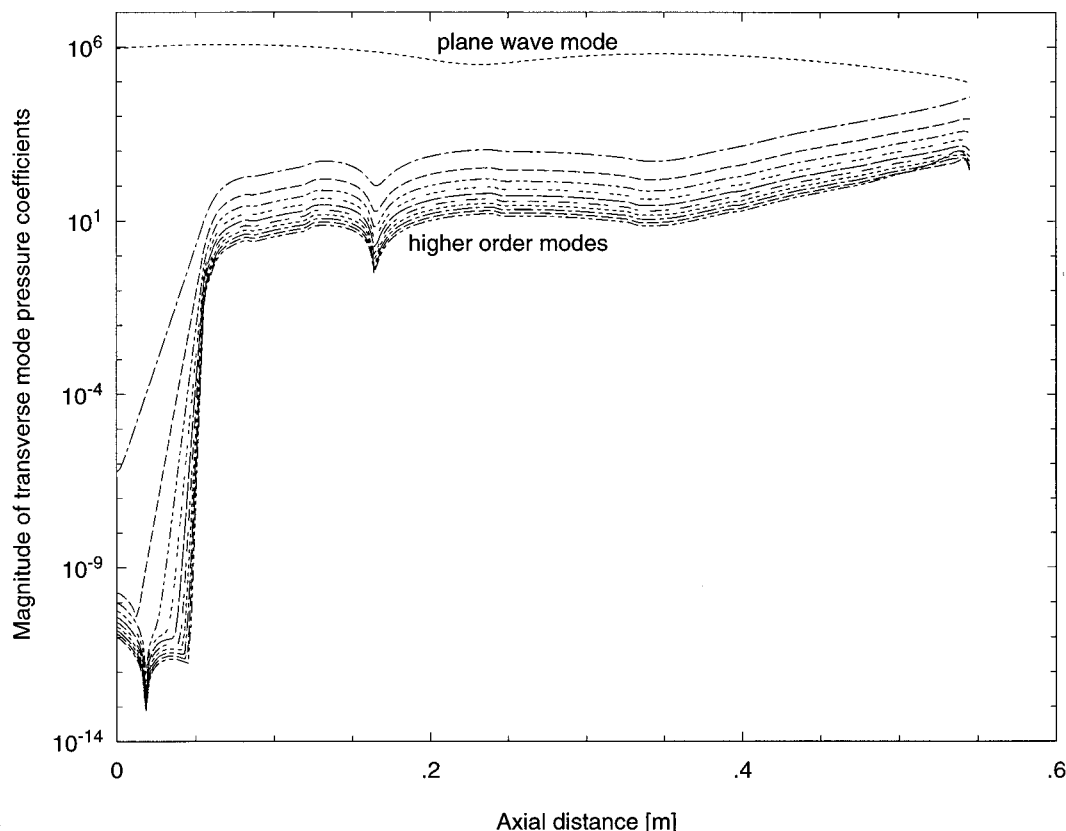


FIG. 8. Coefficients of the pressure series expansion along the horn. Note that the higher-order modes are significant only near the mouth of the horn.

tained using only the lowest order mode. The validity of the simplest model, with only the plane-wave mode, is confirmed for the trombone to be limited to roughly 500 Hz, as predicted, using simple arguments, by Kergomard.¹⁰ On the other hand we do not obtain perfect agreement, unfortunately. A further study is necessary in order to verify the source of this error.

E. Calculations of the pressure and velocity fields

In this section we present calculations of the pressure and velocity fields along the horn for various frequencies, usually with a plane piston (velocity source) excitation.

Our first example is a plane piston excitation at 600 Hz using 11 modes. In Fig. 8 we see the values of the coefficients \mathbf{P} along the horn. From this figure it is clear that the higher-order modes modify the pressure field only toward the very end of the horn, where the flare is largest. This justifies once more the use of a scalar input impedance at the throat.

Next, in Fig. 9 we show graphs of equipressure surfaces toward the end of the horn, using 2, 6, and 11 modes. Despite the small coefficients in \mathbf{P} for the higher order modes, the differences between the figures are clearly visible. Obviously, to obtain a good visual representation of these surfaces, a relatively large number of modes must be used, even though they are all evanescent.

Raising the frequency can create a much more complicated picture. In Fig. 10 all of the higher order modes are

evanescent. At this frequency the equipressure surfaces are far from being spherical.

One more comparison we made is the axial pressure calculated by modal analysis to the axial pressure obtained with the horn equation, i.e., using only the plane-wave mode. This appears in Fig. 11. A considerable difference is apparent.

III. SECOND WAVEGUIDE: THE ELLIPSOIDAL INDENTATION

In this section we show how the method can be applied to a cylindrical duct with an ellipsoidal indentation. The geometry, which is axisymmetric, is characterized by the radius of the cylindrical duct R_0 , the maximum radius of the indentation R_{\max} , and by the distance between the two focal points L (see Fig. 12).

Our first objective is to observe how the transition from low frequencies to higher frequencies is accompanied by an increase of the focalization at the right focal point M_f . This guide is excited by a monopole located at the left focal point M_0 . Classical analysis of the radiation of monopole source into a straight waveguide can be found in classical works (see Morse and Ingard¹¹ and Doak¹²). The influence of the number of modes selected is also analyzed and comparisons with a finite element method (FEM) code (ANSYS) are made.

In a second part, in the same geometry, the Riccati equation (1) is integrated to find reflection properties of the geometry with two different maximum radius of the indentation, R_{\max} .

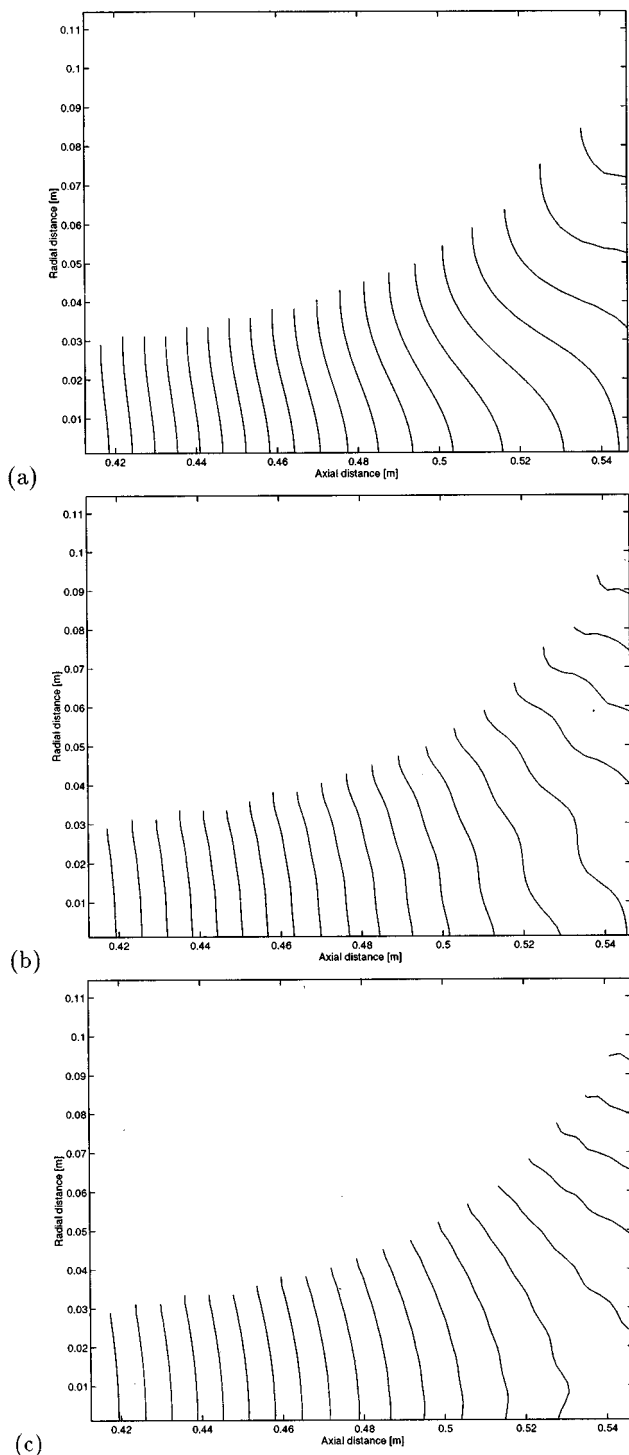


FIG. 9. Equipressure surfaces along the last 14 cm of the horn, at 600 Hz, calculated with: (a) 2 modes; (b) 6 modes; (c) 11 modes.

A. Study of the focalization

1. Procedure

According to the reasoning of geometrical acoustics, if a signal is launched from the left focal point M_0 , a receiver located at the right focal point M_f will observe a focalization. This is due to the fact that each ray emitted from M_0 has to go to M_f with the same phase shift. This is true in a complete ellipsoidal geometry and under the assumption of

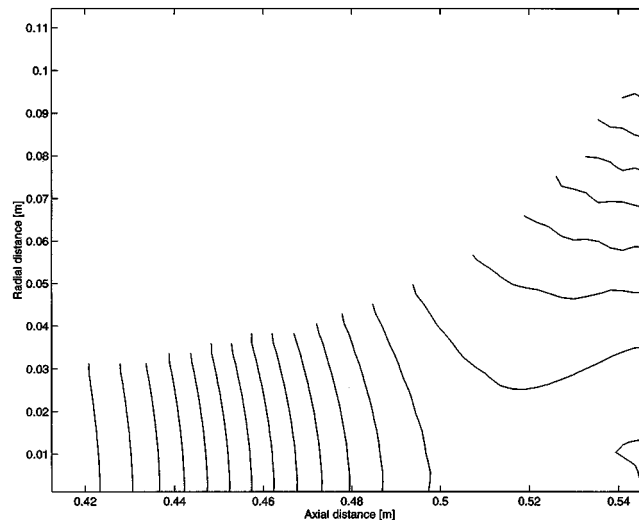


FIG. 10. Equipressure surfaces along the last 14 cm of the horn, at 1200 Hz, calculated with 11 modes.

very high frequency. Here the effects may be modified due to truncation of the ellipsoidal geometry and to finite wavelengths.

Three different frequencies have been selected for the analysis. In the first case, where $kR_0 = 1.85$, only one mode is propagating, therefore it is possible to compare with the classical MPWA (matched plane-wave approximation) calculation and to observe the advantage of a multimodal approach even at low frequency. In the second case, where $kR_0 = 7.4$, three modes are propagating. In the third case, where $kR_0 = 14.8$, five modes are propagating in the cylindrical duct. The peculiar interest here is that a sixth mode can propagate in the indentation only, making this mode a localized one. We can therefore examine whether our method is still efficient in such a case.

Boundary conditions at the right and at the left end of the guide are those of infinite propagation without reflection, i.e., $Z = Z_c$ (resp. $Z = -Z_c$) at the right (resp. left) termination. Calculations were made using the following procedure; first, starting from the right boundary condition ($Z = Z_c$), the

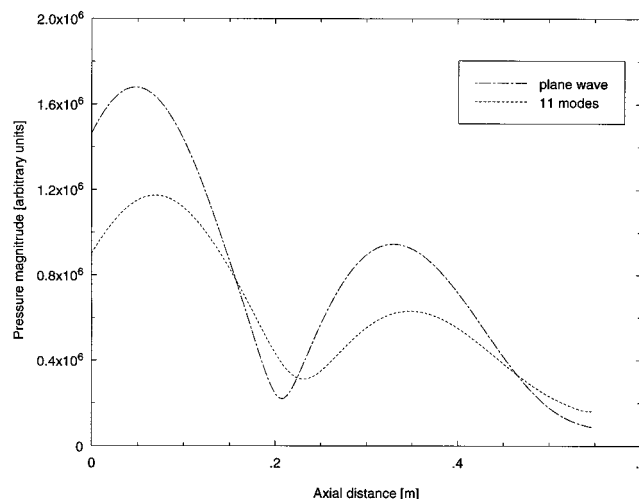


FIG. 11. Pressure amplitude along the axis of the horn, at 600 Hz, calculated with 1 and 11 modes.

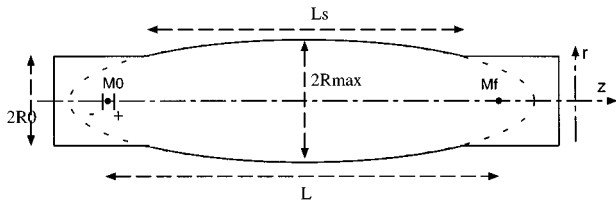


FIG. 12. Geometry with ellipsoidal indentation. The two points, M_0 and M_f , are focal points located at positions $\pm 1.2R_0$.

Riccati equation is integrated down to the point M_0 . Then we have $P_0^+ = ZU_0^+$ valid at the right of M_0 , where + and - will be thereafter related to quantities at the right and left limit of M_0 . On the other hand, knowing that the left termination is infinite, we have $P_0^- = -Z_c U_0^-$ valid at the left of M_0 . Finally, it is possible to get the velocity coefficients of the source by applying the conservation of pressure and flowrate,

$$P_0^+ - P_0^- = 0$$

and

$$U_0^+ - U_0^- = Q,$$

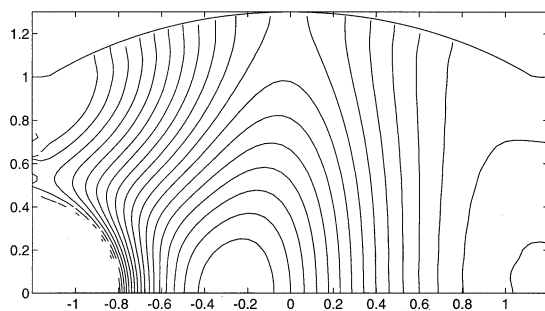
where vector Q contains the coefficient $Q_j = q_0 \psi_j(0)$ of the projection of the monopole of strength q_0 (flowrate source) on the function ψ_m . This gives us

$$U_0^+ = (Z_c + Z)^{-1} Z_c Q.$$

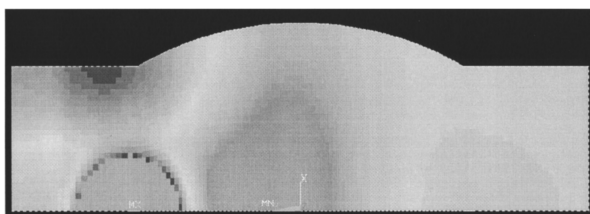
Finally, Eq. (2) is integrated to yield the field across the waveguide.

2. Frequency below the first cutoff of the waveguide

Figure 13 shows the modulus of the pressure for such a low frequency. Our method [Fig. 13(a)] achieves very good agreement with the pattern obtained from the FEM code

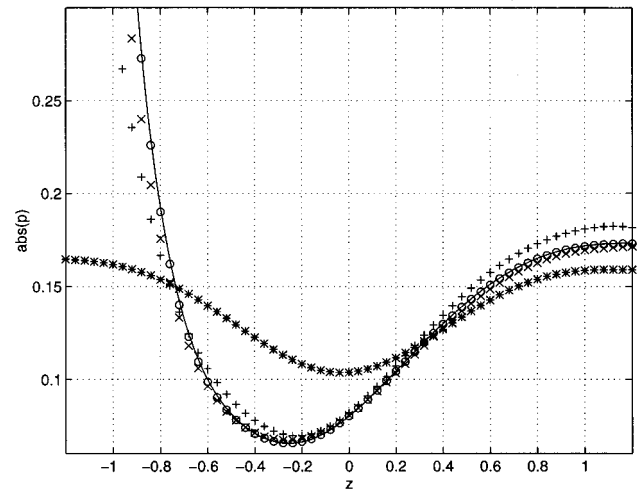


(a) multimodal

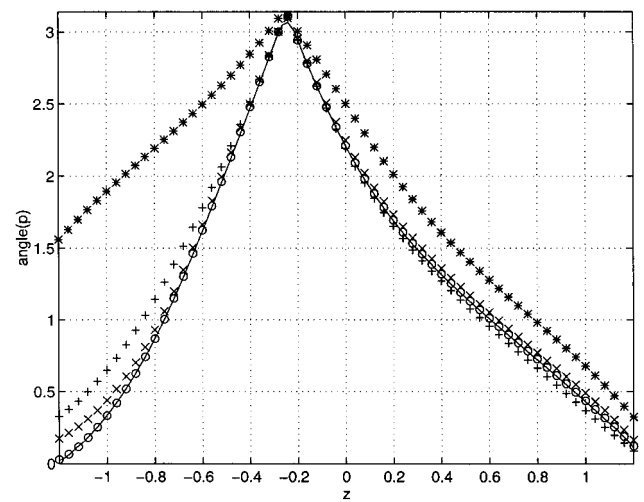


(b) FEM code

FIG. 13. Modulus of the pressure, $kR_0 = 1.85$.



(a) modulus of the pressure



(b) phase of the pressure

FIG. 14. Modulus and phase of the pressure along the central line $r=0$, $kR_0 = 1.85$. (*): 1 mode; (+): 2 modes; (×): 3 modes; (○): 15 modes; plain line: FEM method.

[Fig. 13(b)]. In the two figures, a light focalization can be observed even at this rather low frequency. In order to quantitatively compare the two computations, we have plotted the profiles of the modulus and of the phase of the pressure along the central line ($r=0$) in Fig. 14; four different calculations with different numbers of modes ($m=1, 2, 3$, and 15, where m is the number of modes) are compared to the FEM profile. It is very clear that selecting only one mode ($m=1$, corresponding to the MPWA method) is not sufficient to yield the correct result. Adding just one mode to the modeling ($m=2$) enables to improve the agreement. Eventually, with 15 modes ($m=15$) the agreement between the multimodal and the FEM approaches is almost perfect.

Another view of the effect of the multimodal convergence is seen in Fig. 15, where the pattern of the modulus of the pressure is plotted for 1, 2, 3, and 15 modes. This shows the failure of the MPWA method ($m=1$), but as soon as two modes are used, a roughly correct pattern is obtained. As

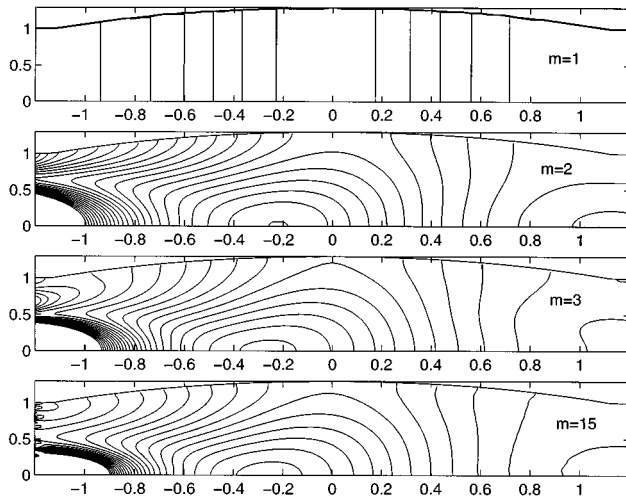


FIG. 15. Modulus of the pressure with 1, 2, 3, and 15 modes, respectively.

m is increased, the pressure tends to achieve the boundary condition at the wall ($\partial p / \partial n = 0$), due to the increasing number of evanescent modes used.

We also measured the rate of convergence of the multimodal method in Fig. 16. The result obtained with 15 modes is chosen as the reference, and three different quantities are proposed for the error ϵ :

$$\epsilon_1 = \frac{\int_V \|p - p_{\text{ref}}\|^2 2\pi r dr dz}{\int_V \|p_{\text{ref}}\|^2 2\pi r dr dz},$$

$$\epsilon_2 = \frac{\int_V \|p - p_{\text{ref}}\|^2 dr}{\int_V \|p_{\text{ref}}\|^2 dr},$$

$$\epsilon_3 = \frac{\int_{r=0} \|p - p_{\text{ref}}\|^2 dz}{\int_{r=0} \|p_{\text{ref}}\|^2 dz}.$$

Here, ϵ_1 represents the actual error in the volume, ϵ_2 is the error ignoring the axisymmetry of the geometry, and thus stressing the difference along the central line ($r=0$), ϵ_3 is the error only along the central line.

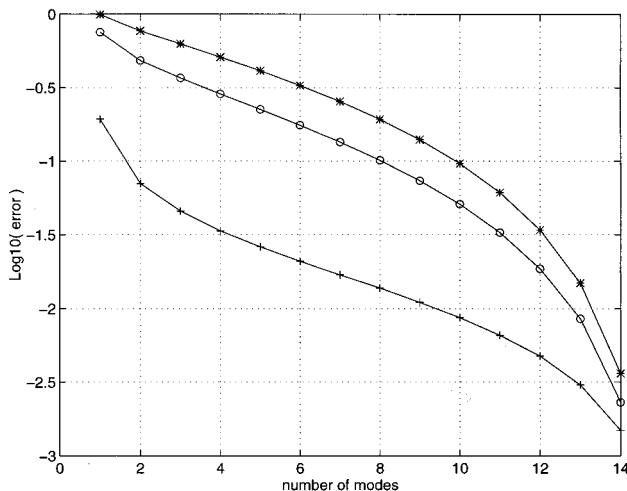
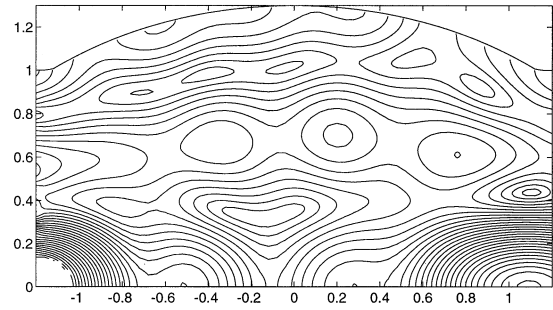
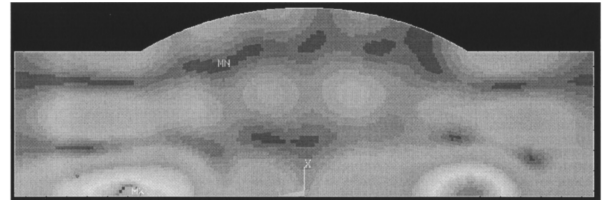


FIG. 16. Error of the method. (+): ϵ_1 , (O): ϵ_2 , (*): ϵ_3 .



(a) multimodal



(b) FEM code

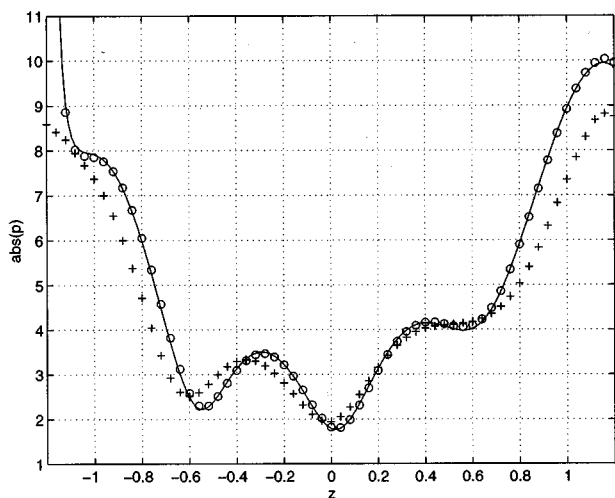
FIG. 17. Modulus of the pressure, $kR_0 = 7.85$.

Figure 16 shows that all three error measures decrease monotonically as more modes are introduced in the calculation. The increased values of ϵ_1 , ϵ_2 , and ϵ_3 , respectively, are due to the fact that the source is a monopole located on the central line; ϵ_1 , ϵ_2 , and ϵ_3 take this progressively more into account.

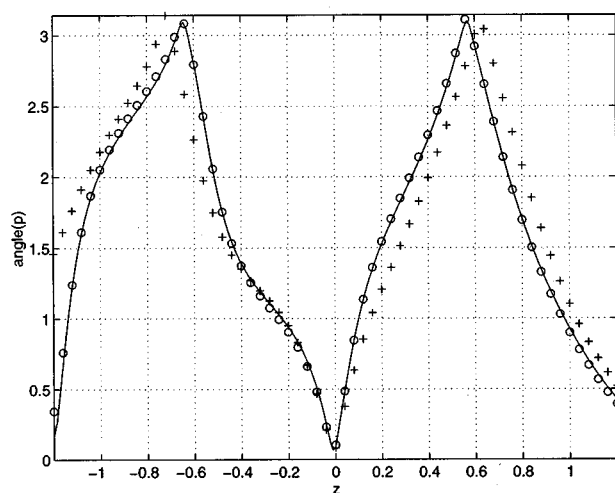
The behavior of ϵ_1 is almost exponential, except between $m=1$ and $m=2$ where the error is reduced. That shows, once more, that even with only two modes a large degree of accuracy is gained with the multimodal method.

3. Higher frequencies

When the frequency is increased to $kR_0 = 7.85$, three modes are propagating and the focalization increases as can be seen in Fig. 17. Figure 17(a) and (b) show the agreement between the multimodal and the FEM computations, as far as the pattern is concerned. On the other hand, when the profile of the modulus of the pressure along the central line ($r=0$) is plotted, we discover some discrepancies. In fact, the latter are an artifact of the FEM code since it is unable to model the correct infinite termination condition at the extremities of the guide. We therefore repeated the comparison (Fig. 18), but this time applying in the multimodal calculation the same condition as in the FEM code ($p = \rho_0 c u_z$) at the extremities ($z = \pm 2R_0$). This is equivalent to imposing $Z_{nm} = \delta_{nm} \rho_0 c / S$ for the matrix impedance, rather than the correct infinite termination condition which is $Z_{nm} = Z_{cnm} = \delta_{nm} \rho_0 c k / (S k_n)$. The three profiles corresponding to the correct and incorrect termination conditions and to the FEM code are plotted in Fig. 18. We see that incorporating the same boundary condition as the FEM code in our multimodal method enables us to obtain an excellent agreement. Nevertheless, one has to keep in mind that the actual correct solution is the multimodal one with correct boundary condition. Obviously this difficulty of the FEM could be overcome by



(a) modulus of the pressure



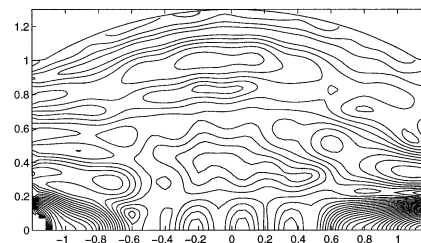
(b) phase of the pressure

FIG. 18. Modulus and phase of the pressure along the central line $r=0$, $kR_0=7.85$. (+): correct boundary condition; (O): incorrect boundary condition; plain line: FEM method.

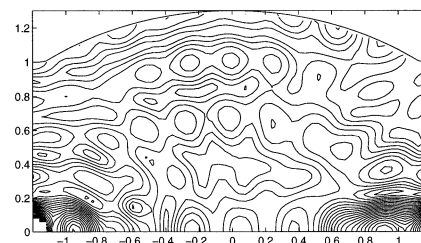
using an integral formulation at the extremities, like the “boundary element method” or the “infinite element method” (see Ref. 13).

The last frequency studied ($kR_0=15.7$) corresponds to five propagating modes in the straight portions with a sixth mode becoming propagating in the indentation. Apart from the increasing frequency, this case is also interesting because the sixth mode has a turning point and we wish to know if our method is able to deal with such a situation. Figure 19 shows the pattern of the modulus of the pressure compared to the FEM solution; the focalization has become more pronounced than with the two preceding frequencies. As in the previous case, incorrect termination conditions for the impedance have to be taken into account in order to get a very good agreement with the FEM code results.

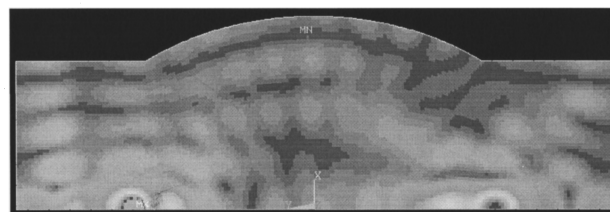
In Fig. 20, the modulus and phase profiles along the central line show excellent agreement with the FEM solution when the incorrect infinite termination condition is taken.



(a) multimodal



(b) multimodal, incorrect



(c) FEM code

FIG. 19. Modulus of the pressure, $kR_0=15.7$.

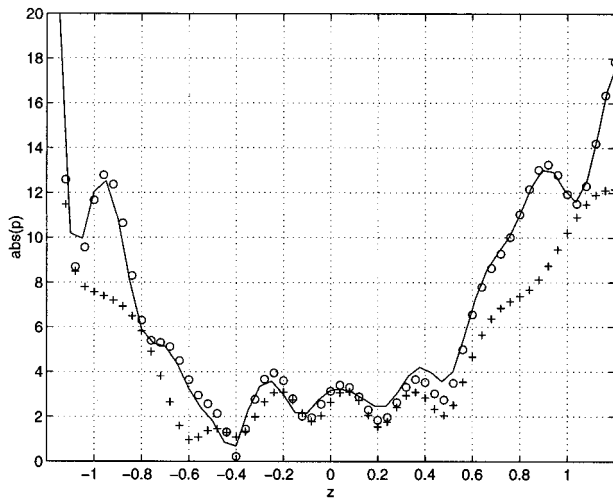
B. Reflection coefficient

Results for the reflection coefficient (defined in Part I) are studied in an ellipsoidal geometry where $L=10$ and $R_{\max}=1.1$ (geometry 1) or $R_{\max}=1.3$ (geometry 2). A common way to characterize a singularity in a waveguide, here the indentation, in a waveguide is to study the reflection coefficient as a function of the frequency.

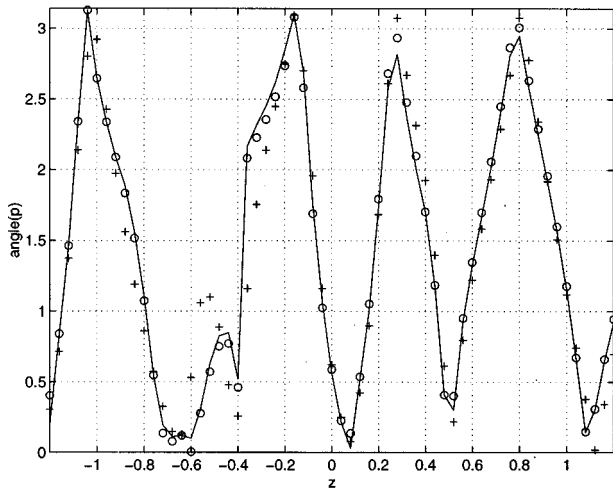
In this case, our method is especially convenient, since it is sufficient to calculate the impedance matrix from the left end to the beginning of the duct to get the reflection coefficient. We do not have to integrate the velocity or pressure equations. On the other hand, by using a FEM code, for instance, we would have to specify exact properties of the source, and the results obtained would have to be transformed to a suitable form to get the reflection coefficient.

In the two following examples, the boundary condition at the right of the indentation is taken as the characteristic matrix impedance, modeling the correct infinite termination. It has been assumed that a plane wave is incident at the left. By knowing the impedance at the left of the indentation, it is possible to obtain the energy reflection coefficient R which is defined as $R=W_r/W_i$, where W_r and W_i are, respectively, the reflected and incident energy flux.

Results for the two geometries considered are shown in Fig. 21. They display the behavior of R against f/f_{c1} where



(a) modulus of the pressure



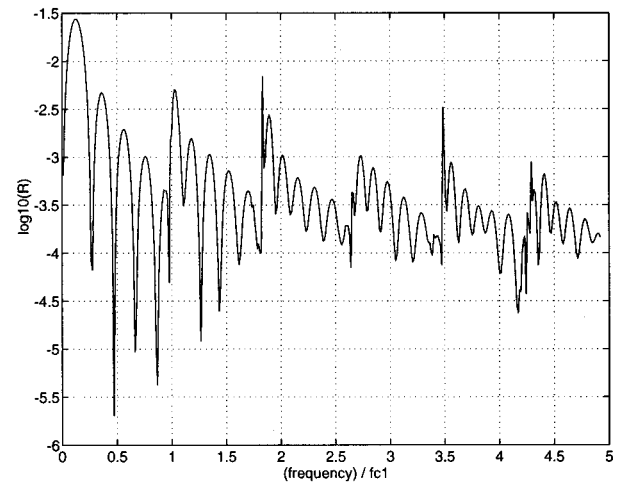
(b) phase of the pressure

FIG. 20. Modulus and phase of the pressure along the central line $r=0$, $kR_0=15.7$. (+): correct boundary condition; (O): incorrect boundary condition; plain line: FEM method.

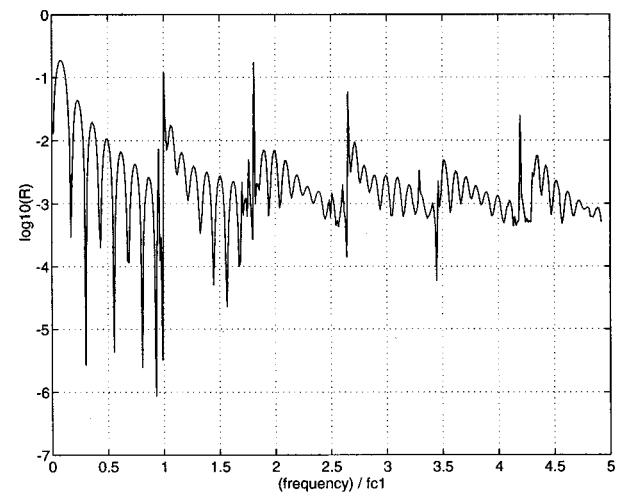
f is the frequency and $f_{c1}=(c\gamma_1)/(2\pi R_0)$ is the cutoff frequency of the first higher-order mode of the duct. The two geometries produce the same kind of results. The first major difference is the lower value of R in the geometry that has a lower radius R_{\max} in the indentation. It seems logical because R_{\max} measures the strength of the indentation. The second difference is in the number of oscillations between each cut-off frequency. This is due to the relation of these oscillations to the length of the indentations L_S ; $L_S=4.4$ for geometry one and $L_S=6.6$ for geometry two. In fact, $L_S=2\sqrt{L^2+4R_{\max}^2}\sqrt{1-1/R_{\max}^2}$.

Three principal features can be observed for the global behavior of the reflection coefficient.

First, the oscillations are due to quasiresonance of the elliptical indentation viewed as an open cavity. In an expansion chamber of length L_S they correspond to the condition $kL_S=n\pi$ under the plane-wave approximation. This condi-



(a) $R_{\max}=1.1$



(b) $R_{\max}=1.3$

FIG. 21. Reflection coefficient as a function of normalized frequency f/f_{c1} (f_{c1} : cutoff frequency of the first nonplane mode) for geometry 1 and 2.

tion expresses the transparency of the cavity since the impedance is then the same at the two extremities of the cavity. It yields five and eight oscillations for geometry 1 ($R_{\max}=1.1$) and geometry 2 ($R_{\max}=1.3$), respectively. From Fig. 21 it can be seen that it is roughly the number of oscillations we can observe.

Second, peaks correspond to each cutoff frequency. They can be understood by remarking that when $f=f_{cn}$ for some mode n , the n th component of the characteristic matrix goes to infinity. Then the end condition for this mode is equivalent to a wall condition and the reflection is enhanced.

Finally, the decrease of R between each cutoff frequency can be understood by drawing an analogy with a potential barrier in quantum mechanics; for the plane-wave approximation the equation for the pressure is

$$(Rp)'' + \left(k^2 - \frac{R''}{R}\right)(Rp) = 0,$$

and the indentation is equivalent to a potential $V(z) = R''/R$. So when the frequency increases, so does the energy k^2 , and the wave is more capable to overcome the potential V . This kind of analogy was developed by Benade and Jansson.¹⁴ Nevertheless, we notice that it is entirely valid if the junctions between the straight duct and the portion with varying cross section satisfy the continuity of both the function $R(z)$ and its first derivative $R'(z)$.¹⁵

IV. CONCLUSION

In this paper, we have shown how the multimodal method can be used to solve various problems concerning waveguides with varying cross section, with no restriction on the flare.

The convergence of the method has been investigated. The limit of the matched plane-wave approximation has been observed, and it appears that adding only a few modes, or even one mode in some cases, can be very effective in improving the accuracy of the modal approach.

The comparison with experiments for an baffled trombone shows better agreement than the plane-wave approximation, and we think that the remaining discrepancies could be due to the finite size of the baffle in the experiments. We plan to study this effect.

We also show that the multimodal method is able to deal with fairly high frequencies (five modes propagating) and to yield high-frequency results, like the focalization.

In conclusion, the multimodal method we have presented in Part I and Part II of this series of papers seems to be well adapted to the general form of waveguides with varying cross section in a large range of frequencies.

ACKNOWLEDGMENTS

One of the authors (V.P.) would like to thank the Gas Engineering Montrouge team of Schlumberger Industries for

scientific and part of financial supports. We also thank Nicolas Joly (CTTM) for interesting discussions on the trombone bell problem.

- ¹V. Pagneux, N. Amir, and J. Kergomard, "A study of wave propagation in varying cross-section waveguides by modal decomposition. Part I: Theory and validation," *J. Acoust. Soc. Am.* **100**, 2034–2048 (1996).
- ²V. Pagneux, J. Kergomard, and G. Marquette, *Propagation multimodale en géométrie complexe* (Congrès Français de Mécanique, Lille, 1993), pp. 1045–1049.
- ³J. Kergomard, "Une equation matricielle des pavillons acoustiques," *C. R. Acad. Sci. Paris Ser. II* **316**, 1691–1694 (1993).
- ⁴A. M. Bruneau, M. Bruneau, Ph. Herzog, and J. Kergomard, "Boundary layer attenuation of higher order modes in waveguides," *J. Sound Vib.* **119**, 15–27 (1987).
- ⁵W. E. Zorumski, "Generalized radiation impedances and reflection coefficients of circular and annular ducts," *J. Acoust. Soc. Am.* **54**, 1667–1673 (1972).
- ⁶Y. Kagawa, T. Yamabuchi, and T. Yoshikawa, "Finite element approach to acoustic transmission-radiation systems and application to horn and silencer design," *J. Acoust. Soc. Am.* **69**, 207–228 (1980).
- ⁷E. G. Williams, "Numerical evaluation of the radiation from unbaffled, finite plates using the FFT," *J. Acoust. Soc. Am.* **74**, 343–347 (1983).
- ⁸J. P. Dalmont and A. M. Bruneau, "Acoustic impedance measurement: plane-wave mode and first helical mode contributions," *J. Acoust. Soc. Am.* **91**, 3026–3033 (1992).
- ⁹N. Amir, G. Rosenhouse, and U. Shimoni, "Modal solution for the sound wave equation in spherical coordinates for non-canonical horns," *Acustica* **82**, 1–8 (1996).
- ¹⁰J. Kergomard, "Champ interne et champ externe des instruments à vent," Doctoral thesis, Université Paris 6 (1981).
- ¹¹P. Morse and U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).
- ¹²P. E. Doak, "Excitation, transmission and radiation of sound from distribution source in hard-wall ducts of finite length," *J. Sound Vib.* **31**, 137–174 (1971).
- ¹³T. W. Dawson and J. A. Fawcett, "A boundary integral equation method for acoustic scattering matrix in a waveguide with nonplanar surface," *J. Acoust. Soc. Am.* **87**, 1110–1125 (1990).
- ¹⁴A. H. Benade and E. V. Jansson, "On plane wave and spherical waves in horns with nonuniform flare. Part I: Theory," *Acustica* **31**, 79–98 (1974).
- ¹⁵C. Depollier, J. Kergomard, and F. Laloe, "Localisation d'Anderson des ondes dans les réseaux acoustiques unidimensionnels aléatoires," *Ann. Phys.* **11**, 457–492 (1986).

Rational operators for filtering

R. J. Cederberg and Michael D. Collins

Naval Research Laboratory, Washington, DC 20375

Henrik Schmidt

Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

William L. Siegmann

Rensselaer Polytechnic Institute, Troy, New York 12180

(Received 7 August 1996; accepted for publication 7 January 1997)

The parabolic equation (PE) method has been applied to approximate the wave equation, construct radiation conditions, solve scattering problems, construct initial conditions, and derive energy flux conditions. A new application of the PE method, filters based on rational approximations of depth operators, is described and tested. The rational approximation is designed to act as the identity operator on the desired part of the spectrum of the depth separated wave equation and to annihilate other parts of the spectrum. This approach does not require explicit knowledge of the spectrum. The applications of the filter include directly solving eigenvalue problems, annihilating components of the wave-number spectrum, generating initial conditions at the source range, and eliminating Gibbs oscillations that arise in energy-conserving PE solutions. © 1997 Acoustical Society of America. [S0001-4966(97)04905-9]

PACS numbers: 43.20.Mv, 43.30.Dr [ANN]

INTRODUCTION

Parabolic equation (PE) techniques are useful for solving propagation and scattering problems in waveguides that can be regarded as locally stratified. There are several mathematically different applications of the PE method, including approximating the wave equation in the interior of the domain,¹⁻⁴ advancing the field outward in range,⁵ constructing radiation boundary conditions for elliptic⁶ and parabolic⁷ wave equations, solving scattering problems,^{8,9} constructing initial conditions,¹⁰ deriving interface conditions that conserve energy,¹¹ and obtaining the solution of the wave equation on a vertical array.¹² This paper describes a new application of the PE method, filters that do not require explicit knowledge of the spectrum of the depth separated wave equation.

The filtering operator is based on a rational approximation of a step function. Other PE applications involve rational approximations of square roots,¹³⁻¹⁶ exponentials of square roots,⁵ and other functions. Applications of the filter include implementing the self-starter¹⁰ at the source range, eliminating Gibbs oscillations associated with energy-conservation corrections,^{11,17,18} solving the eigenvalue problem associated with the normal mode solution,¹⁹⁻²¹ and annihilating components of the wave-number spectrum.²²⁻²⁴ PE techniques and the implementation of the filter are described in Sec. I. Applications of the filter are illustrated in Sec. II.

I. FILTERING OPERATORS

In this section, we discuss the PE method and describe filters based on rational approximations. For simplicity, we consider the case of a horizontally stratified acoustic waveguide in which k is the wave number and ρ is the density. The generalization to other cases, such as problems involv-

ing elastic layers,^{10,11} is straightforward. We work in plane geometry, where z is the depth below a pressure release boundary and the range x is the horizontal distance from a line source.

The complex acoustic pressure p satisfies the frequency domain wave equation,

$$\frac{\partial^2 p}{\partial x^2} + k_0^2(1+X)p = 2\delta(x)\delta(z-z_0), \quad (1)$$

where k_0 is a representative wave number, z_0 is the source depth, and

$$X = k_0^{-2} \left(\rho \frac{\partial}{\partial z} \frac{1}{\rho} \frac{\partial}{\partial z} + k^2 - k_0^2 \right). \quad (2)$$

The normal mode solution¹⁹⁻²¹ of Eq. (1) is

$$p(x, z) = \sum_m k_m^{-1} \phi_m(z_0) \phi_m(z) \exp(ik_m x), \quad (3)$$

where the m th mode $\phi_m(z)$ and eigenvalue k_m^2 satisfy

$$k_0^2(1+X)\phi_m = k_m^2 \phi_m. \quad (4)$$

Since the operator X reduces to scalar multiplication in mode space, we treat X as a complex number in the discussion of rational approximations.

The PE solution is based on factoring the operator in Eq. (1) for $x > 0$ to obtain

$$\left(\frac{\partial}{\partial x} + ik_0(1+X)^{1/2} \right) \left(\frac{\partial}{\partial x} - ik_0(1+X)^{1/2} \right) p = 0. \quad (5)$$

Applying an outgoing radiation condition, we obtain the outgoing wave equation,

$$\frac{\partial p}{\partial x} = ik_0(1+X)^{1/2} p. \quad (6)$$

TABLE I. Functions that have been approximated by rational functions for PE applications.

Function	Application	References
$(1+X)^{1/2}$	parabolic wave equation	13–16
$(1+X)^{-1/2}$	two-way PE	9
$(1+X)^{1/4}$	energy conservation	11
$(1+X)^{-1/4}$	self-starter	10
$(1+X)^{3/4}$	self-starter implementation	10
$\exp[ik_0\Delta r(1+X)^{1/2}]$	split-step Padé algorithm	5
$(1+X)^{7/4} \exp[ik_0r_0(1+X)^{1/2}]$	solution on a vertical array	12
step functions	filtering	this paper

Approximating the operator square root in Eq. (6) with the rational function,

$$R(X) = \sum_{j=1}^n \frac{\alpha_{j,n}X}{1 + \beta_{j,n}X} = \prod_{j=1}^n \frac{1 + \gamma_{j,n}X}{1 + \beta_{j,n}X}, \quad (7)$$

we obtain an equation that may be solved with standard numerical techniques,¹⁶

$$\frac{\partial p}{\partial x} = ik_0 R(X)p. \quad (8)$$

The sum and product expressions for $R(X)$ are both useful for PE applications. The coefficients $\alpha_{j,n}$ and $\beta_{j,n}$ are selected so that the propagating spectrum of X is handled accurately and the evanescent spectrum is annihilated. The coefficients may be obtained efficiently by first solving a linear least-squares problem for the coefficients of the ratio of two polynomials expressed in terms of powers of X and then finding the roots of the polynomials.¹² Generalizing the PE solution to gradually range-dependent media involves the assumption that outgoing energy dominates incoming energy and a correction to guarantee that energy flux is conserved.^{11,17,18}

As Table I indicates, rational functions have been used to approximate a wide variety of functions associated with operators that arise in wave propagation and scattering problems. Since rational approximations are very robust,²⁵ a high level of accuracy is often achieved by using a relatively small number of terms. It is rather remarkable that rational functions can even be used to approximate step functions. Filters are based on rational approximations of the function,

$$f(X) = \begin{cases} 1, & \text{for } X \in \Omega, \\ 0, & \text{for } X \notin \Omega, \end{cases} \quad (9)$$

where Ω corresponds to a subset of the propagating modes. The case in which Ω includes only the i th mode provides a direct method for solving the eigenvalue problem. The case in which Ω includes the entire propagating spectrum provides a filter that is useful for improving stability for problems involving the self-starter and energy-conservation corrections. To simplify the design of the filter, we consider cases in which Ω is part of the real axis. This approach is effective when attenuation is relatively weak and the eigenvalues are close to the real axis.

After determining an interval $\Omega = (X_a, X_b)$ that includes only the i th mode, we operate on the delta function,

$$\delta(z - z_0) = \sum_m \phi_m(z_0) \phi_m(z), \quad (10)$$

with $f(X)$ to obtain

$$f(X) \delta(z - z_0) = f(X) \sum_m \phi_m(z_0) \phi_m(z) = \phi_i(z_0) \phi_i(z). \quad (11)$$

To avoid numerical difficulties associated with the discontinuities of f , we implement Eq. (11) with an intermediate step. We approximate the continuous function,

$$g(X) = \begin{cases} (X - X_a)(X - X_b), & \text{for } X \in \Omega, \\ 0, & \text{for } X \notin \Omega, \end{cases} \quad (12)$$

by $R(X)$ and obtain

$$R(X) \sigma(z) \equiv \phi_i(z_0) \phi_i(z), \quad (13)$$

$$(X - X_a)(X - X_b) \sigma(z) = \delta(z - z_0). \quad (14)$$

In addition to eliminating numerical difficulties, this approach improves robustness because the intermediate result $\sigma(z)$ is twice differentiable at $z = z_0$. The numerical techniques discussed in Ref. 10 may be applied to solve first Eq. (14) and then Eq. (13).

Although applying the filter to test an entire interval at once may prove to be more efficient than testing one point at a time, developing an efficient search algorithm based on the filter may be relatively complicated. It is necessary to select search intervals. A simple approach involves exhaustively searching with a sequence of small intervals. It is necessary to decide whether $R(X) \sigma(z)$ corresponds to an eigenfunction (or perhaps a combination of eigenfunctions). This is relatively easy because $R(X) \sigma(z)$ is very small when Ω does not contain an eigenvalue. It is necessary to select a value of z_0 that does not correspond to a null of the desired eigenfunction. The rational function must be designed with care. It is necessary that $R(X) \ll 1$ for $X \notin \Omega$ so that undesired modes are annihilated and that $R(X) = O(1)$ for $X \in \Omega$ so that the desired mode is retained. Since the actual value of $R(X)$ is not important for $X \in \Omega$, efficiency may be enhanced by placing a relatively large number of the least-squares constraints outside of Ω . The sum form of the rational approximation is more efficient on computers with parallel processing capability. The product form should be used on computers with a single processor because it is less sensitive to round-off errors.

The filtering operator has two applications associated with the wave-number spectrum,^{22–24} eliminating the trapped modes and eliminating all but the trapped modes. Taking the Fourier transform in x of Eq. (1), we obtain

$$k_0^2(1+X)\hat{p} - h^2\hat{p} = 2\delta(z - z_0), \quad (15)$$

where \hat{p} is the spectrum and the separation constant h is the horizontal wave number. We operate on the right side of Eq. (15) to obtain the filtered spectrum \hat{p}_f , which is defined by

$$k_0^2(1+X)\hat{p}_f - h^2\hat{p}_f = 2R(X)(X - X_c)^{-2}\delta(z - z_0). \quad (16)$$

To eliminate the trapped modes from the spectrum, we take

$$R(X) \equiv \begin{cases} (X - X_c)^2, & \text{for } X \leq X_c, \\ 0, & \text{for } X > X_c, \end{cases} \quad (17)$$

where X_c corresponds to a horizontal phase speed slightly larger than the horizontal phase speed associated with the highest trapped mode. For this choice of $R(X)$, the sharp peaks that correspond to the propagating modes are eliminated from the wave-number spectrum. This approach for reducing the required number of sampling points is an alternative to moving the integration contour off the real axis.²⁴ This approach is based on three techniques that are commonly used in propagation modeling: (1) The filtering is performed with a PE technique; (2) energy trapped in the waveguide is obtained as a finite sum of normal modes; and (3) the remainder of the spectrum is obtained as a wave-number integral involving a relatively smooth spectrum.

To retain only the trapped modes in the spectrum, we reverse the inequalities in Eq. (17). For this case, we have the normal mode representation,

$$\hat{p}_f(h, z) \equiv 2 \sum_{m=1}^n \frac{\phi_m(z_0) \phi_m(z)}{k_m^2 - h^2}, \quad (18)$$

where n is the number of trapped modes. The eigenvalues may be estimated by performing a least-squares fit of a rational function to the filtered spectrum. The estimates of the eigenvalues correspond to the poles of the rational approximation. This approach is very efficient when it is possible to perform the least-squares fit using only $O(n)$ samples of the spectrum. Although it is also possible to fit a rational function to \hat{p} , the contribution of the nontrapped modes can lead to a breakdown in the accuracy of the eigenvalue estimates.

The filtering operator can also be used to derive initial conditions for the outgoing wave equation. Integrating Eq. (1) over an arbitrarily small interval about $x = 0$, we obtain

$$\lim_{x \rightarrow 0^+} \frac{\partial p}{\partial x} = \delta(z - z_0). \quad (19)$$

Using Eq. (6) to replace the range derivative in Eq. (19), we obtain the initial condition,

$$ik_0(1 + X)^{1/2}p(0, z) = \delta(z - z_0). \quad (20)$$

The numerical solution of Eq. (20) must be obtained with care to avoid difficulties associated with the singularity at $z = z_0$. We annihilate the evanescent modes to obtain the filtered self-starter,

$$p_f(0, z) = R(X)\sigma(z), \quad (21)$$

$$ik_0(1 + X)^2\sigma(z) = \delta(z - z_0), \quad (22)$$

where

$$R(X) \equiv \begin{cases} (1 + X)^{3/2}, & \text{for } X \geq -1, \\ 0, & \text{for } X < -1. \end{cases} \quad (23)$$

Both $\sigma(z)$ and $p_f(0, z)$ are bounded. For the case of a point source in cylindrical geometry, the exponent of $3/2$ in Eq. (23) is replaced by $7/4$ (Ref. 12). The filtered self-starter should be useful for problems involving range dependence near the source location.

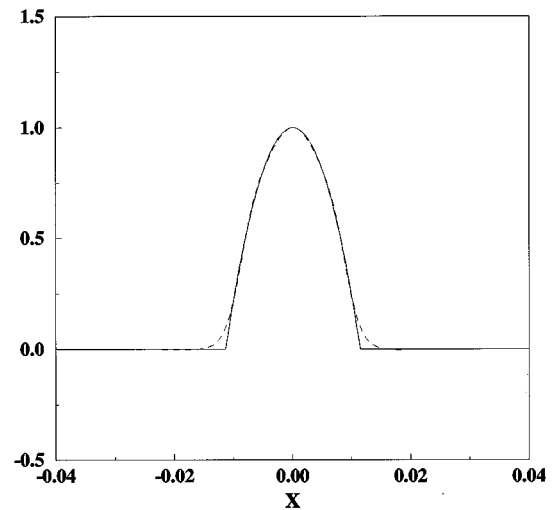


FIG. 1. Rational approximation (dashed curve) of the product of a step function and a quadratic function (solid curve). The six-term rational approximation was constructed using least squares with 60 samples between $X \approx -0.03$ and $X \approx 0.03$. The quadratic function is normalized so that its maximum value is unity.

II. APPLICATIONS

In this section, we discuss applications of the filtering operator and present examples. We first illustrate the rational approximation for the case $\Omega = (X_a, X_b)$, where k_0 corresponds to a horizontal phase speed of 2630 m/s and Ω corresponds to phase speeds between 2615 and 2645 m/s ($X_b \approx -X_a \approx 0.011$ for this problem). We take $n = 6$, and perform the least-squares fit using 60 equally spaced points between $3X_a$ and $3X_b$. The rational approximation is illustrated in Fig. 1. Away from the endpoints, the agreement between the rational function and the quadratic function is good throughout Ω . The rational function nearly vanishes outside of Ω .

The purpose of example A is to illustrate the application of the filter to solving the eigenvalue problem using the approach involving Eqs. (13) and (14). The waveguide consists of a 200-m layer of water overlying an elastic half-space. The frequency is 25 Hz. The speed of sound is 1500 m/s in the water. In the elastic layer, which has a density 1.5 times the density of the water, the compressional and shear wave speeds and attenuations are 3400 and 1800 m/s and 0.1 and 0.2 dB/ λ . We searched for modes by constructing filters corresponding to 50-m/s intervals in phase speed. After obtaining the modes, we estimated the eigenvalues using the Rayleigh quotient. The eigenvalues and modes obtained by

TABLE II. The eigenvalues for example A, which were estimated using the Rayleigh quotient and a filtering operator to extract modes from the delta function as in Eqs. (13) and (14). These estimates are compared with eigenvalues found with KRAKENC. The first eigenvalue corresponds to an interface wave.

Mode	Filtering	KRAKENC
1	(0.119 56, 2.20 $\times 10^{-4}$)	(0.119 25, 2.18 $\times 10^{-4}$)
2	(0.103 43, 3.60 $\times 10^{-6}$)	(0.103 43, 3.60 $\times 10^{-6}$)
3	(0.099 62, 1.50 $\times 10^{-5}$)	(0.099 62, 1.50 $\times 10^{-5}$)

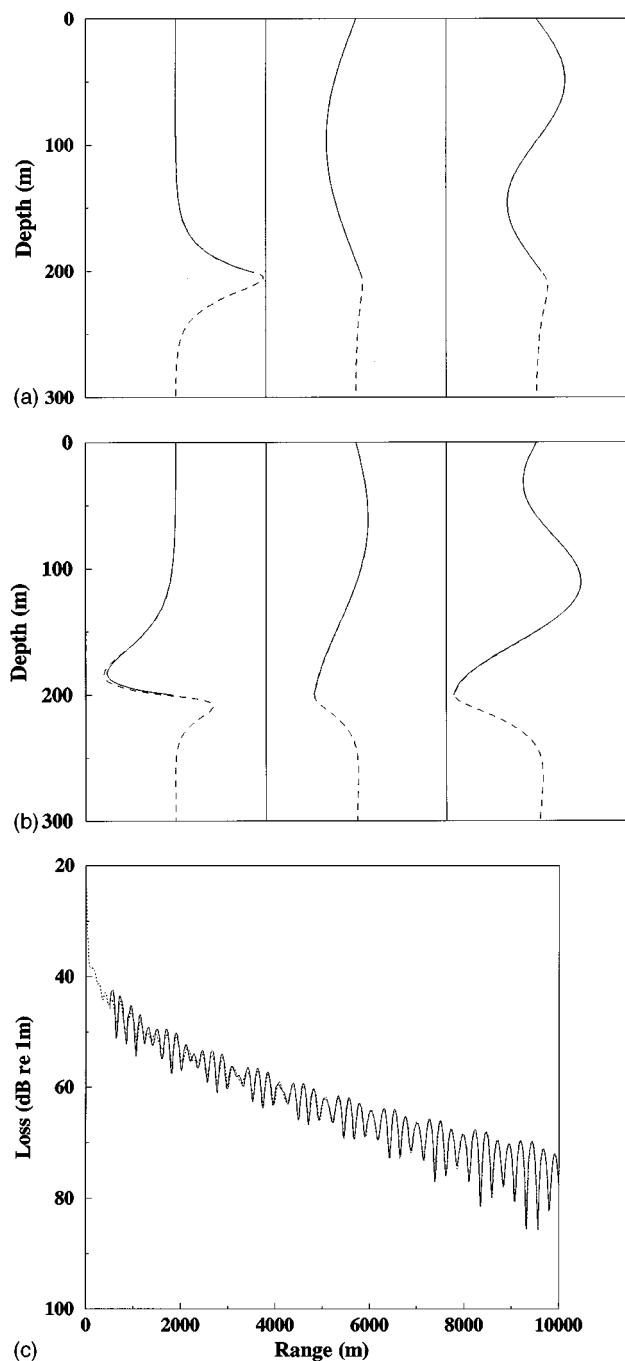


FIG. 2. Results for example A, which illustrates the direct solution of an eigenvalue problem with the filtering operator and involves a 200-m water layer overlying an elastic bottom. Modes obtained by applying a filtering operator to a delta function (dashed curves) compared with modes obtained with KRAKENC (which does not provide the modes in the bottom). (a) The real parts of the modes are nearly identical. (b) The imaginary parts of the modes exhibit only small differences. (c) Transmission loss corresponding to source and receiver depths of 190 m generated with an elastic PE model (solid curve) and by combining the modes and eigenvalues obtained using the filtering operator (dashed curve).

filtering appear in Table II and Fig. 2 and are in agreement with results obtained with the KRAKENC model.²⁶ Transmission loss curves generated with the modes obtained by filtering and with an elastic PE model²⁷ also appear in Fig. 2 and are in agreement.

Example B illustrates the filtered self-starter for a prob-

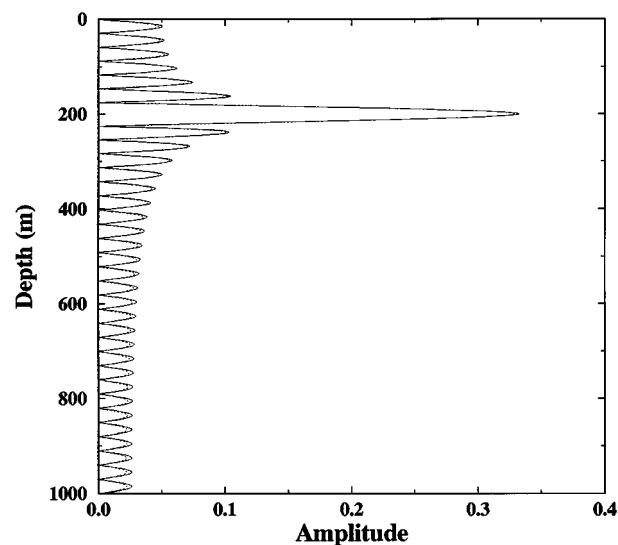


FIG. 3. Results for example B, which illustrates the filtered self-starter and involves a point source at $z = 200$ m in a homogeneous waveguide. The solid curve corresponds to the exact initial condition obtained by summing the propagating modes. The dashed curve was obtained using the filtering operator to generate the self-starter.

lem involving a 25-Hz point source at $z = 200$ m in a waveguide that is homogeneous, lossless, and 1 km thick. The sound speed is 1500 m/s and p vanishes at the top and bottom boundaries. The filtered self-starter appearing in Fig. 3 is in agreement with the exact initial condition constructed from the propagating modes.

Example C illustrates the application of a filtering operator to suppress Gibbs oscillations in an energy-conserving PE solution. This benchmark problem²⁸ involves a 25-Hz point source at $z = 100$ m in a water layer that overlies a sloping ocean bottom. The sound speed is 1500 m/s in the water. In the bottom, the sound speed is 1700 m/s, the density is 1.5 times the density of water, and the attenuation is 0.5 dB/ λ . The depth of the water layer decreases linearly with range from 200 m at the source to zero at a range of 4 km. The sloping ocean bottom is approximated using a sequence of stair steps. For many problems, energy flux is approximately conserved and accurate solutions are obtained by conserving $u = (\rho c)^{-1/2} p$ across vertical interfaces.^{17,18} With this approach, it is necessary to use care to avoid introducing Gibbs oscillations into the solution because u is not continuous across horizontal interfaces. The Gibbs oscillations may be avoided by designing a propagator that annihilates the evanescent spectrum.¹⁸ An alternative approach is to apply a filter at each vertical interface. Although this approach requires some additional effort, it may prove to be useful for resolving the elastic and poro-elastic²⁹ cases, which are relatively prone to stability problems.^{27,29,30} We designed the filter by taking $n = 15$ and $\Omega = (-0.8, 0)$, letting k_0 correspond to 1500 m/s, and using 20 samples in $-0.8 < X < 0$ and 30 samples in $-2 < X < -0.85$. Solutions for example C appear in Fig. 4. The unfiltered solution contains Gibbs oscillations as in Fig. 1(e) of Ref. 18. The solution obtained using complex coefficients in the rational approximation of the square root (as in Ref. 18) contains a

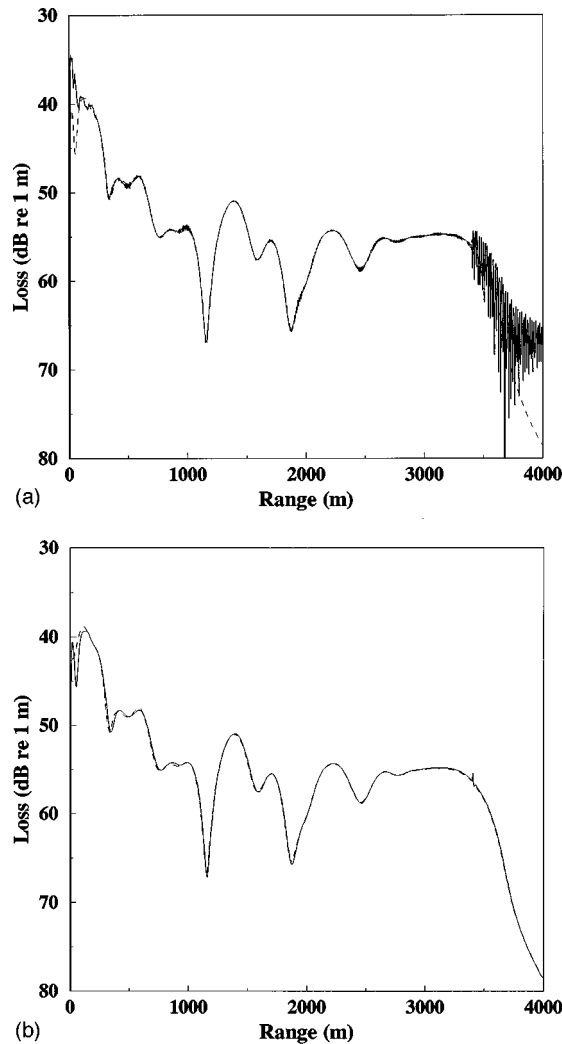


FIG. 4. Results for example C, which illustrates the elimination of Gibbs oscillations with the filtering operator and involves a 25-Hz point source in a water layer overlying a sloping bottom. The receiver is at $z = 30$ m. (a) Transmission loss curves obtained using the energy-conserving PE and a rational approximation involving real coefficients (solid curve) and complex coefficients (dashed curve). (b) Transmission loss curves obtained using the energy-conserving PE and a rational approximation involving complex coefficients (solid curve) and real coefficients plus filtering (dashed curve).

spurious oscillation near $r = 3.4$ km that quickly decays with range. The filtered solution is essentially free of Gibbs oscillations.

Examples D and E illustrate the filtered spectrum de-

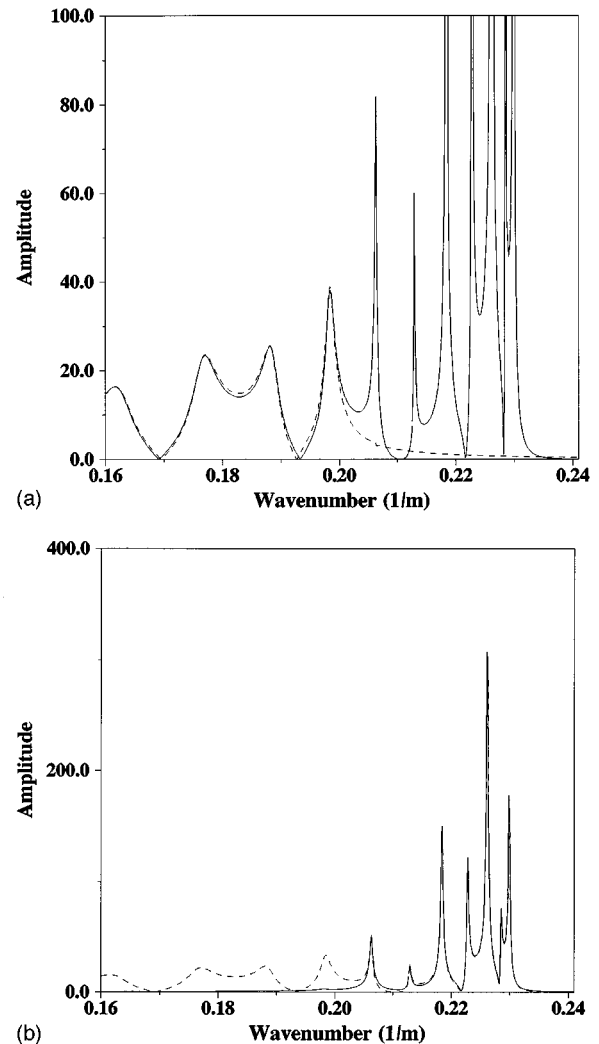


FIG. 5. Results for examples D and E, which illustrate the elimination of components of the wave-number spectrum with the filtering operator and involve a 200-m deep water layer and a 55-Hz source. (a) The filtering operator is used to eliminate the trapped modes (dashed curve); the solid curve is the unfiltered spectrum. (b) The filtering operator is used to eliminate the nontrapped spectrum (solid curve); the dashed curve is the unfiltered spectrum.

fined in Eq. (16). These examples involve the same environmental parameters as example C, with the exception that the depth of the water column is constant at 200 m. A 55-Hz source is placed at $z = 100$ m. We take $n = 12$, use 80

TABLE III. The eigenvalues for example E, which were estimated by fitting a rational function to the unfiltered spectrum and the filtered spectrum defined in Eq. (16). These estimates are compared with eigenvalues found with KRAKEN.

Mode	Unfiltered spectrum	Filtered spectrum	KRAKEN
1	$(0.229\ 62, -7.60 \times 10^{-6})$	$(0.229\ 92, 2.18 \times 10^{-6})$	$(0.229\ 91, 1.93 \times 10^{-6})$
2	$(0.226\ 27, 8.35 \times 10^{-6})$	$(0.228\ 51, 8.22 \times 10^{-6})$	$(0.228\ 50, 7.51 \times 10^{-6})$
3	$(0.223\ 05, 6.69 \times 10^{-7})$	$(0.226\ 15, 1.60 \times 10^{-5})$	$(0.226\ 11, 1.64 \times 10^{-5})$
4	$(0.220\ 50, -8.61 \times 10^{-4})$	$(0.222\ 79, 2.78 \times 10^{-5})$	$(0.222\ 71, 2.85 \times 10^{-5})$
5	$(0.218\ 51, 1.09 \times 10^{-4})$	$(0.218\ 38, 4.40 \times 10^{-5})$	$(0.218\ 24, 4.52 \times 10^{-5})$
6	$(0.212\ 71, 2.30 \times 10^{-6})$	$(0.212\ 89, 7.01 \times 10^{-5})$	$(0.212\ 66, 7.25 \times 10^{-5})$
7	$(0.206\ 24, 1.38 \times 10^{-4})$	$(0.206\ 29, 1.43 \times 10^{-4})$	$(0.205\ 97, 1.52 \times 10^{-4})$

samples in $-2 < X < 5$, and select X_c and k_0 to correspond to horizontal phase speeds of 1720 and 3440 m/s ($X_c = 3$ for this problem). The filtered spectra appear in Fig. 5. The trapped modes are annihilated for example D: The filtered spectrum is relatively smooth; there are small differences between the filtered and unfiltered spectra in the nontrapped region due to contributions from terms corresponding to the trapped modes. Only the trapped modes are retained for example E: The filtered spectrum does not contain smooth peaks in the nontrapped region; there are small differences between the filtered and unfiltered spectra in the trapped region. We fit seven-term rational functions to the filtered and unfiltered spectra of example E and obtained the eigenvalues appearing in Table III. The eigenvalues obtained using the filtered spectrum are in agreement with eigenvalues obtained using the KRAKEN model.²⁶ Some of the eigenvalues obtained using the unfiltered spectrum contain significant errors, including sign errors in the imaginary parts of two of the eigenvalues.

III. CONCLUSION

Filtering operators are a new application of the PE method that are based on rational approximations of a step function. Filtering operators can be used to annihilate portions of the spectrum without explicit knowledge of the spectrum. Filtering operators, may be used to directly solve eigenvalue problems, eliminate Gibbs oscillations from energy-conserving PE solutions, implement the self-starter at the source range, and eliminate components of the wave-number spectrum. The potential utility of filtering operators was illustrated with examples for each of these applications.

ACKNOWLEDGMENTS

This work was supported by the Ocean Acoustics Program, Code 3210A, of the Office of Naval Research. This work was performed while R. J. Cederberg held a National Research Council-NRL Research Associateship.

¹M. A. Leontovich and V. A. Fock, "Solution of the problem of propagation of electromagnetic waves along the earth's surface by the method of parabolic equation," *J. Exp. Theor. Phys.* **16**, 557–573 (1946).

²V. A. Fock, *Electromagnetic Diffraction and Propagation Problems* (Pergamon, New York, 1965), pp. 213–234.

³F. D. Tappert, "The parabolic approximation method," in *Wave Propagation and Underwater Acoustics*, edited by J. B. Keller and J. S. Papadakis, Lecture Notes in Physics **70** (Springer-Verlag, New York, 1977), pp. 224–280.

⁴F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, New York, 1994), pp. 343–412.

⁵M. D. Collins, "A split-step Padé solution for the parabolic equation method," *J. Acoust. Soc. Am.* **93**, 1736–1742 (1993).

⁶B. Engquist and A. Majda, "Absorbing boundary conditions for the numerical simulation of waves," *Math. Comput.* **31**, 629–651 (1977).

⁷R. W. Clayton and B. Engquist, "Absorbing boundary conditions for wave-equation migration," *Geophysics* **45**, 895–904 (1980).

⁸G. A. Kriegsmann, A. Taflove, and K. R. Umashankar, "A new formulation of electromagnetic wave scattering using an on-surface radiation boundary condition approach," *IEEE Trans. Antennas Propag.* **35**, 153–161 (1987).

⁹M. D. Collins and R. B. Evans, "A two-way parabolic equation for acoustic backscattering in the ocean," *J. Acoust. Soc. Am.* **91**, 1357–1368 (1992).

¹⁰M. D. Collins, "A self-starter for the parabolic equation method," *J. Acoust. Soc. Am.* **92**, 2069–2074 (1992).

¹¹M. D. Collins, "An energy-conserving parabolic equation for elastic media," *J. Acoust. Soc. Am.* **94**, 975–982 (1993).

¹²R. J. Cederberg and M. D. Collins, "Application of an improved self-starter to geoacoustic inversion," *IEEE J. Ocean Eng.* **22**, 102–109 (1997).

¹³J. F. Claerbout, *Fundamentals of Geophysical Data Processing* (McGraw-Hill, New York, 1976), pp. 206–207.

¹⁴R. R. Greene, "The rational approximation to the acoustic wave equation with bottom interaction," *J. Acoust. Soc. Am.* **76**, 1764–1773 (1984).

¹⁵A. Bamberger, B. Engquist, L. Halpern, and P. Joly, "Higher order paraxial wave equation approximations in heterogeneous media," *SIAM J. Appl. Math.* **48**, 129–154 (1988).

¹⁶M. D. Collins, "Applications and time-domain solution of higher-order parabolic equations in underwater acoustics," *J. Acoust. Soc. Am.* **86**, 1097–1102 (1989).

¹⁷M. B. Porter, F. B. Jensen, and C. M. Ferla, "The problem of energy conservation in one-way models," *J. Acoust. Soc. Am.* **89**, 1058–1067 (1991).

¹⁸M. D. Collins and E. K. Westwood, "A higher-order energy-conserving parabolic equation for range-dependent ocean depth, sound speed, and density," *J. Acoust. Soc. Am.* **89**, 1068–1075 (1991).

¹⁹R. B. Evans, "A coupled mode solution for acoustic propagation in a waveguide with stepwise depth variations of a penetrable bottom," *J. Acoust. Soc. Am.* **74**, 188–195 (1983).

²⁰M. B. Porter and E. L. Reiss, "A numerical method for ocean acoustic modes," *J. Acoust. Soc. Am.* **76**, 244–252 (1984).

²¹Reference 4, pp. 271–337.

²²F. R. DiNapoli, "Theoretical and numerical Green's function solution in a plane multilayered medium," *J. Acoust. Soc. Am.* **67**, 92–105 (1980).

²³H. Schmidt and F. B. Jensen, "A full wave solution for propagation in multilayered viscoelastic media with application to Gaussian beam reflection at fluid-solid interfaces," *J. Acoust. Soc. Am.* **77**, 813–825 (1985).

²⁴Reference 4, pp. 203–270.

²⁵F. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN* (Cambridge U. P., Cambridge, 1992), 2nd ed., pp. 194–196.

²⁶M. Porter, "The KRAKEN normal mode program," SACLANT Undersea Research Centre, La Spezia, Italy, SACLANTCEN SM-245, 1991.

²⁷M. D. Collins, "Higher-order Padé approximations for accurate and stable elastic parabolic equations with application to interface wave propagation," *J. Acoust. Soc. Am.* **89**, 1050–1057 (1991).

²⁸F. B. Jensen and C. M. Ferla, "Numerical solutions of range-dependent benchmark problems," *J. Acoust. Soc. Am.* **87**, 1499–1510 (1990).

²⁹M. D. Collins, W. A. Kuperman, and W. L. Siegmann, "A parabolic equation for poroelastic media," *J. Acoust. Soc. Am.* **98**, 1645–1656 (1995).

³⁰B. T. R. Wetton and G. H. Brooke, "One-way wave equations for seismoacoustic propagation in elastic waveguides," *J. Acoust. Soc. Am.* **87**, 624–632 (1990).

Prediction of the vibro-acoustic transmission loss of planar hose-pipe systems

M. L. Munjal

Department of Mechanical Engineering, Indian Institute of Science, Bangalore 560 012, India

P. T. Thawani

Climate Control Operations, Ford Motor Company, 15031 S. Commerce Drive, Dearborn, Michigan 48120

(Received 15 March 1996; revised 1 December 1996; accepted 6 January 1997)

Vibro-acoustic energy travels through hose walls as longitudinal waves and flexural waves, apart from the sound waves through the fluid medium inside. Longitudinal waves in the hose wall are coupled to the sound waves inside by means of the hose-wall Poisson's ratio. Both in turn get coupled to bending or flexural waves because of the energy transfer or interaction at the bends. For any of these three types of waves incident on one end of a hose, waves of all the three types may be transmitted on the other end because of their dynamical coupling with one another. Therefore, in the present paper, expressions have been derived for the 3×3 transmission loss matrix for a two-dimensional or planar piping system in terms of elements of the overall 8×8 transfer matrix of the system. These expressions have then been used in a comprehensive computer program to evaluate the vibro-acoustic performance of hoses, with particular application to the automotive climate control systems with gaseous as well as liquid media. Finally, parametric studies have been made that have led to some general design guidelines. © 1997 Acoustical Society of America. [S0001-4966(97)03105-6]

PACS numbers: 43.20.Mv, 43.20.Jr, 43.50.Gf [JEG]

INTRODUCTION

Rubber hoses have been used for several decades for transport of fluids so as to isolate vibrating machinery from accessories like exhaust and intake mufflers or receiver bottles/plenums. Hoses are capable of executing flexural vibrations and torsional vibrations like a beam or shaft. The hose wall can conduct oscillating longitudinal forces that may interact with waves in the fluid inside as well as outside. Flexural forces and moments get coupled to longitudinal forces and acoustic pressures because of bends in the piping system. If the fluid inside were a liquid, its inertia would accentuate the role of the bend.

Wiggert *et al.*¹ considered Poisson coupling in a pipe with two bends. Based on linearized assumptions and periodic motion, a simultaneous solution of the coupled fluid-structure system equations in waveform were presented recently by Lesmez *et al.*² They have derived a composite 14×14 transfer matrix consisting of submatrices for Poisson-coupled axial stress waves in the wall and pressure waves in the fluid column, flexural waves in the two normal planes, and torsional waves about the axis of the pipe or hose (see Appendix A for the transfer matrix relations and Fig. 1 for the forces and displacements). In their analysis, compliance or yielding of the hose wall is built into a modified bulk modulus for the fluid, and structural damping is accounted for by means of a modified bulk modulus.^{3,4} At higher-frequencies, however, inertia of the wall plays a significant role and must be accounted for by means of wall impedance. The concept of wall impedance, incidentally, enables one to predict transverse transmission loss and thence the breakout noise too.⁵

The concept of transmission loss (TL), which has been

used extensively as a measure of the performance of an acoustical filter⁶ or muffler, may also be applied to the structure-borne sound that travels in the form of longitudinal waves and flexural waves.⁷ TL of a one-dimensional acoustical filter has been expressed in terms of the elements of the overall transfer matrix of the filter in the literature (see Ref. 6 for example), but no such expressions are to be found for structure-borne sound. The present paper seeks to fill this gap.

Structure-borne longitudinal waves, flexural waves, and fluid-borne sound waves are dynamically coupled. Hence, a wave of any of these three types when incident on a planar hose-pipe system will in general result in reflection and transmission of all three types of waves. Thus, complete knowledge of the transmission characteristics of a hose system will in general be given by a 3×3 matrix of transmission loss (TL) values. In this paper, these TL values have been derived in terms of elements of the overall 8×8 matrix of the hose-pipe system, adapted from Ref. 2 (see Appendix A) wherein the effects of shear deformation, rotary inertia, and the fluid load have been incorporated. These expressions have then been used for parametric studies related to vibro-acoustics of hoses. Some general design guidelines based on the parametric studies have been arrived at for visco-elastic hoses, with or without bends, with gases or liquids as the media of propagation.

I. BASIC GOVERNING EQUATIONS

Longitudinal waves in the hose wall and one-dimensional sound waves in the medium inside are governed by the following coupled equations:

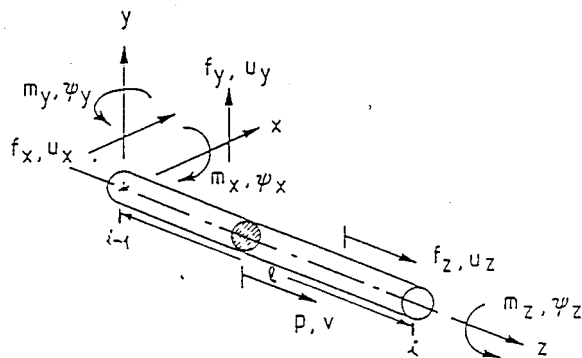


FIG. 1. A straight hose-pipe reach with state variables.

$$f_z - A_p \nu \frac{r}{e} p - EA_p \frac{\partial u_z}{\partial z} = 0, \quad (1)$$

$$\frac{\partial f_z}{\partial z} - \rho_p A_p \frac{\partial^2 u_z}{\partial t^2} = 0, \quad (2)$$

$$p + K^* \frac{\partial v}{\partial z} - 2\nu K^* \frac{\partial u_z}{\partial z} = 0, \quad (3)$$

$$\frac{\partial p}{\partial z} + \rho_f \frac{\partial^2 v}{\partial t^2} = 0. \quad (4)$$

Flexural (or bending) waves are governed by the following equations for a beam incorporating shear deformation, rotary inertia, and fluid inertia:

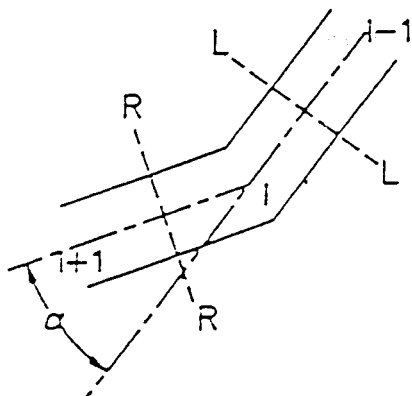


FIG. 2. A sharp bend.

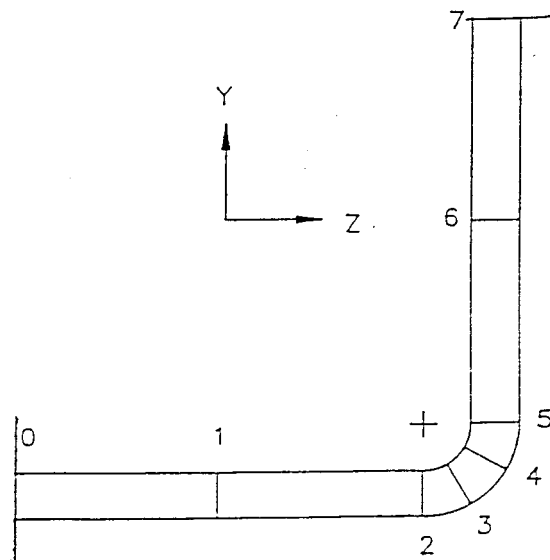


FIG. 3. A hose-pipe with a smooth bend.

$$\text{shear angle: } \beta_x = \frac{\partial u_y}{\partial z} + \psi_x, \quad (5)$$

$$\text{shear force: } f_y = GA_p k_s \beta_x, \quad (6)$$

$$\text{bending moment: } m_x = EI_p \frac{\partial \psi_x}{\partial z}, \quad (7)$$

$$\text{force equilibrium: } \frac{\partial f_y}{\partial z} - (\rho_p A_p + \rho_f A_f) \frac{\partial^2 u_y}{\partial t^2} = 0, \quad (8)$$

$$\text{moment equilibrium: } \frac{\partial m_x}{\partial z} - f_y - (\rho_p I_p + \rho_f I_f) \frac{\partial^2 \psi}{\partial t^2} = 0, \quad (9)$$

where f_z , u_z , p , v , m_x , ψ_x , f_y , and u_y are state variables as shown in Fig. 1. Other notations are as follows: A_p is the cross-sectional area of the pipe, ν is the Poisson's ratio of the pipe material, r is radius of the pipe cross section, e is the pipe wall thickness, E is Young's modulus (generally complex), z is the axial coordinate as shown in Fig. 1, ρ_p is density of the pipe material, ρ_f is density of the fluid inside the pipe, G is the shear modulus of the pipe material, k_s is the shear shape factor, I_p is the moment of inertia of the pipe, I_f is the moment of inertia of the fluid inside the pipe, K^* is the modified bulk modulus; for thin-walled pipes,

$$K^* = \frac{K}{1 + K(1 - \nu^2)r/Er}, \quad (10)$$

where K is the bulk modulus of the fluid.

II. EVALUATION OF THE TL MATRIX

For a one-dimensional hose-line system configured in the y - z plane (the plane of the paper), field matrices (A6) for straight pipes (Fig. 1) and point matrices (A8) of Appendix A for sharp bands (Fig. 2) may be multiplied successively as per configuration so as to obtain the overall or product transfer matrix relation connecting the state vector²

$$\{U_z \ P \ V \ F_z \ U_y \ \Psi_x \ M_x \ F_y\}^T \quad (11)$$

on the right-hand (or downstream) side with that on the left hand (or upstream) side (see Fig. 1). Here U , P , V , F , Ψ , and M denote complex amplitudes of displacement, acoustic pressure, acoustic particle displacement, force, flexural angular rotation (or slope), and bending moment, respectively. Field matrix of the constituent hose-pipe elements would be made up of the 4×4 submatrices T_{fp} and T_{yz} only. Before multiplying out, all transfer matrix elements have to be multiplied with the dimensional factors associated with the state variables (A7), so that state variables have their physical dimensions as in the state vector (11). It may be noted that for a smooth bend (Fig. 3) the bent portion will have to be partitioned and the angular changes will be lumped at sections 2, 3, 4, and 5 (15° , 30° , 30° , 15°) and the intermediate portions will be treated as curved pipes with modified flexural rigidity EI .²

In the absence of any bends, the resultant transfer matrix would consist of two uncoupled 4×4 submatrices, one representing the product of the individual T_{fp} matrices and the other of the T_{yz} matrices. Further, if we put Poisson's ratio ν hypothetically equal to zero in the T_{fp} matrix (that is, if we neglect Poisson coupling), then it would separate into two uncoupled 2×2 submatrices corresponding to longitudinal waves in the hose wall and acoustic waves in the fluid medium inside the hose. Thus,

$$\begin{bmatrix} U_z \\ F_z \end{bmatrix}_d = \begin{bmatrix} T_{11} & T_{14} \\ T_{41} & T_{44} \end{bmatrix} \begin{bmatrix} U_z \\ F_z \end{bmatrix}_u \quad (12)$$

$$\begin{bmatrix} P \\ V \end{bmatrix}_d = \begin{bmatrix} T_{22} & T_{23} \\ T_{32} & T_{33} \end{bmatrix} \begin{bmatrix} P \\ V \end{bmatrix}_u \quad (13)$$

and

$$\begin{bmatrix} U_y \\ \Psi_x \\ M_x \\ F_y \end{bmatrix}_d = \begin{bmatrix} T_{55} & T_{56} & T_{57} & T_{58} \\ T_{65} & T_{66} & T_{67} & T_{68} \\ T_{75} & T_{76} & T_{77} & T_{78} \\ T_{85} & T_{86} & T_{87} & T_{88} \end{bmatrix} \begin{bmatrix} U_y \\ \Psi_x \\ M_x \\ F_y \end{bmatrix}_u \quad (14)$$

where $[T]$ is the overall transfer matrix, and subscripts d and u denote downstream and upstream variables, respectively.

In general, and particularly in the presence of bends, one type of input (incident wave) will result in an output (transmitted wave) of not only the same kind, but also of other kinds. So, one should try to predict a set of nine TLs corresponding to three types of incident waves and three types of transmitted waves, making use of the entire 8×8 transfer matrix of a planar hose-pipe system [see Eqs. (46)–(52)].

As TL is a symmetric function for stationary medium, state vectors subscripted u and d may be interchanged without having to invert the overall transfer matrix $[T]$. Thus,

$$\begin{bmatrix} U_z \\ P \\ V \\ F_z \\ U_y \\ \psi_x \\ M_x \\ F_y \end{bmatrix}_u = \begin{bmatrix} 8 \times 8 \\ \text{overall transfer matrix} \\ \text{of the hose-pipe system} \\ T \end{bmatrix} \begin{bmatrix} U_z \\ P \\ V \\ F_z \\ U_y \\ \psi_x \\ M_x \\ F_y \end{bmatrix}_d \quad (15)$$

For an incoming or incident progressive wave, there will be one reflected wave and one transmitted wave in the case of sound waves and longitudinal waves (in the hose-pipe wall), and two reflected waves and two transmitted waves, one propagating and one evanescent (exponentially decaying) in the case of flexural waves. Thus, in general, there will be four reflected waves and four transmitted waves (anechoic termination is presumed on the downstream side as implied in the definition of transmission loss). Substituting these into the state variables in Eq. (15) would yield a set of eight inhomogeneous equations. These may be solved to obtain the three propagating type progressive waves (evanescent flexural wave carries no wave energy) in terms of the incident wave, and thence the corresponding three transmission loss values corresponding to this particular type of incident wave. The process can be repeated for the other two types of incident waves to obtain all nine elements of the desired 3×3 TL matrix.

In the following analysis, the (known) incident wave amplitude is denoted by A and the (unknown) reflected and transmitted wave amplitudes (complex in general) by B .

Substituting

$$(u_z, f_z, p, v) = (U_z, F_z, P, V) e^{\lambda z/l} e^{j\omega t} \quad (16)$$

into the homogeneous set of Eqs. (1)–(4), and applying the compatibility criterion, yields the characteristic equation

$$\lambda^4 + (\tau + \sigma + \gamma)\lambda^2 + \sigma\tau = 0, \quad (17)$$

where τ , σ , and γ are defined by the identities (A1a) to (A1c) of Appendix A. Equation (17) has roots

$$\lambda = \pm j\lambda_1, \pm j\lambda_2, \quad (18)$$

where λ_1 and λ_2 are given by Eqs. (A1l) and (A1m), and represent wave numbers for the longitudinal wave and sound wave, respectively. Therefore, for prediction of transmission loss of longitudinal waves and sound waves, the standing wave solutions may be written as

$$u_z = \{A_1 e^{-j\lambda_1 z/l} + B_1 e^{+j\lambda_1 z/l}\} e^{j\omega t}, \quad (19)$$

$$F_z = \{Q(-j\lambda_1)A_1 e^{-j\lambda_1 z/l} + Q(j\lambda_1)B_1 e^{j\lambda_1 z/l}\} e^{j\omega t}, \quad (20)$$

$$v = \{A_2 e^{-j\lambda_2 z/l} + B_2 e^{j\lambda_2 z/l}\} e^{j\omega t}, \quad (21)$$

$$p = \{H(-j\lambda_2)A_2 e^{-j\lambda_2 z/l} + H(j\lambda_2)B_2 e^{j\lambda_2 z/l}\} e^{j\omega t}, \quad (22)$$

where functions $Q(\beta)$ and $H(\beta)$ are given by Eqs. (2) and (4) as

$$Q(\beta) = -\frac{\rho_p A_p \omega^2}{\beta}, \quad \beta \equiv \frac{d}{dz} = \mp j\lambda_1/l = \mp jk_1, \quad (23)$$

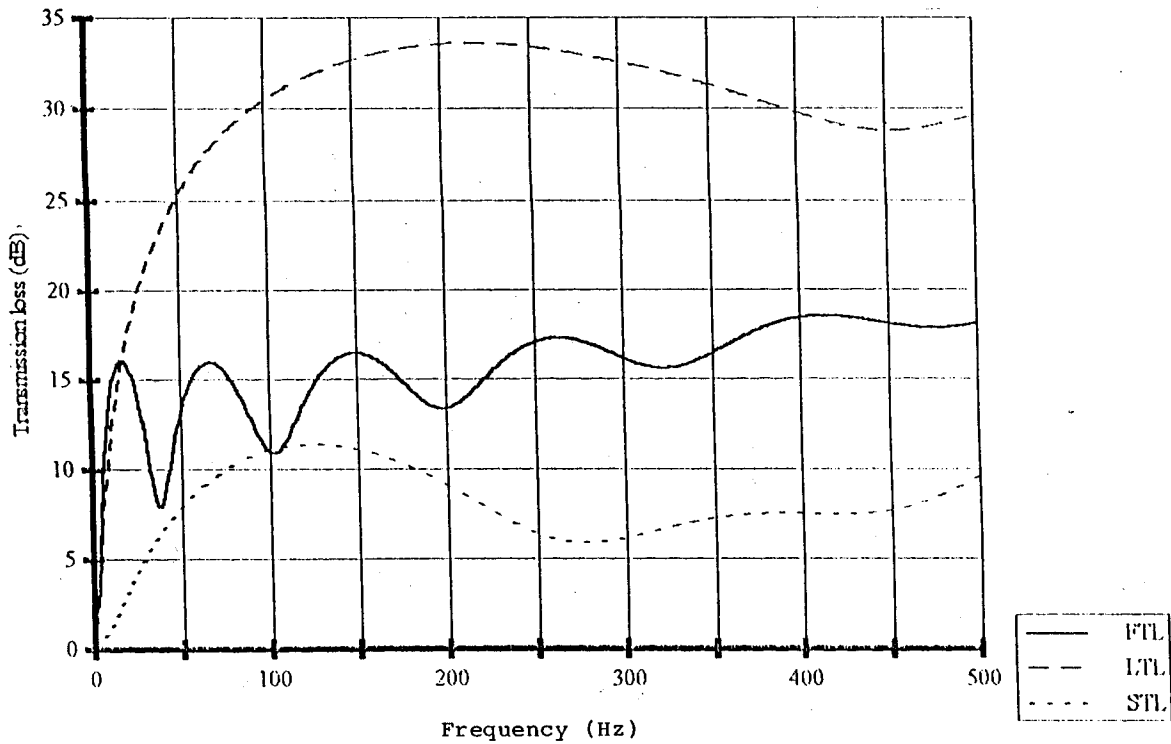


FIG. 4. Vibro-acoustic isolation of a composite rubber hose with liquid medium.

$$H(\beta) = \frac{\rho_f \omega^2}{\beta}, \quad \beta \equiv \frac{d}{dz} = \mp j\lambda_2/l = \mp jk_2. \quad (24)$$

Putting $z=0$ arbitrarily at the upstream as well as downstream end of the hose-pipe system, and assuming the downstream termination to be anechoic as required by definition of TL, one gets

$$U_{z,u} = A_1 + B_1, \quad U_{z,d} = B_2, \quad (25)$$

$$V_u = A_2 + B_3, \quad V_d = B_4, \quad (26)$$

$$P_u = H(-jk_{u,2})A_2 + H(jk_{u,2})B_3, \quad P_d = B_4 H(-jk_{d,2}), \quad (27)$$

$$F_{z,u} = Q(-jk_{u,1})A_1 + Q(jk_{u,1})B_1, \quad (28)$$

$$F_{z,d} = B_2 Q(-jk_{d,1}).$$

The corresponding relations for flexural waves may be obtained as follows.

Substituting

$$(u_y, \psi_x, m_x, f_y) = (U_y, \Psi_x, M_x, F_y) e^{\lambda z/l} e^{j\omega t} \quad (29)$$

into Eqs. (5)–(9), eliminating shear angle β_x , and applying the compatibility criterion, yields the characteristic equation

$$\lambda^4 + (\sigma + \tau)\lambda^2 - (\gamma - \sigma\tau) = 0, \quad (30)$$

where τ , σ , and γ are defined by identities (A3a) to (A3c) of Appendix A. Equation (28) has roots

$$\lambda = \pm \lambda_3, \pm j\lambda_4, \quad (31)$$

where λ_3 and λ_4 are given by Eqs. (A3i) and (A3j) of Appendix A, and represent evanescent waves and propagating waves, respectively. Therefore, for prediction of flexural

transmission loss, the general solutions may be written in terms of λ_3 and $j\lambda_4$ as a sum of four progressive waves (two evanescent and two propagating) similar to Eqs. (19) and (20).

Putting $z=0$ arbitrarily at the upstream as well as downstream end of the hose-pipe system, and assuming the downstream termination to be anechoic as required by definition of TL, yields

$$U_{y,u} = A_3 + B_5 + B_6, \quad U_{y,d} = B_7 + B_8, \quad (32)$$

$$\Psi_{x,u} = E(-jk_{u,4})A_3 + E(jk_{u,4})B_5 + E(k_{u,3})B_6, \quad (33)$$

$$\Psi_{x,d} = E(-jk_{d,4})B_7 + E(-k_{d,3})B_8,$$

$$M_{x,u} = F(-jk_{u,4})A_3 + F(jk_{u,4})B_5 + F(k_{u,3})B_6, \quad (34)$$

$$M_{x,d} = F(-jk_{d,4})B_7 + F(-k_{d,3})B_8,$$

$$F_{y,u} = G(-jk_{u,4})A_3 + G(jk_{u,4})B_5 + G(k_{u,3})B_6, \quad (35)$$

$$F_{y,d} = G(-jk_{d,4})B_7 + G(-k_{d,3})B_8,$$

where $E(\beta)$, $F(\beta)$, and $G(\beta)$ are calculated from Eqs. (5)–(9) as

$$E(\beta) = \frac{-\omega^2 C/\beta}{EI\beta^2 + D\omega^2}, \quad (36)$$

$$F(\beta) = -\frac{EI\omega^2/C}{EI\beta^2 + D\omega^2}, \quad (37)$$

$$G(\beta) = -C\omega^2/\beta, \quad (38)$$

$$\beta \equiv \frac{d}{dz} = \mp \lambda_3/l, \quad \mp j\lambda_4/l = \mp k_3, \mp jk_4, \quad (39)$$

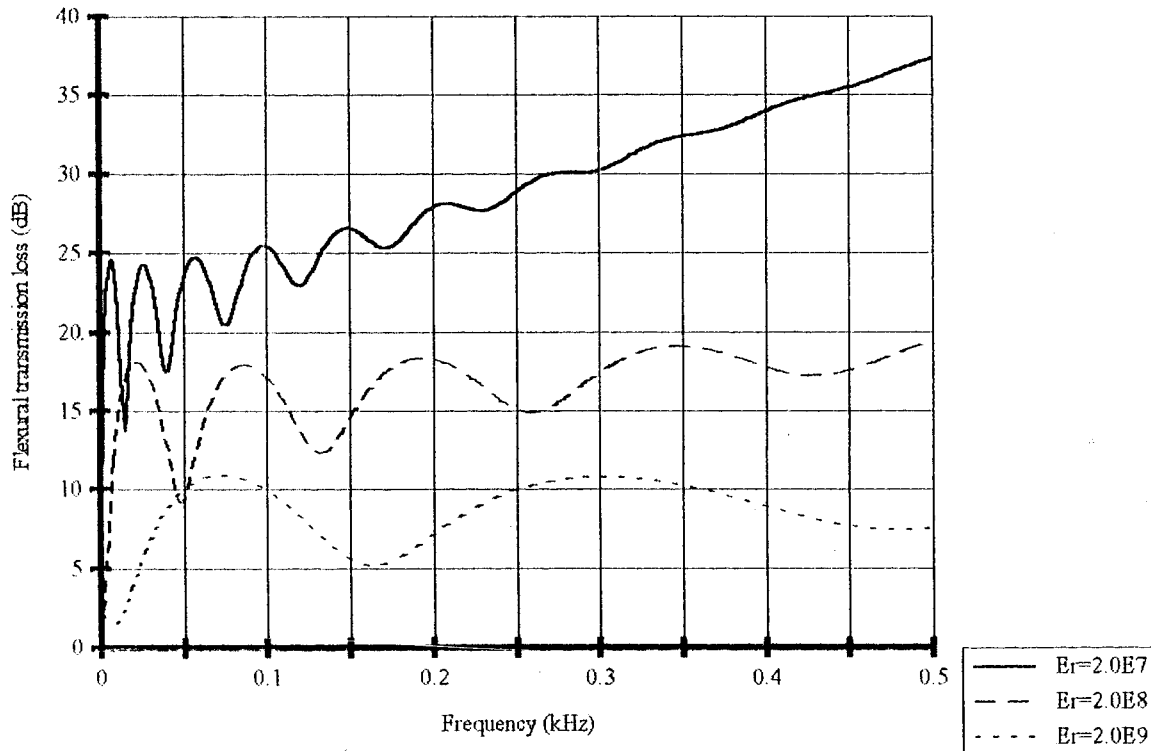


FIG. 5. Effect of the wall storage modulus of a hose on its vibration isolation.

$$C \equiv \rho_p A_p + \rho_f A_f, \quad (40)$$

$$D \equiv \rho_p I_p + \rho_f I_f. \quad (41)$$

Substituting Eqs. (25)–(28) and (32)–(35) into the transfer matrix equations (15), and rearranging, yields a set of eight inhomogeneous equations for the eight unknowns B_1 to B_8 :

$$[A]\{B\} = \{S\}, \quad (42)$$

where

$$\{B\} = [B_1 \ B_2 \ B_3 \ B_4 \ B_5 \ B_6 \ B_7 \ B_8]^T,$$

$$A_{11} = -1, \quad A_{23} = -H(jk_{u,2}), \quad A_{33} = -1,$$

$$A_{41} = -Q(jk_{u,1}),$$

$$A_{55} = -1, \quad A_{56} = -1, \quad A_{65} = -E(jk_{u,4}),$$

$$A_{66} = -E(k_{u,3}),$$

$$A_{75} = -F(jk_{u,4}), \quad A_{76} = -F(k_{u,3}),$$

$$A_{85} = -G(k_{u,4}), \quad A_{86} = -G(k_{u,3}),$$

$$A_{i2} = T_{i1} + Q(-jk_{d,1})T_{i4}, \quad i = 1, 2, \dots, 8,$$

$$A_{i4} = T_{i2} + Q(-jk_{d,2})T_{i3}, \quad i = 1, 2, \dots, 8,$$

$$A_{i7} = T_{i5} + E(-jk_{d,4})T_{i6} + F(-jk_{d,4})T_{i7} \\ + G(-jk_{d,4})T_{i8}, \quad i = 1, 2, \dots, 8.$$

The rest of the elements of matrix $[A]$ are equal to zero:

$$S_1 = A_1, \quad S_2 = H(-jk_{u,2})A_2, \quad S_3 = A_2,$$

$$S_4 = Q(-jk_{u,1})A_1,$$

$$S_5 = a_3, \quad S_6 = H(-jk_{u,4})A_3,$$

$$S_7 = F(-jk_{u,4})A_3, \quad S_8 = G(-jk_{u,4})A_3.$$

The matrix equations (42) are solved by means of a standard subroutine for vector $\{B\}$ for given (say unity) values of the incident waves A_1 , A_2 , and A_3 representing longitudinal, acoustic, and flexural waves, respectively.

Now, the expressions for power flux associated with different types of waves can be given by longitudinal waves:

$$\text{longitudinal waves: } W_l = \frac{\omega}{2} \text{Re}[F_z \cdot U_z^*], \quad (43)$$

$$\text{sound waves: } W_s = \frac{\omega}{2} A_f \text{Re}[P \cdot V^*] \quad (44)$$

$$\text{flexural waves: } W_f = \frac{\omega}{4} \text{Re}[F_y \cdot U_y^* + M_x \cdot \Psi_x^*], \quad (45)$$

where the superscript asterisk indicates complex conjugate.

Applying these expressions to incident waves and transmitted (propagating) waves of the three types, yields

$$W_{l,i} \equiv W_{l,i} = \frac{\omega}{2} \text{Re}[Q(-jk_{u,1})]|A_1|^2, \quad (46)$$

$$W_{l,t} \equiv W_{l,t} = \frac{\omega}{2} \text{Re}[Q(-jk_{d,1})]|B_2|^2, \quad (47)$$

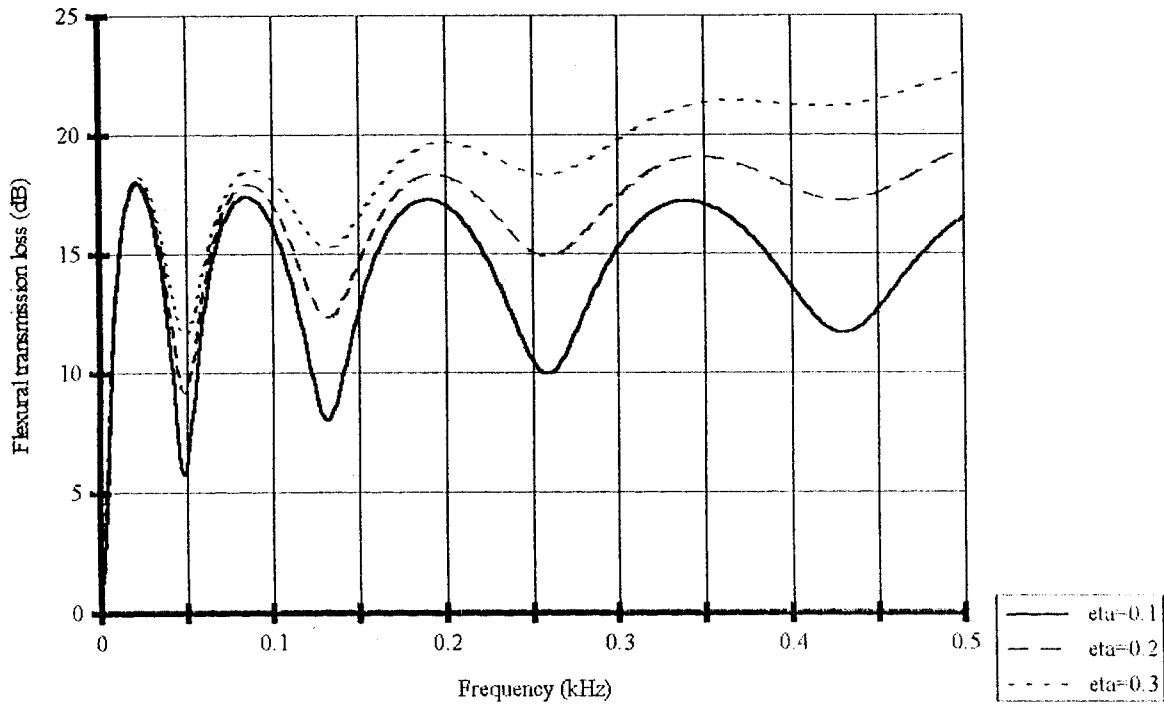


FIG. 6. Effect of the structural damping (loss factor) of a hose on its vibration isolation.

$$W_{2,i} \equiv W_{s,i} = \frac{\omega}{2} A_{f,u} \operatorname{Re}[H(-jk_{u,2})] |A_4|^2, \quad (48)$$

$$\text{TLM}(1,1) = \text{TL}_l \equiv \text{LTL}$$

(TL of longitudinal waves in the hose wall),

(53)

$$W_{2,t} \equiv W_{s,t} = \frac{\omega}{2} A_{f,d} \operatorname{Re}[H(-jk_{d,2})] |B_4|^2, \quad (49)$$

$$\text{TLM}(2,2) = \text{TL}_s \equiv \text{STL}$$

(TL of longitudinal waves in the fluid inside),

(54)

$$W_{3,i} \equiv W_{f,i} = \frac{\omega}{4} \operatorname{Re}[G(-jk_{u,4}) + F(-jk_{u,4}) \times E^*(-jk_{u,4})] |A_3|^2, \quad (50)$$

$$\text{TLM}(3,3) = \text{TL}_f \equiv \text{FTL} \quad (\text{TL of flexural waves}). \quad (55)$$

In the presence of bends, however, coupling terms of the TL matrix may also become relevant because of cross-mode energy transfer.

$$W_{3,t} \equiv W_{f,t} = \frac{\omega}{4} [G(-jk_{d,4}) + F(-jk_{d,4}) \times E^*(-jk_{d,4})] |B_7|^2. \quad (51)$$

Finally, elements of the 3×3 matrix of transmission loss may be calculated by means of the equation:

$$\text{TLM}(j1, j2) = 10 \log \frac{W_{j1,i}}{W_{j2,t}} \quad (\text{dB}), \quad j1, j2 = 1, 2, \text{ and } 3. \quad (52)$$

For straight hose-pipe systems, there is no coupling between flexural waves and the other two waves. Therefore, for systems without any bends, $\text{TLM}(1,3) = \text{TLM}(2,3) = \text{TLM}(3,1) = \text{TLM}(3,2) = 0$. Similarly, if one neglects Poisson coupling between longitudinal waves in the hose wall and the sound waves in the fluid inside, then $\text{TLM}(1,2) = \text{TLM}(2,1) = 0$. For most practical applications, only the diagonal elements of the TL matrix are of relevance. These are

III. TYPICAL RESULTS AND CONCLUSIONS

A comprehensive computer program has been prepared in FORTRAN for prediction of the transmission loss values for longitudinal waves, sound waves, and flexural waves (LTL, STL, and FTL), making use of the expressions derived above. Parametric studies were made for different materials and geometries as listed in Appendix B, for flexural waves (TL_f or FTL).

Typical results are shown in Fig. 4 for the default configuration of a 500-mm long, 19-mm nominal diameter, composite-rubber hose with steel-pipe terminations, but with a liquid medium (power-steering system oil: see Appendix B for specifications). Dips in the TL curves correspond to resonance frequencies that roughly correspond to $kl = n\pi$, $n = 1, 2, 3, \dots$, where

$$k = \frac{\omega}{(K^*/\rho_f)^{1/2}} \quad \text{for sound waves}, \quad (56)$$

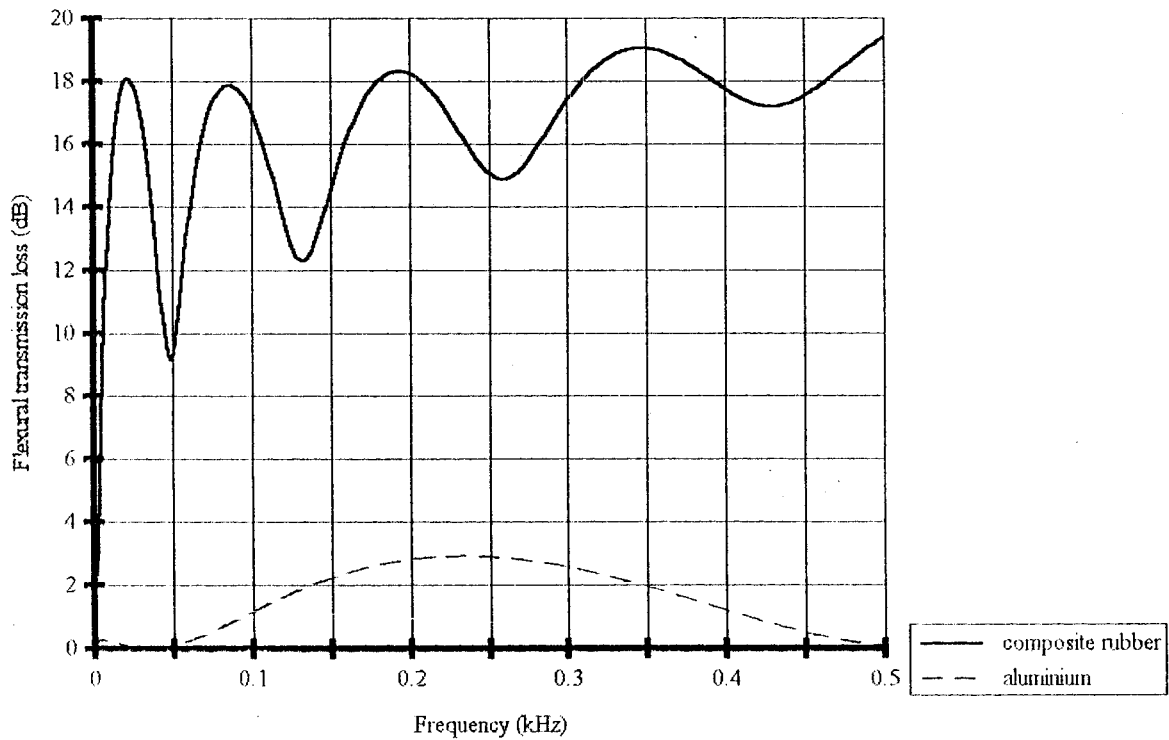


FIG. 7. Effect of the material of a hose on its vibration isolation.

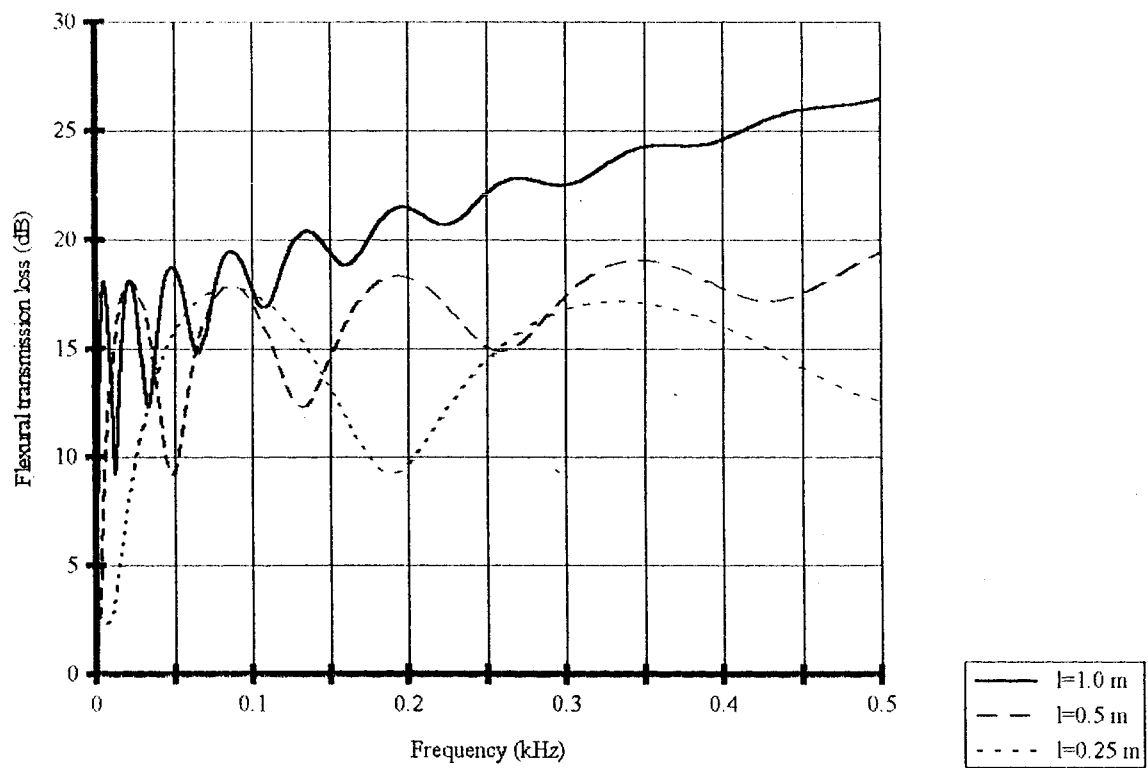


FIG. 8. Effect of the length of a hose on its vibration isolation.

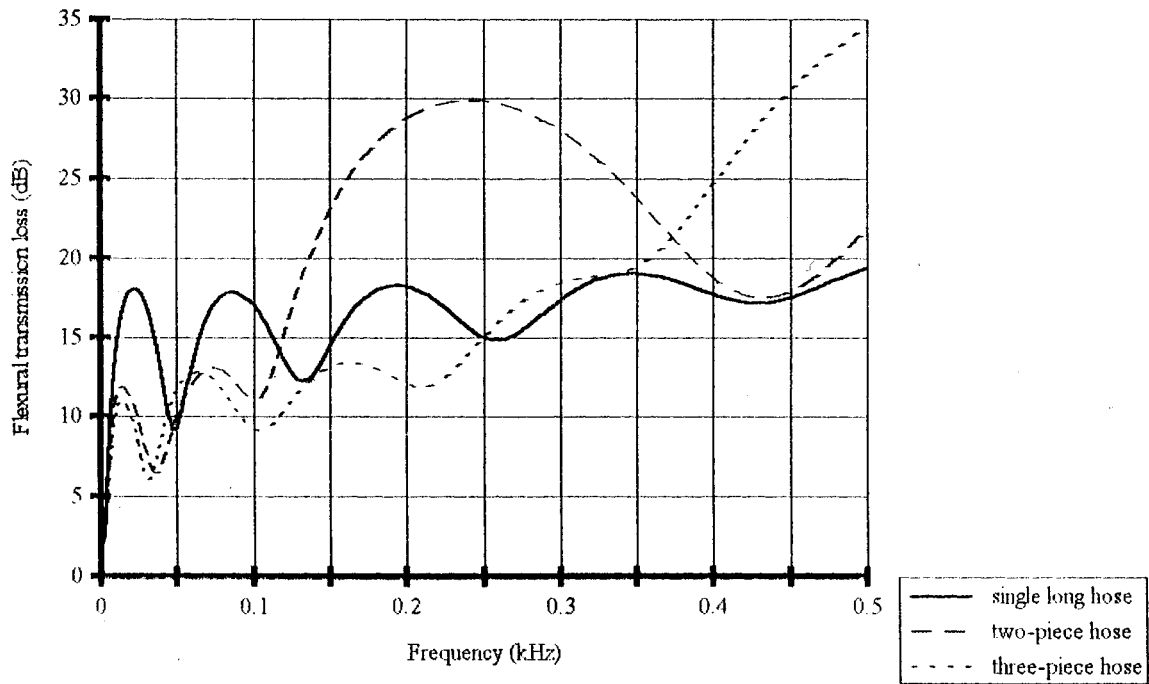


FIG. 9. Effect of the number of hose pieces within the same overall length on vibration isolation.

and

$$k = \frac{\omega}{(E_r/\rho_p)^{1/2}} \text{ for longitudinal waves.} \quad (57)$$

The observed values are within 20% of those given by the approximate relationships (56) and (57).

However, this simple relation ($kl = n\pi$) does not hold for flexural waves for which

$$k = \left(\frac{\rho_p A_p \omega^2}{E_r I_p} \right)^{1/4}. \quad (58)$$

If the medium were a gas, LTL and FTL curves would re-

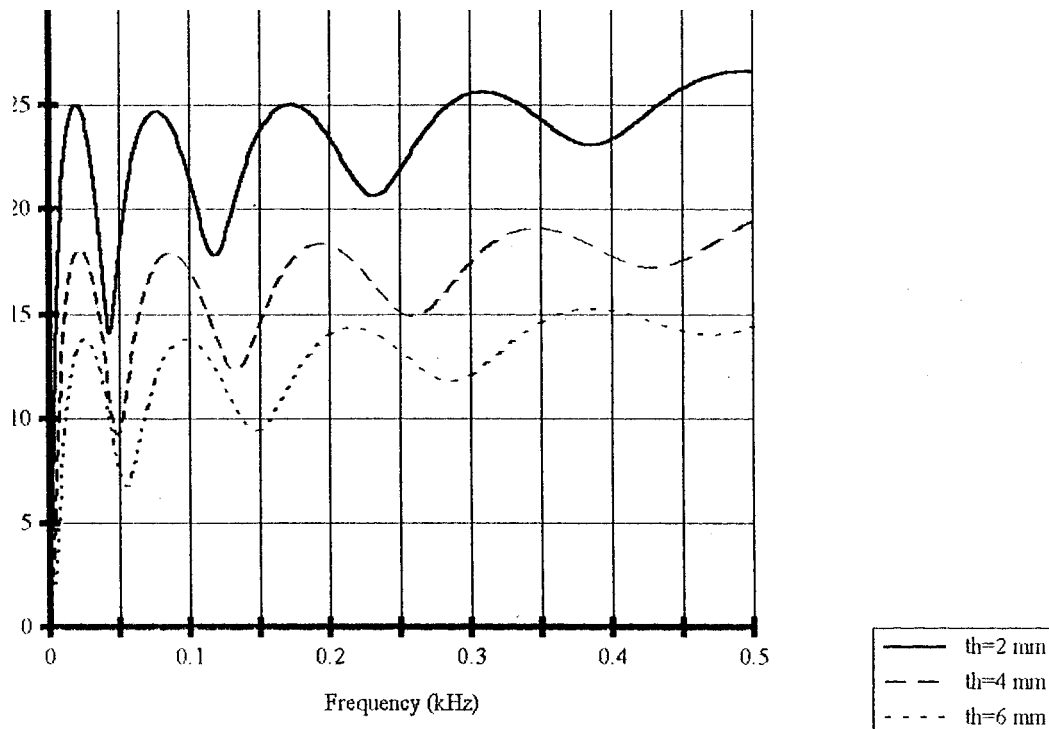


FIG. 10. Effect of wall thickness of a hose on its vibration isolation.

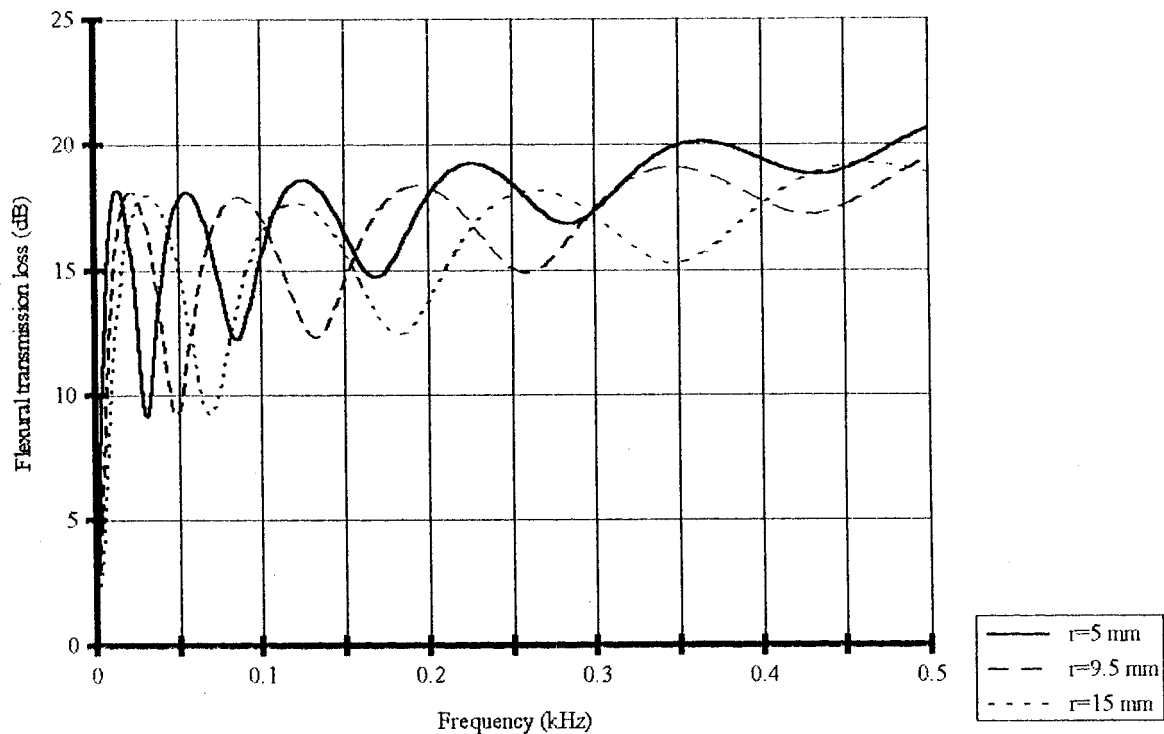


FIG. 11. Effect of the inner radius of a hose on its vibration isolation.

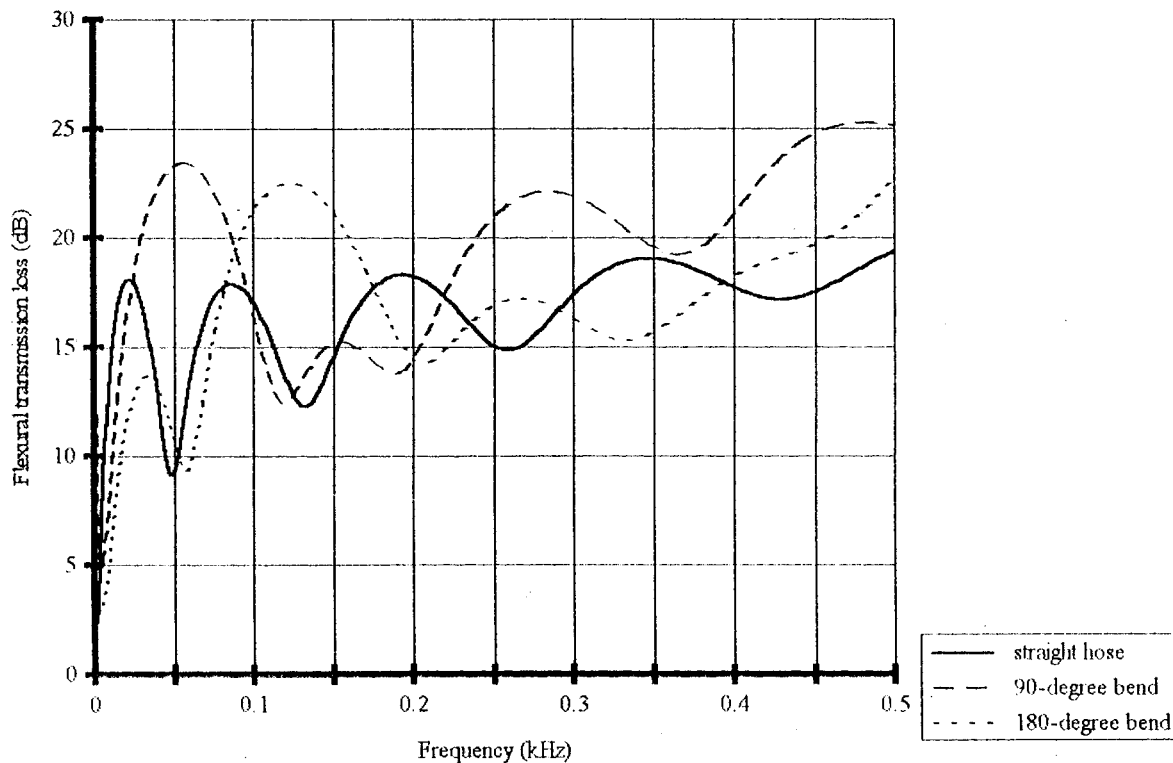


FIG. 12. Effect of bend angle on vibration isolation of a hose.

main more or less unchanged, but STL would reduce to zero (less than 0.5 dB). This is because longitudinal and flexural waves travel in the wall material whereas sound waves travel in the fluid medium inside. For air medium, impedance mismatch between air and hose material is so strong that the wall behaves more or less like a rigid boundary. For liquid medium, however, the impedance mismatch is rather weak.

Poisson coupling is not of much significance because its neglect (putting $\nu=0$ in the T_{fp} transfer matrix) alters LTL and STL by less than 0.5 dB even at 500 Hz; the difference is much less at lower frequencies. Therefore, use of uncoupled transfer matrices for evaluation of LTL and STL would suffice.

FTL values computed with the classical beam assumptions (neglecting shear deformation, rotary inertia, and fluid inertia) turn out to be substantially lower (by 4–5 dB) at higher frequencies (~ 500 Hz). Therefore, the exact theory (that is, without the classical beam assumptions) has been used here for the parametric studies. These studies for gaseous medium (refrigerant vapor) indicate that vibration isolation [represented by flexural transmission loss (FTL)] would improve with:

- (i) Softer hoses (lower storage modulus, E_r) as indicated by Fig. 5,
- (ii) lossier hoses (higher loss factor, η) as is obvious from Fig. 6,
- (iii) rubber hoses instead of aluminum as is clear from Fig. 7,
- (iv) longer hoses as is borne out by Fig. 8,
- (v) more (smaller) hoses in series with metallic (steel) joints as indicated by Fig. 9,
- (vi) thinner walls as is obvious from Fig. 10, and
- (vii) lower internal radius as is borne out by Fig. 11.

In particular, Fig. 9 shows that the effect of a multiple hose is similar to that of a multiple expansion chamber muffler: better at higher frequencies, but worse at lower frequencies.

Bends have a mixed but generally beneficial effect on vibration isolation (see Fig. 12), and therefore may be used as necessitated by logistics.

All these observations, viz., (i)–(vii), are in fact independent of the fluid medium inside the hose pipe, inasmuch as they have been found to hold for liquid medium well.

Finally, it may be noted that this paper has been concerned primarily with vibration isolation of hoses. For applications dealing primarily with sound waves along hoses or hose mufflers like power-steering systems,^{3,4,8,9} the theory developed in Ref. 5 with that of Ref. 6 as the base would be more useful inasmuch as it would help in evaluating the breakout noise or transverse transmission loss (TL_{tp}) that imposes a limit on effectiveness of axial sound TL (STL) realization.

ACKNOWLEDGMENTS

The work reported here was done at Ford Motor Company when the first author was on sabbatical leave from Indian Institute of Science, Bangalore. The authors thankfully acknowledge the support and permission to publish by the Climate Control Operations of the Automotive Components Division of Ford Motor Company.

APPENDIX A: TRANSFER MATRICES

The field transfer matrix of a beam element is composed of submatrices representing different types of waves. For Poisson coupled axial stress waves in the hose wall and pressure waves in the fluid inside, the nondimensional representation of the field transfer matrix and state vector are²

$$[T_{fp}] = \begin{bmatrix} \sigma C_2 - C_0 & \frac{\nu b}{h} [C_1 - (\sigma + \tau + \gamma) C_3] & \frac{\nu b}{h} \tau C_2 & -C_1 + (\sigma + \gamma) C_3 \\ 2\nu\sigma\tau C_3 & (\tau + \gamma) C_2 - C_0 & \tau[(\tau + \gamma) C_3 - C_1] & -2\nu\tau C_2 \\ 2\nu\sigma C_2 & \frac{1}{\tau} ((\tau + \gamma) C_1 - [(\tau + \gamma)^2 + \sigma\gamma] C_3) & (\tau + \gamma) C_2 - C_0 & 2\nu[(\sigma + \tau + \gamma) C_3 - C_1] \\ \sigma(C_1 - \sigma C_3) & -\frac{\nu b}{h} \sigma C_2 & -\frac{\nu b}{h} \sigma\tau C_3 & \sigma C_2 - C_0 \end{bmatrix}, \quad (A1)$$

where subscript fp indicates sound waves in the fluid and longitudinal waves along the pipe wall,

$$\tau = \frac{\omega^2 l^2}{a_f^2}, \quad (A1a)$$

$$\sigma = \frac{\omega^2 l^2}{a_p^2}, \quad (A1b)$$

$$\gamma = 2\nu^2 \frac{b\sigma}{d}, \quad (A1c)$$

$$b = \frac{\tau}{e}, \quad (A1d)$$

$$d = \frac{\rho_p}{\rho_f}, \quad (A1e)$$

$$h = \frac{E}{K^*}, \quad (A1f)$$

$$C_0 = \Delta[\lambda_2^2 \cos(\lambda_1) - \lambda_1^2 \cos(\lambda_2)], \quad (A1g)$$

$$C_1 = \Delta \left[\frac{\lambda_2^2}{\lambda_1} \sin(\lambda_1) - \frac{\lambda_1^2}{\lambda_2} \sin(\lambda_2) \right], \quad (\text{A1h})$$

$$C_2 = \Delta [\cos(\lambda_1) - \cos(\lambda_2)], \quad (\text{A1i})$$

$$C_3 = \Delta \left[\frac{1}{\lambda_1} \sin(\lambda_1) - \frac{1}{\lambda_2} \sin(\lambda_2) \right], \quad (\text{A1j})$$

$$\Delta = [\lambda_1^2 - \lambda_2^2]^{-1}, \quad (\text{A1k})$$

$$\lambda_1^2 = \frac{1}{2}((\tau + \sigma + \gamma) - [(\tau + \sigma + \gamma)^2 - 4\sigma\tau]^{1/2}), \quad (\text{A1l})$$

$$\lambda_2^2 = \frac{1}{2}((\tau + \sigma + \gamma) + [(\tau + \sigma + \gamma)^2 - 4\sigma\tau]^{1/2}), \quad (\text{A1m})$$

and the nondimensional state vector at location i (the right-hand end) in Fig. 1 is

$$Z_i = \left(\frac{U_z}{l} \frac{p}{K^*} \frac{V}{l} \frac{F_z}{A_p E} \right)_i^T. \quad (\text{A2})$$

For flexural vibration of the hose in the $y-z$ plane, the non-dimensional representation of the field transfer matrix and state vector are²

$$[T_{yz}] = \begin{bmatrix} C_0 - \sigma C_2 & -[C_1 - (\sigma + \tau)C_3] & -C_2 & -\frac{1}{\gamma}[-\sigma C_1 + (\gamma + \sigma^2)C_3] \\ -\gamma C_3 & C_0 - \tau C_2 & C_1 - \tau C_3 & C_2 \\ -\gamma C_2 & (\gamma + \tau^2)C_3 - \tau C_1 & C_0 - \tau C_2 & [C_1 - (\sigma + \tau)C_3] \\ -\gamma(C_1 - \sigma C_3) & -\gamma C_2 & \gamma C_3 & C_0 - \sigma C_2 \end{bmatrix}, \quad (\text{A3})$$

$$\sigma = \frac{(\rho_p A_p + \rho_f A_f)}{G A_p k_s} \omega^2 l^2, \quad (\text{A3a})$$

$$\tau = \frac{(\rho_p I_p + \rho_f I_f)}{E I_p} \omega^2 l^2, \quad (\text{A3b})$$

$$\gamma = \frac{(\rho_p A_p + \rho_f A_f)}{E I_p} \omega^2 l^4, \quad (\text{A3c})$$

$$C_0 = \Delta [\lambda_4^2 \cosh(\lambda_3) + \lambda_3^2 \cos(\lambda_4)], \quad (\text{A3d})$$

$$C_1 = \Delta \left[\frac{\lambda_4^2}{\lambda_3} \sinh(\lambda_3) + \frac{\lambda_3^2}{\lambda_4} \sin(\lambda_4) \right], \quad (\text{A3e})$$

$$C_2 = \Delta [\cosh(\lambda_3) - \cos(\lambda_4)], \quad (\text{A3f})$$

$$C_3 = \Delta \left[\frac{1}{\lambda_3} \sinh(\lambda_3) - \frac{1}{\lambda_4} \sin(\lambda_4) \right], \quad (\text{A3g})$$

$$\Delta = [\lambda_3^2 + \lambda_4^2]^{-1}, \quad (\text{A3h})$$

$$\lambda_3^2 = (\gamma + \frac{1}{4}(\sigma - \tau)^2)^{1/2} - \frac{1}{2}(\sigma + \tau), \quad (\text{A3i})$$

$$\lambda_4^2 = (\gamma + \frac{1}{4}(\sigma - \tau)^2)^{1/2} + \frac{1}{2}(\sigma + \tau). \quad (\text{A3j})$$

The state vector in the $y-z$ plane at location i (right-hand end) in Fig. 1 is

$$Z_i = \left(\frac{U_y}{l} \psi_x \frac{M_x l}{E I_x} \frac{F_y l^2}{E I_p} \right)_i^T. \quad (\text{A4})$$

1. General field transfer matrix

The field transfer matrix for a single straight pipe reach shown in Fig. 1 is composed of two submatrices: longitudinal vibration of the liquid and pipe wall and transverse vibration in the $y-z$ plane (neglecting torsion). Their expressions were given in Eqs. (A1) and (A3), respectively. The state vectors have eight dependent variables: two for each of the forces, moments, displacements, and rotations of the pipe wall, and one each for acoustic pressure and particle displacement of the fluid inside. The equation below shows these arrangements:

$$Z_i = [T_L] Z_{i-1}, \quad (\text{A5})$$

where $[T_L]$ is the field transfer matrix for a pipe reach of length l in the local coordinate system. This 8×8 matrix may be partitioned as²

$$[T_L] = \begin{bmatrix} [T_{fp}] & 0 \\ 0 & [T_{yz}] \end{bmatrix}. \quad (\text{A6})$$

The state vector at location i in Fig. 1 is

$$Z_i = \left(\frac{U_z}{l} \frac{P}{K^*} \frac{V}{l} \frac{F_z}{A_p E} \frac{U_y}{l} \Psi_x \frac{M_x l}{E I_p} \frac{F_y l^2}{E I_p} \right)^T. \quad (\text{A7})$$

2. Transfer matrix for a bend

The point transfer matrix relation for a sharp bend (see Fig. 2) is²

$$\begin{pmatrix} \frac{U_z}{l} \\ \frac{P}{K^*} \\ \frac{V}{l} \\ \frac{F_z}{A_p E} \\ \frac{U_y}{l} \\ \Psi_x \\ \frac{M_x l}{EI_p} \\ \frac{F_y l^2}{EI_p} \end{pmatrix}_i^R = \begin{bmatrix} \cos \alpha & 0 & 0 & 0 & \sin \alpha & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -(1 - \cos \alpha) & 0 & 1 & 0 & \sin \alpha & 0 & 0 & 0 \\ 0 & bq(1 - \cos \alpha) & 0 & \cos \alpha & 0 & 0 & 0 & \frac{1}{g} \sin \alpha \\ -\sin \alpha & 0 & 0 & 0 & \cos \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & gqb \sin \alpha & 0 & -g \sin \alpha & 0 & 0 & 0 & \cos \alpha \end{bmatrix}_i \begin{pmatrix} \frac{U_z}{l} \\ \frac{P}{K^*} \\ \frac{V}{l} \\ \frac{F_z}{A_p E} \\ \frac{U_y}{l} \\ \Psi_x \\ \frac{M_x l}{EI_p} \\ \frac{F_y l^2}{EI_p} \end{pmatrix}_i^L \quad (A8)$$

or

$$Z_i^R = [P_L^B]_i Z_i^L, \quad (A9)$$

where superscripts R , B , and L denote right, bend, and left respectively, and

$$g = \frac{A_p l^2}{I_p}, \quad (A9a)$$

$$q = \frac{A_f}{A_p}, \quad (A9b)$$

$$b = \frac{K^*}{E}. \quad (A9c)$$

APPENDIX B: HOSE DIMENSIONS AND MATERIAL PROPERTIES USED IN THE PARAMETRIC STUDIES

Inside fluid: default: HFC-134a refrigerant vapor
density = 22 kg/m³, bulk modulus = 4.436 × 10⁵ Pa
alternative: power-steering system oil
density = 834 kg/m³, bulk modulus = 1.5 × 10⁹ Pa

Hose material: default: 3/4 in. hose used in climate control system

Elastic modulus $E = E_r(1 + j\eta)$
Storage modulus,
 $E_r = 2.0 \times 10^8 (1 + f/1000)$, Pa
Loss factor, $\eta = 0.2(1 + 0.1f/1000)$
Poisson's ratio, $\nu = 0.48$
Internal radius, $r = 9.5$ mm
Wall thickness, $e = 4$ mm
Density, $\rho = 1196$ kg/m³
alternative: aluminum hose
Storage modulus, $E_r = 6.9 \times 10^{10}$ Pa
Loss factor, $\eta = 0.0001$

Poisson's ratio, $\nu = 0.33$

Density, $\rho = 2710$ kg/m³

Internal radius, $r = 8$ mm

Wall thickness, $e = 2$ mm

Terminal material: mild steel

Storage modulus, $E_r = 2.1 \times 10^{11}$ Pa

Loss factor, $\eta = 0.0001$

Poisson's ratio, $\nu = 0.29$

Density, $\rho = 7800$ kg/m³

Internal radius, $r =$ hose radius

Wall thickness, $e =$ hose wall thickness

Flexural transmission loss (FTL) has been selected for parametric studies on vibration isolation.

- ¹D. C. Wiggert, R. S. Otwell, and F. J. Hatfield, "The effect of elbow restraint on pressure transients," *ASME J. Fluids Eng.* **107**, 402–406 (1995).
- ²M. W. Lesmez, D. C. Wiggert, and F. J. Hatfield, "Modal analysis of vibrations in liquid-filled piping systems," *ASME J. Fluids Eng.* **112**, 311–318 (1990).
- ³L. Suo and E. B. Wylie, "Complex wave speed and hydraulic transients in viscoelastic pipes," *ASME J. Fluids Eng.* **112**, 496–500 (1990).
- ⁴M. C. Hastings and Chuan-Chiang Chen, "Analysis of tuning cables for reduction of fluid borne noise in automotive power steering hydraulic line," SAE paper No. 931295, Noise and Vibration Conference in Traverse City, 1994.
- ⁵M. L. Munjal and P. T. Thawani, "A simple model for wave propagation along and across a hose," *Proceedings of Inter-Noise 95* (1995), pp. 361–366.
- ⁶M. L. Munjal, *Acoustics of Ducts and Mufflers* (Wiley-Interscience, New York, 1987).
- ⁷L. Cremer, M. Heckl, and E. E. Ungar, *Structure-borne Sound* (Springer-Verlag, Berlin, 1988), 2nd ed.
- ⁸D. K. Longmore and A. Schlesinger, "Transmission of Vibration and pressure fluctuations through hydraulic hoses," *Proc. Inst. Mech. Eng.* **205**, 97–104 (1991).
- ⁹Jean Botti, G. Venizelos, and N. Benkaza, "Optimization of power steering systems vibration reduction in passenger cars," in *SAE Proceedings of the 1995 Noise and Vibration Conference*, pp. 1179–1201, paper No. 951253.

On the onset of transient cavitation in gassy liquids

L. Gaete-Garretón, Y. Vargas-Hernández, and R. Vargas-Herrera

Laboratorio de Ultrasonidos, Universidad de Santiago de Chile, Casilla 307, Santiago-2, Chile

J. A. Gallego-Juárez

Instituto de Acústica, Consejo Superior de Investigaciones Científicas, Serrano 144, Madrid 28006, Spain

F. Montoya-Vitini

Instituto de Física Aplicada, Consejo Superior de Investigaciones Científicas, Serrano 144, Madrid 28006, Spain

(Received 17 May 1995; revised 14 February 1996; accepted 18 December 1996)

This paper deals with a numerical and experimental study of ultrasound induced cavitation. On the basis of a computational analysis of Gilmore's equation for the individual bubble dynamics, a slight decrease in the pressure radiated by a single bubble immediately before reaching the transient cavitation regime is predicted. Experiments presented in this paper confirm this observation. Therefore, a new method to detect the inception of transient cavitation is proposed. The numerical analysis of a single bubble behavior, carried out for the different bubble sizes, suggests that two distinct "types" of transient cavitation may occur in a cavitation field in connection with the relative number of bubbles involved in the process. The experiments seem to confirm the theoretical hypothesis. The experimental work consisted mainly of insonifying distilled water, with a gas content of about 75% of saturation, at 20 kHz and analyzing the resulting acoustic emissions. © 1997 Acoustical Society of America. [S0001-4966(97)06304-2]

PACS numbers: 43.25.Yw [MAB]

INTRODUCTION

One of the most controversial points in relation to cavitation in a gassy liquid is the establishment of clear criteria to determine the real threshold for the onset of transient acoustic cavitation. The analysis of acoustic emissions from the cavitation field could be an adequate way to formulate a modern and objective procedure to detect the inception of cavitation. Nevertheless, the interpretation of the cavitation signals picked up from a liquid where bubbles of different sizes are present is not straightforward. Generally, the emission spectrum has a strong content of harmonics and subharmonics.^{1,2}

The idea, which has been suggested several times, is to link the frequency content of the emission with the cavitation state of the liquid. The nonlinear oscillation of a bubble in stable cavitation is a widely studied subject and frequency response curves have been given for different pressure amplitudes. However, when the bubble collapses, no steady-state response can be obtained. Therefore it becomes difficult to establish the relation between the transient cavitation threshold and the emission signals. Some authors have associated the threshold of transient cavitation with a sudden increase of the subharmonic signals.¹ Nevertheless, the subharmonic cannot be accepted as a clear feature for the onset of transient cavitation because it may also appear in stable cavitation.¹ In an interesting work, Lauterborn² proposed ultraharmonics of different kinds as the source of the first subharmonic.

In summary, the complex situation described above shows that, even though precise mathematical definitions about dynamical thresholds for cavitation exist,³ much work has to be done to establish clear relations between bubble

motions and measurable experimental observations. In addition, to study the cavitation field in a gassy liquid, the relations for a single bubble should be extended to a population of differently sized bubbles.

The present paper deals with a numerical and experimental study of cavitation in a gassy liquid with the purpose of defining the onset of transient cavitation and establishing practical methods for detecting the corresponding threshold. The numerical study is based on the Gilmore model for individual bubble dynamics,⁴ where the effects of compressibility of the liquid are included. The experimental study is focused on the collective behavior of different sized bubbles under the action of an acoustic field. Experimental results, in connection with numerical predictions, reveal the existence of two types of transient cavitation in a gassy liquid: a weak and a strong cavitation. A simple method to detect the inception of the strong transient cavitation has been formulated.

I. EQUATIONS AND NUMERICAL ANALYSIS

A suitable framework to study acoustic cavitation in water is furnished by the Gilmore equation describing the dynamics of a single cavity under an acoustic field:⁴

$$R \left(1 - \frac{U}{C} \right) \frac{d^2 R}{dt^2} + \frac{3}{2} \left(1 - \frac{U}{3C} \right) \left(\frac{dR}{dt} \right)^2 - \left(1 + \frac{U}{C} \right) H - \frac{U}{C} \left(1 - \frac{U}{C} \right) R \frac{dH}{dR} = 0, \quad (1)$$

where R is the bubble radius, $U = dR/dt$, C is the sound velocity, and H is the specific free enthalpy in the bubble surface. If the process is considered adiabatic, the expression for the free enthalpy will be

$$H = \frac{n}{n-1} \frac{A^{1/n}}{\rho_0} \left\{ \left[\left(P_0 2 \frac{s}{R_0} \right) \left(\frac{R_0}{R} \right)^{3\gamma} - 2 \frac{s}{R} + B \right]^{(n-1)/n} - [P_0 - P_n \sin(\omega t) + B] \right\} \quad (2)$$

and the sound velocity,

$$C = [C_0 + (n-1)H]^{1/2}, \quad (3)$$

where A , B , and n are constants (for water $A=3001$ atm, $B=3000$ atm, $n=7$), P_n is the pressure amplitude of the applied acoustic field, P_0 is the static pressure, R_0 is the equilibrium bubble radius, ρ_0 is the equilibrium density of the liquid, C_0 is the sound velocity in the unperturbed liquid, ω is the angular frequency, σ is the surface tension, and γ is the ratio of the specific heats. These equations take account of the compressibility of the medium, but viscosity and mass transfer are not included.

The numerical solution of Eqs. (1), (2), and (3), using the Runge–Kutta fourth-order algorithm, will determine the motion of the bubble and the conditions under which transient cavities will be produced. The values used for the various parameters are: $\sigma=75$ dyn/cm², $\gamma=4/3$, $\omega=2\pi \times 20\,000$ rad/s, $C_0=1550$ m/s, and $P_0=1.013 \times 10^5$ Pa. A typical calculation lasts for about seven acoustic periods.

The study of cavitation in a gassy liquid containing initially a wide distribution of small bubbles that serve as cavitation nuclei will be made considering the bubble resonance radius as a critical parameter. The resonance radius R_r , determined by the acoustic field frequency and the liquid characteristics, can be calculated with the following expression obtained from the linearized equations:⁴

$$\rho \omega_r^2 R_r^2 = 3\gamma \left(P_0 + 2 \frac{\sigma}{R_r} \right) - \frac{\sigma}{R_r}, \quad (4)$$

where ω is the frequency of the applied sound field, and ρ the liquid density. As an example, for an applied acoustic field of 20 kHz in water at 20 °C the resonance radius will be 0.016 cm.

Some results from the numerical solution of Eqs. (1)–(3) are shown in Fig. 1. The radiated sound pressure near the cavity wall for different bubble sizes was computed by using the equation of state for water in the form $P=A(\rho/\rho_0)^n-B$, and following the analysis presented in Ref. 4. It is displayed as a function of the applied sound pressure at a frequency of 20 kHz. It can be observed that the slopes of the curves suddenly change at a certain point, tending to infinity. The applied acoustic pressure which produces such a change for each bubble size can be identified as the transient cavitation threshold for the corresponding bubble. In fact, if we calculate the bubble wall velocity at this point, values of the order or even higher than sound velocity are obtained, which represents another indication for transient cavitation.⁵ As an example, Fig. 2 shows the wall velocity of a bubble of radius 0.004 cm. It is apparent that for an applied acoustic field of about 0.9 atm the wall velocity suddenly increases and tends to infinity. In such conditions, the bubble collapses and a shock wave will be formed.¹ It should be noted that immediately before the cavitation threshold pressure is reached, a

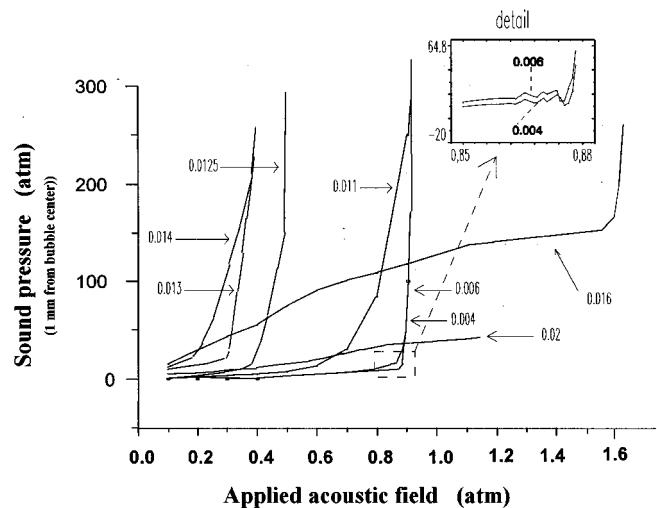


FIG. 1. Sound pressure near the bubble wall versus applied acoustic pressure at 20 kHz (from the Gilmore equation). The numbers attached to each curve represent the equilibrium bubble radii in cm.

small drop in the emitted sound pressure is observed. This effect is shown in the detail of Fig. 1 for two bubbles of 0.004 and 0.006 cm in radii, respectively.

Figure 3 presents the thresholds computed for single bubbles of radii below and above the resonance radius. The criterion used to produce this figure was to take the value at which the bubble wall velocity reaches the sound velocity value in the unperturbed liquid for the pressure threshold. It can be observed that bubbles with radii within the range 0.001–0.012 cm exhibit a similar threshold pressure for the onset of transient cavitation (ca. 0.9 atm). Therefore the population of bubbles included within this wide size range could be the cause of a great cavitation activity in a gassy liquid as they will act as nuclei to produce a simultaneous collective transient cavitation when the threshold pressure is reached. In this paper this type of cavitation is called “strong transient cavitation.” There is another characteristic group of bubbles, with radii just below the resonance radius, where the threshold pressure is different for each size of bubble and generally of a lower value than the threshold for the so-called

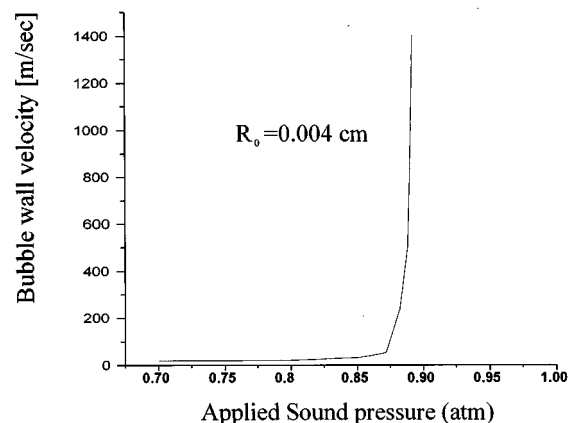


FIG. 2. Bubble wall velocity as a function of applied acoustic field for a bubble of $R_0=0.004$ cm (from the Gilmore equation).

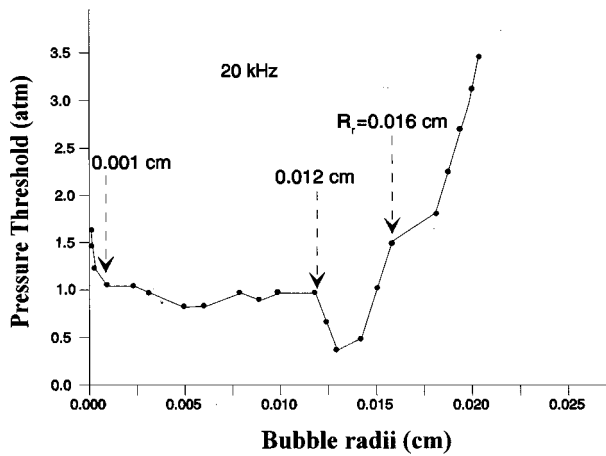


FIG. 3. Cavitation thresholds computed for single bubbles of different radii.

“strong cavitation.” The bubble radii of this group for the example here considered, range between 0.012 and 0.016 cm. At the lower threshold pressure (about 0.4 atm) the number of collapsing bubbles is small and increases if the applied pressure rises. Therefore, it seems evident that in a real liquid, where a relatively homogeneous distribution of bubbles of radii between 10^{-3} and 2×10^{-2} cm, can be expected,⁶ the cavitation activity produced by the collapsing bubbles within the 0.012–0.016 cm range is generally smaller than that produced by the collapsing bubbles within the 0.001–0.012 cm range. As a consequence, we believe we have identified a new kind of cavitation regime, the “weak transient cavitation” in a liquid. The threshold pressure for this type of cavitation is lower than that required for “strong cavitation.” Finally, there is a third group of bubbles, with radii above the resonance radius, which need a much higher applied acoustic pressure to produce cavitation.

In summary, in a gassy liquid with a wide bubble size distribution, two main cavitation processes are expected. At relatively low pressures (in the example here reported between approximately 0.3 and 0.8 atm) a weak transient cavitation will be produced. At a higher level (about 0.9 atm in this case) strong cavitation will occur because of the large number of transient cavities involved. The collective nature of this latter process allows the threshold to be well-defined. In this next section, experimental evidence of these events will be presented.

II. EXPERIMENTAL WORK

The experimental work consisted in the insonification at 20 kHz of distilled water with a certain air content and in measuring and analyzing the acoustic emission from the cavitation field. The experimental setup is shown schematically in Fig. 4. A piezoelectric piston source 30 mm in diameter radiates at 20 kHz in an anechoic cylindrical container (25 cm in diameter and 26 cm in height) filled with distilled gassy water (air content of about 75% saturated). The emission signal was picked up by a probe specially designed to work under high acoustic pressures in a cavitation environment.⁷ The applied power was controlled by an electronic system that keeps the output power constant during the

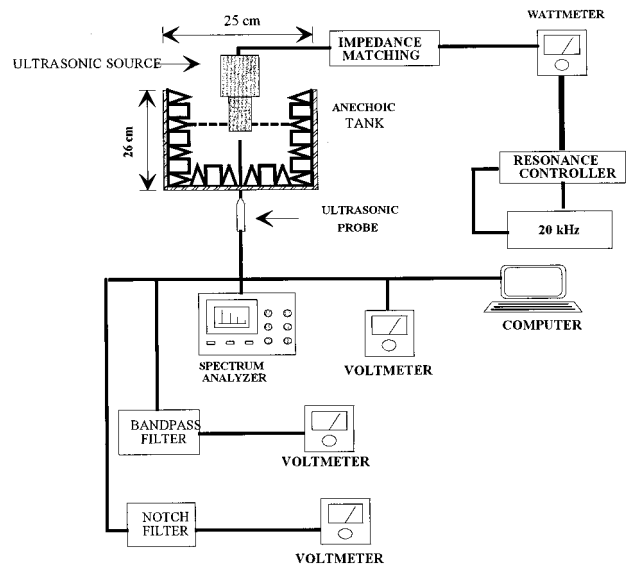


FIG. 4. Experimental setup to study the emission from a cavitation field.

experiment.⁸ The received signals were analyzed by using a superheterodyne spectrum analyzer, 40-Mhz bandwidth and 100-dB dynamic range. In the present measurements the operating band used was 10–200 kHz with a frequency resolution of 10 Hz and an amplitude accuracy of 0.2 dB.

Figure 5 shows the rms values of the total pressure received with the probe and measured with the voltmeter, versus the fundamental component of the same signal. It can be observed that for a pressure value of about 0.86 atm, the amplitude of the received signal shows a drop just before the onset of strong cavitation. Therefore it seems advisable to use this point, where the change of the slope of the curve is very clear, as the indication of the onset of cavitation. To that purpose, the excitation signal was filtered out from the received signal by introducing a notch-filter (6-dB cutoff frequency 2 kHz) (Fig. 4). As is well-known a notch-filter is an electronic device used to eliminate a narrow band in a wide-

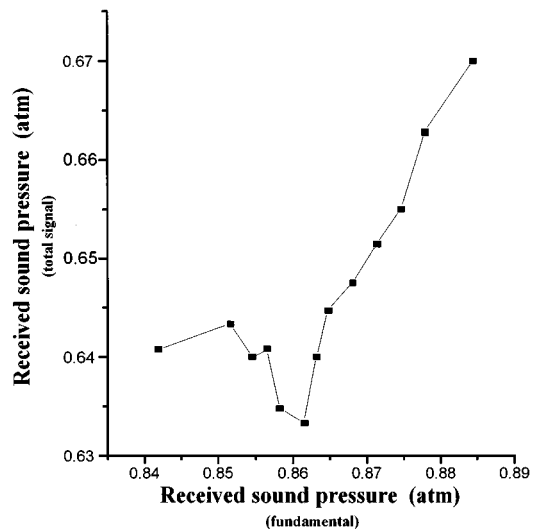


FIG. 5. Total sound pressure received from the cavitation field as a function of the fundamental signal.

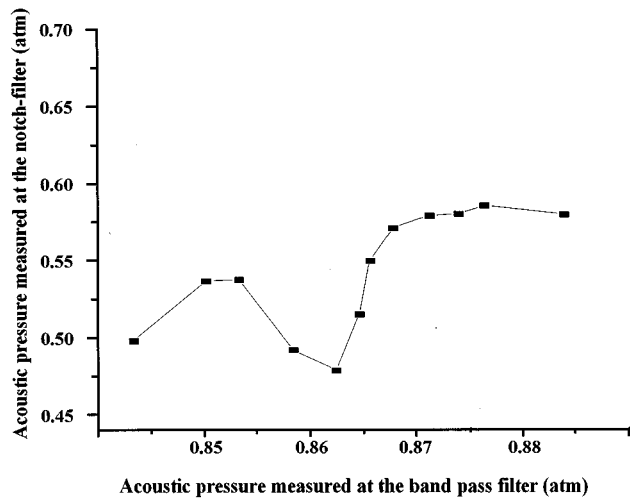


FIG. 6. Acoustic pressure measured at the notch-filter versus acoustic pressure measured at the bandpass filter.

band signal. In such a way the driving signal was removed from the emission spectrum. The fundamental component of the received signal was measured by using a bandpass filter (6-dB cutoff frequency 1 kHz) (Fig. 4). Representing the signal amplitude from the notch-filter output as a function of the signal amplitude from the bandpass filter output, the curve of Fig. 6 was obtained. It is evident that the curve drop is emphasized by this measuring procedure.

The onset of weak cavitation cannot be so well-defined because it depends on the bubble size and consequently is also linked to the collapse of a small number of cavities. Experimentally we have detected the onset of this cavitation by exposing a thin aluminum film to an acoustic field of about 0.34 atm. The photograph of Fig. 7(b) shows the weak erosion effect produced. The corresponding emission spectrum does not contain subharmonic components. A similar test carried out at an applied acoustic pressure of about 1 atm shows the strong erosion effect produced on the metallic film exposed to the cavitation field [Fig. 7(c)]. In this case, the received spectrum clearly shows subharmonic content.

III. DISCUSSION AND CONCLUSIONS

The results obtained numerically from the Gilmore's equation for a single cavity seem to be a good basis for explaining the production of acoustic cavitation in a gassy liquid. Nevertheless, there are a few points which still are not well-explained. For instance, the decrease that the signal amplitude received from the cavitation field suffers just before the onset of the strong cavitation. A possible interpretation of this behavior can be made in terms of the numerical results of the Gilmore equation. In the curves of Fig. 1, it was observed that the sound pressure emitted for each different size of bubble drops slightly just before reaching the threshold point (see the detail in Fig. 1). Therefore, as there are a large number of bubbles that show this effect for about the same level of acoustic pressure, the collective behavior could result in a more pronounced drop in the slope of the curve. The physical explanation of this effect is not so apparent, though

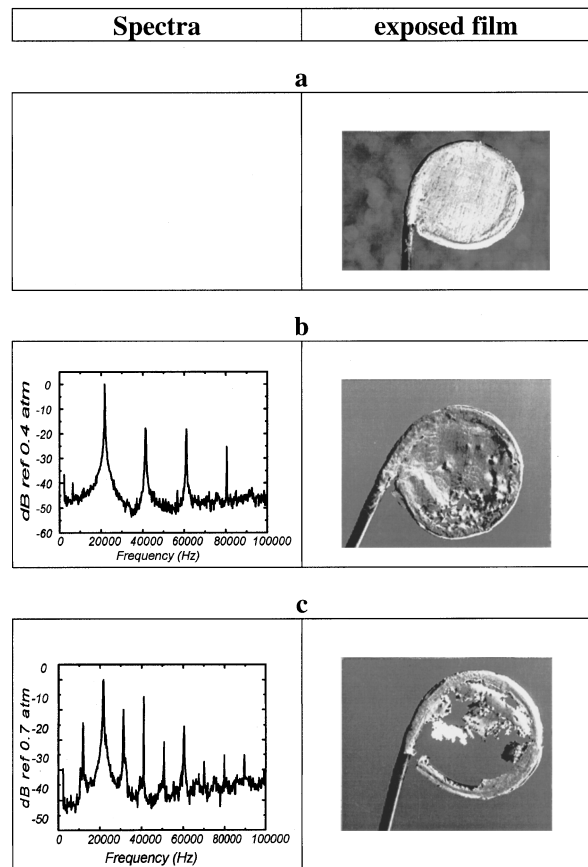


FIG. 7. Emission spectra from weak (b) and strong (c) transient cavitation fields and photographs of the corresponding erosion effects on aluminum films.

it could be attributed to an increase of the dissipative mechanism when the bubble motion approaches that of the transient cavity.

In conclusion, the present numerical and experimental study of transient cavitation in gassy liquids has shown the existence of two types of transient cavitation: Strong cavitation, which can be attributed to the collective behavior of a large number of bubbles, and weak cavitation, which occurs at lower acoustic pressures and seems to be produced by a restricted number of cavities.

ACKNOWLEDGMENT

We acknowledge the financial support from the European Commission, project CII*-CT92-0032, FONDECYT project 105-92, and CICYT project, TAP93-230.

- ¹E. A. Neppiras, "Acoustic Cavitation," *Phys. Rep.* **16**, 161–251 (1980).
- ²W. Lauterborn, "Numerical investigation of nonlinear oscillation of gas bubbles in liquids," *J. Acoust. Soc. Am.* **38**, 283–293 (1976).
- ³H. G. Flynn, "Cavitation dynamics. II. Free pulsation and models for cavitation bubbles," *J. Acoust. Soc. Am.* **58**, 1160–1170 (1975).
- ⁴V. A. Akulichev, "Pulsation of cavitation voids" in *High Intensity Ultrasonic Fields*, edited by L. D. Rozenberg (Plenum, New York, 1971), pp. 203–259.
- ⁵P. N. Vaughan and S. Leeman, "Acoustic cavitation revisited," *Acustica* **69**, 109–119 (1989).
- ⁶M. G. Sirotyuk, "Experimental investigations of ultrasonic cavitation," in

High Intensity Ultrasonic Fields, edited by L. D. Rozenberg (Plenum, New York, 1971), pp. 263–343.

⁷L. Gaete-Garretón, Y. Vargas Hernández, S. Pino Dubreuil, and F. Montoya-Vitini, “Ultrasonic detectors for high-intensity acoustic fields,”

Sens. Actuators A **37-38**, 410–414 (1993).

⁸A. Ramos Fernández, F. Montoya Vitini, and J. A. Gallego-Juárez, “Automatic system for dynamic control of resonance in high power and high Q ultrasonic transducers,” *Ultrasonics* **23**, 151–156 (1985).

Modeling the environmental influence on the vertical directionality of ambient noise in shallow water

T. C. Yang and Kwang Yoo

Naval Research Laboratory, Washington, DC 20375

(Received 8 April 1996; accepted for publication 4 November 1996)

The vertical directionality of acoustic ambient noise has been a subject of much interest in the past. It is a well-defined physical quantity that can be measured experimentally with a vertical array. It possesses certain deterministic features that can be modeled theoretically with environmental acoustic and source data. Ambient noise in shallow waters, including its vertical directionality, is not very well known and is also difficult to model/predict. This is because the acoustic environment varies with time and is location dependent. Thus arises the question, how does the vertical directionality of the ambient noise depend on the acoustic environments (found in typical coastal waters)? Due to the shallow water depth, it is noted that sound (noise) propagation can be significantly influenced by the bottom. The degree of bottom interaction will depend on the sound-speed profile in the water column: whether it is downward refractive or not. Bottom attenuation will in turn determine how far the sound will propagate in the water column. Using a modal representation, a closed-form expression is obtained which can be used to interpret and predict the distant noise vertical directionality as a function of the environmental acoustic parameters. The nearby (overhead) noise is separately modeled and is found less sensitive to environmental changes as the propagation distance is short. As the deterministic features of the noise vertical directionality is controlled by the sound propagation in the channel, it could be used as an acoustic indicator of the acoustic environment in the area. Time variations of the noise directionality will also be discussed. [S0001-4966(97)06104-3]

PACS numbers: 43.30.Nb, 43.30.Bp [MBP]

INTRODUCTION

Acoustic ambient noise in the ocean has been measured and studied over half a century.¹⁻⁴ Still many properties of the ambient noise are unknown and/or beyond our capability to model or predict. Experimentally we note that field measurements⁵⁻¹⁰ are limited by the response (or transfer) function of the sensors (arrays) used, such as the spectral bandwidth and sensitivity of the hydrophones, the array beam pattern, and angular resolution. Theoretically, we note that the ambient noise is a well-defined physical quantity that can be modeled and predicted if there are sufficient environmental acoustic and source data.¹¹⁻²⁵ Specifically, if the source level, radiation pattern, and the spatial/temporal distribution of the noise sources are known (in an ocean), then one can predict the ambient noise at any sensor location within an acoustic waveguide. (It is assumed that sufficient acoustic environmental information is given to model the sound propagation in this waveguide.) This is the underlying principle for noise modeling.

The noise properties of general interest are: the temporal and spatial (site) variation of the noise level, noise level as a function of sensor depth, temporal and spatial coherence of the noise, and vertical and horizontal directionality of the noise.

In reality, many of the inputs required to model the ambient noise are not known or poorly known. Of them, the least known may be the spatial and temporal distribution of the noise source distribution. Noise can be generated by many mechanisms, such as, by sea surface motion, by

surface-wave interactions, by bubbles generated by surface waves, by biological sources, and by ships transversing the oceans. The basic physics of some of the noise generation has been hypothesized but not yet fully verified experimentally.^{3,4} The noise source level and radiation pattern for some processes have been experimentally measured, the data are, however, limited.²⁶⁻²⁸ For this reason, predicting the noise level as a function of time at a given site is a hard and almost impossible task. In reality, one has to settle with a long-term average, e.g., seasonal variability, or a statistical description of the noise level.

Nonetheless, certain properties of the ambient noise can be modeled and predicted with high reliability. These are the properties which are based on the physics of sound propagation and certain characteristics of the noise sources. For example, the azimuthal directionality of the noise beams will peak in the directions of major shipping lanes at frequencies where shipping noise is dominant or in the directions of major storms when wind-induced noise is dominant. At low frequencies (e.g., <1 kHz) the noise source can often be treated as a point source, which in conjunction with the sea surface, produces effectively a dipole source.^{11,14,16} In this case, the vertical directionality of the ambient noise is primarily controlled by the physics of sound propagation in the waveguide,²⁵ the noise source level is not required. In a lossy shallow-water waveguide, the noise vertical coherence length is determined by the critical angle of bottom reflection at low frequencies.²² In the Arctic ocean, the vertical coherence length of the noise is affected by the water depth.²³ In

both cases, the vertical coherence length of the noise controls the span of the vertical arrival angles of a surface-generated noise source.

There can be various levels of sophistication in noise modeling. The most sophisticated noise model requires detailed temporal and spatial distributions of noise sources and could in principle predict everything about the noise. A less sophisticated noise model forgets about the time variation and only attempts to model the spatial properties of the ambient noise. This is a stationary noise model which is used to explore the properties of the long-term averaged ambient noise. In this category, one can use a distribution of noise sources based on real information (e.g., the location of a major storm) or on an historical survey (e.g., average density of ships over shipping lanes) or on an educated guess (e.g., distribution of biological species) or one can simply assume a uniform distribution of noise sources over the (entire) ocean surface.

There are various approaches to modeling the noise vertical directionality. Many noise models employ numerical integration of noise beam power from all noise sources.^{19,20} This type of model will be referred to as the numerical noise model. All that is required for this model is a propagation code given the noise source function. One notes that the numerical noise model can handle a complicated noise source distribution in a range-dependent environment. It can in principle also be used to model the time variation of the noise field.²¹ However, the numerical noise model is computationally intensive. Its use in practical applications is limited by the computation time required. The physics of noise propagation is not apparent from the numerical integration approach.

There exists a closed-form expression for noise covariance/coherence function by Kuperman and Ingenito,¹⁴ assuming a uniform distribution of surface noise sources, which can be used to calculate the noise directionality in shallow water.¹⁶ This model is limited to a range-independent environment and stationary noise. The closed-form expression can be used to calculate the noise vertical directionality.¹⁸ The extension of Kuperman and Ingenito's work to a three-dimensional environment assuming adiabatic normal mode approximation leads to a complicated formula for the noise coherence.²⁴ For a review of these works, the reader is referred to Ref. 25.

Previous studies on ambient noise have concentrated on noise directionality in deep water.¹⁻²⁵ Ambient noise in shallow waters, including its vertical directionality, is not very well known and is difficult to model/predict. This is because the sound speed in shallow water and the bottom properties can vary substantially with the location.²⁹ The former can vary significantly with time. Thus arises the question, how does the vertical directionality of the ambient noise depend on the acoustic environments (found in typical coastal waters)? We note that due to the shallow-water depth, sound (noise) propagation can be significantly influenced by the bottom. The degree of bottom interaction will depend on the sound-speed profile.

For example, previous works^{10,16} have shown the existence of a notch in the noise vertical directionality in shallow

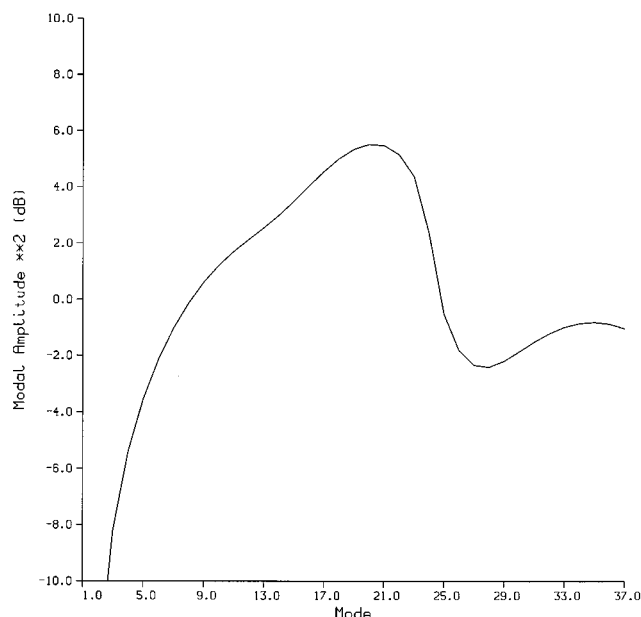


FIG. 1. The noise modal intensity n_m at 500 Hz for a constant sound-speed profile with a soft bottom.

water with a downward refractive sound-speed profile. The question “When will the ambient noise vertical directionality exhibit a notch?” has been asked by many people. Can one predict it? Is there a rule of thumb? To address this question, one could use a numerical computer program, such as RANDI,^{18,19} to model the noise vertical directionality and check if there exists a notch for each given environment. It can take hours to make one run for a given shallow-water environment on a UNIX work station depending on the frequency. Each time one changes the array configuration, one needs to run it again. There is no physical interpretation provided. There is no rule provided which enables one to predict if the noise notch exists. One can also use the Kuperman–Ingenito model. But how does one physically interpret the noise coherence formula and relate to the acoustic environment? Can one predict the existence of the noise notch based on the vertical coherence length? The answer is probably no.

In this paper we present a modal analysis of the ambient noise in shallow water. This noise model provides a physical interpretation of the noise directionality and a capability to predict the noise vertical directionality. It sheds light on the questions whether the noise notch exists for a given acoustic waveguide and when to expect one.

The noise vertical directionality is expressed in terms of the noise modal intensities multiplied by the modal beam patterns of the vertical array. The latter represents the array response to the modal noise expressed in terms of the array beam pattern. A closed-form expression is given for the noise modal intensities associated with distant noise. The nearby (overhead) noise is separately modeled and is found less sensitive to environmental changes as the propagation distance is short.

For an environment, the modal wave numbers and attenuation coefficients calculated by a normal mode program, e.g., KRAKEN,³⁰ are used to calculate the noise modal intensities. The noise modal intensities (see Fig. 1) determines if a

notch exists in the noise vertical directionality for a given environment. (The actual noise directionality can be obtained from Fig. 1 with a translation of coordinates.) Figure 1 reveals how the noise vertical directionality is affected by the sound-speed profile in the water column (i.e., mode excitation by noise sources) and by the sound interaction with the bottom (modal attenuation due the bottom absorption).

The modeling of the noise vertical directionality is quick using the closed form expression. If one changes the array configuration, one only has to fold in a different beam pattern (without making another run). This feature could be useful to the system design people in their study of various system response to the ambient noise. One notes that different vertical directionality distributions are obtained in the same area by different vertical arrays at different frequencies. When the sound speeds change, one finds that the noise vertical directionality also changes. All these features can be understood by tracking the modal arrival angles and modal beam patterns. One can obtain the same answer using a numerical noise model or using the Kuperman–Ingenito model. However, no clear physical interpretation is given by either model on what is happening.

The notion of modal beam patterns of the noise is new.²³ What is the value of this? One notes that the measured noise directionality is the convolution of the array beam pattern with the true noise directionality. Methods to deconvolve the beam pattern from the received noise data are often ill conditioned as there exists more degrees of freedom in the noise field than the number of hydrophones in the vertical array. Thus mathematically the deconvolution solution is not unique. A meaningful solution can be obtained if one assumes a normal mode model for noise propagation. The (inverse) problem is, however, complicated due to the contribution of the discrete versus the continuum if the deconvolution is done by brute force. The modal representation to be described below provides a framework to sort out/deduce the noise modal intensities which represents the true noise directionality.

Another advantage of the modal representation is that it can be used to reduce the field data of noise vertical directionality to a set of coefficients as shall be discussed later in Sec. IV.

The objective of this paper is to present the modal representation of noise vertical directionality and use it to model the environmental acoustic influence on the noise vertical directionality. For this purpose, we shall concentrate on range-independent environments. But as discussed below, this model can also be applied to a range-dependent environment where an adiabatic normal model is valid. It can also be used to model the time variation in the noise field. Other applications of this approach will be presented elsewhere. In Ref. 31, it was pointed out that the effects of the internal waves on the noise vertical directionality can now be modeled in a reasonable time (hours) using the modal representation presented in this paper. If one were to model the noise vertical directionality in an internal wave field by numerical integration over the noise sources, it would take one-half to one month on an 8 processor IRIS work station for 100 realizations of the internal wave field.

The noise model is presented in Sec. I. The noise model assumes that the modal covariance of the noise is diagonal. With this assumption, the noise vertical directionality for the distant noise can be expressed in a simple closed form. In Sec. II, we conduct a comparison study of the noise vertical directionality predicted by this noise model with that of a numerical noise model in which the noise beam power from different noise sources are numerically summed. The closed-form solution results in more than order of magnitude reduction in computation time compared with the numerical noise model. The closed form expression of the noise vertical directionality allows one to investigate and physically interpret the effects of propagation under various acoustic conditions. This will be covered in Sec. III. In Sec. IV, we illustrate how the model can be generalized to include time varying and range-dependent noise sources. Section V contains the summary and concluding remarks.

I. VERTICAL DIRECTIONALITY OF THE AMBIENT NOISE

A. Ambient noise model

Ambient noise sources, such as wind-generated noise and ship noise, are assumed to be (or approximated by) monopoles close to the ocean surface at frequencies below 1 kHz. Due to the reflection from the pressure release surface, these monopoles produce effectively a dipole radiation pattern. Experimentally, the dipole radiation pattern has been observed for wind-generated noise at 200–4000 Hz frequencies.^{9,28} At high frequencies, the noise sources may have complicated radiation patterns. For example, the radiation pattern for oscillating bubbles trapped in a surface wave will depend on the bubble structure and the wave shape.

Our model will be presented in the frequency domain. Assuming monopole sources, the pressure field for a receiver located at the center of the coordinate system, i.e., $r=0$, at a depth z_j can be written as the sum of pressure fields from individual sources located at (\mathbf{r}_l, z_0) ,

$$p(z_j) = \sum_l S(\mathbf{r}_l) G(r_l, z_0, z_j) e^{i\phi_l}, \quad (1)$$

where $G(r_l, z_0, z_j)$ is the Green's function from the source to the receiver and S is the noise source spectral amplitude at a fixed frequency. The phase of the noise source has been explicitly denoted as ϕ_l .

The vertical directionality of the ambient noise can be measured using a vertical array of hydrophones at depth z_j , $j = 1, \dots, N$. The beam power for a given angle θ can be obtained using

$$B(\theta) = \left| \sum_j p(z_j) e^{-ikz_j \sin \theta} \right|^2, \quad (2)$$

where θ represents the grazing angle. It is commonly assumed that noise sources are random and incoherent, hence, one has

$$B(\theta) = \sum_l S^2(\mathbf{r}_l) \left| \sum_j G(r_l, z_0, z_j) e^{-ikz_j \sin \theta} \right|^2. \quad (3)$$

The distributed noise sources are both time and space dependent. In Eqs. (1) and (3), the time dependence is implicit; it can be viewed as a snapshot representation of (a fixed frequency component of) the noise field at a certain time. We note that a vertical array cannot determine the azimuthal distribution of the noise field. Thus we shall divide the ocean surface into circular rings around the origin. Denoting $S(r)$ as the azimuthally averaged noise source spectral density (per unit area) at range r , i.e., $S(r) = \int S(r, \theta) d\theta/2\pi$, one has

$$B(\theta) = \int 2\pi r S^2(r) \left| \sum_j G(r, z_0, z_j) e^{-ikz_j \sin \theta} \right|^2 dr, \quad (4)$$

where for convenience we have replaced summation by integration.

We note that Eqs. (1)–(4) can also be used to represent a time-averaged noise field. The ambient noise is normally assumed to be stationary, i.e., the noise has well-defined statistical behavior. Thus, averaging the noise directionality over many time samples is equivalent to averaging the noise source spectral level at each location over a long time period. One expects over a basin area the noise source level, averaged over a long period of time, shall be location independent since there will be equal probability for noise to be generated at any place (unless there is an environmental reason why it should not be so). Hence in many noise models noise sources are assumed to be uniformly distributed over the surface.

The range integration in Eq. (4) can be divided into three ranges; $0-r_1$, r_1-r_2 , and $r_2-\infty$. For ranges less than r_1 (hundreds of meters), the noise is coming directly above the vertical array. The vertical arrival angles of this “overhead noise” can be straightforwardly estimated using a ray acoustic and/or a full-field propagation model. The discussion of the “overhead noise” will be deferred to Sec. II. We shall concentrate (for the remainder of this section) on the noise sources at ranges $>r_1$, which are referred to as distant noise sources. The noise originating from distant noise sources will be called “distant noise” to contrast with “overhead noise.”

For source ranges $>r_1$, we assume that the Green’s function can be represented in terms of (discrete) normal modes, since the continuum contribution dies out with range. The range r_2 is used to denote the size of the ambient noise source area, i.e., r_2 is defined by either the basin boundary or the range beyond which the contribution is negligible compared with that from ranges $<r_2$. As seen below, the size of the noise source area depends on the modal attenuation with range and is an environment-dependent parameter.

Given a propagation model, Eq. (4) can be used to calculate numerically the vertical directionality of the ambient noise. This was referred to earlier as the numerical noise model. The numerical noise model is computationally intensive and generally takes several hours to run on a UNIX work station for an ocean of 10-km radius or more. The computation of the vertical directionality can be speeded up by an order of magnitude using the modal representation of

the ambient noise. We shall review the modal representation of the ambient noise next in Sec. I B and discuss its interpretation in Sec. I C. Numerical examples presented later will show that the vertical directionality of the modal representation agrees very well with the results of the numerical noise models.

B. Modal representation of the distant noise field

Let us first express Eq. (2) in terms of matrix notation. One can write Eq. (2) as

$$B(\theta) = \mathbf{s}^\dagger \mathbf{R} \mathbf{s}, \quad (5)$$

where \mathbf{R} is the covariance matrix of the ambient noise, defined as the ensemble (or time) average of the pressure field (a boldface letter denotes either a vector or a matrix),

$$\mathbf{R}_{ij} = \langle p(z_i) p^*(z_j) \rangle \quad (6)$$

and \mathbf{s} is the steering vector given by $s_j = e^{-ikz_j \sin \theta}$. Using the normal mode representation of the Green’s function

$$G(r, z_j, z_0) = \sum_{m=1}^M \sqrt{\frac{2\pi}{k_m r}} \psi_m(z_j) \psi_m(z_0) e^{-ik_m r}, \quad (7)$$

where ψ_m is the mode depth function and k_m is the complex mode wave number. We express the pressure field from each individual noise source in Eq. (2) in terms of mode amplitudes a_m

$$p(z_j) = \sum_m a_m \psi_m(z_j) \equiv \sum_m E_{jm} a_m, \quad (8)$$

where

$$E_{jm} \equiv \psi_m(z_j),$$

and

$$a_m = S(r) \sqrt{\frac{2\pi}{k_m r}} \psi_m(z_0) e^{-ik_m r}. \quad (9)$$

Substituting Eq. (8) into Eq. (6), the noise covariance matrix can be expressed in terms of its modal covariance matrix,

$$\mathbf{R} = \mathbf{E}^\dagger \mathbf{N} \mathbf{E}, \quad (10)$$

where $\mathbf{N} = \mathbf{a} \mathbf{a}^\dagger$ summed/integrated over all noise sources. The diagonal elements of the noise modal covariance matrix \mathbf{N} is obtained using Eq. (9):

$$\begin{aligned} n_m &= \int 2\pi r dr \langle |a_m|^2 \rangle \\ &= \frac{4\pi^2}{k_m} \int_{r_1}^{r_2} S^2(r) |\psi_m(z_0)|^2 e^{-2\alpha_m r} dr, \end{aligned} \quad (11)$$

where α_m is the modal attenuation coefficient.

The basic assumption for our noise model is that the modal covariance matrix of the ambient noise is diagonal. The physical reasons for this assumption will be discussed below.

With the assumption that the noise modal covariance matrix is diagonal, one obtains a simple closed-form expression for the vertical directionality of the ambient noise,

$$B(\theta) = \sum_{m=1}^M n_m B_m(\theta), \quad (12)$$

where

$$B_m(\theta) = \left| \sum_j e^{-ikz_j \sin \theta} \psi_m(z_j) \right|^2 \quad (13)$$

is the modal beam pattern, i.e., the vertical array response to the m th mode. The noise from distant noise sources is expressed in this model as summation of modal noise arrivals.

The “overhead noise” will be treated separately in Sec. II. (The overhead noise includes the continuum normal modes which can also be treated, in principle, like the discrete normal modes. But existing normal mode computer codes do not handle the continuum modes contribution.) As discussed below, the overhead noise arrivals will be represented by $B_0(\theta)$. Equation (12) including the $m=0$ term will constitute a complete representation (parametrization) of the vertical directionality of ambient noise.

C. Discussion and interpretation

Why should the modal covariance matrix be diagonal? We note that a common assumption for ambient noise is that the (time-averaged) noise coherence between two horizontally separated phones should be a function of the phone separation and not a function of the absolute position of the phones.^{14,32,33} This is true in a range-independent ocean (translational invariance) but is also assumed for more general cases. It was shown in Ref. 33 that this result cannot be true in a range-independent environment if the modal covariance of the noise field is not diagonal. We thus make the assumption that the modal covariance matrix be diagonal as the basis for our noise model even for more general cases (see Sec. IV).

A comment is in order here. We note that the modal covariance of the noise specifies uniquely the vertical distribution of the noise. Given the modal covariance, one determines the noise vertical directionality, the vertical coherence of the noise, and the noise level versus depth. The reverse is not necessarily true. This is due to the fact that the normal modes are not always resolved by a vertical array in practice. Consequently the modal covariance matrix measured by a vertical array needs not be diagonal due to its deficiency in mode decomposition. It also needs to be pointed out that the modal covariance matrix associated with short time data needs not be diagonal due to inadequate statistical samples.

A popular model for ambient noise is the Kuperman–Ingenito model.¹⁴ It can be shown that our model yields the same result obtained by Kuperman and Ingenito (for uncorrelated noise sources). For a range-independent environment, with uniform source density, an analytical expression for the noise spatial covariance has been obtained by Kuperman and Ingenito. To see that our model agrees with theirs, we take Eq. (30) from Ref. 14, which expresses the covariance of the noise for uncorrelated (monopole) sources between two vertically separated phones as

$$R(z_i, z_j) \sim S^2 \sum_m \frac{|\psi_m(z_0)|^2 \psi_m(z_i) \psi_m(z_j)}{k_m \alpha_m}, \quad (14)$$

where an overall normalization factor (due to the difference in mode convention) has been dropped. Substituting Eq. (14) into Eq. (5), we obtain

$$B(\theta) \sim S^2 \sum_m B_m(\theta) \frac{|\psi_m(z_0)|^2}{k_m \alpha_m},$$

where $B_m(\theta)$ has been given above. We obtain the same expression for the vertical directionality using our model, i.e., Eqs. (11) and (12) by integrating over all noise sources ($r_1=0$, $r_2=\infty$). We thus see that the two models agree for the case considered. We note that the Kuperman–Ingenito model is strictly for a range-independent environment. It deals with stationary noise. As discussed in Sec. IV below, the model presented in this paper can be extended to a range-dependent environment. Temporal variation of the ambient noise can also be incorporated into this model (Sec. IV).

We note that in the Kuperman–Ingenito model, an argument was made that “the noise field can be approximated by an incoherent sum of modes,” as the modal interference terms are small in this range-independent model. For our model, we start with the assumption that the modal covariance matrix of the noise is diagonal. As remarked above, our assumption is based on the observation that the horizontal noise coherence should be a function of phone separation but not a function of the (absolute) phone location. This observation is approximately true in range-dependent environments.

Equation (12) will be used in this paper for investigating/interpreting the impact of the acoustic environment (sound speed, bottom depth, bottom attenuation, etc.) on the vertical directionality of the noise.

At this point, it is worthwhile to define what is referred to as the *theoretical* distribution of the noise vertical directionality. This is to be contrasted with what is measured in data. We note that experimental measurements of the ambient noise vertical directionality is limited by the angular resolution of the vertical array which is a function of the vertical array aperture. Theoretically, if we have a vertical array which can resolve all the normal modes, then the modal beam pattern $B_m(\theta)$ will peak at the modal arrival angles. For that case, Eq. (12) becomes basically

$$B(\theta) \sim \sum_{m=1}^M n_m \delta(\sin \theta - \sin \theta_m). \quad (15)$$

This defines the theoretical ambient noise vertical directionality which can be obtained by calculating the noise modal intensity n_m and the modal arrival angles θ_m .

Examining Eq. (11), we note that the noise modal intensity is determined by the mode coupling to the noise sources $\psi_m(z_0)$ and the modal attenuation coefficients α_m . These two parameters control how the acoustic environment impacts on the vertical directionality of the noise. (For discrete modes, the mode wave number k_m in the denominator of Eq. (11) can be approximated by its real value which increases

TABLE I. Bottom parameters used in the noise model.¹⁶

Bottom type	Relative density		Comp. speed (m/s)		Comp. attenu. (dB/ λ)		Shear speed (m/s)	Shear attenu. (dB/ λ)
	Sediment	Bottom	Sediment	Bottom	Sediment	Bottom		
Hard	...	2.0	...	1800	...	0.8	400	2.0
Soft	1.5	1.8	1470–1500	1600	0.06	0.15		
Sediment thickness: 6 m								

gradually with the mode number.) Qualitative prediction of the noise vertical directionality can be given for a waveguide based on the estimated mode amplitudes.

To compare with experimental measurements, one needs to multiply the theoretical noise directionality with the vertical response for each modal arrival [see Eq. (12)]. The modal beam pattern can be easily calculated for various arrays of different configurations. This yields an easy method to compare the performance of various vertical arrays. In contrast, we note that for existing numerical noise models which use Eq. (3), one has to repeat the numerical integration over the noise sources each time the vertical array configuration is changed. For these brute force models, estimating the response of different vertical arrays in a noise field can be time consuming.

II. NUMERICAL IMPLEMENTATION

Before we proceed with the analysis of noise vertical directionality under various acoustic environments we need to extend the model to include the overhead noise. We would like to compare the model prediction of the noise vertical directionality with numerical noise models which integrate numerically the noise beam power over noise sources using various propagation models.

The acoustic environment to be studied involves a waveguide with a constant sound speed of 1500 m/s. (We use the test cases previously modeled in Ref. 16.) The water depth is 100 m. We use the soft bottom with sound speed and density given in Table I. We consider ambient noise at 500 Hz with noise sources at a depth of 0.1 m so that the noise source is effectively a dipole.¹⁶

Figure 1 shows the noise modal intensity n_m as a function of mode number [Eq. (11)] for this environment at 500 Hz using a soft bottom. (Source density equals unity.) Figure 2 plots the theoretical vertical directionality of the distant

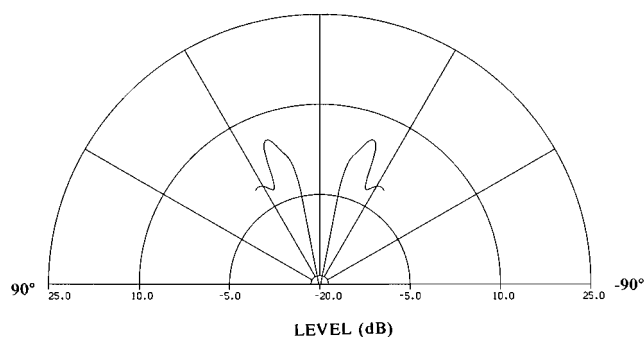


FIG. 2. The theoretical vertical directionality of the ambient noise at 500 Hz for a constant sound-speed profile and a soft bottom.

noise (range >500 m) using Eq. (15).

Next we consider the directional response of a vertical array with respect to the distant noise. According to Eq. (12), the directional response of a vertical array is the product of the modal intensity and the modal response of the vertical array. Three vertical arrays with aperture of 8λ , 16λ , and 24λ deployed at the ocean bottom are considered. Figure 3 plots the noise beam power measured by the three different vertical arrays at 500 Hz. We note that the wider beam width of the shorter vertical array has diminished the (relative) depth of the null compared with the peak noise (beam) power.

To understand the noise field, we investigate the vertical directionality of the noise as a function of the source range. For this purpose, we use a vertical array of 30λ (61 element) such that it fills the water column. Figure 4 plots the noise directionality as a function of the noise source range from 0 to 10 km at 500 Hz. Figure 4 is obtained using Eqs. (11) and (12) at various ranges r_1 with a fixed range integration ($r_2 - r_1 = 50$ m). We also plot the noise as a function of the source range using the numerical noise model [Eq. (3)] which adds up the noise beam power from noise sources within a circular ring of 50 m. Figures 5 and 6 plot the noise directionality as a function of noise source range using KRAKEN³⁰ and SAFARI³⁴ propagation codes, respectively. Compared with Fig. 4, we note that the vertical directionality in both Figs. 5 and 6 fluctuates with the source range. This is due to the discretation of the noise sources in the numerical integration. In other words, the noise sources have not been completely randomized.

Comparing Fig. 4 with Figs. 5 and 6, we observe that the modal representation of the distant noise using Eqs. (11) and

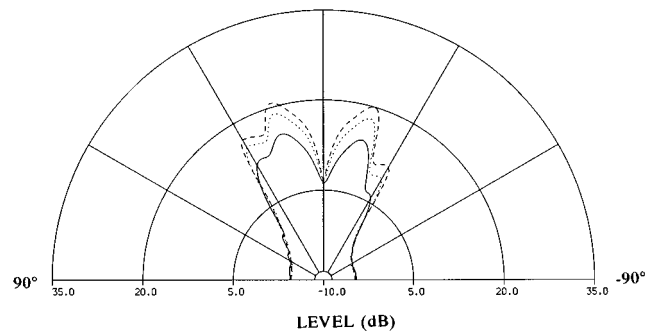


FIG. 3. The noise beam power at 500 Hz measured by three vertical arrays of 8, 16, and 24 wavelength aperture. — 16 phones, 32 phones, ---- 48 phones.

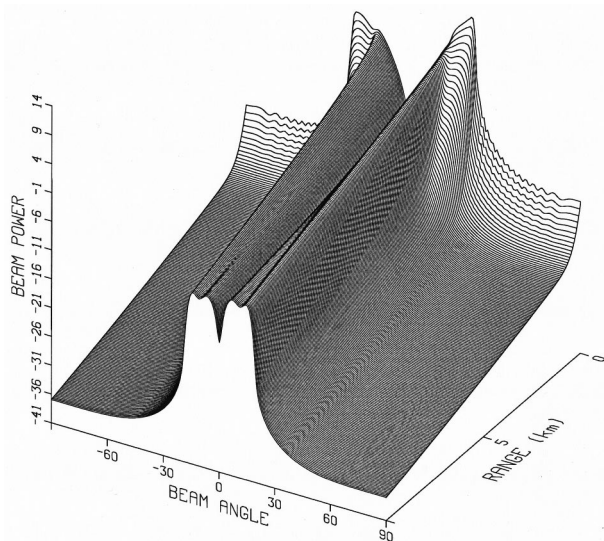


FIG. 4. The noise vertical directionality as a function of the noise source range using the modal intensity representation.

(12) agrees rather well with the numerical integration approach. This result is not surprising. We thus conclude that the vertical directionality of the distant noise can be modeled with a simple closed formula which is fast to calculate. We note that the distant noise directionality can be easily determined for different aperture vertical arrays by inputting the modal beam pattern $B_m(\theta)$ of that array. Using the numerical integration approach, for a different vertical array the noise source integration has to be repeated all over again. The latter is a time-consuming process.

Next we investigate the overhead noise. The contribution of the overhead noise is illustrated by comparing the difference between the noise vertical directionality results using KRAKEN and SAFARI (Figs. 5 and 6). The difference between Figs. 5 and 6 is plotted in Fig. 7. We note that the KRAKEN program includes only the discrete normal modes and neglects the continuum contribution. The SAFARI model,

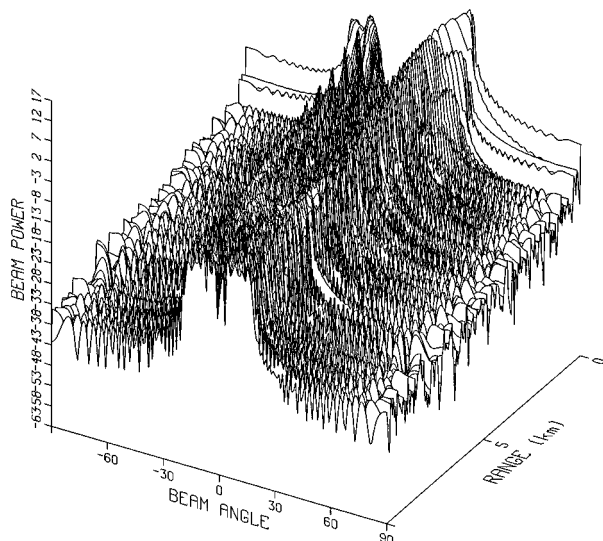


FIG. 5. The noise vertical directionality as a function of the noise source range using the KRAKEN propagation code.

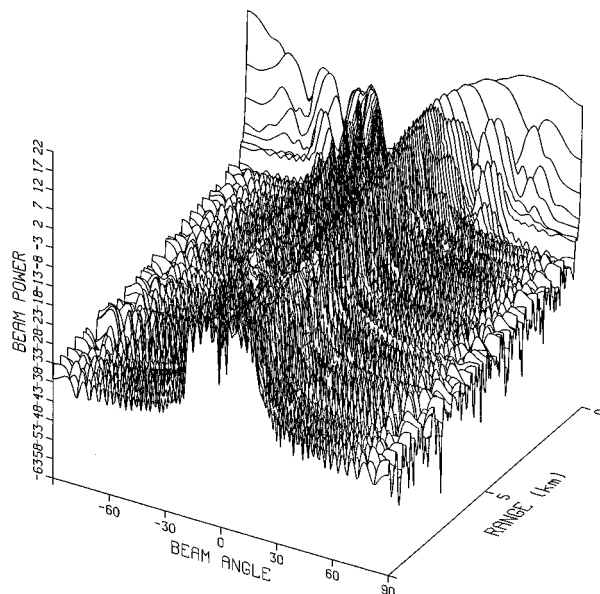


FIG. 6. The noise vertical directionality as a function of the noise source range using the SAFARI propagation code.

a fast field program, includes both. Thus Fig. 7 represents the continuum contribution which originates primarily from short-distance noise sources which we have called the overhead noise. We note that the continuum significantly affects the noise directionality at high grazing angles. This result implies that the short-range (overhead) noise needs to be modeled using a full field propagation code such as SAFARI. Figure 7 also implies that the distant noise is adequately modeled using the discrete normal modes.

Let the noise originating from a surface area of a 500-m radius from the receiver be modeled by SAFARI and the distant noise ($r > 500$ m) be modeled using the closed-form modal representation [Eq. (12)]. The sum of the two is plotted in Fig. 8 and compared with the SAFARI calculation for noise integrated over a range of 10 km. We find a very close agreement between the two calculations.

This result suggests that a hybrid model could be used to describe the noise vertical directionality. Let $B_0(\theta)$ denotes the directionality pattern of the overhead noise, we add a term $n_0 B_0(\theta)$ to Eq. (12) and obtain

$$B(\theta) = \sum_{m=0}^M n_m B_m(\theta), \quad (16)$$

where $B_m(\theta)$ with $m \geq 1$ describes as before the distant noise contribution. This hybrid model is attractive since it takes an order of magnitude less computer time to run than the numerical noise model using the SAFARI code. By separating the overhead noise from the distant noise, we have simplified the problem. Since $B_0(\theta)$ term represents predominantly the continuum mode contribution, we will still refer to Eq. (16) as the modal representation of the noise.

We note that for the overhead noise, since the propagation distance is short, the environment has limited impact on the noise directionality pattern. The distant noise directionality, on the other hand, could be severely modified by the propagation condition, as the noise field will interact repeat-

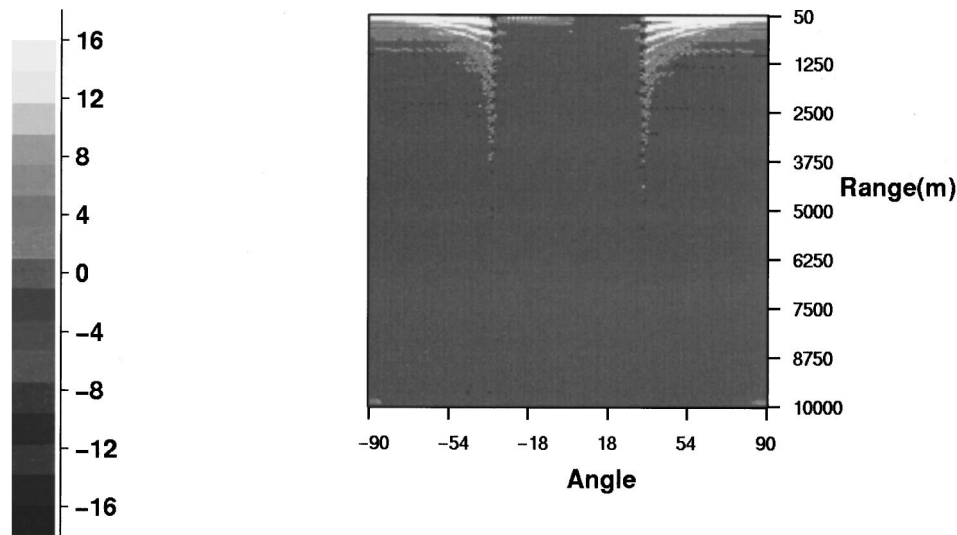


FIG. 7. The difference between the KRAKEN and SAFARI results.

edly with the surface and bottom. The closed form expressions, Eqs. (11) and (12), give us a simple tool to predict and interpret the vertical directionality of the distant noise under various acoustic environments.

We have also compared our model calculation with numerical integration of noise sources using the FEPE code³⁵ and reach similar conclusions. One has to pay attention to the fine differences in the way the different models handle the bottom in order to get consistent results.

III. ENVIRONMENTAL INFLUENCE

In this section we shall use the closed-form expression to illustrate how the acoustic environment affects the noise vertical directionality in shallow water. We shall consider three different (constant, upward and downward refractive) sound-speed profiles shown in Fig. 9 and two different (hard and soft) bottoms described in Table I. Similar studies have been previously conducted,¹⁶ the results can now be understood easily on a physical basis. As shown below, our analysis can be easily applied to more general cases.

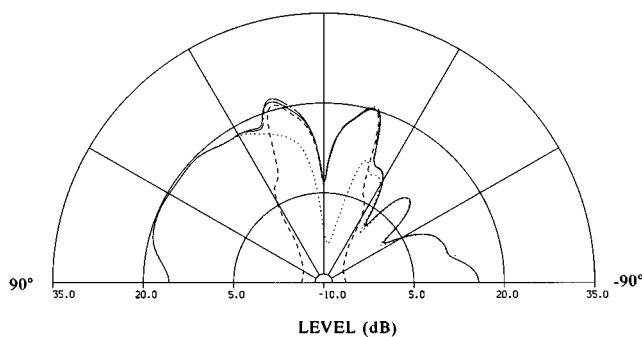


FIG. 8. The vertical beam intensity distribution of the distant noise (dashed line), the overhead noise (dotted line), their sum (dot-dashed line), and finally the numerical integration result of noise sources using SAFARI (solid line) under constant sound-speed environment. Soft bottom with noise sources up to 10 km.

A range-independent environment with uniformly distributed noise sources is assumed so that the bathymetry and source distribution effects are not a factor in the modeling of the noise directionality. Unless specifically mentioned, we will concentrate on the vertical directionality associated with the distant noise since the environmental impact is of greatest interest. In the simulation study, the noise sources integration is up to 10 km unless mentioned otherwise.

A. Dependence on bottom

Figure 10(a) shows the vertical directionality of the noise for the hard bottom case compared with the soft bottom result presented above in Fig. 8. The water sound speed is constant at 1500 m/s. The frequency remains at 500 Hz.

The hard bottom has a critical grazing angle at 34° . Modes arriving at angles $< 34^\circ$ have no loss. The beam intensities for the soft bottom are weaker than the beam intensities for the hard bottom at angles $< 34^\circ$ [Fig. 10(a)]. This difference is attributed to the difference in the noise modal intensities for the two cases as shown in Fig. 10(b). Similar analysis can be carried out to determine the bottom effects under summer and winter sound-speed environments.

The soft bottom also reduces the beam intensities at high ($> 40^\circ$) angles which are dominated by the overhead noise compared with the hard bottom case. Positive high-angle (up looking) beams are associated with direct, surface, and bottom-surface-bounced rays, the latter are responsible for the difference in the beam intensities near 90° . Negative high-angle (down looking) beams are associated with bottom bounced rays. The differences in the down looking beam intensities reflects the differences in the bottom reflectivities.

We note that the effect of the bottom is small for noise arriving at angles $< 15^\circ$. The effect is not negligible for noise arriving at angles $> 15^\circ$.

B. Dependence on sound-speed profile

Next we consider the noise vertical directionality at 500 Hz for a winter sound-speed profile which contains a surface

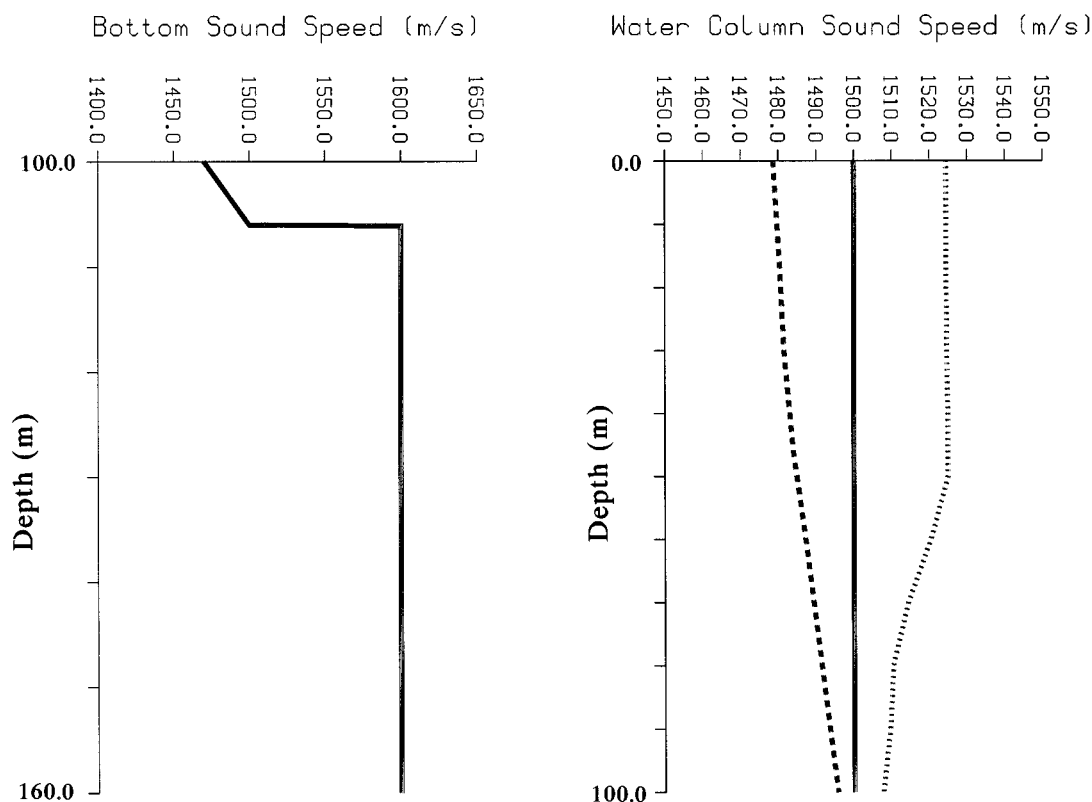


FIG. 9. The constant, winter and summer sound-speed profiles used in the simulation: solid line for the constant sound-speed profile, the dashed line for the winter sound-speed profile, and the dotted line for the summer sound-speed profile. The sound-speed profile for a soft bottom is depicted in the left figure.

duct (see Fig. 9). Figure 11 shows the beam intensities of the distant noise (up to 10 km in range), the overhead noise, their sum, and finally the numerical integration results of noise sources using SAFARI. We see once again that the hybrid model yields practically the same result as that by integrating the noise sources using SAFARI.

Figure 12 compares the noise vertical directionality of the winter sound-speed profile with that of the constant sound-speed profile for noise sources up to 10 km. We note that the effect of the surface duct is the enhanced beam intensities at angles $< 10^\circ$. This result is expected since the lower-order modes are confined in the surface duct, consequently they have stronger coupling to the noise sources than in the constant sound-speed profile case. For the winter profile at 500 Hz, the noise vertical directionality does not have a notch at near horizontal direction. As a matter of fact, distant noise beyond 10 km will make the noise directionality peak in the horizontal direction (see below).

Figure 13 shows the noise vertical directionality under the summer sound-speed profile versus the constant sound-speed profile for the soft bottom. We find that noise vertical directionality for the two cases are very similar despite the difference in the sound-speed profile. Note that the low-order modes are confined to the lower water column for the summer profile. For the constant sound-speed profile, the low modes tend to concentrate in the middle of the water column, depending on the bottom properties. The difference in the sound-speed profiles causes the low-order mode intensities (generated by the surface noise) to be somewhat weaker for the summer sound-speed profile than the constant sound-

speed profile. But the difference has been smeared by the beam patterns associated with the low-order modes. Similar effect is observed for the hard bottom case.

C. Effective area for overhead and distant noise

For the purposes of comparing the effects due to different sound-speed profiles and bottoms, we have fixed the radius of the surface area for noise sources to be 10 km in the above analyses. In this section we investigate how big an area one needs to consider for noise modeling in a shallow-water area. We also address what is the effective area for the overhead noise. In other words, what should one use for r_2 and r_1 ?

Figure 14(a) shows the noise beam outputs for a surface area of 10-km radius versus infinity for the constant sound-speed profile and soft bottom case. We see that due to the bottom loss, noise beyond 10 km contributes little to the noise beam power except decreasing the null at horizontal direction. (The steep rays with grazing angle $> 15^\circ$ have been attenuated at ranges > 10 km. For the sediment/bottom attenuation see Table I.) For the constant sound-speed profile and hard bottom case the distant noise beyond 10 km could completely remove the null in the noise vertical directionality as shown in Fig. 14(b).

Figure 15(a) and (b) show that distant noise could produce a prominent peak at the broadside of the vertical array under the winter (upward refractive) condition for soft and hard bottoms. This is due to the fact that distant noise is trapped in the surface duct and can travel long distances.

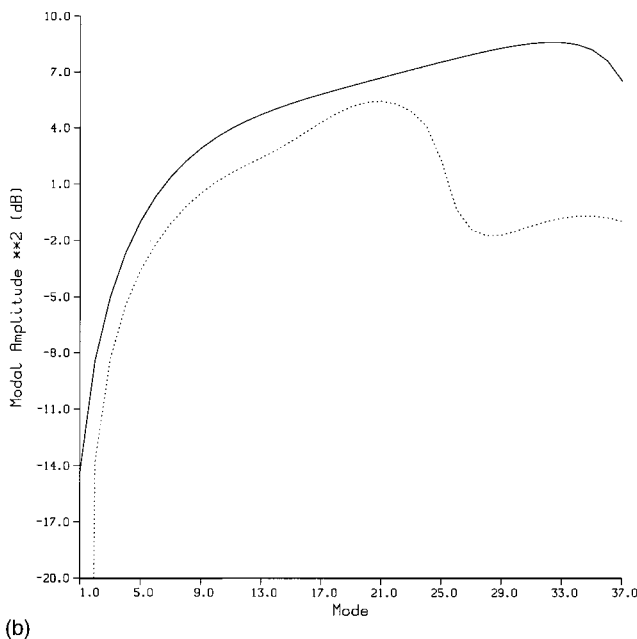
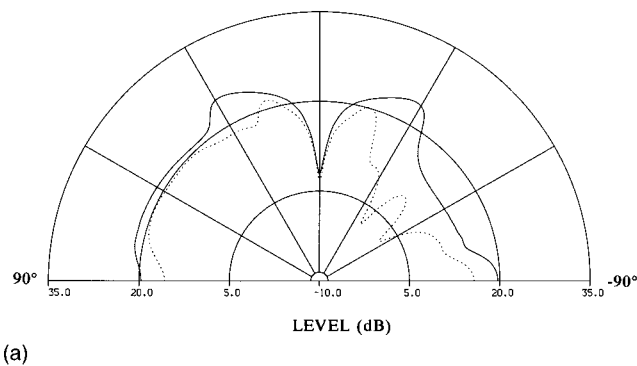


FIG. 10. The noise vertical directionality (a) and noise modal intensities (b) for the hard (solid line) and soft (dashed line) bottom cases. Constant sound speed profile.

For the summer sound-speed profile we find that distant noise beyond 10 km contributes little to the noise vertical beam power for either soft or hard bottom. Repeated bounces from the bottom in the downward refractive environment

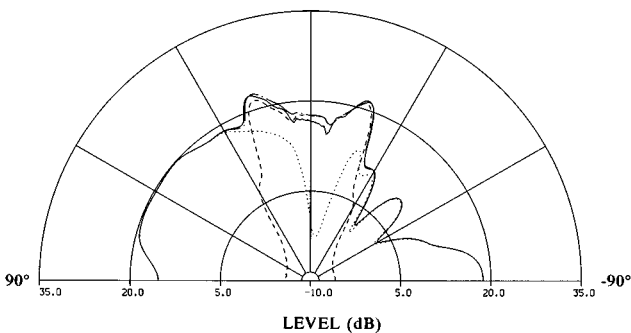


FIG. 11. The vertical beam intensity distribution of the distant noise (dashed line), the overhead noise (dotted line), their sum (dot-dashed line), and finally the numerical integration result of noise sources using SAFARI (solid line) under winter sound-speed environment.

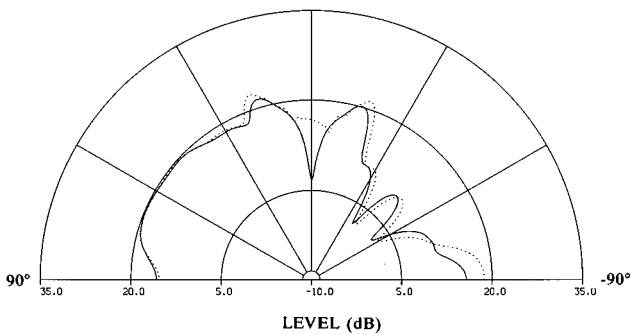


FIG. 12. The noise vertical directionality under winter sound-speed profile (dotted line) versus constant sound-speed profile (solid line). Soft bottom with noise sources up to 10 km.

induce high loss for the distant noise. The notch in the noise vertical directionality remains.

Thus for the summer sound-speed profile, it may be adequate to consider r_2 in the range of 10–20 km at 500 Hz. For a constant and winter sound-speed profile, one should include all the long-distant noise sources as allowed by the environment propagation conditions.

We note that the effective size of the distant noise is determined basically by the modal attenuation at high frequencies. At low frequencies the modal attenuation could be small depending on the environment. In that case, the basin boundary will be the limit.

The effective size of the overhead noise area is determined by r_1 such that the overhead noise plus the distant noise effectively reproduce the results by (brute force) numerical integration. We find that at 500 Hz, r_1 of 0.5 and 1 km for soft and hard bottom, respectively, are good numbers to use. The effective area of the overhead noise tends to be slightly larger for a winter than for a summer sound-speed profile. Except at low (<100–200 Hz) frequencies it is determined practically by the area covered by direct path arrivals, which is implicitly determined by the water depth.

D. Dependence on frequency

Figure 16 shows the noise vertical beam distributions for four frequencies ranging from 50 to 500 Hz for the constant sound speed and soft bottom case. For this figure, the noise source level is constant, independent of frequency. We see

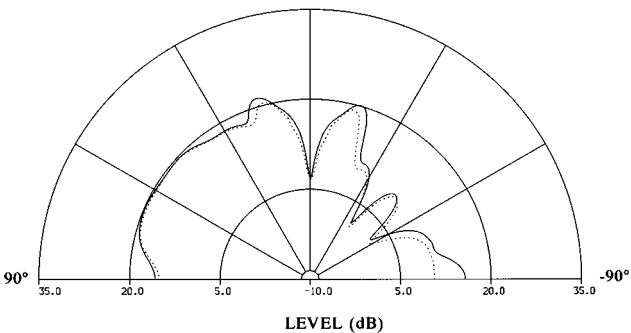


FIG. 13. The noise vertical directionality under the summer sound-speed profile (dotted line) versus the constant sound-speed profile (solid line) for the soft bottom. Noise sources up to 10 km.

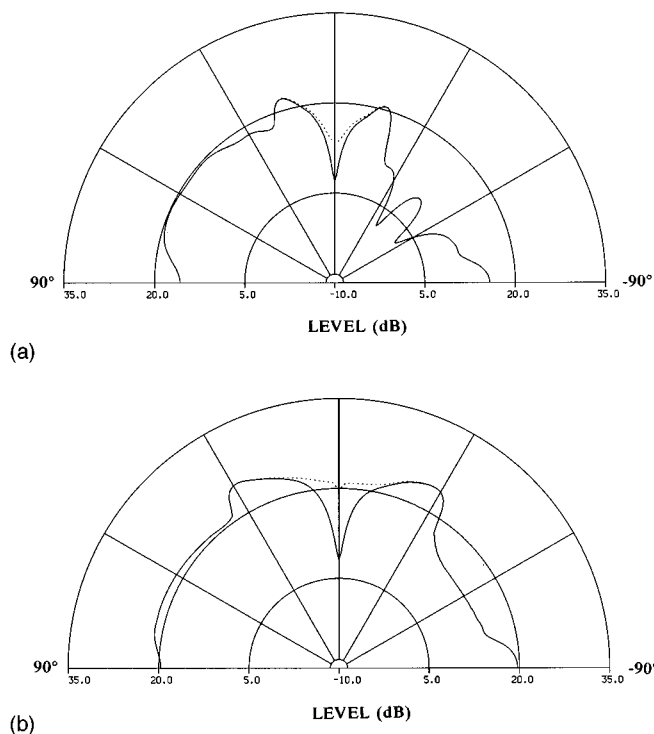


FIG. 14. The noise vertical beam outputs for a surface area of 10 km (solid line) radius versus infinity (dotted line) for the constant sound-speed profile with soft (a) and hard (b) bottoms.

that the noise power decreases with frequency. For distant noise, the beam power at angles $< 30^\circ$ decreases with frequency because the waveguide supports a smaller number of modes at lower frequencies (4 at 50 Hz vs 39 at 500 Hz). The overhead noise beam power also decreases with frequency since the acoustic field penetrates deeper into the bottom and suffers higher loss at lower frequencies.

Since the acoustic energy penetrates deeper into the bottom at lower frequencies, the bottom sound speed and attenuation will have a bigger impact on noise beam power distribution at lower frequencies. At high frequencies, only the top sediment layer will play a role and the area of the distant noise will be limited by the transmission loss. The notch observed in Fig. 16 will persist at higher (> 500 Hz) frequencies since the surface noise sources couple poorly with the low-order modes and the number of higher-order modes increases with frequency. The frequency above which the vertical notch sets in will depend on the environment. The set-in frequency can be estimated by using Eqs. (11) and (12).

IV. TEMPORAL AND SPATIAL CHARACTERIZATION

Assuming that the modal covariance matrix of the noise is diagonal, the current model can be extended to a range-dependent environment for which adiabatic normal mode approximation is valid. The time variation of the noise field can also be incorporated into this model. As discussed above, we expect that the modal covariance matrix of the noise to be (approximately) diagonal assuming that noise is the accumulation of many random processes. The modal cross terms for

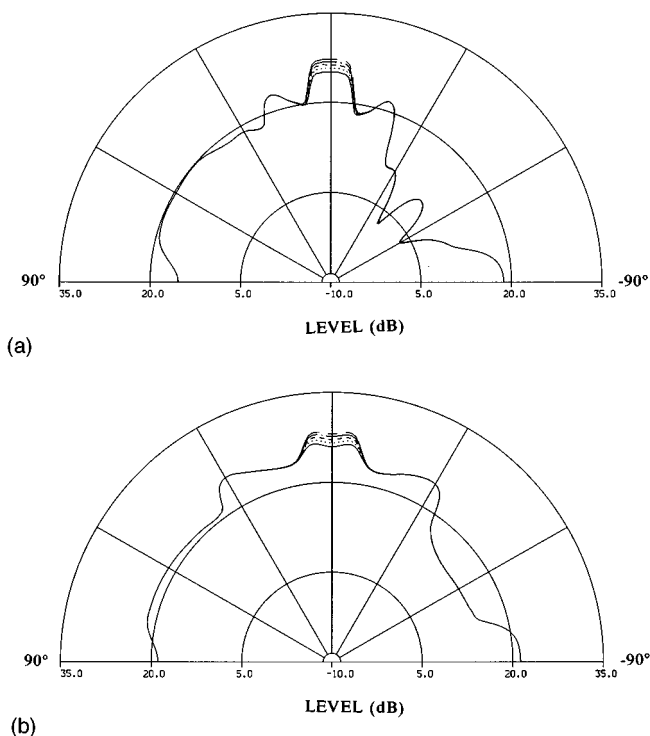


FIG. 15. The noise vertical beam outputs for a surface area of 10 km (solid line) radius versus 60–100 km radius (dotted lines) for the winter sound-speed profile with soft (a) and hard (b) bottoms.

a patch of noise sources over, say, a few to tens of km-square area should cancel for range-dependent propagation as for range-independent propagation.

We note that for the real world the ambient noise varies with time. While it can in principle be modeled given sufficient input data, it is difficult and almost impossible to predict the noise ahead of time. Hence, we intend to use this model to characterize the time variation of the noise and do not intend to model (predict) the temporal variation.

A. Temporally and spatially varying ambient noise modeling

For a range-dependent environment, Eq. (11) will be replaced by

$$n_m(t) = \int r \, dr \, d\theta \langle |a_m(t)|^2 \rangle, \quad (17)$$

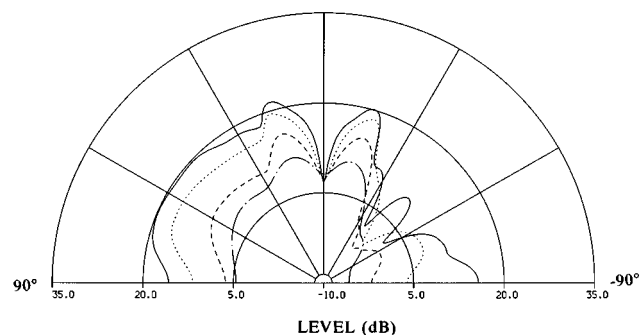


FIG. 16. The noise vertical beam distributions for four frequencies ranging from 50 to 500 Hz for the constant sound-speed and soft bottom case. — 500 Hz, ···· 250 Hz, --- 100 Hz, —·— 50 Hz.

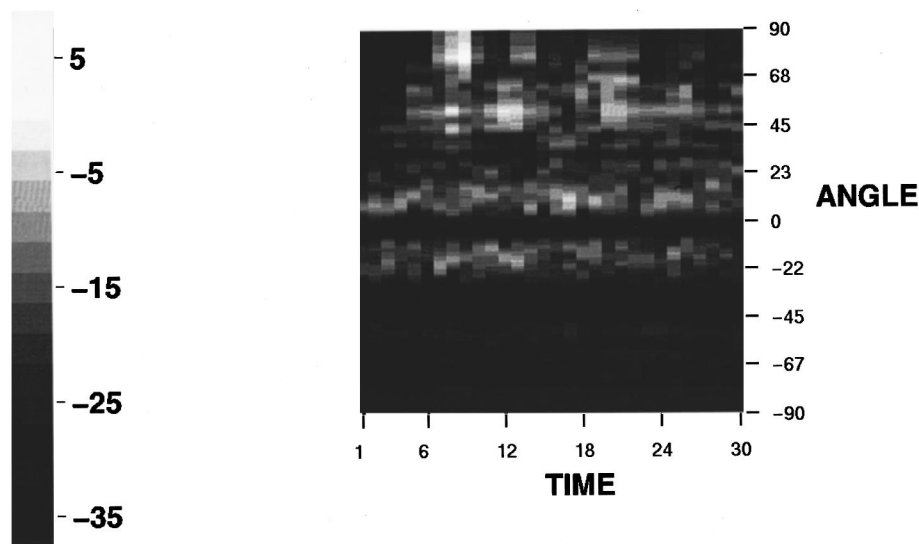


FIG. 17. The noise vertical directionality at 500 Hz as a function of time taken from Ref. 21.

where

$$a_m(t) = S(r, \theta, t) \sqrt{\frac{2\pi}{k_m r}} \psi_m(z_0) e^{-i \int_0^r k_m(r) dr} \quad (18)$$

and

$$\bar{k}_m = \frac{1}{r} \int_0^r k_m(r) dr.$$

Note that the noise source density and source level are now both location dependent. The overhead noise is represented by $n_0(t)$ as before. Using the adiabatic approximation, this noise model takes much less time to run than existing models which integrate the noise sources by brute force.

Using the adiabatic normal modes, Perkins *et al.* has extended the Kuperman–Ingenito formula of the vertical coherence of noise to a range-dependent environment.²⁴ A complicated expression was obtained. In comparison, the above expression of the noise directionality for a range-dependent environment is relatively simple. The reason is that the directionality relates to noise intensities whereas the vertical coherence contains the interference of the spatially separated noise fields.

We note that in certain shallow-water areas, the acoustic environment can be highly variable (e.g., range-dependent bathymetry) such that the adiabatic approximation is no longer valid. For such environments the numerical noise model, while more time consuming, will be more accurate. For some applications, an approximate noise vertical directionality may be sufficient. This model can be used for that purpose, and will be approximately correct if the correction due to mode coupling is small compared with the total noise field. For example, if bathymetry is the cause of mode coupling, only noise sources along the bearing of strongly varying bathymetry will have mode coupling corrections. So if the noise from relatively flat areas dominates, mode coupling from strong bathymetry variations will be a small correction.

B. Characterization of noise vertical directionality data

We propose to use the modal representation to characterize the time-dependent noise field data. For time-dependent noise, Eq. (16) is replaced by

$$B(\theta, t) = \sum_{m=0}^M n_m(t) B_m(\theta). \quad (19)$$

In the above equation, the noise modal intensities are time dependent. The noise modal intensities can be extracted from field data. The extraction of noise modal intensities requires a procedure which deconvolves the array beam pattern from the data. As discussed above, the modal intensities represent the true vertical distribution of the noise.

A word on the normalization is in order. Since $B_m(\theta)$ are normalized to a constant by

$$\int B_m(\theta) \cos \theta d\theta = \frac{1}{k} \sum_j \psi^*(z_j) \psi(z_j) = \text{const} \quad (20)$$

for $m \geq 1$, we will use the same normalization for the overhead noise beam pattern $B_0(\theta)$ as well.

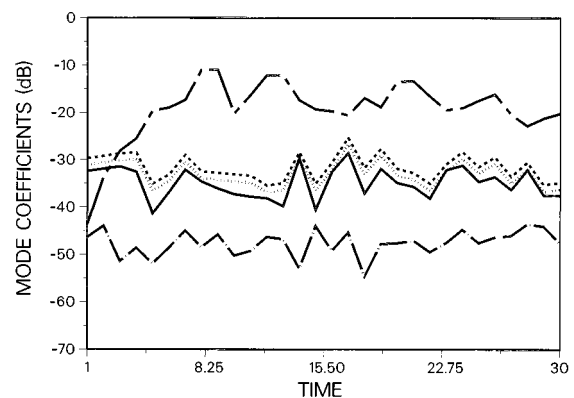


FIG. 18. The time variation of representative modal intensities, $n_m(t)$, obtained by least-mean-square fit to the data in Fig. 17. The top curve is for n_0 , the bottom four curves are for n_6 to n_9 .

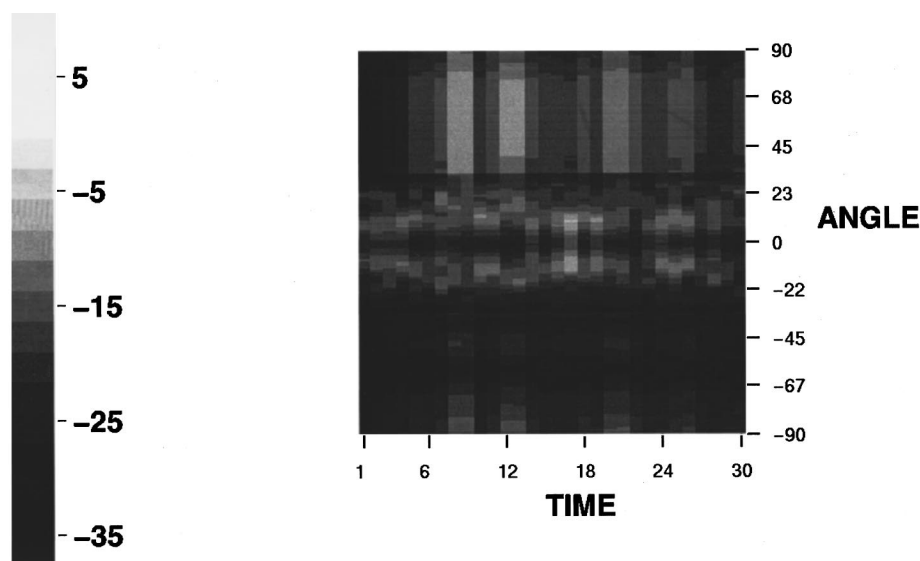


FIG. 19. The noise vertical directionality at 500 Hz regenerated using the extracted modal intensities.

Figure 17 displays the noise vertical directionality at 500 Hz as a function of time as reported in Ref. 21. It represented a modeled temporal fluctuation of the vertical directionality over a period of ~ 30 s based on features observed in experimental data. We are going to represent this (simulated) data in terms of noise modal intensities, Eq. (19). Figure 18 shows the time variation of the (extracted) modal intensities, $n_m(t)$, for a few (representative) modes, obtained by a crude least-mean-square fit to the data. A small additive white noise was added to data in Fig. 17 as the deep nulls in Fig. 17 are not seen in real data. The temporal variations of the noise vertical directionality are now characterized by the variation in noise modal intensities. Using the estimated modal intensities, one can regenerate the noise vertical directionality as shown in Fig. 19. Figure 19 shows a slightly smoother angular distribution than Fig. 17 since the model has assumed smooth noise source distribution. We see that Fig. 19 has captured the essential features of the noise vertical directionality distribution in the data. Equation (19) is a way to reduce a large amount of beam data in terms of a set of coefficients. The usefulness of this equation requires more study.

V. SUMMARY

In this paper, we offer a modal representation for the noise vertical directionality.³⁶ The noise vertical directionality measured by a vertical array is represented by the noise modal intensities, multiplied by the modal beam patterns of the vertical array. The former is independent of the vertical array configuration but the latter will be. A closed form expression is given for the modal intensities associated with distant noise. The overhead noise is separately modeled/calculated using a full field propagation code, such as SAFARI or FEPE.

The modal representation, Eq. (19), offers a physical interpretation of the vertical directionality in terms of modal arrivals. It can be used to reduce field data of noise vertical directionality in terms of a set of modal (intensity) coeffi-

cients which offers insight on how the noise field is changing with time. The modal representation lets one conduct a quick evaluation on how a different vertical array will respond to the noise field as the array aperture and phone spacing are changed.

The modal intensity model, Eq. (11) or (17) offers a simple tool to evaluate how the noise vertical directionality will change as the acoustic environment (i.e., sound-speed profile and bottom) is changed. The change in noise vertical directionality is directly related to the change in noise modal intensities which are controlled by the modal attenuation.

We illustrated the noise model using a constant, winter, and summer sound-speed profile with two types of bottom. We addressed the issue of how the sound speed and bottom types affect the vertical directionality of the noise, particularly when and whether the vertical directionality of the noise exhibits a notch near the horizontal direction. We conducted a comparison test of this noise model with the model results which integrate the noise power numerically over all noise sources. We discussed the effective surface area size for the overhead and distant noise which contribute dominantly to the noise vertical beam power. We touched briefly on the frequency dependence of the noise vertical directionality. We discussed how this model can be generalized to a range-dependent environment and temporally varying noise field.

ACKNOWLEDGMENTS

This work is supported by funds from the Office of Naval Research managed by the Naval Research Laboratory.

¹R. J. Urick, *Ambient Noise in the Sea* (Peninsula, Los Altos, CA, 1986).

²Special issue on sea surface-generated ambient noise 20–2000 Hz, IEEE J. Ocean. Eng. **15**, 268–340 (1990).

³*Sea Surface Sound*, edited by B. R. Kerman (Kluwer Academic, Dordrecht, 1988).

⁴*Natural Physical Sources of Underwater Sound*, edited by B. R. Kerman (Kluwer Academic, 1993).

⁵G. R. Fox, “Ambient noise directivity measurements,” J. Acoust. Soc. Am. **36**, 1537–1540 (1964).

- ⁶V. C. Anderson, "Variation of vertical directionality of noise in the North Pacific," *J. Acoust. Soc. Am.* **66**, 1446–1452 (1979).
- ⁷S. C. Wales and O. I. Diachok, "Ambient noise vertical directionality in the Northwest Atlantic," *J. Acoust. Soc. Am.* **70**, 577–582 (1981).
- ⁸W. S. Hodgkiss, Jr. and F. H. Fisher, "Vertical directionality of ambient noise at 32 °N as a function of longitude and wind speed," *IEEE J. Ocean. Eng.* **15**, 335–339 (1990).
- ⁹R. M. Kennedy and T. V. Goodnow, "Measuring the vertical directional spectra caused by sea surface sound," *IEEE J. Ocean. Eng.* **15**, 229–310 (1990).
- ¹⁰D. Wilson and R. Kneipfer, "Preliminary results of experiment to measure vertical directionality of ambient noise in shallow water," NUWC Tech. Report, 1995 (unpublished).
- ¹¹B. F. Cron and C. H. Sherman, "Spatial correlation functions for various noise models," *J. Acoust. Soc. Am.* **34**, 1732–1736 (1962).
- ¹²R. J. Talman, "Ambient sea noise model," *J. Acoust. Soc. Am.* **36**, 1541–1544 (1964).
- ¹³H. Cox, "Spatial correlation in arbitrary noise fields with application to ambient sea noise," *J. Acoust. Soc. Am.* **54**, 1289–1301 (1973).
- ¹⁴W. Kuperman and F. Ingenito, "Spatial correlation of surface generated noise in a stratified-ocean," *J. Acoust. Soc. Am.* **67**, 1988–1996 (1980).
- ¹⁵R. Dashen and W. Munk, "Three models of global ocean noise," *J. Acoust. Soc. Am.* **76**, 540–554 (1984).
- ¹⁶R. M. Hamson, "The theoretical responses of vertical and horizontal line arrays to wind-induced noise in shallow water," *J. Acoust. Soc. Am.* **78**, 1702–1712 (1985).
- ¹⁷W. M. Carey, R. B. Evans, J. Davis, and G. Botseas, "Down-slope propagation and vertical directionality of wind noise," *IEEE J. Ocean. Eng.* **15**, 324–334 (1990).
- ¹⁸R. M. Hamson, "Vertical array response to shipping noise: Model/measurement comparisons for range-dependent Mediterranean sites," *J. Acoust. Soc. Am.* **94**, 386–395 (1985).
- ¹⁹R. M. Hamson and R. A. Wagstaff, "An ambient noise model that includes coherent hydrophone summation for sonar system performance in shallow water," SACLANT Undersea Research Centre SACLANTCEN SR-70, La Spezia, Italy, 1983 (unpublished).
- ²⁰P. Scrimger, R. M. Heitmeyer, and P. Boulon, "A computer model of merchant shipping in the Mediterranean sea," SACLANT Undersea Research Centre SACLANTCEN SR-164, La Spezia, Italy, 1990 (unpublished).
- ²¹S. Finette and R. M. Heitmeyer, "Angle-time-frequency resolution of the noise field generated by wind-induced breaking waves," *J. Acoust. Soc. Am.* (to be published).
- ²²M. J. Buckingham and S. A. S. Jones, "A new shallow-ocean technique for determining the critical angle of the seabed from the vertical directionality of the ambient noise in the water column," *J. Acoust. Soc. Am.* **81**, 938–946 (1987).
- ²³T. C. Yang and G. Giellis, "Normal mode modeling of low frequency Arctic ambient noise: Comparison with FRAM IV data below 50 Hz" (1989) (unpublished).
- ²⁴J. S. Perkins, W. A. Kuperman, F. Ingenito, L. T. Fialkowski, and J. Glattetre, "Modeling ambient noise in three-dimensional ocean environments," *J. Acoust. Soc. Am.* **92**, 739–752 (1993).
- ²⁵F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (AIP, New York, 1994), Chap. 9.
- ²⁶J. Wilson, "Low frequency wind-generated noise produced by the impact of spray with the ocean surface," *J. Acoust. Soc. Am.* **68**, 952–956.
- ²⁷L. Ding and D. M. Farmer, "On the dipole acoustic source level of breaking waves," *J. Acoust. Soc. Am.* **96**, 3036–3044 (1994).
- ²⁸R. M. Kennedy, "Sea surface dipole sound source dependence on wave-breaking variables," *J. Acoust. Soc. Am.* **91**, 1974–1982 (1992).
- ²⁹See, for example, E. M. Podeszwa, "Shallow-water environmental data for the U.S. east coast," NUSC Tech. Document 8571, 1991 (unpublished).
- ³⁰M. Porter, "The KRAKEN normal mode program," SACLANTCEN Mem. SM-245, 1991.
- ³¹T. C. Yang and A. Al-Kurd, "The internal waves effect on the noise vertical directionality in shallow water," *Proc. of OCEANS '96*, 9–16 (MTS/IEEE 1996).
- ³²R. Klemm, "Use of generalized resolution methods to locate sources in random dispersive media," *IEEE Proc.* **127**, 34–40 (1980); "Range and depth estimation by line arrays in shallow water," in *Signal Processing* (Reidel, Dordrecht, The Netherlands, 1981), Vol. 3, pp. 333–344.
- ³³T. C. Yang, "Effectiveness of mode filtering: a comparison of matched-field and matched-mode processing," *J. Acoust. Soc. Am.* **87**, 2072–2084 (1990).
- ³⁴H. Schmidt, "SAFARI, seismo-acoustic fast field algorithm for range independent environments, user's guide," Report No. SR-113, SACLANT Undersea Research Centre, San Bartolomeo, Italy, 1988 (unpublished).
- ³⁵M. D. Collins, "A split-step Pade' solution for parabolic equation method," *J. Acoust. Soc. Am.* **93**, 1736–1742 (1992).
- ³⁶The modal representation for the noise vertical directionality is similar to the Legendre polynomial representation for the spherical wave.

Modeling low-frequency reverberation near the Mid-Atlantic Ridge and comparison with ARSRP data

Jerald W. Caruthers and E. J. Yoerger

Naval Research Laboratory, Stennis Space Center, Mississippi 39529

J. C. Novarini

Planning Systems, Inc., Long Beach, Mississippi 39560

(Received 12 June 1995; revised 12 July 1996; accepted 14 October 1996)

The Acoustic Reverberation Special Research Program (ARSRP) of the Office of Naval Research conducted low-frequency acoustic reverberation experiments just west of the Mid-Atlantic Ridge in 1993. Analyses of these data are supported by a bathymetric survey that produced nearly full-coverage bathymetry gridded to a resolution of 200×200 m. At this scale direct-path reverberation data has been successfully modeled using Lambert's law applied to the local grazing angle determined from the high-resolution bathymetry. The analysis suggests that there could be a slightly weaker dependence on grazing angle than sine squared. The good agreement of this simple model with data shows that, for this region, two-dimensional seafloor morphology at the proper scale determines the ability to predict reverberation. For a model, such as Lambert's law, that glosses over the details of seafloor microroughness or texture at the scale the acoustic wavelength ($\lambda = 6$ m), the seafloor morphology scale that is critical is shown to be an order of magnitude or two greater than an acoustic wavelength. An average Lambert coefficient for the region is between -14 and -17 dB, but it is less than -20 dB for deeply sedimented areas and around -12 dB for rugged exposed rock areas. [S0001-4966(97)06605-8]

PACS numbers: 43.30.Gv, 43.30.Hw, 43.30.Ma [JHM]

INTRODUCTION

Sound scattering from rough surfaces is a very complex phenomenon. Although it has been studied extensively, a full understanding of a three-dimensional scattering function is still lacking. In 1989 the Office of Naval Research (ONR) established the Acoustic Reverberation Special Research Program (ARSRP) whose purpose is to improve this understanding for scattering from the sea surface and the seafloor.

As a part of the bottom scattering effort of the ARSRP, researchers from the Naval Research Laboratory (NRL), several universities, and oceanographic institutions have engaged in experimental and theoretical studies of acoustic reverberation in a region just west of the Mid-Atlantic Ridge near 46°W , 26°N , called the ONR Atlantic Natural Laboratory. Extensive, high-quality, bottom reverberation data were collected on two cruises to the area: The Acoustic Reconnaissance Cruise in August 1991¹ (designated ARSRP'91) and The Full-Scale Experiment in July 1993² (designated ARSRP'93). In these experiments both distant and short-range reverberation data were taken at frequencies around 250 and 230 Hz, respectively.

To support analyses of the acoustic data, two geological and geophysical (GG) cruises were also conducted in the Atlantic Natural Laboratory. In 1992 a wide-area GG survey was conducted which established, among other seafloor properties, a 200×200 -m gridded bathymetric dataset.³ In addition to using the measured depth values, we compute for our reverberation analyses other geomorphic parameters such as macroscale slope and bearing. In this paper we demonstrate that a crucial element of modeling the ARSRP data is an adequate two-dimensional (2-D) description of the sea-

floor. Issues relevant to the resolution of the geomorphology are discussed.

The long-range goal of the bottom reverberation and scattering studies of the ARSRP is to achieve a full understanding of the physics of seafloor scattering in the Atlantic Natural Laboratory. Yet the ARSRP acoustics experiments and GG surveys provide an excellent opportunity to test the degree to which simple algorithms (e.g., Lambert's law), applied with the benefit of good two-dimensional geomorphology, can explain the data at the macroscale, and thereby provide a useful, applied prediction tool.

Based on analyses of this GG survey and acoustic results of the Acoustic Reconnaissance Cruise, several regions were selected for study in greater detail. One such site was selected for the character of its short-range reverberation. This site was designated site A by the ARSRP and revisited in 1993. Here we discuss the character of the short-range (7–20 km), direct-path, monostatic reverberation data collected in 1993 at site A. Since we achieved consistent results for the dozen or so pings that we analyzed, we choose to discuss three here in detail. We first describe the GG characteristics of site A.

I. THE GEOLOGY AND GEOPHYSICS OF SITE A

Figure 1 shows the geomorphology in a portion of the survey area (65×50 nm) at a resolution of 200 m. In this area there are three distinct morphological provinces.⁴ In the vicinity of site A, the morphology is typical of "outside corner" crust.⁵ This crust is formed at a midocean ridge segment adjacent to a passive transform fault. The relief is subdued and is dominated by orthogonal, steeply dipping

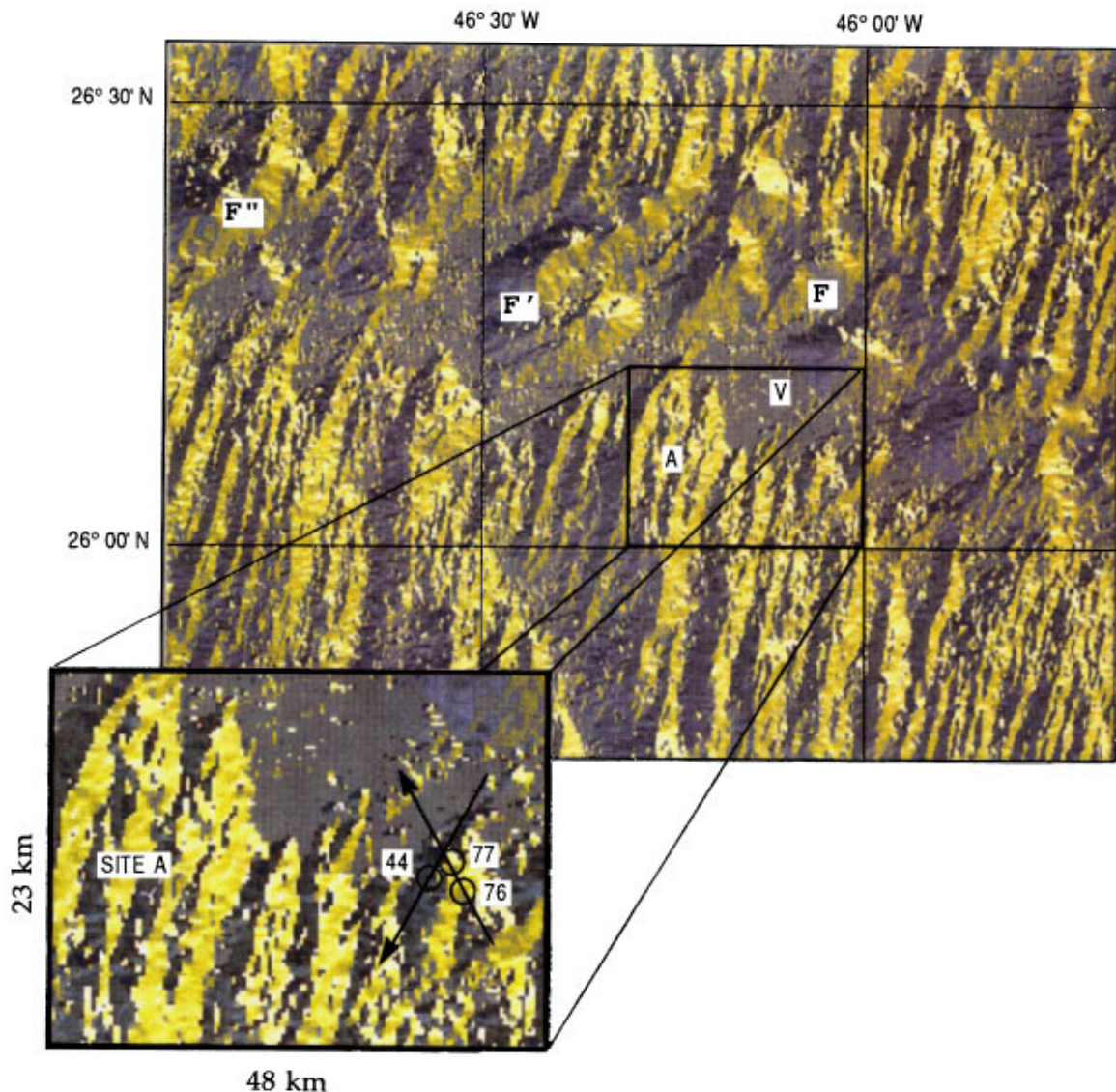


FIG. 1. Geomorphology for a 3000-nm² region in the ONR Atlantic Natural Laboratory as it would appear if lit by a rising sun. Inset shows site A and ping positions and headings for data discussed in text.

normal faults which face towards the valleys and are aligned parallel and perpendicular to the ridge. Regions F, F', and F'' are typical of "inside corner" crust which formed at a ridge adjacent to an active transform fault. The relief is dominated by fault scarps, with variable orientation, but primarily oblique to the ridge axis. The exposed-rock types are more plutonic than volcanic representative of deeper crustal levels in "normal" crust. Between the inside and outside corners is a flat sedimented valley (labeled V), running WNW to ESE, about 4400-m deep.

Figure 2 shows a representation of this seafloor morphology and the tectonics in a slow spreading environment including faults, talus, and sedimentation.³ The sedimentation is very thin or absent on the scarps and tops of blocks where it has been removed by currents or mass wasting. This structure persists through most of the southern half and the northeasternmost quarter of Fig. 1.

Figure 3 shows a typical seismic section taken in the GG cruise of 1992.⁶ Two points that may not show clearly in this figure are a layered structure in the sediment ponds and a

stronger residual relief on the face of the large ridge at site A (outside corner) than on the large ridge opposite the sediment pond (inside corner). Sediment thicknesses through the center of the valleys can reach 400 m.

The scale of the GG bathymetry gridded at 200 m is not a direct indicator of the scattering mechanisms and strengths in the frequency range of a few hundreds of hertz. In fact,

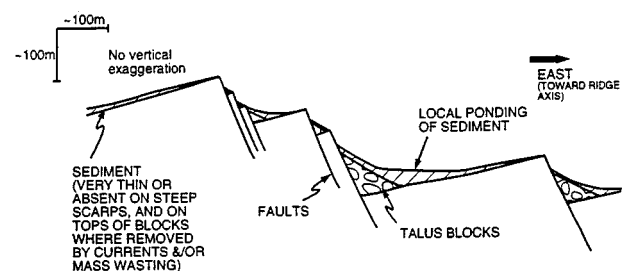


FIG. 2. Representation of seafloor morphology and tectonics in a slow spreading environment (Ref. 3, sketch by B. Tucholke).

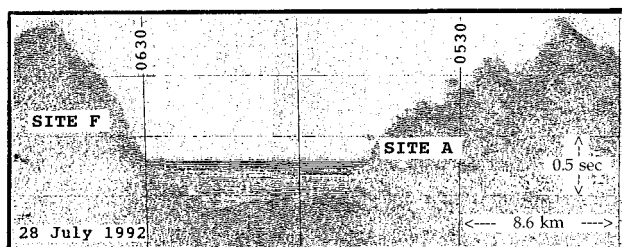


FIG. 3. Seismic profile through sediment pond and including outside (site A) and inside corner (site F) types (Ref. 6).

since we know nothing about features on the scale of an acoustic wavelength, we assume the simplest scattering hypothesis, Lambert's law. Once this is assumed, however, the issues are to determine if scattering patterns can match the geomorphology at measured scales and if Lambert's coefficient takes on reasonable values including any regional variation that may relate to geological provinces that are present.

One of the goals of ARSRP'93 was to study reverberation from these sedimented ponds, ridges, and seamounts; hence, there are many acoustic runs which ensound this area from various angles. The inset in Fig. 1 shows the location of the three ARSRP'93 pings discussed in detail here (labeled 44, 76, and 77). The arrows through these locations indicate the ship's track during recording which, for this analysis, become the array headings.

II. THE ACOUSTICS EXPERIMENTS AND DATA

The experimental setup included a vertical line array (VLA) source and a horizontal line array (HLA) receiver arranged in a quasimonostatic configuration.² For ARSRP'93, the source had ten elements separated by 2.27 m and transmitted a 5-s linear FM waveform centered at 230 Hz with a bandwidth of 55 Hz and a level of 232 dB *re*:1 μ Pa at 1 m. [The proper ARSRP designation of these specific "pings" used here is "wavetrain WT93RP019 of Segment 044," etc.; however, we will simply call it "Ping 44" (Ref. 2, pp. 136–137).] The horizontal receiving array consisted of 128 elements spaced at 2.5 m. The output of the receiving array was formed into 127 beams and match-filtered.

Figure 4 shows an example of the calculated one-way transmission loss to the region of interest on the seafloor. The VLA source was steered 9° down from horizontal to ensound the seafloor at ranges from 7 to 20 km. This downward steering helped reduce the energy reflected from the sea surface, and greatly simplified the transmission-loss character on the bottom. The figure also shows the source beam-pattern and the acoustic axis hitting the bottom at about 12 km.

Data for Pings 44, 76, and 77 are shown in Fig. 5. The displays represent a typical A-scan beam/time presentation used for most of the analyses of the ARSRP data. These plots are waterfall displays of reverberation level as a func-

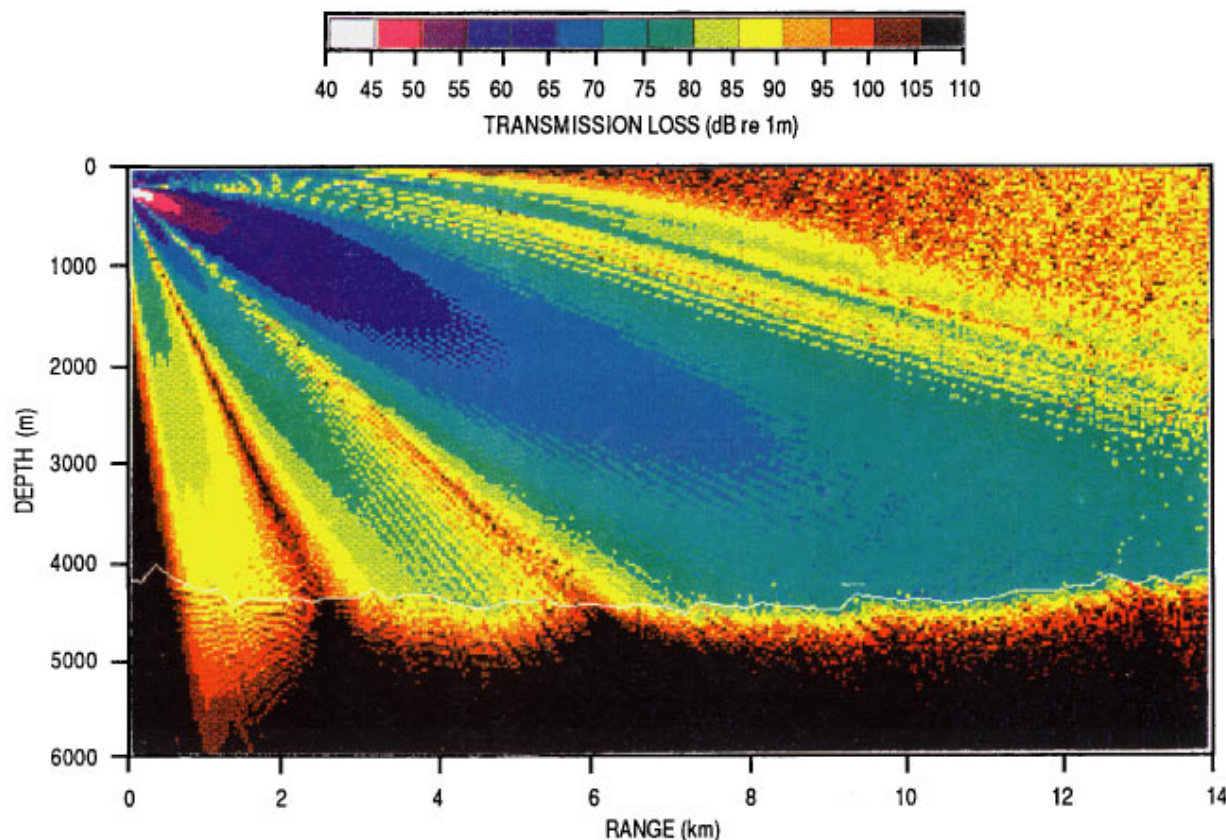


FIG. 4. Transmission loss predicted by UMPE (University of Miami parabolic equation model) for 9° downward steered source impinging on site A (provided by K. Smith).

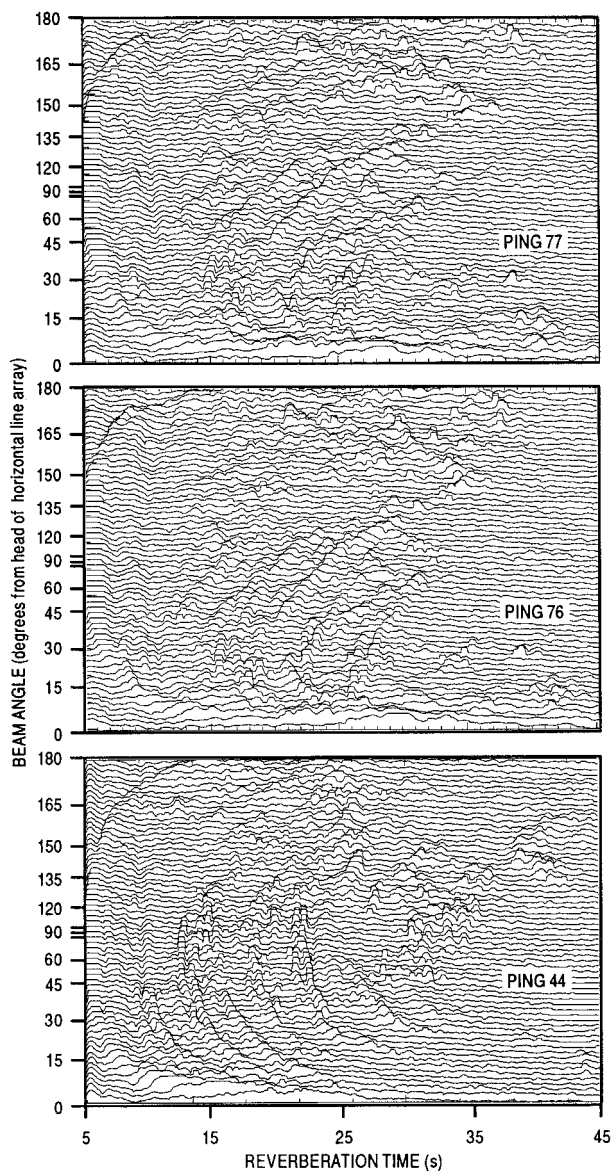


FIG. 5. Reverberation evolutions for ARSRP'93 Pings 44, 76, and 77. Scattering from site A seen in region from 15 to 30 s and 5° to 90°.

tion of time and beam angle. Time is measured from the peak of the match-filtered outgoing ping. Beam angle is measured from the head of the HLA (toward the ship, essentially the ship's heading) to aft (0° to 180°) with 127 discrete angles uniformly spaced in constant increments of cosine. (Sometimes beam number, rather than beam angle, is indicated with #0 being the forward beam, #63 broadside, and #127 aft.) The cross-range resolution (for a 3-dB half-width) was about 200 m at broadside and 12 km in range. Reverberation level is the matched filtered output in each beam, expressed in dB. The theoretical down-range resolution was about 30 m after matched filtering.

The return from the bottom begins at about 5 s for the broadside beams and appears at later times as the beams approach end fire. For all three pings, the ship is at approximately the same position, forming a small triangle about a kilometer on a side. Hence, times of echos from individual points on the seafloor are within a few seconds of each other.

There is evidence in Fig. 5 of the expected strong lineations in the reverberation from the ridges at site A. Reverberation from those ridges occurs between about 15–30 s in the forward half beams (0° to 90°) for each of the diagrams. For Pings 76 and 77, having the same heading, the patterns of scatter are very similar and for Ping 44, having a different heading, the pattern is very different. These observed patterns are consistent with expectation. Of course, because the receiver is a horizontal line array, there is a left/right ambiguity; that is, without *a priori* information one could not tell from which side of the array the echos are coming. With the benefit of known geomorphology in relation to the array headings, we know that site A is in the left forward beams for Pings 76 and 77 and the right forward beams for Ping 44. The ambiguous side from site A for Pings 76 and 77 is the massive (inside corner) block at site F. The ambiguous side from site A for Ping 44 is an outside corner of lower relief which is probably sedimented in the valleys and slightly sedimented on the ridges. Although it is difficult to say from which side any given reverberation value is coming, the pattern of ridges at site A is clearly distinguished in each of the pings in Fig. 5 and the contributions from the ambiguous side for each must, therefore, play a secondary role.

III. MODELING SCATTERING AND REVERBERATION

To model the direct-path, monostatic reverberation data, we have adopted a scattering function of the Lambert form,

$$\text{scattering strength} = \mu + 20 \log(\sin \theta),$$

where μ is the Lambert coefficient (in dB) and θ is the local grazing angle on the bottom. The critical and distinctive feature in this analysis is the use of local grazing angles as determined by a ray propagation model and the local mean slope and bearing of the average 200- \times 200-m planes which are resolved by the bathymetry. The mean slope and bearing of the planes determine the gradient of the bottom in the direction of acoustic propagation and, therefore, the local grazing angle at the elemental plane.

For the initial plots and discussion we have adopted as an initial value $\mu = -17$ dB and concern ourselves with empirical adjustments to μ as needed to describe various geographical regions. A value of μ that is routinely used is -27 dB, which is sometimes called the MacKenzie coefficient since it was determined by MacKenzie as an empirical fit for an abyssal plain region.⁷ ARSRP participants have concluded that for the rugged region of the Atlantic Natural Laboratory the Lambert coefficient should be 10–20 dB stronger than MacKenzie's value.¹ The value of -17 dB chosen here is reasonable in that light and represents the weaker end. As we will see, most of our adjustments will be toward the stronger values, not quite reaching the 20-dB increase at the stronger end.

The particular architecture for the application of Lambert's law is coded in the bistatic scattering strength model (BISSM).^{8–10} This architecture generates a gridded dataset of a quantity we call the instantaneous elemental reverberation level (IERL) and assigns a value to each of the 200- \times 200-m planes in the gridded bathymetry. IERL signifies the 200- \times 200-m element's contribution to the rever-

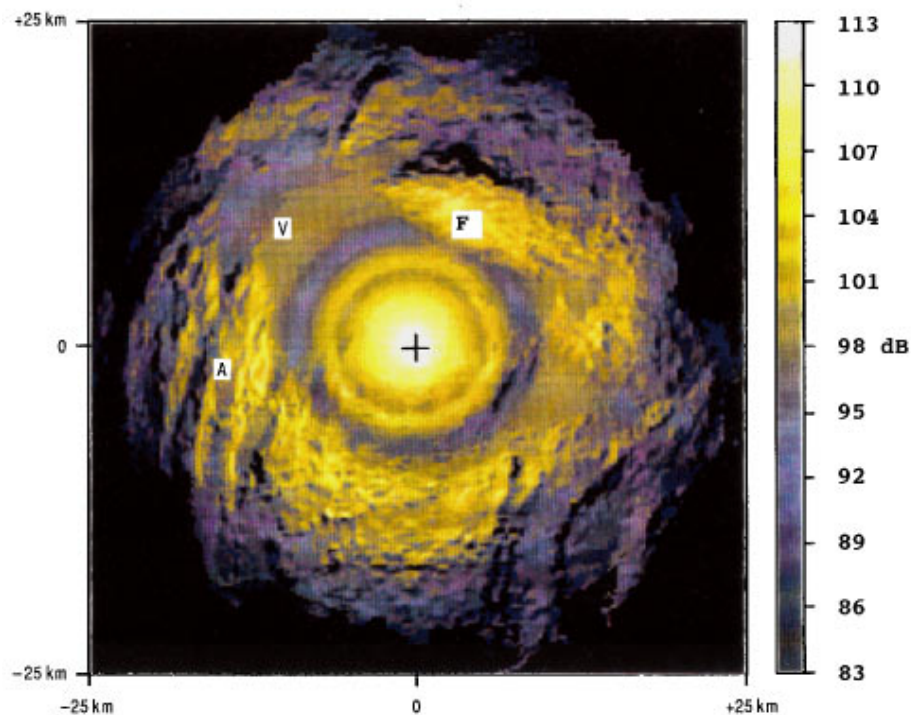


FIG. 6. Instantaneous elemental reverberation levels (for each 200×200 -m patch) predicted by BISSM and mapped onto the seafloor.

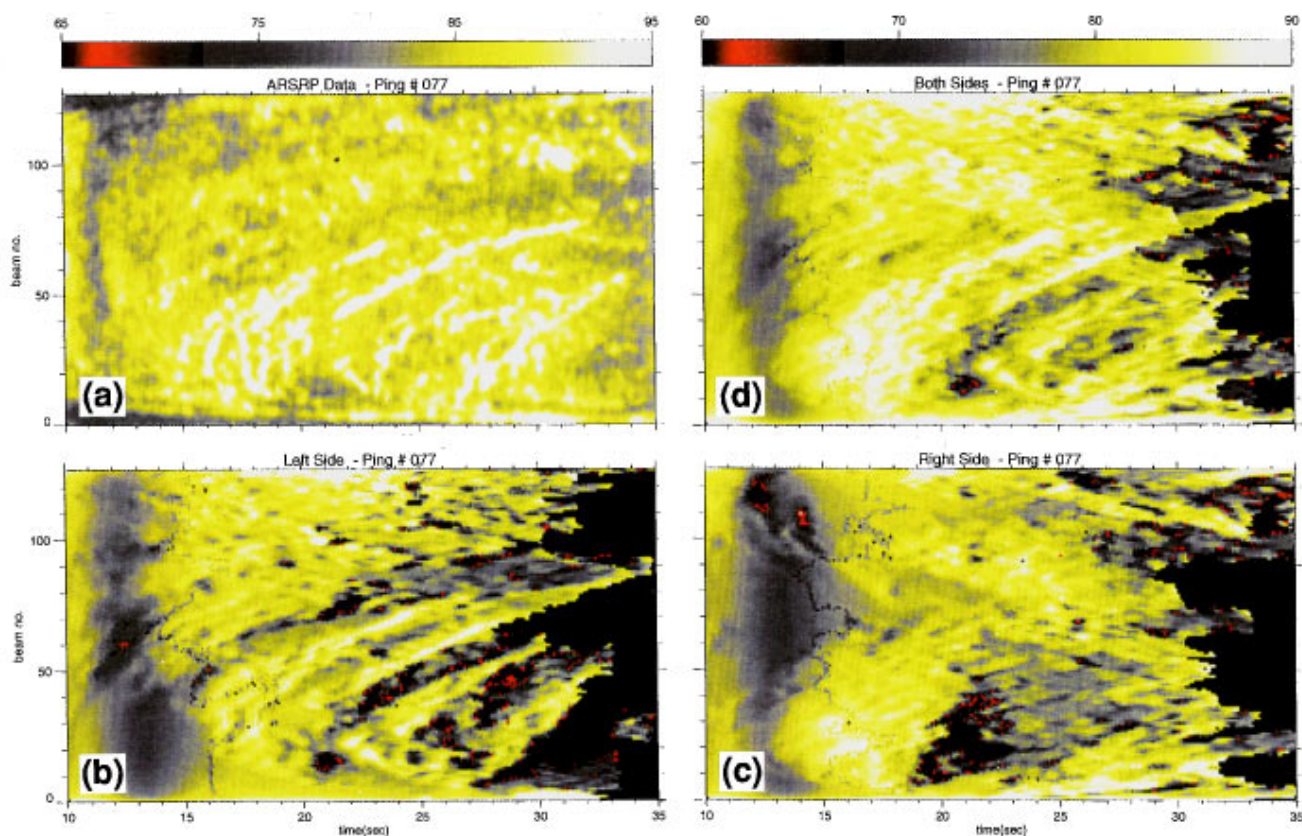


FIG. 7. A-scan plots of reverberation levels for Ping 77: (a) ARSRP data, (b) simulation: left side, (c) simulation: right side, (d) simulation: both sides combined. Scale over (d) applies also to (b) and (c).

beration at the instant (interval) it is ensonified. Thus, the word “instantaneous” indicates that the time element is also captured and assigned to the plane along with the elemental reverberation level.

BISSM was run for several pings at site A. The next section serves to illustrate the details of the simulation using ARSRP’93 Ping 77. In the course of outlining this technique here, the essential features of scattering at site A will also be discussed. In subsequent sections, Pings 76 and 44 of the ARSRP’93 dataset will be used to reinforce the conclusions and develop more subtle features.

A. BISSM simulation architecture and analysis of Ping 77

Figure 6 shows the resulting IERL values for Ping 77 mapped onto the seafloor at their points of origination. The values of IERL in the plot are determined from the sonar equation:

$$\text{IERL} = \text{SL} - 2 \cdot \text{TL} + \text{BL} + \mu + 20 \log(\sin \theta) + 10 \log(A) + 3,$$

where SL is the source level (232 dB *re*:1 μPa at 1 m); $2 \cdot \text{TL}$ is the two-way transmission loss to the point on the bottom (at the point the acoustics axis hits the bottom this is about 161 dB); BL is the beampattern level (on the acoustic axis this is 0 dB); $\mu = -17$ dB (initially); θ is the local grazing angle (on the acoustics axis the nominal value is about 18° , so the term is about -5 dB); A is the elemental area ($200 \times 200 = 40\,000$, so the term is 46 dB); the additional 3 dB is for the surface-reflected, incoherently summed energy on return. This gives a nominal value for the IERL of 98 dB for a flat plane on the bottom on the acoustic axis. (This nominal value is seen to exist in Fig. 6 near the symbol V where the bottom is flat and on the acoustic axis.) The distinctive character of the IERL plot is provided by the variability of TL, BL, and θ . The strongest small-scale variability seen in Fig. 6 is due to variations of the local grazing angle (θ); that is, due to the geomorphology.

In Fig. 6 bright yellow represents regions of the seafloor that are “lit up” by the acoustic energy. The bright spot in the center is directly under the ship at Ping 77 and the annular rings are indicative of the sidelobes and nulls of the VLA source steered down at 9° . The irregularity of the rings is due to geomorphology. The bright region to the NNE is the side of a seamount (site F) of inside corner character, and just beyond is the dark shadow on its far side. Directly under the ship and off to the NW/ESE is a flat, sedimented valley (V) of low backscatter. Continuing counterclockwise around this figure, to the west and WSW is a series of ridges showing bright faces on their forward slopes followed by dark shadows on their far slopes. These are the ridges at site A. The darkness around the diagram beyond the illuminated area is where the acoustic rays bend upward from the bottom and leave the bottom in a shadow zone. (Because of simplifications in our propagation code, the beginning of this shadow zone is more abrupt in the simulation than in reality.)

Subsequent BISSM processing of this rectangularly gridded IERL transforms these values into beam/time space (A-scan) and convolves them with the HLA beampattern. (The beampattern used here consists of simple wedges with

effective widths and steered in the appropriate directions for a given HLA orientation.) Figure 7 shows the comparison between ARSRP data for Ping 77 in an A-scan display [Fig. 7(a)] and the BISSM A-scan displays. Figure 7(b) shows the IERL representation in this A-scan display for the left side and Fig. 7(c) for the right side of the HLA. Figure 7(a) corresponding to the data, includes the unresolvable ambiguity and must be compared to a combination of right- and left-side reverberation levels [Fig. 7(d)]. Although the scattering strengths may be different on each of the two sides, for lack of *a priori* information we will weight them equally for their combination. Our discussions, however, will highlight any recognizable differences that may be attributed to one side or the other.

For Fig. 7 and all subsequent processing a certain amount of smoothing has been done. First, the process of converting the simulation from a map plot to an A-scan involved interpolation and smoothing. This reduces the resolution to about 300 m. Because the data are resolved much more finely in range, we have applied range smoothing to correspond to the range resolution of the simulations. To further mitigate against any other data problems, e.g., effects outside the assumed 3-dB half-widths, we smooth in range and beam angle (for both data and simulation) by three additional points. Therefore, we are now dealing with patch sizes on the bottom of approximately 0.5×0.5 km.

Before proceeding with the analysis, several points must be made. First, since it is not a feature of interest to us, the curved response of the bottom in the data [Figs. 5 and 7(a)] as seen from an HLA in our experimental configuration is not important and has not been simulated [Fig. 7(b)–(d)]. Second, the transformation from a discrete rectangular grid to a discrete polar grid causes a few data dropouts as seen in Fig. 7(b)–(d), but this is of no consequence since it is minor and occurs only for short ranges. Third, the nulls and sidelobes of the VLA appear in slightly different places between the simulation and data displays because we use a theoretical beampattern while the real beampattern is known to be different (but is not sufficiently quantified for our use). Finally, because the source and receiver arrays are horizontally separated (~ 1 km) and because there is some error in locating the source and receiver with respect to the geomorphology, there is a bearing-dependent timing error. This error can amount to nearly 2 s in the worst case. There can be bearing errors due to errors in determining the heading of the array. We do not attempt to calculate these errors, but will shift the simulation to achieve a correlation with the data features. (The analysis is made only if there is a clear correlation and the shift needed is within reason).

Finally, the data show weak scattering from the aft quarter beams which is not apparent in the simulation. The region of weak scatter is the sedimented valley for both sides of the HLA, and, therefore, the scattering strength for the sedimented, smooth valleys must be lower than it is for the ridges. Clearly, local adjustments of μ are required. These adjustments are the subject of the remainder of the paper.

We now make more detailed and quantitative comparisons between the simulated beam reverberation and the ARSRP data. For this we focus on the selected beam/time re-

TABLE I. Beam/time regions selected for further analysis and resulting scattering strength values.

Ping and region no.	Region		Adjmt to - 17 dB (dB)		
	Beam nos.	Time (s)	Median	“AvgErg”	μ (dB)
77/1	30–70	17.0–27.0	2.7	0.4	-13.9
77/2	95–122	14.0–19.0	-3.1	-0.3	-20.4
77/3	4–34	14.0–20.0	1.5	0.4	-15.1
77/4	10–40	22.0–30.0	4.2	-0.1	-12.9
77/5	95–115	28.0–34.0	6.0	-1.1	-12.1
76/1	34–73	16.3–27.2	2.3	0.4	-14.3
76/2	99–123	14.8–20.3	-3.1	-4.4	-19.9
76/3	5–38	12.7–19.4	0.3	0.9	-15.8
76/4	12–43	20.8–29.5	4.3	-0.4	-12.1
76/5	97–116	28.7–35.2	7.3	-2.0	-11.7
44/1	10–40	16.0–25.0	1.4	0.1	-15.4
44/2	32–72	22.0–25.0	3.9	-0.3	-13.4
44/3	84–112	18.0–22.0	-1.5	-2.1	-18.1
44/4	91–119	31.0–35.0	4.2	0.2	-12.6
44/5	10–70	15.0–20.0	2.6	0.3	-14.1
44/6	30–70	26.0–29.0	5.3	0.7	-11.0
44/7	30–60	30.0–35.0	10.4	-3.6	-10.2

gions given in Table I. Figure 8 shows where these regions are in beam/time space for Ping 77.

The following procedure is used in matching the levels between the simulation and the data: First, the median levels for the simulated data (left, right, and both sides) and the actual data for the region are established. The lower half of the data points for each is cast out because these points may have been in shadows or they are in the weaker locations where sidelobe and multibounce problems could arise. We are, therefore, fitting to the stronger returns in a given region. Differences in median values are, in part, due to use of $\mu = -17$ dB, and, therefore, equate to adjustments of scattering strengths for the region (see column labeled “median” in Table I for adjustment values). Next we reference the simulated and actual data to the median values (i.e., sub-

tract the median values to create scattering strengths measured from 0 dB up).

Adjustment to the median value is not the only adjustment required, however. An additional adjustment is needed to bring the average energies in the upper half-median into agreement. We have chosen a “weighted average energy” to be the averaged decibel value. The reason for choosing the dB average is to avoid further emphasizing the highest intensities. This turns out to be the smaller adjustment and not critical anyway. These values are given in Table I as “AvgErg.”

Figures 9 and 10 display the upper half-median values for two of the selected regions of Ping 77 (regions 77/1 and 77/4 of Table I, respectively). The color scale at the left side of each represents 0–10 dB (*re*: median values). The time and beam scales are linear in the intervals given in Table I, but not placed in the figures.

Region 77/1 includes the three ridges in the southern area of site A on the left and a low-scattering region east of site F on the right. There is a clear correlation between data and simulation for the area. The correlation with site A (left side) appears to be better than with the combined side, implying that the right side is weaker than modeled and should be reduced relative to the left before combining. We conclude that ridges at site A are strong scatters with $\mu = -13.9$ dB.

Region 77/4 includes several ridges to the west at site A and the shadow behind the mountain at site F. There is again a clear correlation between the data and simulation of the left side, but some of the finer details of the structure of the ridges appear not to be modeled well. This could be due to misrepresentations in the bathymetry at the finer scales or could relate to specific, strong scattering patches (such as very large microscale faults). The right side must contribute little energy and that energy, appearing in the lower right quadrant of the right-side simulation, must be weaker than modeled, since it does not appear in the data.

Region 77/2 included some weak scattering areas in the

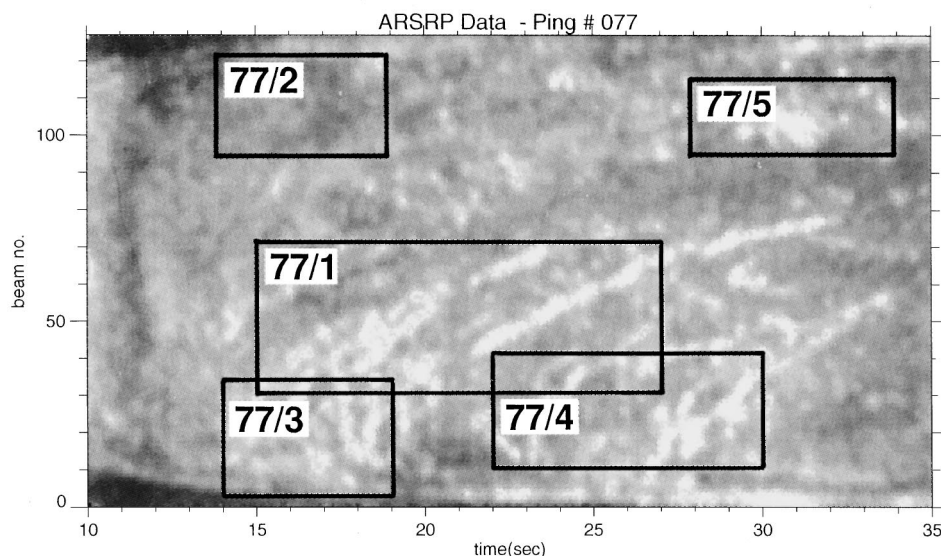


FIG. 8. ARSRP data for Ping 77 in A-scan format showing regions selected (Table I) for detailed analysis.

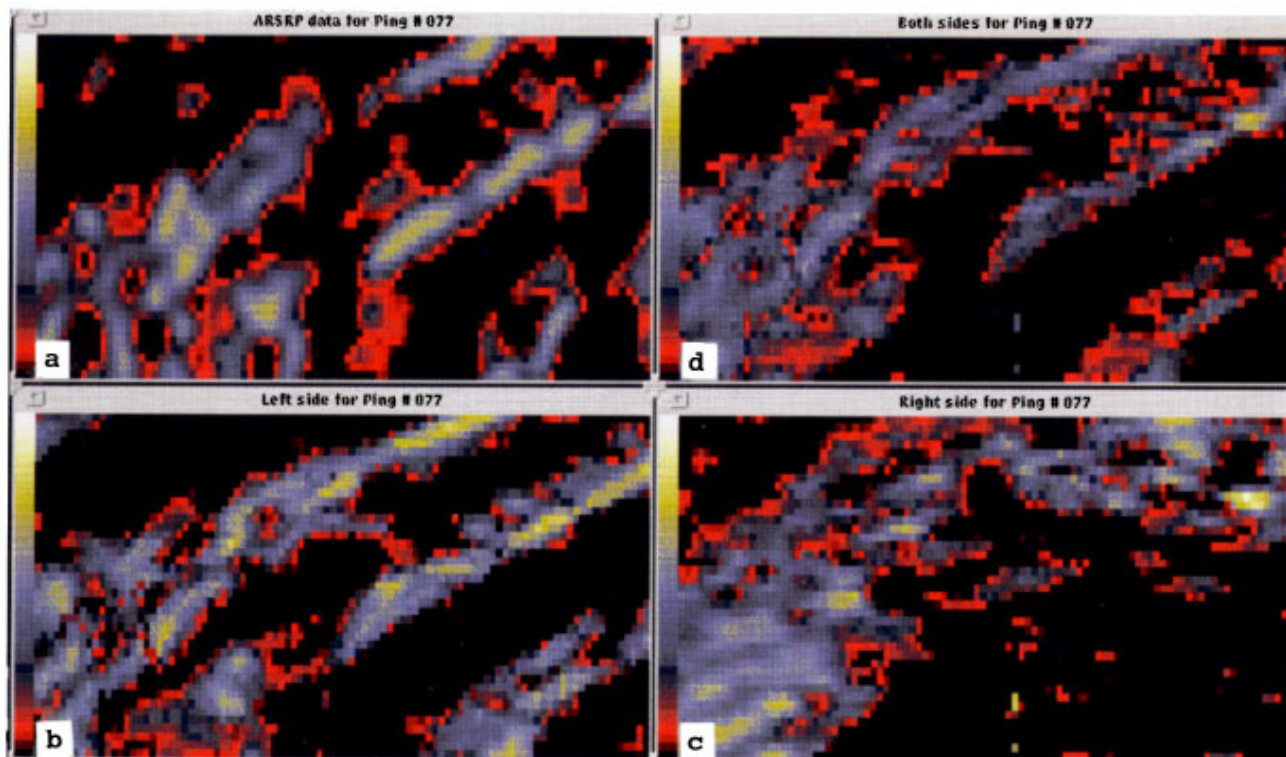


FIG. 9. A-scan plots of reverberation levels for Region 77/1: (a) ARSRP data, (b) simulation: left side, (c) simulation: right side, and (d) simulation: both sides. Time/beam number axes are linear and given in Table I. Intensity scale is strip at left of each and scaled from 0 to 10 dB.

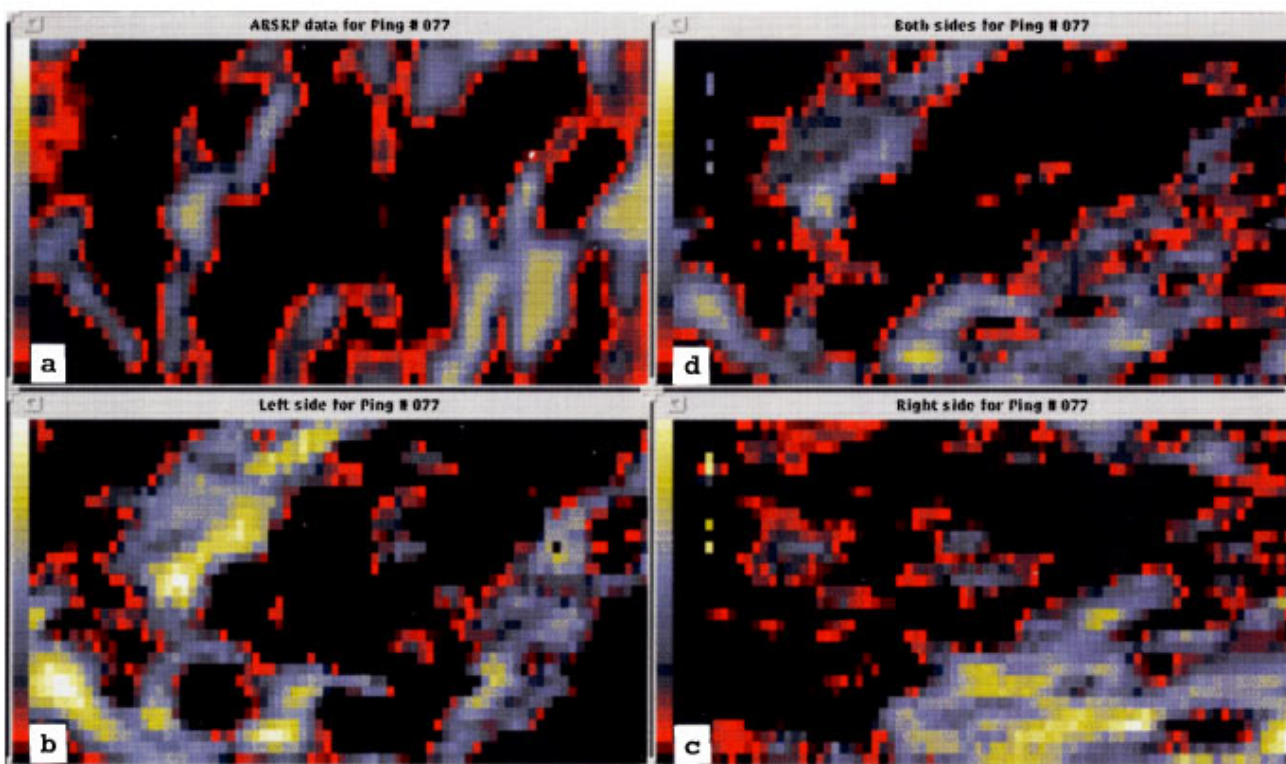


FIG. 10. A-scan plots of reverberation levels for Region 77/4: (a) ARSRP data, (b) simulation: left side, (c) simulation: right side, and (d) simulation: both sides. Time/beam number axes are linear and given in Table I. Intensity scale is strip at left of each and scaled from 0 to 10 dB.

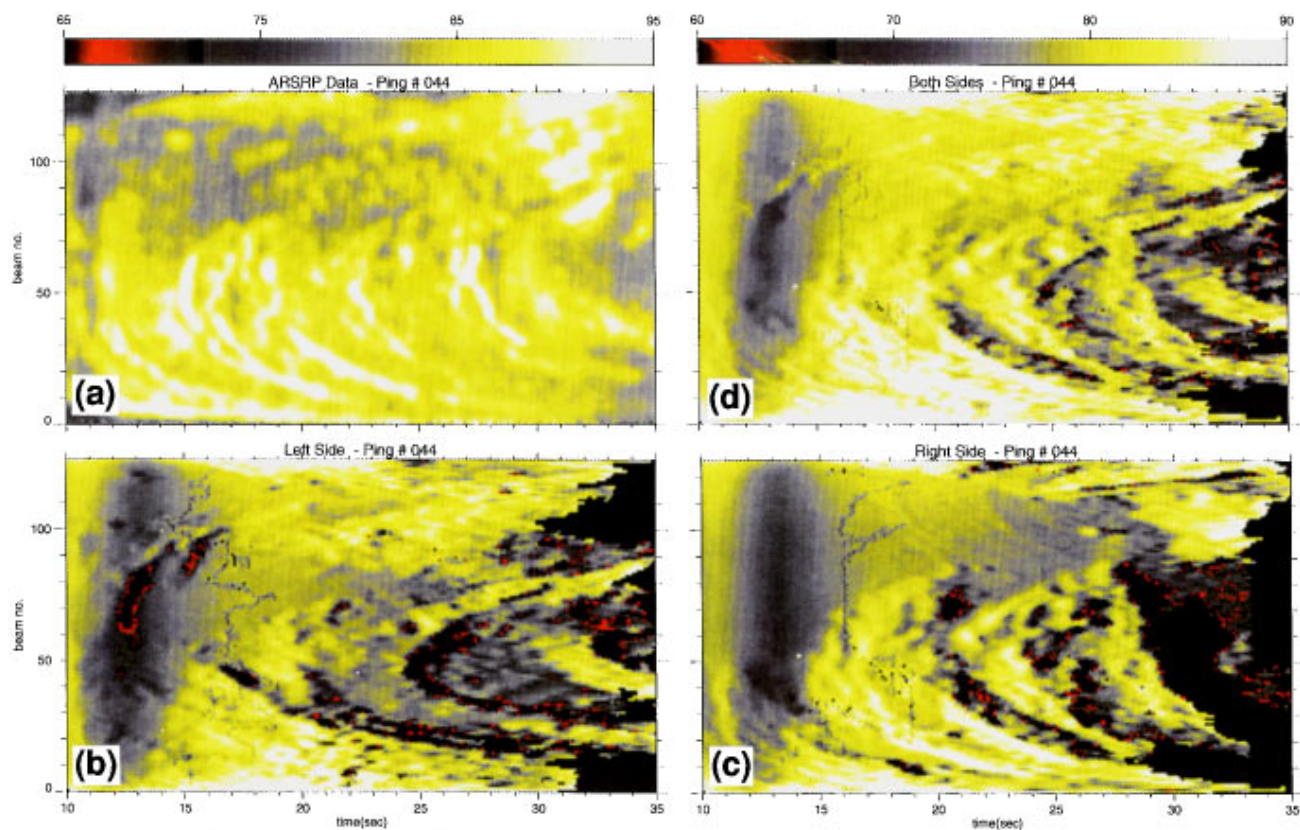


FIG. 11. A-scan plots of reverberation levels for Ping 44: (a) ARSRP data, (b) simulation: left side, (c) simulation: right side, (d) simulation: both sides combined. Scale over (d) applies also to (b) and (c).

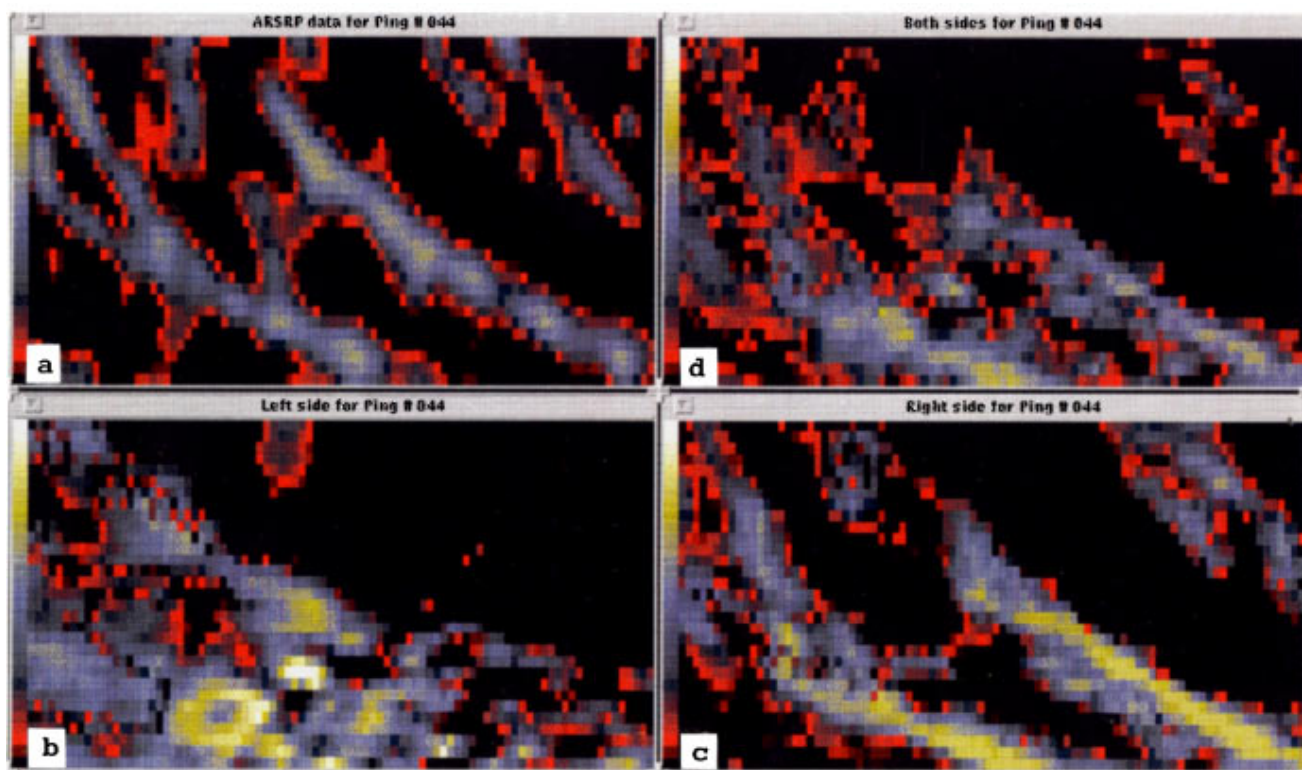


FIG. 12. A-scan plots of reverberation levels for Region 44/1: (a) ARSRP data, (b) simulation: left side, (c) simulation: right side, and (d) simulation: both sides. Time/beam number axes are linear and given in Table I. Intensity scale is strip at left of each and scaled from 0 to 10 dB.

aft beams for ping 77 and to achieve a match we reduced the Lambert coefficient to -20.4 dB. Other than that weak area, the data from Table I indicates relative strong scattering with μ value around -14 dB.

B. Analysis of Ping 76

We use Ping 76 only to confirm the results of Ping 77. Comparisons of Lambert coefficients can be made using Table I. (The slight shift in regions provide for the same geographical coverage.) The mean, rms, and maximum differences between the two pings for the five regions are 0.1, 0.3, and 0.8 dB, respectively. The results are clearly repeatable within a half a dB.

Displays of the data and simulations for Ping 76 are not visibly different from that of Ping 77 (Fig. 7) and are, therefore, not shown. An analysis of detailed differences reveals precisely the expected results for two pings separated by about 1 km. Equivalent figures to Figs. 9 and 10 confirm this, but are not presented here because of their similarity with those figures.

C. Analysis of Ping 44

Ping 44 is from nearly the same location as Pings 76 and 77, but there is a significant difference in the fact that the receiver array is oriented differently. Figure 11 shows the relationship between data and simulation for Ping 44. Site A is in the right forward half of the beams for Ping 44. The ambiguous side for site A is also an outside corner (and geologically similar) and their geographic placement with respect to the array orientation is very nearly the same. Their effects should be difficult to distinguish in the data. However, the data do appear to be more characteristic of the right side (site A). There are only two differences we can discern at this point: (1) the energy is scattered at the ridges on the left at a more oblique angle than the right side thereby reducing the local grazing angle and (2) the energy is scattered from the ridges' eastern faces rather than their western faces, which might have a different character in the microscale. Therefore, the right side, being a stronger scatterer, appears to dominate.

In the aft beams there is evidence of weaker scatter from sedimented areas. On the right aft side the large, deep sedimented pond in front of site A is probably a very weak scatterer. On the left aft side is a shallower sedimented pond, apparently with slightly outcropping ridges, that appears to dominate the scattering. In the aft beams at a much later time (30–35 s), after the signal has traversed the pond on the right side, there is evidence of strong scattering from the mountain at site F' (Fig. 1).

Table I shows selected regions of analysis for Ping 44. The results are similar to those for Pings 76 and 77. Region 44/1 (Fig. 12) was selected to match as closely as possible region 77/1 and 76/1 (average $\mu = -14.1$ dB). Region 44/1 is the mirror image of the other two because of the differences in the orientation of the receiving array. The Lambert coefficient is -15.4 dB. Although this difference is not large (1.3 dB), less than a dB was expected considering a maximum difference for any region between Pings 76 and 77 is

0.8 dB. It is likely that the ambiguous side contributed measurable energy, even though its character cannot be distinguished in the data. We might assign about 0.5 dB to this additive energy.

Region 44/2 includes sedimented ponds on both sides and has a Lambert coefficient of -18.4 dB. The pond off site A is deeply sedimented and probably a relatively weak scatterer compared to the ambiguous side. The pond on the ambiguous side contains outcropping ridges which contribute to its stronger scattering strengths.

IV. SUMMARY, CONCLUSIONS, AND IMPLICATIONS

Early ARSRP analyses determined that, if one were to fit the ARSRP data with Lambert's law, one would find relatively large values for the Lambert coefficient ranging from -17 dB up to -7 dB.¹ We have analyzed some of the data in the vicinity of site A in greater detail applying Lambert's law with a local grazing angle determined by a ray propagation code and local slopes and bearings of 200×200 -m planes. We found that Lambert's law can describe the data analyzed, with an exception noted below, and that regional values of the coefficients were repeatable and closely associated with geological provinces. Outside corners, the distinctly ridged regions, were very strong scatterers with a value of μ around -14 to -15 dB. Inside corners, more seamount like, were slightly weaker scatterers. Sedimented regions had values probably less than -20 dB.

We hold one reservation concerning the applicability of Lambert's law: There appears to be a slight trend toward stronger coefficients for the longer times (longer ranges and shallower grazing angles) in Table I. Since the coefficient should not depend on range, we should seek other sources for the apparent trend. This could indicate a breakdown of Lambert's law, the propagation code is providing a transmission loss biased in range, or the multiple-bounce contamination is becoming more serious as ranges increases. At this point, it would seem that the simplest explanation would be that the scattering strength dependence on grazing angle may be weaker than sine squared. To confirm this, analysis on more pings in greater detail would be required. In the meantime, we believe that Lambert's law is adequate for practical prediction for the Atlantic Natural Laboratory.

Finally, we note that a calculation of the overall average scattering strength for the region could be as much as 3 dB lower because one would be using the full area (we considered only the upper half of the data and therefore half the area) with less energy contributed by the lower half of the data. Thus, $\mu = -17$ dB appears to be a reasonable average.

Seafloor morphology is clearly a controlling factor in the broadscale scattering observed in a rugged region like the Atlantic Natural Laboratory. Prominent features were found to be repeatable for pings from nearby positions, suggesting that variations in the transmission loss is not the cause of the features observed. Lambert's law glosses over the details of unresolved microroughness but, as successfully applied here, requires bathymetry measured at scales an order of magnitude or two above an acoustic wavelength. Because the Lambert coefficient varies over the region, it is likely that the degree and form of surfacial roughness at the microscale

(i.e., microroughness or texture) is also a factor in defining the character of the echos. A successful microscale theoretical scattering model should conform to Lambert's law (or possibly a slightly weaker dependence on grazing angle than sine squared), over a grazing angular range of a few degrees up to about 20°, in the macroscale and explain regional differences. An example of one such theory (presumed to be a conjecture at this point) developed for the Atlantic Natural Laboratory can be found in the paper by Dyer *et al.*¹¹ Their results produce a value for scattering strength at 5° grazing that is consistent with our results using a Lambert coefficient of -14 dB.

ACKNOWLEDGMENTS

We wish to recognize the contributions of the following scientists: Dr. Jeff Simmen, ONR Program Manager, ARSRP, Dr. Brian Tucholke, for collecting, analyzing, and providing the seafloor morphological, Dr. Kevin Smith for UMPE (University of Miami parabolic equation) calculations, and Dr. Ralph Stephen for his comments, extracted and modified from Ref. 4, relevant to the geological interpretation of site A. We also acknowledge the leadership of Dr. John Orcutt and Dr. Arthur Baggeroer that helped make ARSRP'91 and ARSRP'93 significant successes.

¹Scientific Party ARSRP'91, *Initial Report of ARSRP Reconnaissance Cruise: 19 August 1991*, edited by John Orcutt and Arthur Baggeroer (Scripps Institution of Oceanography, La Jolla, CA, 1991).

²Scientific Party ARSRP'93, *Initial Report of ARSRP Acoustics Experiment: 5-26 July 1993*, edited by John Orcutt and Arthur Baggeroer (Scripps Institution of Oceanography, La Jolla, CA, 1993).

³B. E. Tucholke, M. C. Kleinrock, and W. K. Stewart, "Geological and geophysical characteristics of the Acoustic Reverberation Corridor, and their relevance to the conduct of G&G fine-scale surveys," *Proceedings of the Spring '93 Symposium of the ARSRP*, edited by John Orcutt (Scripps Institution of Oceanography, La Jolla, CA, 1993).

⁴J. W. Caruthers, J. R. Fricke, and R. A. Stephen, "Acoustic reverberation at selected sites in the Mid-Atlantic Ridge Region," in *1994 NRL Review* (Naval Research Laboratory, Washington, DC, 1994), pp. 118-122.

⁵J. A. Karson and H. J. B. Dick, "Tectonics of ridge-transform intersections at the Kane Fracture Zone," *Marine Geophys. Res.* **6**, 51-98 (1983).

⁶J. W. Caruthers and B. E. Tucholke, "Geology and geophysics at sites A and C," *Proceedings of the Spring '93 Symposium of the ARSRP*, edited by John Orcutt (Scripps Institution of Oceanography, La Jolla, CA, 1993).

⁷K. V. MacKenzie, "Bottom reverberation for 530 and 1030 cps sound in deep water," *J. Acoust. Soc. Am.* **33**, 1498-1504 (1961).

⁸J. W. Caruthers and J. C. Novarini, "Modeling bistatic bottom scattering strength including a forward scattering lobe," *IEEE J. Oceanic Eng.* **18**, 100-107 (1993).

⁹J. W. Caruthers and E. J. Yoerger, "ARSRP reconnaissance results and BISSM modeling of direct path scatter," *Ocean Reverberation*, edited by D. D. Ellis, J. R. Preston, and H. G. Urban (Kluwer, Dordrecht, The Netherlands, 1993), pp. 203-208.

¹⁰E. J. Yoerger and J. W. Caruthers, "Comparison of low-frequency acoustic backscatter data with model predictions using fine-scale geomorphology," *Proceedings of the Spring '93 Symposium of the ARSRP*, edited by John Orcutt (Scripps Institution of Oceanography, La Jolla, CA, 1993).

¹¹I. Dyer, A. B. Baggeroer, H. Schmidt, J. R. Fricke, N. Ozluer, and D. Giannoni, "Discrete backscatter can be dominant in rough bottom reverberation," *Ocean Reverberation*, edited by D. D. Ellis, J. R. Preston, and H. G. Urban (Kluwer, Dordrecht, The Netherlands, 1993), pp. 51-57.

Coupled-amplitude approach to solving the Helmholtz equation

Dalcio K. Dacol

Naval Research Laboratory, Washington, DC 20375-5350

(Received 12 September 1996; accepted for publication 11 November 1996)

The coupled-amplitude technique for solving the Helmholtz equation has been developed in the context of coupled normal modes by researchers working in a variety of wave propagation problems. In this article it is shown that this approach is not dependent on modal expansions and first-order differential equations in range for the coupled amplitudes are derived without invoking normal-mode expansions. The relationship of this exact transformation to the parabolic approximation is analyzed and numerical methods for solving the coupled-amplitude equations are discussed. The usual range-stepping algorithms used to obtain an approximate solution to the Helmholtz equation are based on the parabolic approximation and restricted to the forward propagating component of the solution. A complete solution of the Helmholtz equation in an inhomogeneous medium must also include backpropagating waves, that is, waves scattered towards the source by inhomogeneities. The inclusion of such effects in a numerically feasible full-wave approach to acoustic propagation is a problem of continual interest in the acoustics of inhomogeneous media and in ocean acoustics. The method discussed in this article addresses this difficult problem. [S0001-4966(97)00805-9]

PACS numbers: 43.30.Ft, 43.20.Bi, 03.40.Kf [JHM]

INTRODUCTION

A complete solution of the Helmholtz equation in an inhomogeneous medium must also include backpropagating waves, that is, waves scattered towards the source by inhomogeneities. This is a problem of enduring interest in acoustics, specially ocean acoustics. The forward propagation problem, where the backscattered energy is considered negligible, has been successfully attacked by approaches based on the parabolic approximation.¹ Progress in this area has been quite extensive and nowadays there are algorithms that can handle quite rough boundaries as well as other strong inhomogeneities in the medium.² This creates a bit of a paradox since the basic tenet of these approximations is the negligibility of backscattering and, obviously, the rougher the boundaries and the stronger the inhomogeneities the more energy will be backscattered. Thus eventually further refinements of forward propagation algorithms will reach (or may already have reached) a point of diminishing returns. In this article the coupled-amplitude technique for solving the Helmholtz equation is explored as a possible way to address the difficult problem of including such effects in a numerically feasible full-wave approach to acoustic propagation.

The coupled-amplitude technique for solving the Helmholtz equation has been developed in the context of coupled normal modes by researchers working in a variety of wave propagation problems. For example, this method has been applied to the problem of optical mode coupling in waveguide optics³ and to the problem of seabed induced mode coupling in ocean acoustics.⁴ In this article it is shown how one can derive first-order differential equations in range for the coupled amplitudes without invoking normal-mode expansions. The relationship of this exact transformation to the parabolic approximation is analyzed and numerical methods for solving the coupled-amplitude equations are discussed.

In Sec. I the formulation of the coupled-amplitude

method is discussed using the case of an oceanic waveguide as an example. The range-independent case, where a complete solution can be obtained in terms of normal modes is discussed in Sec. II. Section III addresses a possible numerical implementation of the coupled-amplitude method in the case of an oceanic waveguide. Finally, Sec. IV summarizes the article and indicates possible new developments.

I. THE COUPLED-AMPLITUDE METHOD

The proposed approach will be illustrated with the case of an oceanic wave guide. For simplicity we neglect the azimuthal dependence here. It can easily be incorporated in most of mathematical developments carried out in this paper. Consider the Helmholtz equation for the pressure field in range and depth with a cylindrically symmetric point source,

$$\left[\partial_r^2 + \frac{1}{r} \partial_r + \partial_z^2 + k(r, z)^2 \right] P(r, z) = \frac{1}{2\pi r} \delta(r) \delta(z - z_s). \quad (1)$$

In the above equation $k(r, z) = \omega/c(r, z)$, where $c(r, z)$ is the sound-speed profile.

The radial Green's function $G(r, r')$, where

$$\left(\partial_r^2 + \frac{1}{r} \partial_r + k_0^2 \right) G(r, r') = \frac{1}{2\pi r} \delta(r - r') \quad (2)$$

can then be used to rewrite the Helmholtz equation as a pair of coupled partial differential equations. The quantity k_0 is the reference wave number, of magnitude typical of the sound channel wave number. The radial Green's function has a simple expression in terms of Bessel and Hankel functions:

$$G(r, r') = \frac{1}{4i} H_0^{(1)}(k_0 r_>) J_0(k_0 r_<). \quad (3)$$

A standard procedure,⁵ allows one to use the explicit form of the radial Green's function to write the pressure as a superposition of two coupled waves:

$$P(r, z) = \frac{1}{4i} H_0^{(1)}(k_0 r) \alpha(r, z) + \frac{1}{4i} J_0(k_0 r) \beta(r, z). \quad (4)$$

The amplitudes $\alpha(r, z)$ and $\beta(r, z)$ obey the following integro-differential equations:

$$\begin{aligned} \alpha(r, z) = & \alpha(\epsilon, z) - 2\pi \int_{\epsilon}^r r' dr' J_0(k_0 r') \\ & \times [\partial_z^2 + k(r', z)^2 - k_0^2] P(r', z), \end{aligned} \quad (5)$$

$$\begin{aligned} \beta(r, z) = & \beta(R, z) - 2\pi \int_r^R r' dr' H_0^{(1)}(k_0 r') \\ & \times [\partial_z^2 + k(r', z)^2 - k_0^2] P(r', z). \end{aligned} \quad (6)$$

The point-source condition implies $\alpha(0, z) = \delta(z - z_s)$. The long-range limit of the amplitude $\beta(r, z)$, $\beta(R, z)$, must be determined such that the radiation boundary condition is satisfied as $R \rightarrow \infty$. A method for implementing the radiation boundary condition for very large ranges is discussed further below in this section.

Differential equations for the amplitudes $\alpha(r, z)$ and $\beta(r, z)$ are obtained by differentiating the integral equations with respect to r . At this point it is convenient to introduce two new amplitudes

$$a = \alpha + \beta/2, \quad b = -\beta/2, \quad (7)$$

which allow for the pressure to be written in a more symmetrical way in terms of Hankel functions. The differential equations for $a(r, z)$ and $b(r, z)$ also have a more symmetrical form:

$$\partial_r a = i\gamma(r) [\partial_z^2 + k(r, z)^2 - k_0^2] (a - e^{-i\phi(r)} b), \quad (8)$$

$$\partial_r b = i\gamma(r) [\partial_z^2 + k(r, z)^2 - k_0^2] (e^{i\phi(r)} a - b). \quad (9)$$

In the above equations $\gamma(r) = (\pi r/4) |H_0^{(1)}(k_0 r)|^2$ and $\phi(r)$ is the phase of $[H_0^{(1)}(k_0 r)]^2$. Notice that if one neglects the amplitude $b(r, z)$, the one associated with the backpropagating wave $H_0^{(2)}(k_0 r)$, then the above equations reduce to the parabolic approximation¹ for the forward amplitude $a(r, z)$.

It is also useful to write the coupled amplitudes $a(r, z)$ and $b(r, z)$ in terms of $P(r, z)$:

$$a(r, z) = 4\pi i r \{ [\partial_r f(k_0 r)]^* P(r, z) - f(k_0 r)^* \partial_r P(r, z) \}, \quad (10)$$

$$b(r, z) = -4\pi i r \{ [\partial_r f(k_0 r)] P(r, z) - f(k_0 r) \partial_r P(r, z) \}, \quad (11)$$

where $f(x) = (1/4i) H_0^{(1)}(x)$.

It must be pointed out that the amplitudes $a(r, z)$ and $b(r, z)$ cannot, in general, be identified with forward and backpropagating waves in the medium. This is clear if one constructs the range-independent solution. In such a case both amplitudes $a(r, z)$ and $b(r, z)$ are nonzero and yet there are no actual backpropagating waves. This is not a paradox, the fact is that in an inhomogeneous media the waves will not have the same wave number. For example, in the strati-

fied medium case, although the environment is range independent the modal character of the propagation implies a set of wave numbers and not just one. Nonetheless, a correct determination of $a(r, z)$ and $b(r, z)$ yields the correct solution to the Helmholtz equation.

In this context it is interesting to observe that the flow of energy per unit azimuthal angle, per unit of length in depth, and per unit of time along a radial direction,

$$J_r = \frac{r}{2i\omega\rho} [P^* \partial_r P - P \partial_r P^*], \quad (12)$$

where ω is the angular frequency and ρ is the density, has a simple expression in terms of a and b , namely,

$$J_r(r, z) = \frac{1}{8\pi\omega\rho} [|a(r, z)|^2 - |b(r, z)|^2]. \quad (13)$$

It follows that if $k(r, z)^2$ is real, that is, if there is no attenuation, then $\partial_r J_r(r, z) = 0$.

The differential equations for the coupled amplitudes can be formally solved using propagation operators. In vector notation, with the range r being the propagation parameter, and with

$$\Psi(r) = \begin{bmatrix} a(r) \\ b(r) \end{bmatrix} \quad (14)$$

being a complex column vector, one has the following equation for the range evolution of $\Psi(r)$:

$$\Psi(r) = T(r, r') \Psi(r') \quad (15)$$

or, reintroducing the coordinate z to label the components of $\Psi(r)$,

$$\Psi(r, z) = \int dz' T(r, z | r', z') \Psi(r', z'). \quad (16)$$

In the above equations $T(r, r')$ is the propagation operator between ranges r' and r . From this definition of the propagation operator it follows that it satisfies the following equation:

$$\partial_r T(r, r') = iM(r) T(r, r'), \quad (17)$$

with the initial condition

$$T(r', r') = I, \quad (18)$$

where I is the identity operator. The operator $M(r)$ is defined as

$$M(r) = \begin{bmatrix} k_0^2 \gamma(r) X(r) & -e^{-i\phi(r)} k_0^2 \gamma(r) X(r) \\ e^{i\phi(r)} k_0^2 \gamma(r) X(r) & -k_0^2 \gamma(r) X(r) \end{bmatrix}, \quad (19)$$

where the functions $\phi(r)$ and $\gamma(r)$ were defined before and $X(r)$ is the operator $(1/k_0^2) [\partial_z^2 + k(r, z)^2 - k_0^2]$.

A noteworthy property of the propagator $T(r, r')$ is that its determinant is equal to the determinant of the identity operator which, by definition, is 1. See Ref. 6 for details on trace and determinants of operators. One starts from the identity

$$\text{Det}[T(r, r')] = \exp \text{Tr}[\ln(T(r, r'))], \quad (20)$$

and one must remember that r and r' are just range parameters as far as the operator manipulations are concerned. From the above equation it follows that

$$\partial_r \text{Det}[T(r, r')] = (\text{Det}[T(r, r')]) \text{Tr}[(\partial_r T(r, r')) T(r, r')^{-1}], \quad (21)$$

and so,

$$\partial_r \text{Det}[T(r, r')] = i(\text{Det}[T(r, r')]) \text{Tr}[M(r)] = 0. \quad (22)$$

Thus, $\text{Det}[T(r, r')] = \text{Det}[T(r', r')] = \text{Det}[I] = 1$. This property assures that the incoming flow of acoustic energy into the region delimited by the ranges r' and r and the outgoing flow are equal in the absence of attenuation.

In order to obtain the pressure field from the propagation operator one has to impose the point-source condition and the radiation boundary condition. It was pointed out earlier in this section that the point-source condition implies $a(0, z) = \delta(z - z_s)$, or

$$a(0, z) + b(0, z) = \delta(z - z_s). \quad (23)$$

The above equation can be used to eliminate $a(0, z)$ from the expression for $\Psi(0, z)$, leaving $b(0, z)$ as the only unknown quantity at range $r=0$. The choice of keeping $b(0, z)$ as the unknown is motivated by the fact that $\beta(r, z) = -2b(r, z)$ is the coefficient of the Bessel function in the equation for the pressure in terms of the amplitudes, Eq. (4), and thus $\beta(r, z)$ should be a well-behaved function near the source. That is, only $a(0, z)$ contains the singular function $\delta(z - z_s)$.

A practical way of imposing the radiation boundary condition for very large ranges is to assume that the environment is range independent for $r \geq R$ where R is a suitably large range value. In range-independent environments one can introduce depth-dependent normal modes $u_m(z)$ with associated eigenwave number k_m . The asymptotic pressure field for $r \geq R$ can be written as

$$P_a(r, z) = \sum_m A_m f(k_m r) u_m(z), \quad (24)$$

where A_m is an unknown amplitude. From the above expression one can then find $a(R, z)$ and $b(R, z)$ in terms of the A_m using Eqs. (10) and (11). From the definition of the propagator, see Eq. (15), one obtains

$$\Psi(R) = T(R, 0) \Psi(0). \quad (25)$$

The above equation and the source condition, Eq. (23), allow one to determine $\Psi(0)$ such that one can use Eq. (15) to generate the pressure field due to a point source which obeys the Sommerfeld radiation condition for $r \geq R$. Using the modal expression for $\Psi(R, z)$ one ends up with an integral equation for $b(0, z)$. Using the completeness property of normal mode sets one can show that such integral equation does not depend explicitly on normal modes, only on the outgoing Green's function for the asymptotically range-independent medium. The integral equation satisfied by $b(0, z)$ is the following

$$\int dz' K(z, z') b(0, z') = \Phi(z), \quad (26)$$

where

$$\begin{aligned} \Phi(z) = \int dz' [A(z, z') T_{21}(R, z' | 0, z_s) \\ - B(z, z') T_{11}(R, z' | 0, z_s)], \end{aligned} \quad (27)$$

$$\begin{aligned} K(z, z') = \int dz'' \{ B(z, z'') [T_{12}(R, z'' | 0, z') \\ - T_{11}(R, z'' | 0, z')] - A(z, z'') [T_{22}(R, z'' | 0, z') \\ - T_{21}(R, z'' | 0, z')] \}, \end{aligned} \quad (28)$$

with $A(z, z')$ and $B(z, z')$ being related to the Helmholtz equation Green's function in the asymptotic range-independent environment as follows:

$$\begin{aligned} A(z, z') = 4\pi i r [(\partial_r f(k_0 r))^* \\ - f(k_0 r)^* \partial_r] G_a(r, z | 0, z')|_{r=R} \end{aligned} \quad (29)$$

and

$$\begin{aligned} B(z, z') = -4\pi i r [(\partial_r f(k_0 r)) \\ - f(k_0 r) \partial_r] G_a(r, z | 0, z')|_{r=R}. \end{aligned} \quad (30)$$

In terms of normal modes the Green's function $G_a(r, z | 0, z')$ can be written as

$$G_a(r, z | 0, z') = \sum_m f(k_m r) u_m(z) u_m(z'), \quad (31)$$

Once one has obtained $b(0, z)$ by solving the above equation, which implies having solving the equations for the propagator, one then can use the propagation equation, Eq. (15), to generate the acoustic field for all desired ranges. The method described here is an extension of a well-known procedure for solving one-dimensional boundary value problems, see Ref. 7, to the ocean acoustic propagation problem. Similar procedures are described in Refs. 8 and 9.

II. THE RANGE-INDEPENDENT CASE

In the case of a range-independent environment both $\Psi(r, z)$ and $T(r + \epsilon, z | r, z')$ can then be expressed in terms of the normal modes:

$$\Psi(r, z) = \sum_m \Psi_m(r) u_m(z), \quad (32)$$

and

$$T(r + \epsilon, z | r, z') = \sum_m T^m(r + \epsilon, r) u_m(z) u_m(z'), \quad (33)$$

where now Ψ_m is just a two-dimensional complex vector and $T^m(r + \epsilon, r)$ is just a 2×2 complex matrix.

In this case the range equations can be solved in terms of Hankel functions and explicit expressions are obtained for $T^m(r + \epsilon, r)$. Using the far-field (large $k_m r$) approximations for the Hankel functions one obtains

$$T_{1,1}^m = e^{-ik_0\epsilon} \left[\cos(k_m\epsilon) + i \frac{k_m^2 + k_0^2}{2k_mk_0} \sin(k_m\epsilon) \right], \quad (34)$$

$$T_{1,2}^m = e^{-ik_0(2r+\epsilon)} \frac{k_m^2 - k_0^2}{2k_mk_0} \sin(k_m\epsilon), \quad (35)$$

$$T_{2,1}^m = e^{ik_0(2r+\epsilon)} \frac{k_m^2 - k_0^2}{2k_mk_0} \sin(k_m\epsilon), \quad (36)$$

$$T_{2,2}^m = e^{ik_0\epsilon} \left[\cos(k_m\epsilon) - i \frac{k_m^2 + k_0^2}{2k_mk_0} \sin(k_m\epsilon) \right]. \quad (37)$$

Notice that $\text{Det}[T^m] = 1$, this is also true for the exact expression in terms of Hankel functions, not just for the far-field expressions. This, of course, is a general property of the propagator as shown in Sec. I.

The above expressions for the propagator operator can be used as the foundation for an alternative formulation of a numerical scheme for coupled normal mode propagation as is discussed in the next section.

III. THE RANGE-DEPENDENT CASE

In the case of a range-dependent environment a numerical solution is the only possibility. The basic approach (see, for example, Ref. 2) consists in choosing ϵ small enough so that in the restricted range between $r + \epsilon$ and r the medium is approximately homogeneous and one can then use the range-independent approximation for $T(r + \epsilon, r)$. The propagator for longer ranges and the pressure field are then obtained by repeated and iterative application of $T(r + \epsilon, r)$.

To further develop the above idea it is convenient to introduce another operator related to the propagator. As noticed in the previous section, the operator $T(r, 0)$ plays a special role when the source is located at range 0. Thus let us define the operator $U(r)$ as

$$U(r) = T(r, 0). \quad (38)$$

Obviously this operator yields $\Psi(r)$ given $\Psi(0)$:

$$\Psi(r) = U(r)\Psi(0). \quad (39)$$

The operator $U(r)$ satisfies the following equation

$$U(r) = T(r, r')U(r'), \quad (40)$$

with the obvious initial condition

$$U(0) = I, \quad (41)$$

where I is the identity operator. Thus $U(R)$ can be obtained, in an approximation scheme, by iteration:

$$U(r + \epsilon) = T(r + \epsilon, r)U(r), \quad (42)$$

starting with $U(0) = 1$ and a suitable approximation for $T(r + \epsilon, r)$, like the one mentioned above that assumes that for small ϵ the environment between ranges r and $r + \epsilon$ is range independent. Once $U(R)$ is computed then $\Psi(0)$ is obtained in the way outlined in Sec. I. Notice that in Eqs. (27) and (28), the $T_{ij}(R, z''|0, z')$ are just the kernels of $U(R)$.

An approximation for $T(r + \epsilon, r)$ can be obtained, for example, by using local normal modes and the expression for

the propagator in terms of normal modes derived in Sec. II. One ends up with an alternative formulation of the coupled normal-mode propagation model.⁸ This numerical approach can be implemented in the following way: At range r the normal modes and the acoustic field $\Psi(r)$ are available, then the acoustic field at $r + \epsilon$ is obtained by using

$$\Psi(r + \epsilon, z) = \int dz' T(r + \epsilon, z|r, z')\Psi(r, z'), \quad (43)$$

with $T(r + \epsilon, z|r, z')$ approximated by its normal-mode expansion

$$T(r + \epsilon, z|r, z') = \sum_m T^m(r + \epsilon, r)u_m(z; r)u_m(z'; r), \quad (44)$$

where $u_m(z; r)$ indicates the mode function at depth z and range r . The modal components of $\Psi(r, z)$ are given by

$$\Psi_m(r) = \int dz u_m(z; r)\Psi(r, z). \quad (45)$$

Thus Eq. (43) leads to the following expression for $\Psi_m(r + \epsilon)$, the modal coefficients for the acoustical field $\Psi(r + \epsilon, z)$ in terms of the mode functions at range $r + \epsilon$, $u_m(z; r + \epsilon)$:

$$\Psi_m(r + \epsilon) = \sum_n C_{m,n}(r)T^m(r + \epsilon, r)\Psi_n(r), \quad (46)$$

where the mode coupling coefficients $C_{m,n}$ are given by

$$C_{m,n}(r) = \int dz u_m(z; r + \epsilon)u_n(z; r). \quad (47)$$

Now one repeats the cycle with $r \rightarrow r + \epsilon$, $r + \epsilon \rightarrow r + 2\epsilon$. Analogous equations hold for computing $U(R)$.

A more flexible approach, and one that should be particularly well-suited to the case of rough boundaries and interfaces can be obtained by following in this new framework the approach outlined by Ref. 2 in the case of the parabolic approximation. One uses the small range increment approximation (small ϵ) for the same reasons stated above but instead of using the modal expansion for $T(r + \epsilon, r)$ one writes this operator in terms of the depth operator X :

$$X = \frac{1}{k_0^2} (\partial_z^2 + k(r, z)^2 - k_0^2). \quad (48)$$

The matrix elements of the operator $T(r + \epsilon, r)$ are then

$$T_{1,1} = e^{-ik_0\epsilon} \left[\cos(k_0\epsilon\sqrt{1+X}) + i \left(\frac{2+X}{2\sqrt{1+X}} \right) \sin(k_0\epsilon\sqrt{1+X}) \right], \quad (49)$$

$$T_{1,2} = e^{-ik_0(2r+\epsilon)} \left(\frac{X}{2\sqrt{1+X}} \right) \sin(k_0\epsilon\sqrt{1+X}). \quad (50)$$

One obtains $T_{2,2}$ from $T_{1,1}$ and $T_{2,1}$ from $T_{1,2}$ by letting $i \rightarrow -i$. Notice that this is not just a complex conjugation operation. In particular, the operator $\sqrt{1+X}$ can be complex and it is implicit in the above formulas that one uses the

same branch of the square root in all the expressions.

The next step here would be to develop Padé approximants for the elements of the T operator as is done in Ref. 2 for the operators

$$Z_{\pm} = e^{\pm ik_0 \epsilon \sqrt{1+X}}. \quad (51)$$

The cosine terms can be easily expressed in terms of the Z_{\pm} operators:

$$\cos(k_0 \epsilon \sqrt{1+X}) = \frac{Z_+ + Z_-}{2}. \quad (52)$$

The terms involving sines should be treated together with the $\sqrt{1+X}$ in the denominator

$$\frac{\sin(k_0 \epsilon \sqrt{1+X})}{\sqrt{1+X}} = \frac{W_+ - W_-}{2i}, \quad (53)$$

where

$$W_{\pm} = \frac{Z_{\pm} - 1}{\sqrt{1+X}}, \quad (54)$$

then Padé approximants for W_{\pm} can be used. In this way a rational function approximant is obtained for the operator $T(r + \epsilon, r)$.

Due to the presence of evanescent waves and their counterpart of exponentially increasing solutions the action of the operator Z_+ may be vanishingly small while the action of the operator Z_- may lead to exponentially increasing solutions. This is to be expected in a numerical solution of hyperbolic differential equations such as the Helmholtz equation. This presents a technical difficulty that has to be handled in an approximate way. Thus, in order to use the approach of Ref. 2 one has to use the following approximations:

$$Z_- \Psi \approx (Z_+ \Psi^*)^*, \quad (55)$$

and, similarly, $W_- \Psi \approx (W_+ \Psi^*)^*$. This will eliminate the exponentially increasing solutions from the numerical treatment.

IV. SUMMARY AND CONCLUSIONS

The coupled amplitude method for the Helmholtz equation developed in Sec. I is not predicated on an existing normal-mode expansion. This is different from previous applications of this method in acoustics and optics which used it to solve the ordinary differential equations for coupled modes. In particular, this formulation allows one to use state-of-the-art techniques developed for the well-known parabolic approximation to solve the Helmholtz equation without neglecting the backscattered waves.

The Green's function method that was used in Sec. I is a well-known technique that can be employed to replace a second-order differential with a pair of coupled first-order

differential equations. Thus, for example, in Ref. 10, this approach is used to convert the one-dimensional Helmholtz equation into a pair of coupled amplitude differential equations. In the same article the higher dimensional problems are initially reduced to one-dimensional problems by modal expansions before applying the coupled wave method (see also Ref. 11). Thus the approach outlined in that article is another example of the usual application of the coupled-amplitude method to solve coupled normal-mode problems and quite distinct in spirit from what is proposed in this article.

There are other methods of transforming the Helmholtz equation into an equivalent pair of coupled first-order differential equations in range. Most of those approaches involve the introduction of pseudodifferential operators and in fact result in integro-differential equations for the coupled components of the pressure field. An interesting and rigorous approach along these lines is discussed in Ref. 9, for example. The structure of these integro-differential equations, as far as propagation in range is concerned, is similar to the coupled-amplitude equations. However, one must notice that the coupled-amplitude equations derived in Sec. I are just coupled linear and local partial differential equations of first order in the range variable. All the terms in the coupled-amplitude equation have a simple and explicit expression for them thus facilitating the analysis.

Further work will involve a study of the treatment of rough boundaries and interfaces as well as the implementation of the numerical algorithms based on the Padé approximants and finite differences for the depth operator as outlined in Sec. III.

¹F. D. Tappert, "The Parabolic Equation Method," in *Wave Propagation and Underwater Acoustics*, edited by J. B. Keller and J. S. Papadakis (Springer-Verlag, Berlin, 1977).

²M. D. Collins, "A split-step Padé solution for the parabolic equation method," *J. Acoust. Soc. Am.* **93**, 1736–1742 (1993).

³H. Kogelnik, "Theory of Dielectric Waveguides," in *Integrated Optics*, edited by T. Tamir (Springer-Verlag, Berlin, 1979).

⁴S. T. McDaniel, "Mode coupling due to interaction with the seabed," *J. Acoust. Soc. Am.* **72**, 916–923 (1982).

⁵P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vol. I, Chap. 7.

⁶M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (Academic, New York, 1978), Vol. IV.

⁷N. Fröman and P. O. Fröman, *JWKB Approximation* (North-Holland, Amsterdam, 1965).

⁸R. B. Evans, "A coupled mode solution for acoustic propagation in a waveguide with stepwise depth variations of a penetrable bottom," *J. Acoust. Soc. Am.* **74**, 188–195 (1983).

⁹L. Fishman, "One-way wave propagation methods in direct and inverse scalar wave propagation modeling," *Radio Sci.* **28**, 865–876 (1993).

¹⁰D. G. Hall, "A comment on the coupled-mode equations used in guided-wave optics," *Opt. Commun.* **82**, 453–455 (1991).

¹¹D. G. Hall, "Coupled-amplitude equations via a Green's-function technique," *Am. J. Phys.* **61**, 44–49 (1993).

Low-frequency acoustic scattering by a cube: Experimental measurements and theoretical predictions

Paul A. Chinnery, Victor F. Humphrey, and Jingdong Zhang
School of Physics, University of Bath, Bath BA2 7AY, United Kingdom

(Received 27 July 1996; accepted for publication 15 January 1997)

The acoustic backscattering from a submerged metal (iron) cube has been investigated as a function of frequency and angle of incidence in the low-frequency regime $ka < 5$, where k is the acoustic wave number and a is a characteristic dimension of the cube. The experimental measurements were made in the laboratory using a parametric array as an acoustic source at frequencies between 8 and 120 kHz and a cube of side length 20 mm. The backscattered field is expressed as a dimensionless "form function" and normalized in terms of the average projected cross-sectional area of the cube. The scattered field is seen to vary considerably with the orientation of the cube. The scattering behavior of a rigid cube is modeled numerically using a boundary element code. The results for a cube are compared with calculations of the form function for an elastic sphere. When averaged over cube orientations the form function has characteristics reminiscent of a spherical scatterer, suggesting the propagation of creeping waves. © 1997 Acoustical Society of America. [S0001-4966(97)04105-2]

PACS numbers: 43.30.Gv, 43.20.Fn [ANN]

INTRODUCTION

The use of acoustic backscattering and attenuation measurements to monitor the concentration and transport of suspended sediment in the marine environment has recently been studied by a number of authors¹⁻⁹ and instrumentation based on multi-frequency techniques has been successfully developed.^{4,7} One important advantage of such techniques is that they are noninvasive, permitting measurements to be made remotely, without disturbing the sediment or the natural environment in which the sediment exists. All such acoustic techniques depend on the interaction between the sediment particles and the acoustic wave. There are many factors which influence the scattering process, including the distribution, shape, size, and material properties of the particles. Of particular interest at present is the relationship between the particle shape and the scattered field. Until recently particles have been assumed, for simplicity, to be spherical. However, realistic sediment particles tend to be irregular, having facets and corners, unlike the relatively simple spherical scatterers considered in the past. Although there are certain circumstances under which sediments can be modeled as a distribution of elastic spheres, laboratory studies have shown many sediment particles to be ellipsoidal, cubic, octahedral, or even flat in geometry.¹⁰ Acquiring a better understanding of the backscattering properties of particles having similar shapes to natural sand will be of great benefit in advancing the ultrasonic spectroscopy techniques used in sedimentology.

The first stage in a detailed investigation of scattering by suspensions of nonspherical particles is to develop an understanding of how a single scattering particle behaves. The special case of an elastic spherical scatterer is now well understood. The frequency dependence of the scattered field, expressed as the form function, has a structure which is readily explained, using resonance scattering theory,¹¹ in terms of rigid body scattering, surface waves, and elastic

resonances. Many questions arise when nonspherical scattering objects are considered, in particular with regard to the propagation of surface waves, and the influence of edges and facets. There is, therefore, a need for accurate controlled measurements and validated theoretical predictions for scattering from objects of greater complexity.

Theoretical investigations of nonspherical scatterers have been severely limited in the past by the fact that the wave equation is only separable in a few special geometries. Low-frequency or high-frequency approximations have had to be made to facilitate solution, severely restricting the range of applicability of many theories. However, with the increasing speed and power of modern computers, such complex geometries are becoming amenable to solution by both analytic and numerical methods.

A limited number of theoretical approaches are available for studying scattering by particles of complex geometry. Some aspects can be investigated in two dimensions, using, for example, conformal transformations. It is worth mentioning that DiPerna and Stanton¹² have recently introduced a new technique, referred to as the Fourier matching method (FMM), based on such conformal mappings. The authors of that paper have presented results for cylinders of square cross section that are of particular interest in the context of the present study. However, it seems unlikely that such methods can be generalized to higher dimensions, and the only realistic solutions in three dimensions are currently numerical.

This paper presents the results of a parallel experimental and theoretical study of the backscattering properties of a solid cube. The experimental program was performed on relatively large cubes, rather than small sediment scale particles, in order to facilitate accurate measurements over a range of angles. The measurements were undertaken using a 20-mm iron cube over the frequency range 8–120 kHz, corresponding to a ka range of approximately 0.5–6 (where a is

a characteristic dimension of the cube; see Sec. II C). The variation of backscattered pressure with both the frequency and the orientation of the cube is expressed using the “form function” approach originally developed for spherical and cylindrical scatterers. A form function for the cube is defined in terms of the average projected cross-sectional area of the cube, a quantity which, we believe, correctly satisfies energy considerations in the total scattering problem. The correct definition of form function is important if comparisons are to be made with the more familiar spherical scattering. The measurements are compared with numerical predictions of the backscattered form function for a rigid cube. The numerical approach uses a combined finite element and boundary element code (PAFEC¹³). In the low-frequency range of interest in the present study it was not considered necessary to use finite elements to model the particle.

In Sec. I the experimental system and measurement technique are described. The average projected cross-sectional area of the cube is calculated in Sec. II and used to define a dimensionless normalized form function for the cubic scatterer. Numerical details are presented in Sec. III, and the results of a parallel experimental and theoretical investigation into scattering from the cube are given in Sec. IV.

I. EXPERIMENTAL SYSTEM AND MEASUREMENT TECHNIQUE

The measurements presented in this paper were made using a laboratory facility incorporating a parametric array.^{14,15} A parametric array utilizes the nonlinear propagation of high amplitude waves to produce signals at low frequencies.^{16,17} A conventional transducer is used to generate high amplitude, high-frequency waves which propagate nonlinearly through the water and interact to produce additional frequency components. If the primary wave consists of two high-frequency components that are close together, then a low-frequency component is generated at the difference frequency (between primary wave components). This nonlinear generation is a continuous process and the low-frequency wave can be considered to come from secondary sources distributed throughout the interaction region of the primary waves. The distributed nature of the secondary sources, and their phases, produces a behavior similar to that of an end-fire array, and results in an exceptionally narrow beam. The use of this source permits low ka measurements to be made on model scattering objects having a reasonable size. In addition, the parametric source can be used to generate a broadband pulse whose frequency content is easily adjusted by changing the primary frequencies transmitted by the transducer. This allows measurements to be made over a broad frequency (ka) range with a single transmitting transducer and a single cube. Basic details of the experimental system and measurement technique are given in this section; further details can be found in Refs. 14 and 15.

A. Experimental arrangement

The experimental arrangement is shown in Fig. 1. The measurements were made in a large water-filled tank (approximately $2.7 \text{ m} \times 1.3 \text{ m} \times 1.5 \text{ m}$) using a parametric array as the acoustic source. The parametric array was truncated by

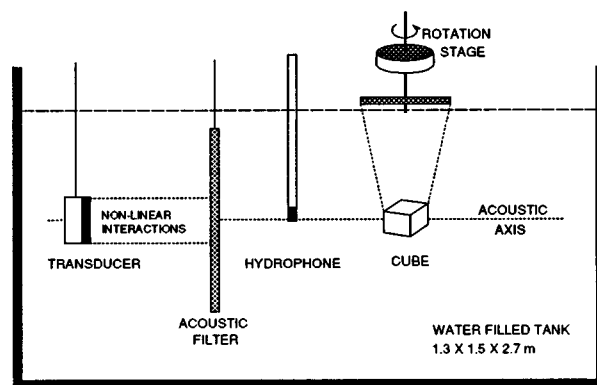


FIG. 1. Laboratory system incorporating a parametric array.

an acoustic filter in order to isolate a source-free region in the tank in which measurements could be made without problems associated with the nonlinear response of the hydrophone^{18,19} and the interaction of the primary waves with the test particle itself. The transducer, acoustic filter, hydrophone, and test particle were all suspended from an optical bench spanning the length of the tank, parallel to the acoustic axis. In this way the positions of the components could easily be varied.

The system electronics are shown in Fig. 2. Instead of transmitting two distinct frequencies, an amplitude modulated tone burst [Fig. 3(b)] was used to drive the transducer. This was generated by amplitude modulating a continuous sine wave (972 kHz) with an “envelope” signal [Fig. 3(a)] produced by a waveform generator (Phillips PM5133). The resulting demodulated pulse (c) generated by the parametric array has a spectrum which is easily altered by adjusting the length and shape of the modulating envelope, allowing an

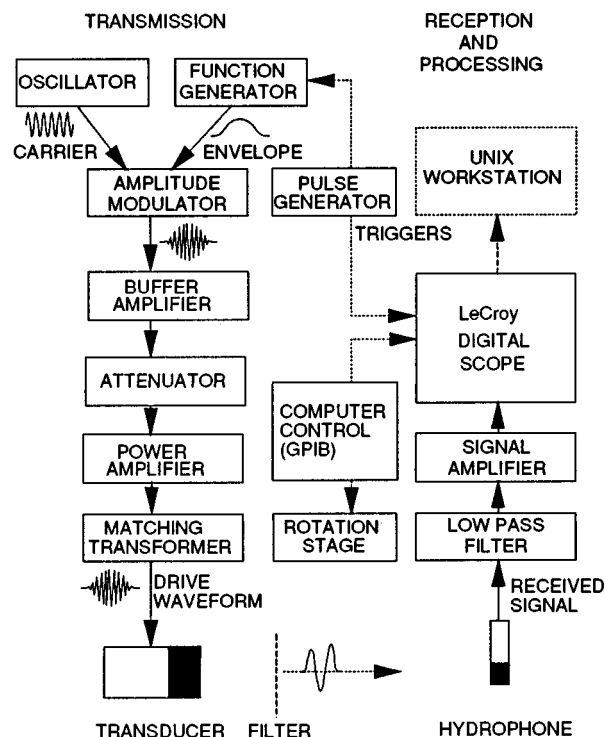


FIG. 2. System electronics and generation of the parametric array.

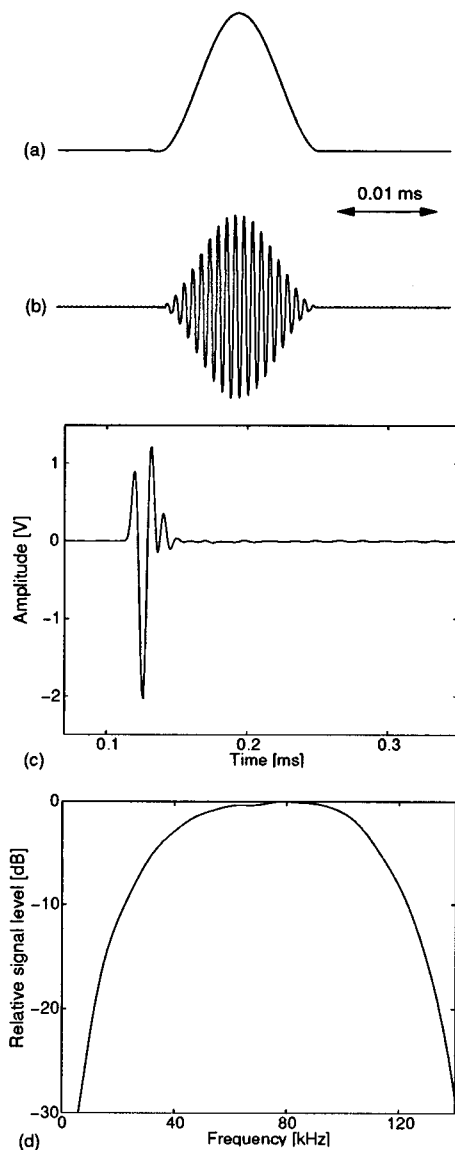


FIG. 3. Generation of the low-frequency pulse: (a) raised cosine-bell modulating envelope based upon a single cycle of 50 kHz; (b) modulated tone burst applied to the transducer; (c) demodulated pulse (after passage through receiving system); (d) spectrum of (c).

extended frequency range to be covered. The modulating envelope used in the present study consisted of a “raised cosine-bell” [Fig. 3(a)] obtained by adding a dc offset to a single cycle, phase shifted, cosine wave. The modulated pulse (b) was amplified and passed, via a matching transformer, to a conventional piezoelectric transducer, 50 mm in diameter and resonant at 972 kHz.

The signal detected by the hydrophone (Brüel & Kjær 8103) was passed through a passive low-pass (250-kHz) filter, a signal amplifier (Brookdeal 9452), and digitized by a LeCroy 9310A digital oscilloscope. Typically 500 μ s of signal was captured using a 1- μ s sampling interval. To improve the signal to noise ratio the LeCroy typically averaged 500–1000 incoming signals. The waveforms were stored on a floppy disk and transferred to a UNIX workstation for further processing. The pulse repetition frequency of 20 Hz was controlled by a pulse generator triggering the function generator and oscilloscopes.

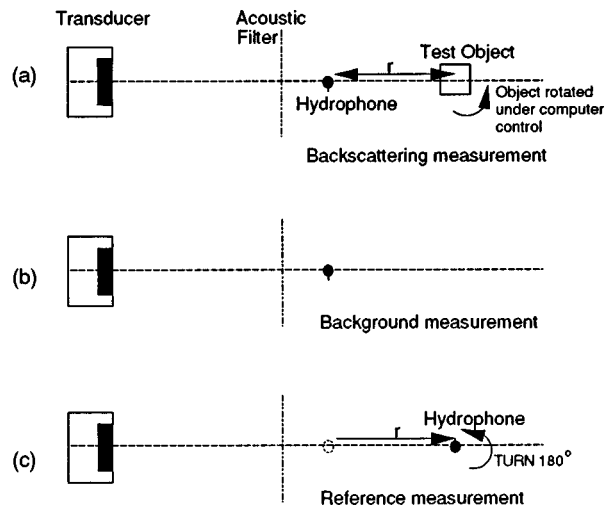


FIG. 4. Scattering geometry and measurement technique: (a) total backscattered field recorded with test object in different configurations; (b) background reverberation measured with test object absent; (c) reference measurement made at position of test object.

An example of the resulting demodulated pulse (after passage through the receiving system) and its spectra is shown in Fig. 3(c) and (d). The width of the low-frequency beam incident on the scatterer is a function of frequency, but was many times greater than the dimensions of the cube used in this study. The required frequency range of 8–120 kHz was covered by using two pulses based on modulating envelopes of 20 and 50 kHz.

B. Measurement technique

The test particle was suspended from a rotation stage by nylon thread and careful vertical and horizontal adjustments were used to position and orientate the particle accurately on the acoustic axis.

For backscattering measurements the hydrophone was typically placed at a distance of 275 mm from the center of the test particle. This distance was sufficiently far from the particle to approximate the far field, and yet close enough to ensure that the scattered signals had a reasonable amplitude. The parametric array was truncated by the acoustic filter at a distance of 53 cm from the transducer and the test particle was suspended at 102 cm. The temperature in the tank varied, but was typically 10 °C.

In order to make accurate measurements of the scattered field it was necessary to remove the effect of any coherent reverberation. This was achieved using a technique illustrated in Fig. 4. First the total field in the backscatter position was measured at a fixed distance (r) from the center of the test particle [Fig. 4(a)]. Then the test particle was removed and the “background” reverberation signal measured with the hydrophone in the same position (b). This background signal was then subtracted from the total signal, leaving only the backscattered signal due to the test particle. Finally the incident (or “reference”) signal was measured by moving the hydrophone to the position of the center of the test par-

ticle (with the test particle absent) and rotating the hydrophone by 180° to compensate for any directionality of the hydrophone (c). In practice a series of backscattering measurements was made by rotating the cube under computer control and storing the waveforms on disk.

The averaged signals were windowed to remove unwanted multiple reflections (from the various components in the tank), and Fourier transformed to obtain their spectra. The resulting backscattered spectrum (with background reverberation subtracted) and reference spectrum were used to calculate the form function. The modulus of the experimentally determined dimensionless form function for a three-dimensional scatterer is given by

$$|f(ka)| = \frac{2r}{a} \frac{|P_{sc}(\nu)|}{|P_{in}(\nu)|}, \quad (1)$$

where ν is the frequency, k is the acoustic wave number, $P_{sc}(\nu)$ and $P_{in}(\nu)$ are the spectra of the scattered and incident signals, respectively, and a is a characteristic dimension of the test particle. The choice of a is discussed in Sec. II.

In backscattering measurements the receiving hydrophone had to be placed in the path of the incident beam and therefore caused a shadowing effect. A detailed discussion of this problem is given in Ref. 15.

As an example of the performance of the system, consider the results obtained for a solid steel sphere shown in Fig. 5. The sphere used was a 25.4-mm-diam ball-bearing. These measurements were obtained for an incident pulse based upon a 50-kHz envelope (see Fig. 3). Figure 5(a) shows the field recorded in the backscattering position with the test particle in place. Figure 5(b) is the reverberant background recorded in the same position after the sphere has been removed. The result of subtracting this background signal from the initial measurement is shown in Fig. 5(c). This is the contribution from the sphere.

C. Measurement configurations

In order to obtain a representative sample of measurements for different incident angles, three initial configurations and subsequent rotations were employed (see Fig. 6). In all configurations the incident wave propagates in the horizontal plane and rotations are about a vertical axis passing through the center of the cube.

For configuration A the cube was initially aligned so that the wave field was normally incident on a face, and then the cube was rotated about an axis through the midpoint of two opposite faces [Fig. 6(a)].

For configuration B the cube was initially aligned so that the wave field was normally incident on an edge and was rotated about an axis through the midpoint of two edges [Fig. 6(b)].

In the third configuration (C) the cube was initially aligned for incidence on a corner, the acoustic axis passing through a body diagonal [Fig. 6(c)].

II. FORM FUNCTION OF A CUBE

The field scattered by an object is usually expressed in a normalized dimensionless form using the “form function”

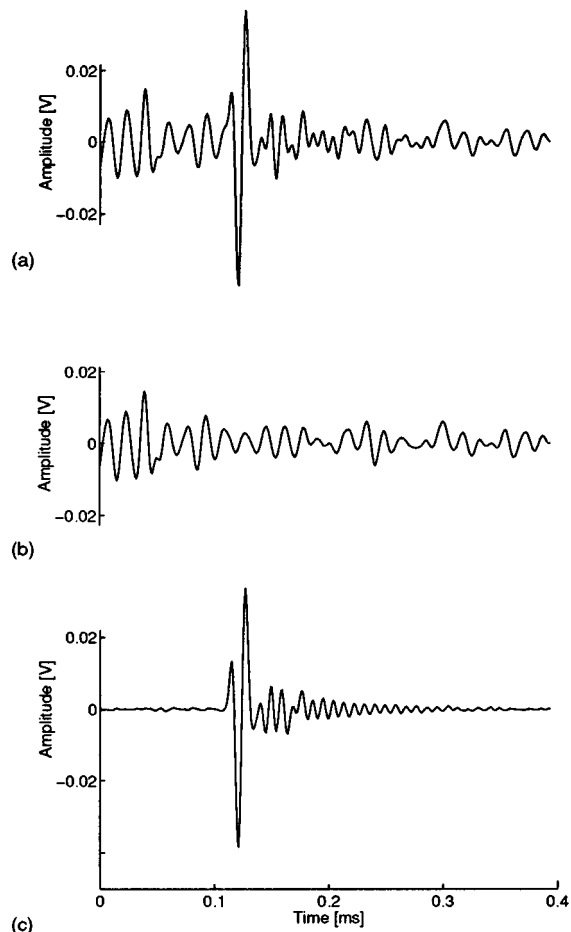


FIG. 5. Experimental waveforms for a steel sphere measured with 50-kHz pulse [Fig. 3(c)]: (a) backscattered signal plus reverberant background; (b) reverberant background; (c) subtraction of (b) from (a): backscattered signal due to sphere.

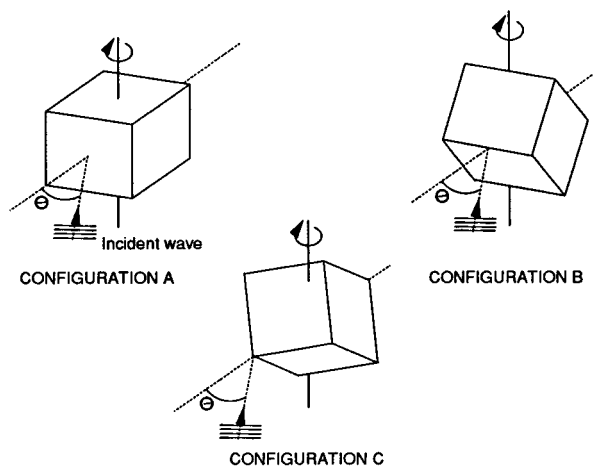


FIG. 6. The three “configurations” of the cube (A, B, and C) are defined by the rotation axes shown. The configurations are such that a plane wave at $\theta = 0^\circ$ is incident normal to a face (A), edge (B), and corner (C). The rotation axes pass through the center of the cube in each case. (In configuration C the rotation axis does not pass through a body diagonal, although the axis of the incident wave does.)

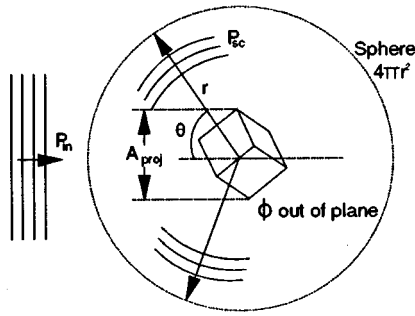


FIG. 7. Scattering of a plane wave by a cube; area of incident wave intercepted (A_{proj}) depends on orientation of the cube (θ, ϕ).

[Eq. (1)]. The form function for spherical scatterers is a well known and much used expression in which the scattered pressure is normalized by the size of the sphere, the observation range, and the incident pressure. The characteristic dimension used for the sphere is its radius. The corresponding dimension to be used for a cubic scatterer is less obvious. Consequently the normalization of the form function for a cube is not obvious and must be deduced by considering the total scattering problem, and the requirement that normalization should be consistent with that for the spherical case in some limit. The following argument is used to derive an expression for the form function.

The scattering situation is shown in Fig. 7. How much of the incident plane wave is intercepted by the cube depends on the orientation of the cube. Consider a plane pressure wave P_{in} (having intensity $I_{\text{in}} \propto P_{\text{in}}^2$) incident upon a target having projected cross-sectional area A_{proj} (parallel projection along the acoustic axis of the incident wave). The scattered pressure P_{sc} (intensity I_{sc}) scatters into a sphere of radius r having an area $4\pi r^2$. If the scattering is isotropic then P_{sc} is independent of position (θ, ϕ). If we equate the power intercepted by the cube ($A_{\text{proj}} I_{\text{in}}$) and the total scattered power ($4\pi r^2 I_{\text{sc}}$) then

$$A_{\text{proj}} P_{\text{in}}^2 = 4\pi r^2 P_{\text{sc}}^2, \quad (2)$$

and the “geometrically” scattered field is given by

$$P_{\text{sc}} = \frac{\sqrt{A_{\text{proj}}}}{2r\sqrt{\pi}} P_{\text{in}}. \quad (3)$$

In reality the scattering is not geometrical and the scattered pressure will vary with θ and ϕ . The relative scattering strength is determined by the form function $f(r, \theta, \phi)$. Thus if we write

$$P_{\text{sc}}(r, \theta, \phi) = f(r, \theta, \phi) \frac{\sqrt{A_{\text{proj}}}}{2r\sqrt{\pi}} P_{\text{in}}, \quad (4)$$

we obtain

$$|f(r, \theta, \phi)| = \frac{2r\sqrt{\pi}}{\sqrt{A_{\text{proj}}}} \frac{|P_{\text{sc}}(r, \theta, \phi)|}{|P_{\text{in}}|}. \quad (5)$$

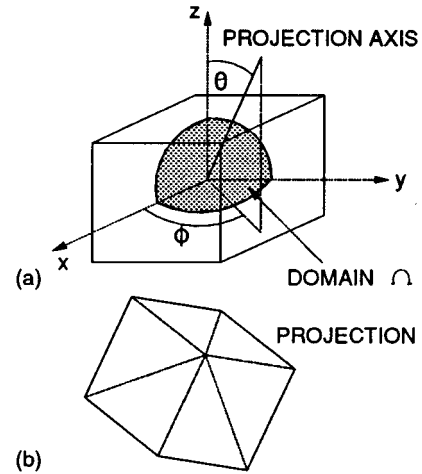


FIG. 8. Parallel projection of the cube along projection axis θ, ϕ ; (a) domain (Ω) of projections θ, ϕ ; (b) typical projection of the cube showing division of area into six triangular regions.

This then is the expression for the modulus of the form function of the scatterer intercepting an area A_{proj} of the incident plane wave.

For a sphere A_{proj} is independent of orientation and equals πa_{spn}^2 , where a_{spn} is the radius of the sphere. In this case the form function reduces to the familiar form shown in Eq. (1). The geometrical limit is recovered when $f=1$, this being the value about which the form function of the rigid spherical scatterer oscillates at high frequencies.

For the cube, the projected area (A_{proj}) varies with the orientation of the cube and so some other means of normalizing is necessary. As our ultimate aim is to obtain the scattered field averaged over all orientations of the cube, then it is sensible to normalize with respect to the *average* parallel projected area of the cube; i.e., the average projection over all orientations of the cube.

The projected cross-sectional area of the cube in any particular orientation can be calculated from geometrical considerations. Consider Fig. 8. The projection of a cube along any axis (a) is, in general, a hexagon (b). This hexagon can be divided into three quadrilaterals having a common vertex. The three quadrilaterals are projections of three of the faces of the cube (each at a different angle) and the common vertex is a corner of the cube. The special cases (orientations) where some edges of the quadrilaterals are obscured by others correspond to one or more of the quadrilaterals having zero area. These constitute a set of measure zero and are of no importance. By calculating the areas of these quadrilaterals the projected area of the cube in any orientation can be found.

However, we do not wish to normalize the form function for the cube in a specific orientation using the projected area in that particular orientation, but rather using the *average* projected area. Summing the projected areas of the cube in all orientations suggests the average projected area is exactly $1.5l^2$ (i.e., a quarter of the surface area). This initially rather surprising result has a simple proof, as given below.

A. Average parallel projection of a cube

It is possible to calculate the average projected area of a cube *without* calculating the individual projections of the cube in each orientation. For this analytical derivation we are grateful to Professor James Davenport²⁰ for advice given. The argument goes as follows.

The symmetry of the cube is such that only orientations within an eighth of the total solid angle need be considered; that is the domain Ω ; $\theta = 0 \rightarrow \pi/2$, $\phi = 0 \rightarrow \pi/2$ [Fig. 8(a)]. For any particular axis of projection in Ω *three* faces of the cube contribute to the projected area. As each face of the cube is identical, its contribution to the average projected area equals the contribution of each of the other two faces and so the average projected area is three times the contribution of one cube face. Consider a face parallel to the xy plane [the top face in Fig. 8(a)]. For a parallel projection along the line θ , ϕ , the cube face of side l (area l^2) has a projected area $S(\theta, \phi)$ given by

$$S(\theta, \phi) = l^2 \cos(\theta), \quad (6)$$

irrespective of ϕ . The face can be seen from all angles in the domain Ω and so the average contribution of this face ($\langle S \rangle$) is found by integrating S over Ω and dividing by the solid angle $\pi/2$:

$$\langle S \rangle = \frac{2}{\pi} \int_{\Omega} S(\theta) \sin(\theta) d\theta d\phi. \quad (7)$$

The average projection of the whole cube is therefore given by

$$\langle A_{\text{proj}} \rangle = 3 \frac{2}{\pi} \int_0^{\pi/2} \int_0^{\pi/2} l^2 \cos(\theta) \sin(\theta) d\theta d\phi = \frac{3l^2}{2}. \quad (8)$$

The average projected area is thus proven to be a quarter of the surface area of the cube ($6l^2$), as is the case for a sphere. Similar arguments can be applied to polyhedra having other numbers of sides. An alternative derivation, applicable to any convex three-dimensional body is given in the Appendix. This theorem would appear to be more familiar in optics, where it is applied to light scattering by randomly orientated particles; see, for example, van de Hulst.²¹

B. Form function of the cube

Using the average projected cross-sectional area of a cube (8) the form function (5) for a cube of side l becomes

$$|f(r, \theta, \phi)| = 2\sqrt{2\pi/3} \frac{r}{l} \frac{|P_{\text{sc}}(r, \theta, \phi)|}{|P_{\text{in}}|}. \quad (9)$$

Alternatively we can express l in terms of the radius of the sphere having the same surface area (sa) as the cube (i.e., $4\pi a_{\text{sa}}^2 = 6l^2$) and obtain

$$|f(r, \theta, \phi)| = \frac{2r}{a_{\text{sa}}} \frac{|P_{\text{sc}}(r, \theta, \phi)|}{|P_{\text{in}}|}, \quad (10)$$

where $a_{\text{sa}} = l\sqrt{3/2\pi}$. Thus, in terms of spherical scatterers, the correct normalization for cubic particles is in terms of the

sphere having the same surface area, and *not* the volume as some previous studies have used. These equations define a correctly normalized form function for the cube. Using the generalization presented in the Appendix, this expression for the form function can be applied to any three-dimensional convex body.

C. Normalization of frequency

Although the average projected cross-sectional area of the cube is the correct quantity to use in any dimensionless normalized form function, it does not necessarily follow that this is the best characteristic dimension of the cube with which to normalize the frequency. The requirement that the scattering behavior of particles of a given shape be scale invariant is satisfied by multiplying the frequency (or wave number) by *any* characteristic length dimension of the particle. Some additional criterion would seem to be needed for particles of different shape. One criteria would be that particles having different shapes but identical volumes have the same basic frequency behavior. This would mean using the radius of the sphere having the same volume in the normalization:

$$a_{\text{vol}} = l \left(\frac{3}{4\pi} \right)^{1/3}. \quad (11)$$

Another possibility would be to normalize the frequency in the same way as the form function is normalized, i.e., in terms of the surface area:

$$a_{\text{sa}} = l \left(\frac{3}{2\pi} \right)^{1/2}. \quad (12)$$

However, there is no obvious justification for doing so. The average radius of the cube is another possible choice:

$$a_{\text{avg}} = \frac{l}{6} \left[\sqrt{3} - \frac{\pi}{6} + 4 \ln \left(\frac{1+\sqrt{3}}{\sqrt{2}} \right) \right]. \quad (13)$$

Perhaps the most important requirement is that the frequency dependence of the form function of any particle approaches that of the sphere in the limit of a large number of facets (tending to infinity).

As the characteristic periodicity in the form function of a sphere is determined by the propagation of surface waves, one might suppose the mean path length around the cube to be the important quantity. Calculating this quantity is not, it would seem, a trivial matter, especially for objects having edges and corners.

It would therefore seem that there is no obvious normalization of frequency that should be used. However, in sediment applications a_{vol} might be the most suitable normalization factor for the frequency because it is easy to evaluate for a large range of shapes. In Sec. IV C we will compare the three frequency normalizations discussed in this section. It should be noted that the normalization of the frequency is unrelated to the normalization of the form function, as discussed above.

III. NUMERICAL CALCULATIONS

For this paper it was decided to theoretically model a rigid cube rather than an elastic cube. This approach is justified because in the low ka range of interest the elastic response of the cube is not thought to be important. Indeed experimental results on the acoustic cross sections of sand particles have not shown obvious elastic resonant phenomena. This may be due to the complicated composition and size variation in the natural sand, leading to a reduction and broadening of the resonant peaks.

From a theoretical viewpoint the cubic shape is of interest because it is the simplest polyhedron to model (other than the sphere). It has also been suggested as the most probable shape of small sand particles (diameter 10–30 μm).¹⁰

A. Theoretical background

There are several methods available for solving the problem of acoustic wave scattering from submerged bodies. However, analytical solutions in terms of standard functions can be found only for a limited number of geometrical shapes such as spheres, infinite cylinders, and spheroids. Numerical methods must therefore be used for solving the problem of acoustic wave scattering from more complex geometries, such as the cube. There are three popular methods in use that are capable of solving the scattering problem: the T -matrix method; the finite element method; and the boundary element method (BEM).

The T -matrix²² is an efficient method for the scattering problem, but it cannot easily deal with corners, edges, or large aspect ratios, due to the poor convergence in these situations.

The finite element method^{23–25} on its own is not very efficient in most scattering problems because it always gives the solution for the whole field, not just at the points of interest. In addition, the wave reflected from the outside boundaries of the finite element model can create errors due to the imperfect absorbing boundary conditions.

The BEM is a very powerful method for scattering problems in which the scatterer size is comparable with the acoustic wavelength. For an extensive bibliography, see Mackerle.²³ It is easier to deal with the scattering from an elastic object if the BEM is combined with a finite element approach for the vibrational behavior of the scatterer. The main problem with the BEM is that the solution is not unique at certain characteristic frequencies. One solution, referred to as CHIEF (combined Helmholtz integral equation formulation), involves overdetermining the BEM equations by forcing the acoustic pressure at a discrete set of interior points to zero. This approach is widely used and considered sufficiently accurate in the low ka range.

B. Procedure for numerical calculations

The commercially available software package PAFEC¹³ was chosen as the numerical tool in this study. PAFEC is capable of dealing with the coupling between the elastic structure and the acoustic wave in a fluid by using a combined finite element/boundary element method. Finite elements are used to describe the vibration of the elastic struc-

ture, and boundary elements are used to describe the radiation, or scattering, from the structure. For the calculations presented here, only the boundary element capability of PAFEC was used to calculate the scattered field from a rigid cube. This can however be extended to the calculation of the scattering from an object having a given surface impedance, or an elastic solid if necessary.

For configurations A and B (Sec. I C, Fig. 6) only half of the cube surface needed to be meshed because of symmetry. The backscattering fields were evaluated as a function of the angle of rotation θ up to 45° for configuration A and 90° for configuration B. The full surface of the cube was meshed for configuration C and the scattered field evaluated for incidence angles in the range 0° – 90° .

The accuracy of the calculated scattered field is dependent on the mesh density. In PAFEC, an acceptable accuracy for the backscattering field of a rigid object can be achieved by using five elements per wavelength if six-noded triangular or eight-noded quadrilateral elements are used. The upper frequency limit of interest in the theoretical study is 100 kHz (for a cube of side 20 mm). In order to reduce the size of the boundary element model, five elements per sidelength were used in the frequency range 3–60 kHz and seven elements per sidelength in the frequency range 60–100 kHz. A frequency step of 2 kHz was used throughout. In calculations, the backscattered pressure was evaluated at a distance of 20 cm from the center of the cube. The density and sound speed of the fluid (water) were taken to be 1000 kg m^{-3} and 1460 m s^{-1} , respectively.

IV. RESULTS

A. Experimental results

For each of the three configurations described in Sec. I C and illustrated in Fig. 6 two series of backscattering measurement have been made; one with a low-frequency pulse based upon a 20-kHz envelope and one based upon 50 kHz. In each case the cube was rotated through 90 degrees in either 5-degree steps (20-kHz pulse) or 1.5-degree steps (50 kHz). The rotation axes are such that a plane wave (in all cases incident normal to the rotation axis) at $\theta = 0^\circ$ is incident normal to a face (A), edge (B), and corner (C). Note, in configuration C the rotation axis does *not* pass through a body diagonal, although the axis of the incident wave does.

Examples of the scattered time waveforms are shown in Fig. 9 for the 50-kHz pulse [Fig. 3(c)]. The strongest reflection is seen for configuration A, 0° ; the case of incidence normal to a cube face. This time waveform is similar to that observed with a spherical target [Fig. 5(c)]. Figure 9(b) shows the scattered time waveform when the wave is incident on a cube edge (configuration B, 0° ; configuration A, 45° ; or configuration C, 90°). This waveform differs from the previous example in several important respects. The amplitude of the signal is much reduced, as would be expected. There is a low amplitude signal arriving before the main portion of the waveform, this presumably being the contribution from the edge itself. There is also a sequence of discrete pulses in the tail of the signal. The third time waveform, illustrated in Fig. 9(c), corresponds to configuration C,

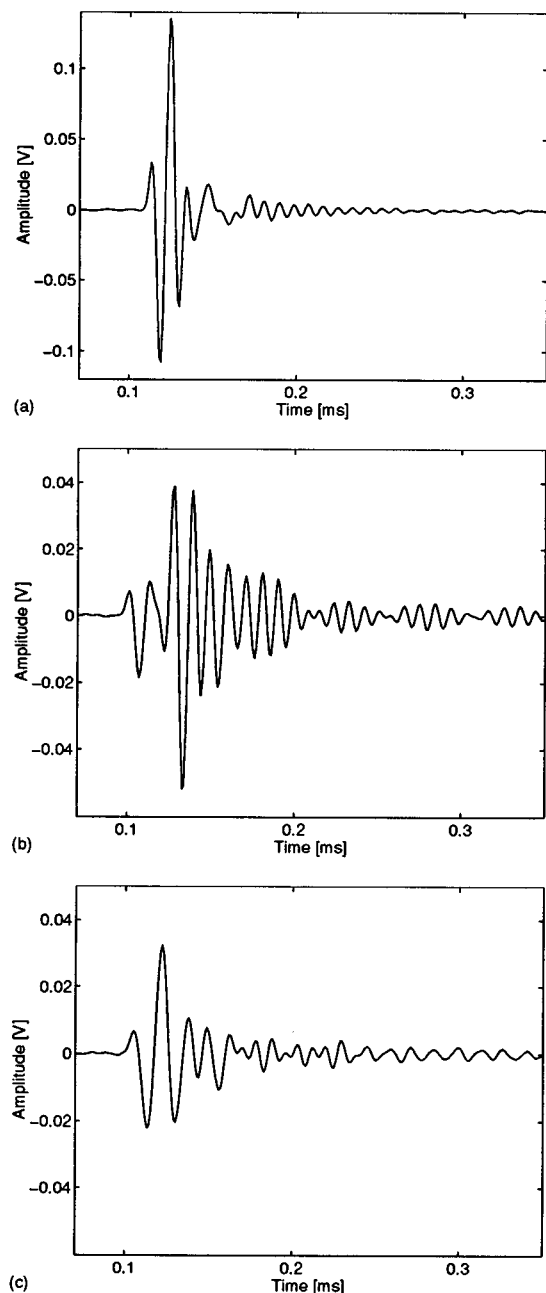


FIG. 9. Measured backscattered waveforms for an iron cube of side 20 mm. Incidence normal to (a) face (configuration A, 0°); (b) edge (configuration B, 0°); and (c) a more general orientation (C, 60°).

60° , which is a more general scattering geometry. This waveform has very little high-frequency content, does not exhibit the delayed pulses of (b), but does have a rather long mid-frequency tail.

The measured time waveforms were processed as described above and the modulus of the form function calculated using Eq. (5). The variation with both frequency and angle of incidence for each of the three configurations are presented in Fig. 10. The regions of high reflectivity are clearly seen, the greatest contributions coming from the faces and edges of the cube. There is also evidence of two features around 75 and 100 kHz which appear at several angles of

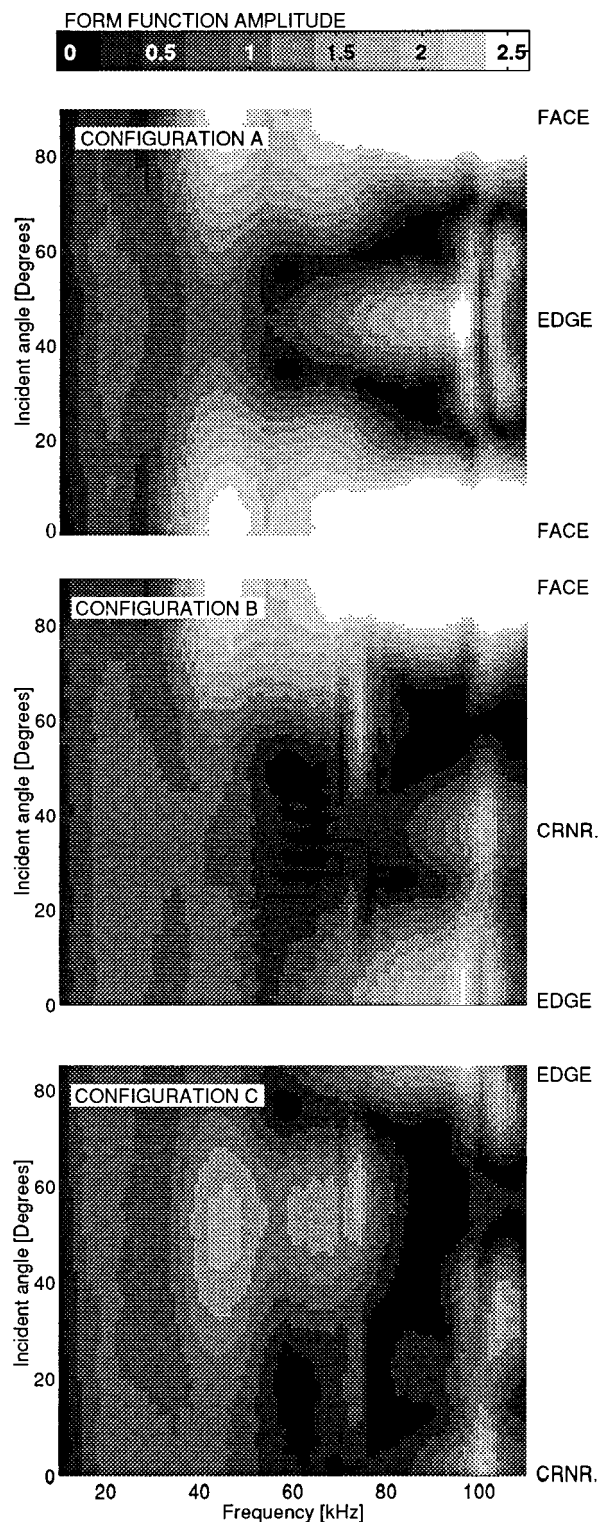


FIG. 10. Variation of experimentally measured form function with both frequency and incident angle (θ) for the three configurations (shown in Fig. 6).

incidence. In all configurations and at most angles of incidence some sort of periodic oscillation can be seen along the frequency axis.

B. Comparison with numerical model

Comparison of measured results with theoretical predictions are made in Figs. 11–13 which show the variation of

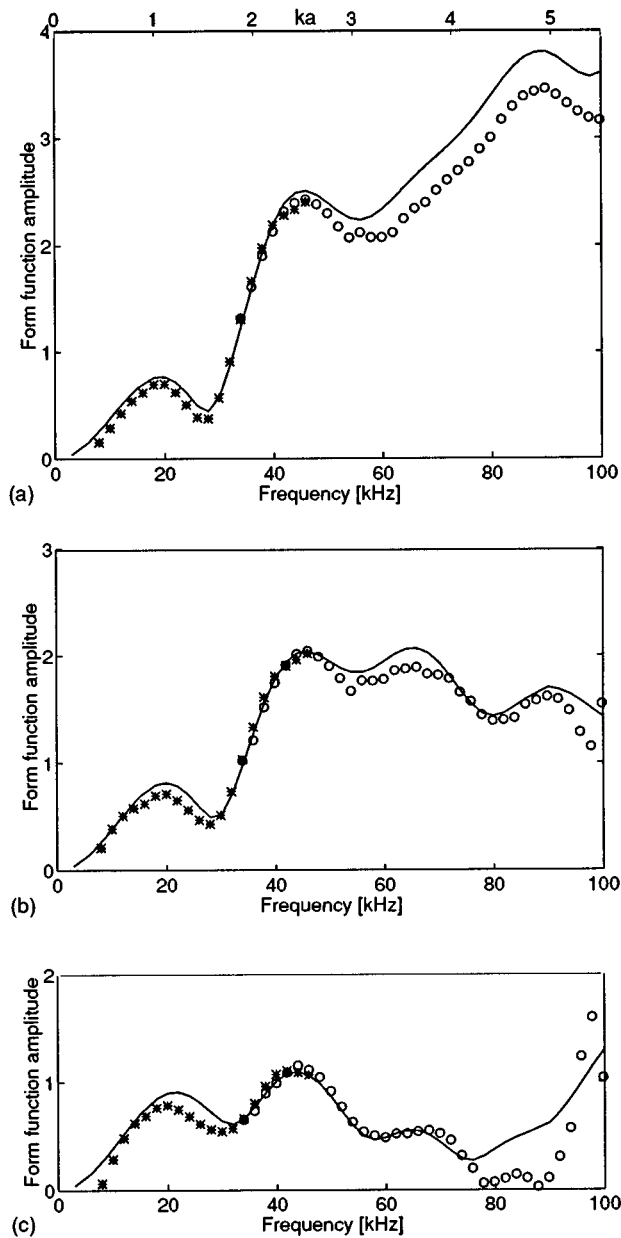


FIG. 11. Variation of form function with frequency for iron cube in configuration A; theory (lines), and experiment (points). [The corresponding ka range is indicated at the top of the figure; a is the average radius of the cube (a_{avg}).] Experimental data obtained with 50-kHz pulse (\circ) and 20-kHz pulse ($*$). (a) 0° (normal incidence); (b) 15° ; (c) 30° . The form function results are normalized using the average projected cross-sectional area of the cube (see Sec. II).

form function with frequency for specific configurations and angles of incidence. Both low- ($*$) and high- (\circ) frequency experimental data are presented along with the calculations (lines). The form function is normalized using the average projected cross-sectional area of the cube (Sec. II). Agreement is generally quite good at all but the highest frequencies where the numerical model is expected to break down due to the elastic nature of the cube. In all orientations a periodic variation in amplitude is seen, this being superimposed upon a more general frequency-dependent background that differs markedly between the orientations. The case of

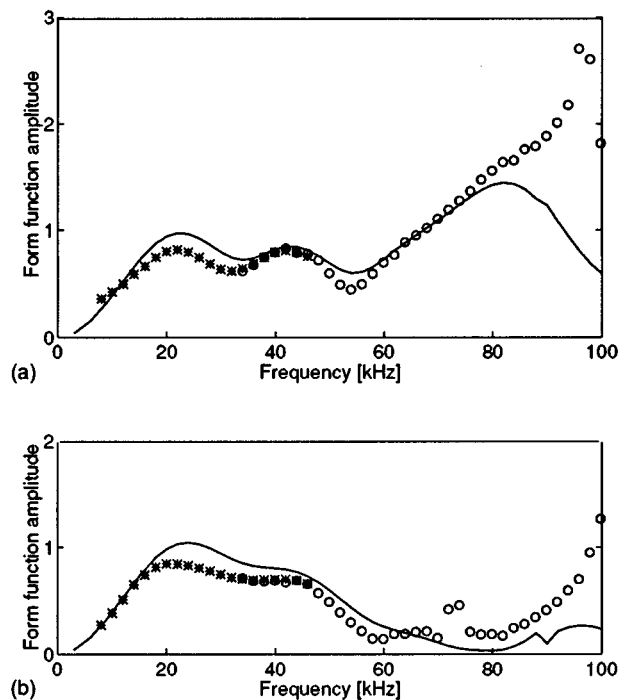


FIG. 12. Form function for iron cube in configuration B; theory (lines), and experiment (points). (a) 0° (edge incidence); (b) 45° . Experimental data obtained with 50-kHz pulse (\circ) and 20-kHz pulse ($*$).

incidence normal to a face, shown in Fig. 11(a), exhibits a steady increase in backscattered pressure with frequency. This is to be expected from the low (i.e., zero) curvature of the face presented to the receiver, and has been observed by other authors.²⁶ The discrepancy between theory and experiment at high frequencies is thought to be due to spherical spreading of the nonplanar wavefronts of the pulse produced by the parametric array. This effect has been considered in

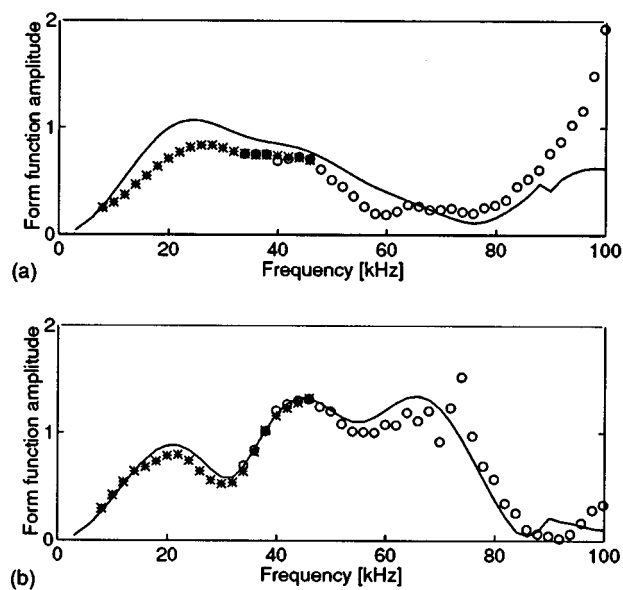


FIG. 13. Form function for iron cube in configuration C, theory (lines) and experiment (points). (a) 0° (corner incidence); (b) 60° . Experimental data obtained with 50-kHz pulse (\circ) and 20-kHz pulse ($*$).

detail by Beckett¹⁵ for cylindrical scatterers. This effect would not be so great for other configurations.

Figure 11(a)–(c) shows that for configuration A, as the cube is rotated away from normal incidence, the high-frequency response (>30 kHz) falls off sharply as expected, whereas the low-frequency regime, below the first minimum (~ 30 kHz), increases slightly due to the increasing projected cross sectional area of the target.

Figure 12 shows examples of the form function for configuration B. Configurations B, 0° corresponds to configuration A, 45° and can also be considered a continuation of the sequence shown in Fig. 11. The same arguments applied to configuration A can be applied here. Examples from configuration C are shown in Fig. 13. In Fig. 13(a) the first maximum in the theoretical form function reaches a value of approximately 1.1 at about 24 kHz. This configuration presents the maximum projected cross-sectional area to the incident wave. The low-frequency behavior is presumed to depend on A_{proj} and can be traced in the continuous sequence of results shown by Figs. 11(a)–(c), 12(a) and (b), and 13(a). In Fig. 13(b) the cross section has decreased again.

Some of the discrepancies between theory and experiment in the low–mid-frequency range may be due to wavefront curvature in the experimental system and alignment of the cube; this was particularly difficult for configurations B and C. Discrepancies in amplitude are also expected due to the fact that the cube is immobile in the model; scattering by a rigid immobile sphere differs in amplitude from that of a mobile rigid sphere or an elastic sphere, and we would expect similar differences here. This is especially true at low frequencies, in the vicinity of the first form function peak. At the higher end of the frequency range the rigid cube approximation is beginning to break down and this is presumed to be the cause of some discrepancy in this region.

C. Averages over angles and configurations

Averages of the form function amplitude for the three configurations A, B, and C are shown in Fig. 14, and the overall average of all three in Fig. 15. In these averages the experimental data are averaged over the same angles as the theoretical data were calculated; that is, at 7.5° intervals in each of the three configurations. The average of the three averages (Fig. 15) is remarkably similar to that observed for a spherical scatterer. However, care must be taken in interpreting this result as some orientations of the cube may be under-represented in the summation due to the fact that the cube was rotated about only three axes. A more democratic sampling of cube orientations is not easily achieved with the present experimental arrangement, and would involve considerable numerical calculations. Despite these considerations, the similarity to spherical scattering, is striking.

The numerically calculated average form function is compared with that for a rigid sphere in Fig. 16. For the cube (broken lines), the variation with ka is shown for three different normalizations of the frequency using the characteristic dimensions (a) discussed in Sec. II C; average radius of cube (a_{avg}), radius of sphere having the same volume (a_{vol}), and radius of the sphere having the same surface area (a_{sa}). [In each case the vertical axis (form function) is nor-

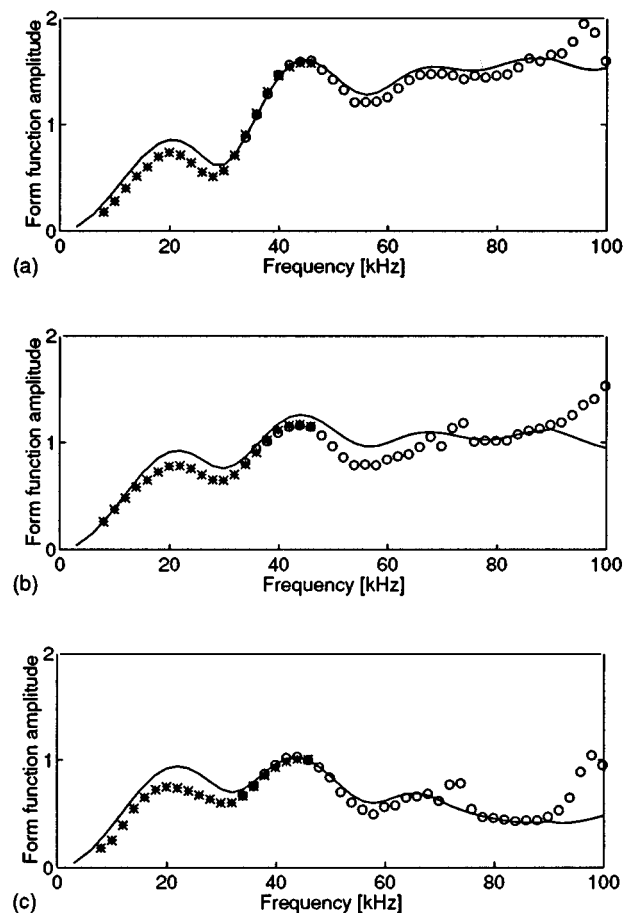


FIG. 14. Form functions for each configuration averaged over all incident angles. (a) Configuration A; (b) configuration B; (c) configuration C.

malized in the same way, using Eq. (9)]. The form function of the rigid sphere (solid line) is seen to have a periodicity that agrees best with the cube predictions plotted against ka_{avg} . However, the variation with ka_{vol} is also reasonable, and this frequency normalization has been used by a number of authors. It should be noted that we here refer to the normalization of frequency (x axis), and not that of the form function itself (y axis).

The average of the experimental results is compared with a theoretical calculation of the form function of a steel sphere in Fig. 17. The theoretical prediction for the elastic sphere is calculated using the normal mode series approach.¹¹ Over the ka range of interest the experimental

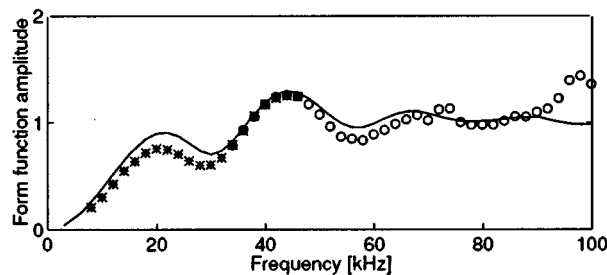


FIG. 15. Average over all three configurations and angles of incidence (i.e., an average of the averages shown in Fig. 14).

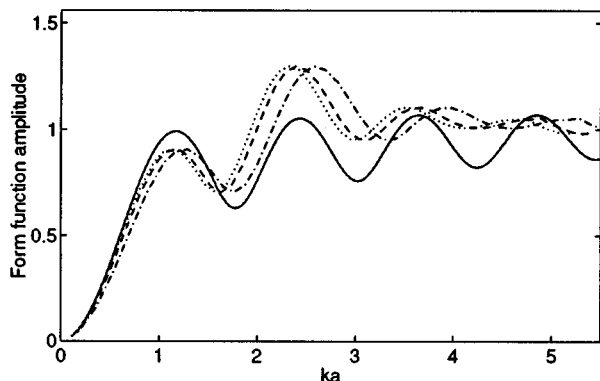


FIG. 16. Normalization of frequency axis: variation of average theoretical form function (broken lines) with ka for different characteristic dimensions (a), and comparison with a rigid sphere (solid line). Characteristic dimensions; a = radius of sphere having a equivalent volume (.....), a = average radius of cube (---), and a = radius of sphere having equivalent surface area (•-•-•-). In each case the same normalization of the form function (vertical scale) is used [Eq. (9)].

results show a remarkable resemblance to the spherical form function.

V. DISCUSSION AND CONCLUSIONS

The backscattering properties of a solid cube have been investigated both experimentally and theoretically in the low ka regime over a range of orientations of the cube. A dimensionless normalized form function for cubic scatterers has been defined in terms of the average projected cross-sectional area (A_{proj}), a quantity related to the surface area of the scatterer.

Results have been presented for several different configurations of the cube and good agreement between experiment and theory observed.

For each individual orientation of the cube there would seem to be several effects responsible for the observed behavior: (i) the angle of the facets, important at high frequency; (ii) the projected cross-sectional area, especially im-

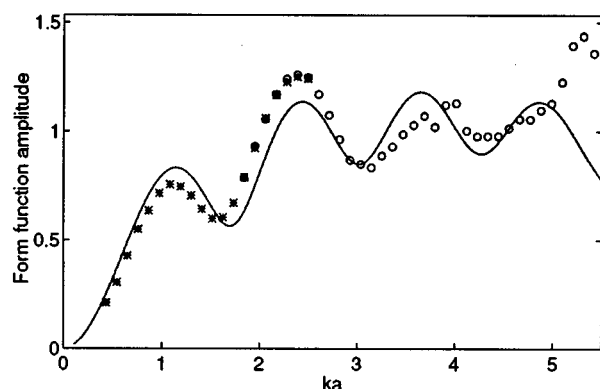


FIG. 17. Variation of measured average form function (points) with ka (a = average radius of cube) and comparison with theoretical prediction for a steel sphere (solid line).

portant at low frequencies; and (iii) the apparent existence of surface waves, at all frequencies.

Averages over all orientations for the calculations and measurements have also been presented; these show that the form function exhibits very similar behavior to that of a spherical scatterer. One notable feature is that the amplitude level is similar when the form function is correctly normalized.

The presence of a pronounced ripple on the averaged form function is quite remarkable. It should be remembered that for a sphere this is due to a creeping wave that continually diffracts around the surface of the sphere. The period of oscillation of the form function agrees well with that of the sphere when the frequency is normalized with the average radius of the cube (a_{avg}) or the radius of the sphere having equivalent volume (a_{vol}).

Although very good agreement between theory and experiment is demonstrated, differences do exist. At low ka there is an amplitude difference which is presumed to occur because the cube is assumed rigid and immobile in the numerical calculations. At higher frequencies resonance effects are occurring in the experimental measurements; these also are not predicted by the rigid model. Measurements of scattering by steel and glass cubes at higher frequencies are compared with calculations for *elastic* cubes in Chinnery *et al.*²⁷ A higher frequency range ($ka = 3-34$) is also dealt with by Thorne *et al.*²⁸ for an iron cube. Backscattering at much higher frequencies, and the mechanisms involved, are dealt with by Gipson and Marston²⁹ for a rectangular parallelepiped block.

We can now make the following interesting observation. The average projected area of a particle (the area the incident wave intersects) is known to be proportional to the surface area of the particle. Thus, for particles of different geometry having the same volume (i.e., equal volumes of sediment) the spherical particles have the lowest average projection (cross section) of all particles and should be expected to give the lowest value of backscattered signal at low frequency where only the cross-sectional area is important. Thus for the following polyhedra having the same volume as the sphere of radius a each particle has the following average projected area relative to that of the sphere:

sphere	1.00
icosahedron	1.06
dodecahedron	1.10
octahedron	1.18
cube (hexahedron)	1.24
tetrahedron	1.49

One might therefore expect the backscattered amplitudes (form function) for different particles to reflect these values. This is perhaps of significance for techniques wishing to use acoustic backscattering to measure the mass of sediment in suspension.

ACKNOWLEDGMENTS

The authors would like to thank Professor James Dav-
enport of the School of Mathematical Sciences, University of

Bath, for assisting with the proof for the average projection of a cube. We also acknowledge the work of Geoffrey Cranch, David Cooper, and Cathy Dyer, who were involved in preliminary experimental investigations of cube scattering. Thanks are also due to Peter Thorne of the Proudman Oceanographic Laboratory, Bidston Observatory, for the interest shown in this study, and for many useful discussions that have a bearing on this work. We are also very grateful to PAFEC Ltd. for assistance in the provision of software for these calculations.

APPENDIX: AVERAGE PROJECTION OF ANY THREE-DIMENSIONAL CONVEX BODY

Consider any three-dimensional convex body and divide the surface into N facets. Each facet (i) can be a different shape and has an area denoted A_i . For smooth objects the argument applies as $N \rightarrow \infty$ and each facet has an infinitesimal area. As before (Sec. II A) each facet has a projected area of $A_i \cos(\theta)$ and can be "seen" from a hemisphere; i.e., the facet is one-sided. The total projection of the facet is then given by

$$S_i^{\text{tot}} = A_i \int \int_{\text{hemi}} \cos(\theta) \sin(\theta) d\theta d\phi = A_i \pi, \quad (\text{A1})$$

and the total projection of the body is

$$A_{\text{proj}}^{\text{tot}} = \sum_i S_i^{\text{tot}} = \sum_{i=1}^N A_i \pi. \quad (\text{A2})$$

The average projection is therefore given by

$$\langle A_{\text{proj}} \rangle = \frac{\sum A_i \pi}{4\pi} = \frac{1}{4} \sum A_i. \quad (\text{A3})$$

But $\sum_{i=1}^N A_i$ simply equals the surface area of the body. Thus it would appear that any convex body has an average projected cross-sectional area equal to a quarter of its surface area. An alternative proof of this theorem is given by van de Hulst²¹ who also states that it cannot be extended to include nonconvex scatterers.

¹J. Sheng and A. E. Hay, "An examination of the spherical scatterer approximation in aqueous suspensions of sand," *J. Acoust. Soc. Am.* **83**, 598–610 (1988).

²A. E. Hay and A. S. Schaafsma, "Resonance scattering in suspensions," *J. Acoust. Soc. Am.* **85**, 1124–1138 (1989).

³A. E. Hay, "Sound scattering from a particle-laden, turbulent jet," *J. Acoust. Soc. Am.* **90**, 2055–2074 (1991).

⁴A. E. Hay and J. Sheng, "Vertical profiles of suspended sand concentration and size from multifrequency acoustic backscatter," *J. Geophys. Res.* **97**, 661–677 (1992).

⁵C. He and A. E. Hay, "Broadband measurements of the acoustic backscatter cross section of sand particles in suspension," *J. Acoust. Soc. Am.* **94**, 2247–2254 (1993).

⁶A. M. Crawford and A. E. Hay, "Determining suspended sand size and concentration from multifrequency acoustic backscatter," *J. Acoust. Soc. Am.* **94**, 3312–3324 (1993).

⁷P. D. Thorne, P. J. Hardcastle, and R. L. Soulsby, "Analysis of acoustic measurements of suspended sediments," *J. Geophys. Res.* **98**, 899–910 (1993).

⁸P. D. Thorne, K. R. Waters, and T. J. Brudner, "Acoustic measurements of scattering by objects of irregular shape," *J. Acoust. Soc. Am.* **97**, 242–251 (1995).

⁹P. D. Thorne, L. Hayhurst, and V. F. Humphrey, "Scattering by nonmetallic spheres," *Ultrasonics* **30**, 15–20 (1992).

¹⁰M. Ramirez and I. K. Bjørnø, "Experimental investigations of acoustic backscattering from particles in water," in *Proceedings of the Second European Conference on Underwater Acoustics, Vol. 2, Copenhagen*, edited by L. Bjørnø (European Commission, Luxembourg, 1994), pp. 1005–1009.

¹¹N. D. Veksler, *Resonance Acoustic Spectroscopy* (Springer-Verlag, Berlin, 1993).

¹²D. T. DiPerna and T. K. Stanton, "Sound scattering by cylinders of non-circular cross section: A conformal mapping approach," *J. Acoust. Soc. Am.* **96**, 3064–3079 (1994).

¹³PAFEC VibroAcoustics, from Pafec Ltd., Strelley Hall, Nottingham N68 6PE, UK.

¹⁴V. F. Humphrey, "The measurement of acoustic properties of limited size panels by use of a parametric source," *J. Sound Vib.* **98**, 67–81 (1985).

¹⁵C. Beckett, "Acoustic backscattering from specific targets," University of Bath internal reports (unpublished, 1986–1989); "Studies of acoustic scattering using a parametric array," Ph.D. thesis, University of Bath, 1992.

¹⁶P. J. Westervelt, "Parametric acoustic array," *J. Acoust. Soc. Am.* **35**, 535–537 (1963).

¹⁷H. O. Berklay, "Possible exploitation of nonlinear acoustics in underwater transmitting applications," *J. Sound Vib.* **2**, 435–461 (1965).

¹⁸M. B. Moffet and J. E. Blue, "Hydrophone nonlinearity measurements," N.U.S.C. Technical memorandum 801150, 1980.

¹⁹V. F. Humphrey and C. H. Hsu, "Nonlinearity of cylindrical hydrophones used for the measurement of parametric arrays," in *Transducers for Sonar Applications: An Institute of Acoustics Conference*, University of Birmingham, December 1980.

²⁰J. Davenport, School of Mathematical Sciences, University of Bath (private communication, 1995).

²¹H. C. van de Hulst, *Light Scattering by Small Particles* (Dover, New York, 1981).

²²*Acoustic, Electromagnetic and Elastic Wave Scattering: Focus on the T-Matrix Approach: International Symposium Held at the Ohio State University, 1979*, edited by V. K. Varadan and V. V. Varadan (Pergamon, New York, 1980).

²³J. Mackerle, "Finite and boundary element techniques in acoustics—A bibliography," *Finite Elements in Analysis and Design* **15**, 263–272 (1994).

²⁴D. B. Woyak, *Acoustics and Fluid-Structure Interaction ANSYS (5.0) Tutorial* (Swanson Analysis System, Inc., 1992).

²⁵*PAFEC Manual (Theory)* from PAFEC Ltd., Strelley Hall, Nottingham NG8 6PE, United Kingdom (1984).

²⁶S. Sun, P. D. Thorne, I. K. Bjørnø, and T. Mazoyer, "Observations of acoustic backscattering by elastic cubes," in *Proceedings of the Third European Conference on Underwater Acoustics, Vol. 1, Crete*, edited by J. S. Papadakis (FORTH-IACM, Crete, 1996), pp. 51–56.

²⁷P. A. Chinnery, J. D. Zhang, and V. F. Humphrey, "Acoustic scattering by nonmetallic and metallic cubes in the elastic resonance regime: Experimental measurements and combined finite element/boundary element modeling," *J. Acoust. Soc. Am.* (in press).

²⁸P. D. Thorne, S. Sun, J. D. Zhang, I. Bjørnø, and T. Mazoyer, "Measurements and analysis of acoustic backscattering by elastic cubes and irregular polyhedra," *J. Acoust. Soc. Am.* (in press).

²⁹K. Gipson and P. L. Marston, "Retroreflective backscattering of ultrasound due to Rayleigh waves on an elastic solid rectangular parallelepiped," *J. Acoust. Soc. Am.* **96**, 3325 (1994).

Simultaneous acoustic and microwave backscattering from the sea surface

Peter H. Dahl and William J. Plant

Applied Physics Laboratory, University of Washington, Seattle, Washington 98105

Bernd Nützel

Forschungsanstalt der Bundeswehr für Wasserschall-und Geophysik, Kiel, Germany

Anke Schmidt

Institut für Meereskunde, Universität Hamburg, Hamburg, Germany

Heinz Herwig

Forschungsanstalt der Bundeswehr für Wasserschall-und Geophysik, Kiel, Germany

Eugene A. Terray

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543

(Received 30 May 1996; accepted for publication 17 November 1996)

Simultaneous and coincident measurements of acoustic and microwave backscatter from the air/sea interface were obtained during Phase II of the SAXON-FPN experiment in December 1992 and again in March 1993. The acoustic and microwave grazing angles were both set to 17° , and the wavelengths were matched, being set to 2.14, 3.00, and 5.66 cm, corresponding to, respectively, acoustic frequencies of 26.5, 50, and 70 kHz and microwave frequencies of 5.3, 10, and 14 GHz. Backscattering cross sections normalized by ensonified area for the acoustic (σ_0^a) and microwave (σ_0^m) returns were determined, and their dependence on wind speed was investigated. The acoustic scattering strength is defined as $10 \log_{10}(\sigma_0^a)$ and the microwave scattering strength is defined as $10 \log_{10}(\sigma_0^m) - 10 \log_{10}(4\pi)$. The results of these experiments show that the two scattering strengths are comparable at wind speeds below about 3 m/s but that the acoustic scattering strength increases much faster than the microwave scattering strength with increasing wind speed until reaching saturation. If these wind-speed dependencies are represented by a power law, U^n , then n is 5–6 for σ_0^a and 2–4 for σ_0^m for wind speeds between 2 and 7 m/s. This difference is ascribed to the effect of bubbles on the acoustic backscatter. The more rapid increase of σ_0^a compared to σ_0^m implies that for our 17° grazing angle acoustic scattering from the surface is negligible at all but the lowest wind speeds. Therefore a simple model is used for bubble scattering to fit the acoustic data as a function of wind speed for all three acoustic frequencies. The bubble densities required to fit the data agree well with previous measurements of near-surface bubble distributions. The model predicts an overshoot of the acoustic scattering strength (above the saturation level) at moderate wind speeds which is clearly seen in the data at 26.5 and 70 kHz. Finally, a composite surface scattering model is utilized for the pure surface scattering component along with the bubble model to predict the wind-speed dependence of the acoustic scattering strength at a 45° grazing angle and compare the results with earlier measurements. © 1997 Acoustical Society of America.

[S0001-4966(97)06005-0]

PACS numbers: 43.30.Gv, 43.30.Ft, 43.30.Hw [JHM]

INTRODUCTION

Microwave and acoustic scattering from a rough surface can be explained using very similar equations, the primary difference being the vector nature of the electric field of the microwaves compared to the scalar nature of the pressure field of the acoustic signals (Bass and Fuks, 1979). In fact, the paradigm of microwave scattering from the sea surface, composite surface theory, was developed simultaneously in the United States and Russia but was based on microwave scattering in wavetanks in the U.S. and on acoustic scattering in wavetanks in Russia (Wright, 1966, 1968; Bass *et al.*, 1968a, 1968b). When considering microwave and acoustic scattering from the actual ocean surface, however, differences are introduced between the two types of scattering by

subsurface bubbles, which attenuate and scatter acoustic signals but are irrelevant to microwave signals. A comparison between microwave and acoustic backscattering from the sea surface should, therefore, help clarify the relative importance of bubbles and the rough interface in the scattering of acoustic waves from the region near the ocean surface. Furthermore, such a comparison helps to clarify under what conditions, and frequencies, microwave remote-sensing methods can be used to assess acoustic environmental conditions.

To our knowledge, only a single comparative study of microwave and acoustic scattering has been carried out to date, that of McDaniel and Gorman (1982). In that study, the importance of bubbles in acoustic scattering from the sea surface at low grazing angles and frequencies above 20 kHz

was suggested by comparing acoustic measurements with the predictions of composite surface theory. Comparisons of acoustic and microwave scattering were also carried out using data sets taken at different times and in different places and under conditions which, for the microwave data, were only marginally well described. These comparisons indicated that at wind speeds of 5 m/s the area-normalized microwave and acoustic scattering cross sections were roughly similar at low grazing angles and the acoustic cross sections dropped below those of microwaves at higher grazing angles. At higher wind speeds, the acoustic cross section increased more rapidly than the microwave cross section at low grazing angles, and the two became comparable at high grazing angles. As demonstrated below, we believe that the definitions of microwave and acoustic cross sections employed in this work were different and that these conclusions may require modification.

In this paper, we report on a study of microwave and acoustic sea surface backscattering in which the microwave and acoustic signals were incident on the same patch of the sea surface at the same time (although from opposite directions) and in which the grazing angles and wavelengths of the two signals were carefully matched. In the following section, we consider the differences in the definition of cross section used by the microwave and acoustic scattering communities and show that the microwave cross section is 4π times the acoustic one. In Sec. II we review microwave scattering from rough water surfaces and acoustic scattering from near-surface bubbles. The experiment is described in Sec. III; in Sec. IV we present our comparisons of the wind-speed dependence of microwave and acoustic backscattering cross sections, while in Sec. V we show that the bubble scattering model outlined in Sec. II fits our data well using reasonable bubble densities. In Sec. VI we hazard a prediction about the behavior of the acoustic cross section at a higher grazing angle by combining the results presented in this paper with a microwave scattering model. A summary and conclusions are presented in Sec. VII.

I. MICROWAVE AND ACOUSTIC TERMINOLOGY

Although microwave and acoustic scattering from rough surfaces are very similar, the two communities have tended to work separately over the years and have therefore developed independent, but equivalent, notations. In attempting to compare microwave and acoustic backscatter from the sea surface, it is essential to understand these notational differences.

Microwave backscatter is most easily described by the "radar equation," which is (Skolnik, 1970)

$$P_r = \frac{P_t \lambda^2 G^2 \sigma^m}{(4\pi)^3 R^4}, \quad (1)$$

where P_r is the power received by the antenna, P_t is the power transmitted by the antenna, λ is microwave length, G is the antenna gain, which is a function of angle off the look direction, σ^m is the microwave cross section, and R is the range to the target. Since the effective area of the antenna aperture can be written

$$A_e = \frac{\lambda^2 G}{4\pi}, \quad (2)$$

Eq. (1) can also be written

$$P_r = \frac{P_t G A_e \sigma^m}{(4\pi)^2 R^4}. \quad (3)$$

Acoustic backscatter (ignoring attenuation) is normally described by the equation (Urlick, 1983; Clay and Medwin, 1977)

$$I_r = \frac{I_1 \sigma^a}{R^2}, \quad (4)$$

where σ^a is the acoustic backscattering cross section, I_r is the intensity received at the transducer, and I_1 is the intensity incident at the scatterer. The received intensity can be written in terms of received power and the effective area of the transducer

$$I_r = P_r / A_e. \quad (5)$$

The incident intensity is normally expressed in terms of the transmitted intensity I_0 , which is not defined at the transducer but 1 m away from it and therefore must be written in terms of transmitted power as follows:

$$I_0 = \frac{P_t G}{4\pi R^2} \bigg|_{R=1 \text{ m}}. \quad (6)$$

If spherical spreading is assumed, I_1 is simply Eq. (6) evaluated at range R , which upon substituting Eqs. (5) and (6) into Eq. (4) yields

$$P_r = \frac{P_t G A_e \sigma^a}{4\pi R^4}, \quad (7)$$

which is the same as Eq. (3) if

$$\sigma^a = \frac{\sigma^m}{4\pi}. \quad (8)$$

Evidence that this is the correct relationship between σ^m and σ^a is provided by the fact that the high-frequency microwave cross section of a sphere of radius a used for backscattering target calibrations is πa^2 (Skolnik, 1970) whereas its acoustic counterpart is $a^2/4$ (Urlick, 1983). Obviously, the microwave definition of cross section is area while that of acoustics is area per unit steradian. We believe that this difference was overlooked by McDaniel and Gorman and that their plots compare area-normalized σ^a with the area-normalized σ^m .

A few other terminology differences occur between the acoustic and microwave communities. When σ^m and σ^a are divided by ensonified area, both communities refer to them as normalized cross sections; here we shall use the symbols σ_0^m and σ_0^a for these normalized quantities. When these quantities are converted to logarithmic scales, however, the two communities differ on the names. The microwave community still refers to $10 \log_{10}(\sigma_0^m)$ as the *normalized radar cross section*, but the acoustic community refers to $10 \log_{10}(\sigma_0^a)$ as the *scattering strength*. In an attempt to avoid confusion in this paper, we shall use the term scatter-

ing strength to refer to $10 \log_{10}(\sigma_0^a)$ and to $10 \log_{10}(\sigma_0^m) - 10 \log_{10}(4\pi)$. Thus microwave and acoustic scattering strengths are directly comparable; cross sections are not. Furthermore, while microwave literature refers to “VV and HH polarized” returns to distinguish whether the electric field is in or out of the plane of incidence, acoustic literature has no need to do so. Therefore, to simplify the terminology we will shorten “HH polarized microwave” to simply “HH” and “VV polarized microwave” to “VV.” Finally, microwave scattering theories and results are typically parameterized by the incidence angle, whereas there is tendency in the acoustic community to use the grazing angle, which we do in this paper.

II. MODELS FOR SCATTERING FROM THE NEAR-SURFACE REGION

In this section, we review microwave backscattering from the rough air/sea interface and acoustic backscattering from bubbles. Since both HH microwave fields and acoustic fields at the air/sea interface must satisfy the Dirichlet boundary condition, we postulate that, except for minor differences noted below, models for predicting microwave scattering from rough surfaces apply equally well to acoustic scattering provided that the radiation wavelengths are matched. Below we briefly review microwave backscattering from the sea surface and summarize a model which we subsequently use to predict the acoustic return associated with backscattering from the rough air/sea interface. Analogous rough surface scattering theories for acoustic scattering are summarized by McDaniel (1993) for sea surface scattering and by Jackson *et al.* (1986) for sea bed scattering. A corollary to the above postulate equating microwave and acoustic rough-surface scattering is that, although both roughness and bubble scattering components can be related to wind speed, they are otherwise uncoupled, with their contributions adding incoherently to the total acoustic scattered field.

A. Microwave backscattering from rough water surfaces

No model exists today that completely explains microwave scattering from the region near the air/sea interface (Keller *et al.*, 1992). Nevertheless, general agreement exists within the microwave scattering community that, except possibly at the very highest wind speeds, scattering of microwaves takes place at the actual density interface between the air and the water. Early attempts to explain microwave backscatter from the sea surface in terms of bubbles or spray above the surface failed to explain either the dependence of the scattering on polarization or the dependence on microwave frequency (Goldstein, 1951). Acoustic scattering, on the other hand, is acknowledged to contain components from both the actual sea surface and from the bubbles in the water column below the surface when the interface is contained in the ensonified volume (McDaniel and Gorman, 1982). The relative contributions of the two types of scattering to the total acoustic scatter from the near-surface region are uncertain at present but must depend on the acoustic wavelength, grazing angle, and wind speed. In general, for the acoustic frequencies of $O(10-100)$ kHz used in this experiment, we

expect that surface scattering will dominate the acoustic backscatter for very large grazing angles and that bubbles will dominate for very small grazing angles. Our primary interest in this paper is with the mid-range of grazing angles where transition from one regime to the other is more likely to be observed.

The most generally accepted model of microwave backscattering from rough air/water interfaces at grazing angles from about 20° to 70° is composite surface theory (Wright, 1966, 1968; Bass *et al.*, 1968a, 1968b). We note that these limits on the range of applicability of this theory are somewhat uncertain and environmentally sensitive. According to present understanding, quasi-specular, or Kirchhoff, scattering begins to contribute to the backscattering at large grazing angles and other, highly speculative, types of surface scattering begin to play a role at small grazing angles. Composite surface theory decomposes the surface into waves of two scales and postulates that the shorter ones cause backscattering according to Bragg resonant theory. Thus when the grazing angle, θ , surface wave number, k_B , and microwave number, k_0 , satisfy the condition

$$k_B = 2k_0 \cos \theta, \quad (9)$$

backscattering occurs. Effectively, the incident microwaves Fourier decompose the small-scale surface since backscattering is negligible if Eq. (9) is not satisfied. Thus σ_0^m is related to the spectral density at only one wave number (within a certain resolution), that of the Bragg wave. The relationship is given by

$$\sigma_0^m = 16\pi k_0^4 |g_p(\theta)|^2 \psi(0, k_B), \quad (10)$$

where ψ is the spectral density of the surface wave normalized so that its integral over all (k_x, k_y) is the mean-square wave height and g is a function determined by boundary conditions which depends on the polarization indicated by the subscript p . For vertically polarized radiation on both transmit and receive,

$$g_v(\theta) = \frac{(\epsilon - 1)[\epsilon(1 + \cos^2 \theta) - \cos^2 \theta] \sin^2 \theta}{(\epsilon \sin^2 \theta + \sqrt{\epsilon - \cos^2 \theta})^2}, \quad (11)$$

while for horizontally polarized radiation on both transmit and receive,

$$g_h(\theta) = \frac{(\epsilon - 1) \sin^2 \theta}{(\sin \theta + \sqrt{\epsilon - \cos^2 \theta})^2}, \quad (12)$$

where ϵ is the relative dielectric constant of the surface. The difference in these two equations stems from the fact that vertically polarized radiation must satisfy Neumann boundary conditions at the surface whereas horizontally polarized radiation must satisfy Dirichlet boundary conditions. According to the composite surface model, the longer waves tilt, advect, and modulate the shorter waves, thus changing the Doppler spectrum and cross section of the backscatter. In the case of acoustic scattering, for example, the effect of longer waves on the Doppler spectrum has been observed by Igarashi and Stern (1971), in measurements made between 15 and 60 kHz. In the case of microwave scattering, this is taken into account in the cross section by allowing the local

grazing angle to vary with the slope of the long waves and allowing this surface tilt to mix the contributions of horizontally and vertically transmitted signals. Thus the local grazing angle becomes

$$\theta = \theta_0 + \phi, \quad (13)$$

where θ_0 is the mean grazing angle and ϕ is the tilt of the local normal in the plane of incidence. Mixing of horizontal and vertical contributions brought about by large-scale surface tilts, α , perpendicular to the plane of incidence is taken into account by modifying g_p as follows:

$$g_v(\theta_0, \phi, \alpha) \doteq g_v(\theta), \quad (14)$$

$$g_h(\theta_0, \phi, \alpha) \doteq g_h(\theta) + (\alpha/\cos \theta)^2 g_v(\theta). \quad (15)$$

Tilts out of the plane of incidence modify horizontally polarized cross sections more than vertically polarized ones. The composite surface cross sections are obtained by simply averaging the Bragg scattering cross section, with these replacements, over the probability distribution, P , of long wave slopes:

$$\sigma_0^m = 16\pi k_0^4 \int |g_p(\theta_0, \phi, \alpha)|^2 \psi[0, 2k_0 \cos(\theta)] \times P(\alpha, \phi) d\alpha d\phi. \quad (16)$$

Note that in the limit of large ϵ , $g_h(\theta) = \sin^2 \theta$, which, when used in Eqs. (8) and (10) (divided by 4π), yields exactly the expression for the acoustic cross section for backscattering from a slightly rough, perfectly free surface such as the air/water interface (Bass and Fuks, 1979, p. 100). Thus acoustic and HH microwave Bragg backscattering yield the same scattering strength for a perfectly free, perfectly conducting, slightly rough air/sea interface. According to composite surface theory, this relationship is disrupted somewhat owing to the admixture of vertically polarized scattering into the horizontally polarized return because of long wave tilts out of the plane of incidence. According to Eq. (15), the difference will become more pronounced for smaller grazing angles. The HH scattering strength is always increased by this effect, however, so according to composite surface theory it will provide an *upper limit* to the acoustic scattering strength if the acoustic scatter comes from the air/sea interface.

The grazing-angle, azimuth-angle, and wind-speed dependence of the microwave backscattering cross section have been extensively investigated (Schroeder *et al.*, 1982; Wentz *et al.*, 1984; Durden and Vesecky, 1985; Plant, 1986, 1987; Donelan and Pierson, 1987; Apel, 1994). All models and data agree that σ_0^m increases with wind speed and grazing angle and varies as the second harmonic of the azimuth angle, i.e., the angle between the antenna pointing direction and the wind. Figure 1 shows the wind-speed, grazing-angle, and azimuth-angle dependence of σ_0^m at K_u band for HH scattering, predicted by the models of Plant (1986) and Wentz *et al.* (1984). Plant's model is based on composite surface theory. We have modified his angular dependence of the short-wave growth rate to make the predictions for cross-wind conditions agree better with measurements, letting it be $(\sin^2 \theta + 0.25)/1.25$ rather than $|\sin \theta|$. We have done this in

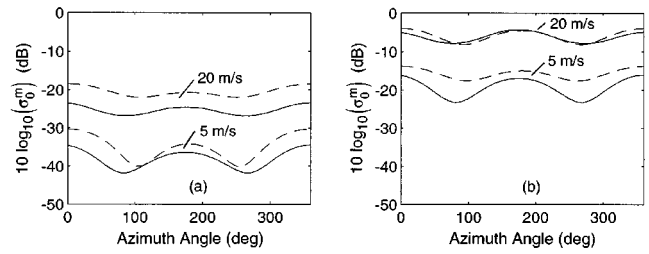


FIG. 1. Comparison of $10 \log_{10}(\sigma_0^m)$ for HH polarization predicted by the composite surface model of Plant (1986) (solid line) and the empirical model of Wentz *et al.* (1984) (dashed line) as a function of azimuth angle. Wind speeds are 5 m/s for the lower sets of lines and 20 m/s for the upper. (a) Grazing angle = 30° . (b) Grazing angle = 60° .

order to be able to predict the azimuthally averaged cross sections at other microwave frequencies and grazing angles (see Sec. VI). The model of Wentz *et al.* (1984) is an empirical one developed for K_u band backscatter at grazing angles above 30° and is not directly applicable to other frequencies and grazing angles.

Apart from the aforementioned difference of 4π and tilting effects, K_u band HH scattering is equivalent to acoustic scattering at 70 kHz. The results discussed in this paper will show that the significant azimuthal dependence due to air/sea interface scattering depicted in Fig. 1 will not be observed in acoustic scattering at grazing angles less than about 45° because of the large scattering contribution from near-surface bubbles, which themselves scatter isotropically. (Note that the bubble field may itself have structure, e.g., lineal elements associated with Langmuir circulation, that could produce a quite different azimuthal dependence.) The exception is at the lowest wind speeds, and here it is important to note the enhanced azimuthal dependence as seen in Fig. 1. Reeves *et al.* (1969), for example, found a difference of 3–4 dB between upwind and cross-wind acoustic measurements made at 60 kHz and at grazing angles between 20° and 40° ; this difference vanished for wind speeds ≥ 5 m/s. Their results, covering an azimuthal range of 90° , are consistent with the first 90° shown in Fig. 1(a). Because of bubble scattering, azimuthal dependence in acoustic scattering from the sea surface is more likely to be observed at higher grazing angles, as shown in Fig. 1(b). Still, azimuthal dependence can be more readily documented in measurements made with airborne microwave systems, which can rapidly sample a full 360° in azimuth angle (e.g., Jones *et al.*, 1977) within a very short time span, a measurement difficult to duplicate with acoustic systems.

B. Acoustic backscattering from near-surface bubbles

Clay and Medwin (1964) first showed quantitatively that only a very low density of near-surface bubbles would be necessary to account for the levels observed in some surface backscattering measurements made at high frequency and at moderate to low grazing angles (in their case, 60 kHz and 45°). They constructed a simple model based on resonant scattering from bubbles which accounted for the attenuation

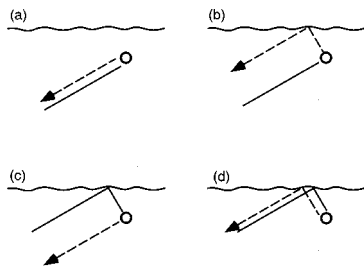


FIG. 2. (a)–(d) The four paths by which backscattering from subsurface bubbles can occur. Solid lines represent incident intensity and dashed lines represent backscattered intensity. Paths (b) and (c) add coherently.

due to bubble scattering and absorption along the transmission path, equivalent to a first-order correction to multiple scattering (Ishimaru, 1978), and for the multiple ensonification of bubbles from rays having been forward reflected from the sea surface.

To see the impact of bubbles on near-surface backscattering, consider the scattering from a single resonant-sized bubble located within a volume V bounded by the surface and unresolvable from the surface return owing to finite pulse width. The acoustical resonance characteristics of bubbles, including relevant equations, are reviewed by Clay and Medwin (1977). We need only the backscattering cross section σ_{bs} (in square meters) for a single bubble at resonance, which is a_R^2/δ^2 , where δ is the total damping constant (evaluated at the resonance frequency) and a_R is the resonant-bubble radius. The resonant damping constant can be approximated by $\delta = 2.55 \times 10^{-3} f^{1/3}$, representing measurements by Devin (1959), and a_R (in meters) is approximately $3.25/f$, where f is frequency (in hertz). The target strength, $10 \log_{10}(\sigma_{bs})$, for a single bubble resonant at 70 kHz ($a_R \sim 46 \mu\text{m}$) is thus -67 dB. Such a bubble would be ensonified via four paths (Fig. 2) because of the reflections from the nearby surface, two of which are reciprocal and therefore add coherently, which increases the scattering by a factor of 6. To convert to the standard σ_0^a , we normalize by the ensonified area, 9 m^2 for the measurements reported below, giving an effective scattering strength of -69 dB due to a single resonant-sized bubble contained within volume V .

The -69 dB value may be readily exceeded if scattering from bubbles whose radii are close to the resonant radius is taken into account. Nevertheless, -69 dB (or -58 dB in terms of σ_0^m) is already comparable to the levels expected for Bragg scattering at a 17° grazing angle for wind speeds near 1–2 m/s, as seen in the K_u band measurements presented in Sec. IV. Any modest increase in bubble concentration caused by wave breaking (e.g., Blanchard and Woodcock, 1957; Thorpe, 1982; Medwin and Breitz, 1989; Dahl and Jessup, 1995) will cause bubble scattering to greatly exceed Bragg scattering. This same qualitative conclusion was reached by McDaniel and Gorman (1982) in their retrospective comparison of acoustic and microwave measurements made at different locations.

To remove the dependence on scattering area and implied scattering volume, consider first the volume-backscattering coefficient s_V (Urlick, 1983), representing the

backscattering from N bubbles per unit volume, and its depth-integrated form, the column-scattering coefficient s_I which is dimensionless. Thus, within single scattering theory (no accounting for bubble attenuation), the part of the surface scattering strength caused by scattering from near-surface bubbles is equal to $10 \log_{10}(6s_I)$, assuming, as usual, that all the bubbles are ensonified. (Note that this assumption is crucial, and not always satisfied. In the Appendix we summarize the relevant geometrical considerations.) Crowther (1980) first expanded on this idea, taking into account the bubble-mediated attenuation occurring along the four paths involved in the multiple ensonification, and developed a model for surface backscattering strength due to the near-surface bubble layer.

Sakar and Prosperetti (1994) have approached the problem of near-surface bubble scattering via an effective media formulation (bubbles/water mixture) after Foldy (1945), which accounts for both coherent and incoherent scattering effects. They also carry out a WKB approximate solution to the effective equations, with both approaches converging in the high-frequency, incoherent limit to Crowther's formulation. The Sakar and Prosperetti result is more general, however, accounting for nonresonant-sized bubbles and not just those with radii at or close to the resonant radius. McDaniel (1993) also discusses Crowther's formulation in context of high-frequency backscattering, i.e., when the bubble layer is thick on a wavelength scale, and, as Sakar and Prosperetti do, makes clear the role of the four scattering pathways involved in the multiple ensonification.

We interpret our backscattering data using an underlying model similar to the one discussed in the above three works, but with some simplifications that facilitate implementing an inversion scheme. First, our primary variable relating to the density of bubbles is β_I , defined as the depth-integrated extinction cross section per unit volume due to scattering and absorption. This is a dimensionless quantity that succinctly characterizes the density of near-surface bubbles vis-à-vis their effect on acoustic scattering and absorption. The depth dependence in bubble density has been integrated out, since depth dependence is not an observable in our backscattering measurements made at a shallow grazing angle. Second, we assume that β_I is equal to $4\pi(\delta/\delta_r)s_I$, where δ_r is the radiation damping coefficient at resonance, taken to be 0.0136. This *hybrid* resonant approximation uses the resonant values for the damping coefficients δ and δ_r but otherwise need not presume contributions to β_I are due solely to resonant-sized bubbles.

Our equation for the normalized acoustic backscattering cross section due to near surface bubbles, σ_0^b , is thus

$$\sigma_0^b = \frac{\sin \theta}{8\pi} \frac{\delta_r}{\delta} (1 + 8\beta e^{-2\beta} - e^{-4\beta}), \quad (17)$$

where $\beta = \beta_I / \sin \theta$. In the high-bubble-density limit, where $\beta_I \gg O(1)$, and for $\theta \lesssim 30^\circ$, Eq. (17) reduces to

$$\sigma_0^b = \frac{\sin \theta}{8\pi} \frac{\delta_r}{\delta}. \quad (18)$$

Here scattering becomes independent of bubble density, as the bubble layer is thick on an absorption scale (Crowther, 1980) and follows a $\sin \theta$ law owing to an increase in the projected scattering area without an equivalent increase in the backscattered return. This effect, known as saturation (McDaniel, 1993), has been verified in a number of experiments (e.g., Lilly and McConnell, 1978; Nützel *et al.*, 1994). In the low-bubble-density limit $\beta_I \ll 1$, Eq. (17) reduces to

$$\sigma_0^b = 1.5 \frac{\beta_I}{\pi} \frac{\delta_r}{\delta}; \quad (19)$$

thus the bubble-mediated surface scattering is directly proportional to the depth-integrated density of bubbles via β_I and is independent of grazing angle θ . [Note that for extremely small grazing angles, where β_I becomes large relative to θ , Eq. (19) will revert to a $\sin \theta$ dependence.] Finally, there are geometrical considerations to be addressed when acoustic surface-scattering measurements are normalized by surface area but are interpreted in terms of Eq. (17), which deals with a *volume* scattering effect. These are discussed in the Appendix.

III. EXPERIMENTAL DESCRIPTION AND ANALYSIS METHODS

Most of the microwave and acoustic data to be discussed in this paper were obtained during the Synthetic Aperture Radar and X Band Ocean Nonlinearities (SAXON)-Forshungsplattform Nordsee (FPN) experiment. The overall SAXON-FPN program consisted of two parts, Phase I, which took place in November of 1990, and Phase II, which took place in November and December of 1991. Both phases were conducted on the German research platform *Nordsee* (FPN) in the German Bight area of the North Sea (54°42'5.8"N, 7°10'9.2"E). While the primary objective of these experiments was to understand synthetic aperture radar imagery of the ocean, a second objective was to investigate acoustic scattering from the near-surface region. During both phases of the experiment, environmental conditions were well monitored. Winds were measured using sonic anemometers deployed on booms extending 20 m away from the platform. Waves were measured using a pitch, roll, heave buoy located within 2 km of the platform. Other environmental parameters such as atmospheric stability, heat and moisture fluxes, and currents were also measured. An overview of the entire SAXON-FPN experimental program is given by Plant and Alpers (1994).

Microwave and acoustic measurements were made at a variety of grazing angles and frequencies during both phases of SAXON-FPN. During Phase II microwave and acoustic measurements were made simultaneously at the same wavelength and grazing angle and looking at the same surface area. The C band microwave system, along with the acoustic and environmental recording instruments, was deployed again on the FPN in March 1992 in the same configuration used in the December 1991 Phase II measurement series, and all of the C band measurements reported here were from

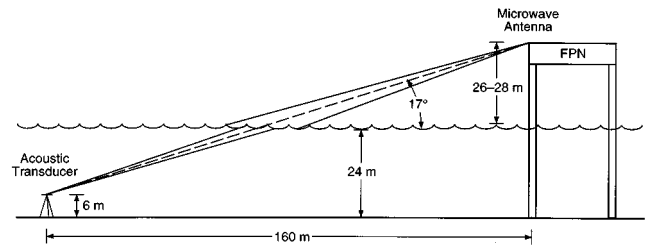


FIG. 3. View of experimental arrangement from the south. The microwave antennas looked west; the acoustic beam was directed east.

these later runs. Thus the total data set reported on here was obtained from December 1–4, 1991, and on March 6 and 9, 1992, except for a set of data collected at a 45° grazing angle during Phase I, which we discuss in Sec. VI.

Figure 3 shows the scattering geometry for the December 1991 and March 1992 acoustic and microwave backscattering measurements made at a grazing angle of 17°. Note that the location of the acoustic transducer and microwave systems required that they be directed in opposite directions. In order to be certain that the microwave and acoustic systems were in fact observing the same surface area, a small boat was deployed that carried targets for both the acoustic and microwave systems. The systems were adjusted in grazing and azimuth angle to maximize the return from these targets.

The acoustic instrumentation and data processing methods were the same during all experiments; details are given by Nützel and Herwig (1995), who discuss Phase I acoustic measurements. Briefly, the acoustic system was an incoherent, narrow-beam system whose center frequency could be tuned over a wide range of values in the kilohertz range. The transducer was mounted on a tower 6 m above the bottom in 30 m of water and located 160 m west of the FPN. For the data reported below, the pulse width was 4 ms and the system was tuned to center frequencies of 26.5, 50, and 70 kHz, with half-power, two-way beamwidths of 5.7°, 2.9°, and 2.0°, respectively.

The microwave systems were all coherent, CW systems of the basic design described by Plant *et al.* (1994). These systems operated at 5.3 GHz (C band), 10 GHz (X band), and 14 GHz (K_u band), corresponding to the same wavelengths as the acoustic systems at 26.5, 50, and 70 kHz, respectively, and measured backscatter for HH and VV polarizations simultaneously. The half-power, two-way beamwidths of these systems were 2.6° for C band, 1.4° for X band, and 6.6° in the E plane and 5.0° in the H plane for K_u band. The C and X band systems were mounted on the FPN 28 m above the mean sea level while the K_u band system was at a height of 26 m. The K_u band system was fully calibrated both before and after the experiment using corner reflectors. The other two systems were not absolutely calibrated, although their relative calibrations were monitored during the experiment. We have set the absolute calibration of these two systems by forcing the X band system to agree with the K_u band during periods when they operated together and by setting the C band cross section at a 45°

TABLE I. Saturation scattering strength levels of σ_0^a at a 17° grazing angle.

Frequency (kHz)	Phase II saturation level, forced (dB)
26.5	-26.9
50	-27.8
70	-28.3

grazing angle and a wind speed of 10 m/s to the level specified by the ERS-1 CMOD4 scatterometer algorithm (Stofelen and Anderson, 1993).

Finally, we note that a calibration uncertainty was discovered in both the Phase II and March 1992 acoustic data. Therefore, to proceed with our analysis we have forced the data to match their respective saturation values (Table I), as predicted by Eq. (18), at very high wind speeds; for this purpose, we selected data taken at $11 < U_{10} \leq 13$ m/s, where U_{10} is wind speed corrected to a 10-m height. Evidence that the predicted saturation values are both correct, and repeatable, can be seen in the 18-kHz data for grazing angles 15° – 30° reported by Nützel *et al.* (1994), which were gathered at this same location in 1985.

IV. DIFFERENCES IN MICROWAVE AND ACOUSTIC RESPONSE TO THE WIND

Figure 4 shows a 3-h time series of 10-m wind speed, 70-kHz acoustic scattering strength, and K_u band microwave HH scattering strength, representing 1-min averages of each variable. Visible are temporally coherent sections (cross hatched and labeled A–D) insofar as both acoustic and microwave scatter generally rise and then fall in the same time scale, which is on the order of 5 min. Within some of these sections there are also oscillations with a shorter time scale that can be identified in both the acoustic and microwave data. Sections A–D correspond to periods when the wind speed exceeded a threshold of 1 m/s and then returned below this value. Stolte's (1994) observations from the FPN point to a threshold wind speed for excitation of small-scale, or capillary, surface waves of around 1.5 m/s. Viewing sections A–D as being consistent with Stolte's data, their coherent properties thus suggest that the acoustic and microwave systems are responding to similar surface manifestations, in this case centimetric Bragg waves. The scattering levels in sections A–D are roughly comparable for both systems, although the microwave levels are generally greater, consistent with HH surface scattering being an upper bound to acoustic surface scattering as noted earlier. The levels are closest for section D, the most energetic one in terms of the mean-square wind speed. At first glance there also appears to be a difference in the dynamic response of the acoustic and microwave systems to wind-speed fluctuations, as noted by the change in scale. The difference in the noise floor for each system (about -78 dB for the acoustic data and -74 dB for the microwave data) may in large part be responsible for this effect.

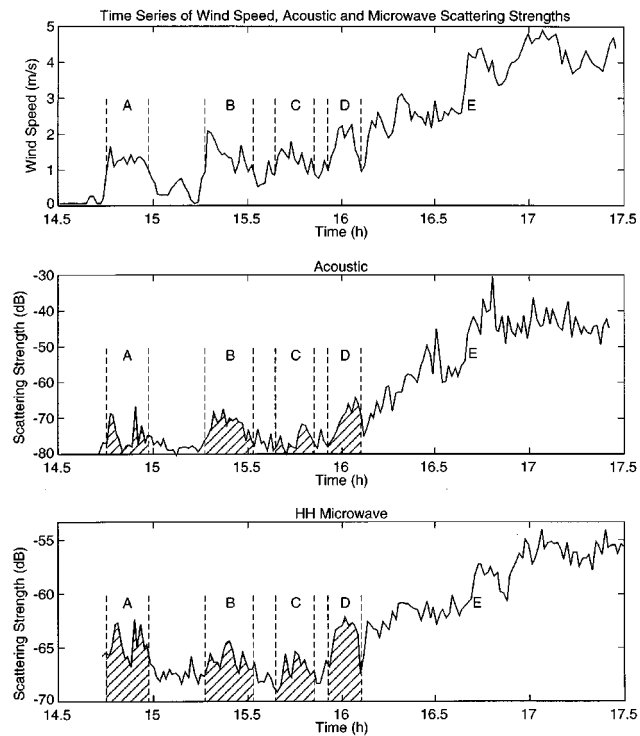


FIG. 4. Time series of wind speed, 70-kHz acoustic scattering strength, and K_u band microwave HH scattering strength, representing 1-min averages of each variable. Wind speeds are at a height of 10 m, and the microwave scattering strength is the cross section minus 11 dB. The cross-hatched areas (labeled A–D) correspond to periods when the wind speed exceeded a 1-m/s threshold and then returned below this value. Note the expanded scale for microwave scattering strength.

In section E, wind-speed oscillations occur on top of a slowly increasing mean, with the wind speed remaining above 2 m/s. Here, it is more difficult to identify fluctuations in acoustic scattering that are coherent with wind-speed fluctuations, whereas the microwave and wind-speed time series maintain a better correlation. We postulate that between wind speeds of about 2 and 3 m/s, bubble production contributes significantly to the acoustic scatter and eventually dominates the scatter originating from the sea surface. This is consistent with Wille and Geyer's (1984) observations of ambient noise (also made from the FPN), which show a threshold wind speed of approximately 2.5 m/s for the onset of sporadic, small-scale wave breaking and splash noise events. In addition to bubbles produced by small-scale wave breaking, bubbles can also be encapsulated by action of capillary waves, a mechanism that may be in effect below a wind-speed threshold for wave breaking. Kolaini *et al.* (1994) have demonstrated in wavetank studies the causal relation between capillary-wave bubble formation and radiated-noise transients from such bubbles. The presence of bubbles, regardless of their origin, if linked to the wind speed will change the acoustic dynamic response to wind-speed fluctuations, and in section E the response is clearly different. For example, starting near the label E, the wind speed increases from about 3 to 5 m/s, producing a 15-dB change in the acoustic scattering level and a 5-dB change in

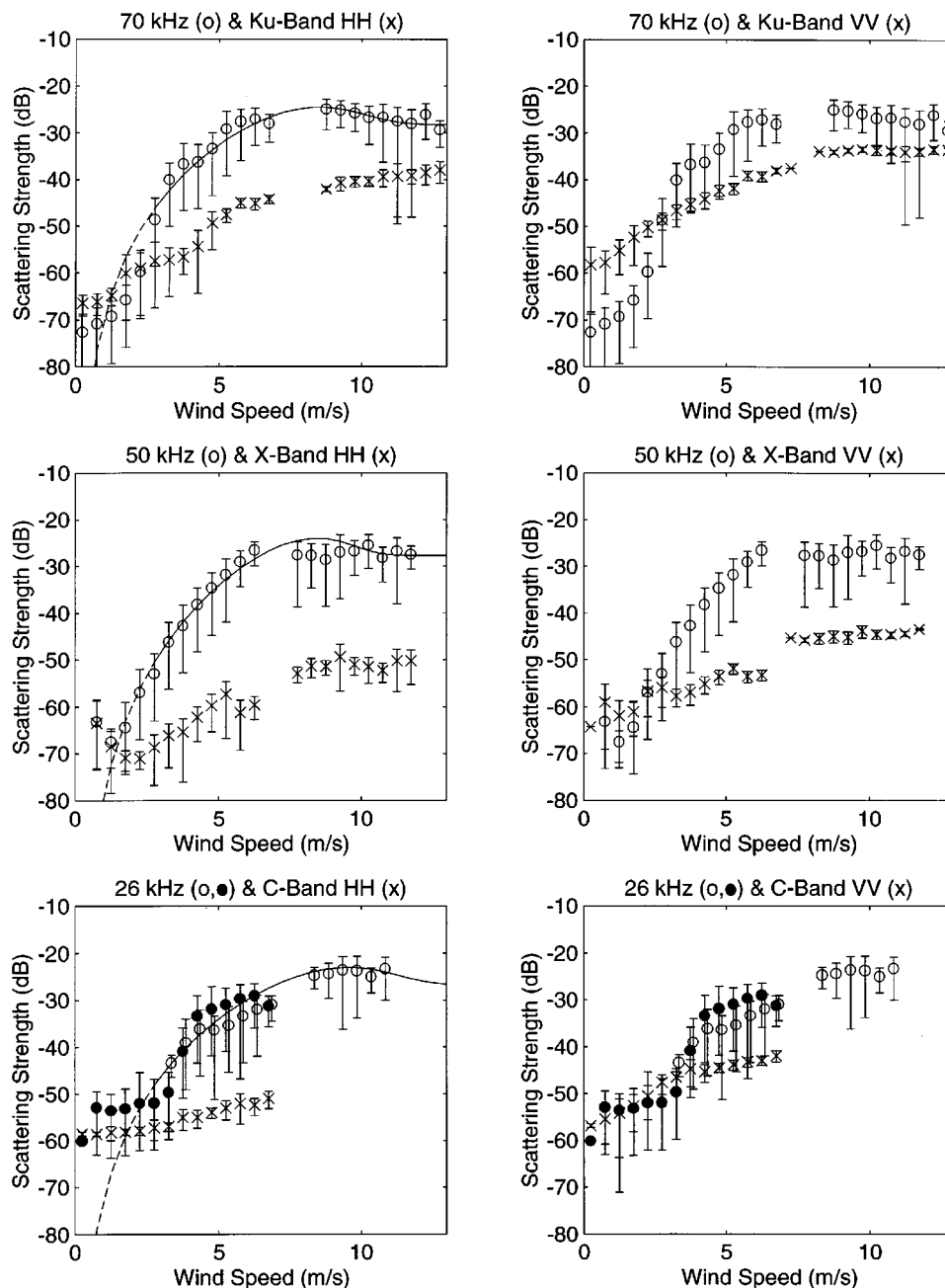


FIG. 5. Measured acoustic and microwave scattering strengths versus wind speed measured at a height of 10 m. The left-hand column compares acoustic with HH scattering strengths. The right-hand column compares the same acoustic data with VV scattering strengths. The systems viewed the same area of the sea surface from opposite directions. All averaging times are 1 min, and the acoustic, microwave, and wind data are simultaneous to within 31 s. Microwave scattering strength is the cross section minus 11 dB. Grazing angles were 17° for all systems. The left-hand column also shows a comparison with Eq. (17), where $\log_{10} \beta_I = a + b \log_{10} U_{10}$ and a and b are given in Table II.

the microwave scattering level. Below, we examine more quantitatively the acoustic and microwave response to a much larger wind-speed range, using averaged data taken over the course of several days.

We have compared our microwave and acoustic backscattering strengths by binning them along with the wind speed and direction into nearly coincident 1-min bins. Because the sampling times and rates were slightly different, these bins were not exactly coincident, but in no case did the centers of the time intervals over which the data were averaged differ by more than 31 s. These coincident data were then matched with the 1-min average of the wind speed mea-

sured within 31 s of the acoustic and microwave data. Figure 5 shows the result of plotting the linearly bin-averaged acoustic, HH, and VV data versus wind speed corrected to a 10-m height. The error bars indicate the standard deviations of the averaged data values in the different wind-speed bins and are representative of the typical spread in scattering strength observed for a given wind speed. The solid lines in the left-hand plots are the result of fitting the data to Eq. (17) and will be discussed in detail in the following section. We have included the VV data primarily for completeness and archival purposes; our comparison with the acoustic results continues using only the HH data. Wind directions in Fig. 5

TABLE II. Frequency, equivalent resonant radius (at surface), and estimated parameters a and b .

Frequency (kHz)	Resonant bubble size (μm)	a	b
26.5	121	-6.16	5.50
50	64	-6.70	6.40
70	46	-5.75	5.35

are unrestricted but, while the wind direction shifted over a full 360° during the measurements, they were generally different at different wind speeds. We attempted to observe a dependence of the scattering strength on the azimuth angle between the antenna or transducer look direction and the wind, but the differences in wind speed in the different directions made this inconclusive.

From the considerations given in Secs. I and II A, we would expect the HH scattering strengths to be an upper limit to the part of the acoustic scattering strength that actually comes from the rough air/sea interface. All of the data in Fig. 5 reach a noise level at wind speeds near or perhaps below 1 m/s, but just above this level the HH scattering strength is comparable to the acoustic one. Above about 3 m/s, the acoustic scattering strength always exceeds the HH scattering strength. If we try to fit the data to a power law in wind speed of the form $\sigma_0^{a,m} \sim U^n$, then n is approximately 5.3, 6.4, and 5.5 for the acoustic data at 70, 50, and 26.5 kHz, respectively, for the wind-speed range of 2–7 m/s. For same wind-speed range and HH data, n is approximately 3.8, 3.3, and 2.1 for the K_u , X, and C bands, respectively, while for the VV data, n is approximately 2.8, 2.3, and 1.8.

From these results we conclude that, at our 17° grazing angle, most of the acoustic scattering from the near-surface region is caused by bubbles very near the surface as long as the wind speed is above about 3 m/s. This bubble population must increase much more rapidly than the spectral densities of centimetric surface waves in order to account for the observations. In the next section, we fit this bubble scattering to the form given in Sec. III B.

V. COMPARISON OF THE ACOUSTIC BUBBLE MODEL WITH THE DATA

Figure 5 also shows the model fit to the data that was obtained by numerically inverting Eq. (17) using Newton–Raphson iteration. Note that we have drawn the model as a dashed line for wind speeds < 2.5 m/s to emphasize the extrapolation into this low-wind-speed regime and that we are less certain of the model's shape here, since surface scattering is likely adding to the bubble scattering. The resulting estimates of β_I versus wind speed have been summarized into the form $\log_{10} \beta_I = a + b \log_{10} U_{10}$. The values of a and b that produced these fits for the three acoustic frequencies are given in Table II along with representative values for the radius of resonant-sized bubbles, which dominate the scatter. Note that, as expected, the three estimates of b reflect the three estimates of n given above for σ_0^a . The data taken at 70 kHz most clearly show the expected overshoot prior to reaching the saturation region caused by bubble attenuation.

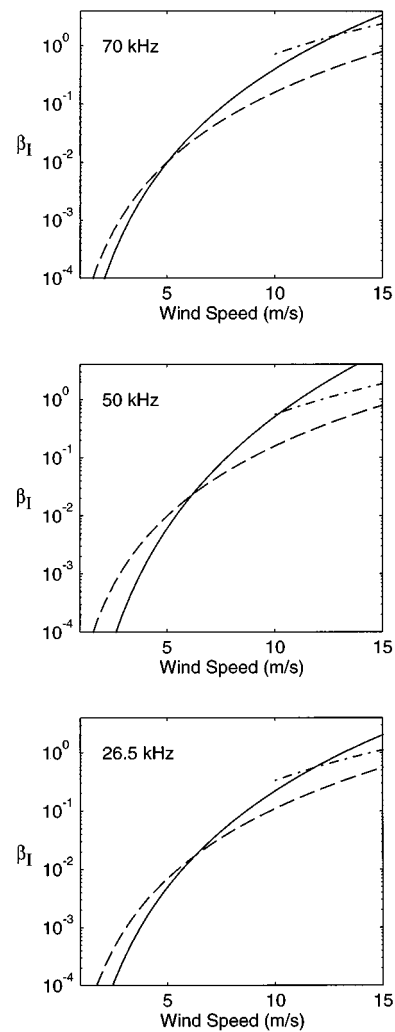


FIG. 6. Comparison of estimates of β_I . The solid lines represent our frequency-dependent expressions for β_I as defined in Table II. The dash-dotted lines represent measurements reported by Farmer and Lemon (1984). These measurements were made at 25 kHz and therefore have been modified (see text) to compare with our results at 50 and 70 kHz. The dashed line is an empirical model given by Crowther *et al.* (1993) that was derived from measurements of β_I .

For all three frequencies the model is nearly ideally encompassed by the error bars, with very few exceptions. The reason for these exceptions is not entirely clear at this time, but may be related to variability in the wind speed.

Figure 6 compares our estimates of β_I as defined in Table II with data representations from two experiments in which β_I was estimated by inverting ambient-noise measurements. The dashed line is an empirical model by Crowther *et al.* (1993), hereafter referred to as CGH. We have extended the portion of their model (via a polynomial fit) that applies to their five frequencies between 20 and 80 kHz in order to compare their results with our three frequencies of 26.5, 50, and 70 kHz. The wind speed used in the CGH model is approximately equivalent to U_{10} . The dashed-dotted line is a representation of β_I estimates reported by Farmer and Lemon (1984), hereafter referred to as FL, in which we represent their data taken at 25 kHz by $\log_{10} \beta_I = -3.5 + 2.9 \log_{10} U$. In this case U is the wind speed reported at two nearby land stations; differences between U

and U_{10} are expected to be small at higher wind speeds. The trend in wind speed in the FL data becomes reliable only for $U > 10$ m/s, which complements our range owing to saturation; below 10 m/s FL observed increased variability without a clear pattern. Our estimates of β_I increase with frequency as approximately f^n , where $n \sim 0.8$; this dependence is most apparent when comparing the 26.5- and 70-kHz model parameters in Table II. We have therefore applied this same frequency dependence to the FL results in order to compare them with our results at 50 and 70 kHz.

The three data representations agree well on a magnitude basis, which is possibly the best we can expect considering the large variance that usually accompanies *in situ* measurements of near-surface bubble populations. Bubble concentration, as parametrized by β_I , generally increases from $O(10^{-3})$ to $O(1)$ within the wind-speed range 3–15 m/s. Our measurements tend to go as U_{10}^b , with $b \sim 6$, and are consistent with other measurements (e.g., McDaniel, 1988; McConnell and Dahl, 1991; and Dahl and Jessup, 1995) covering the range of wind speeds from 4–11 m/s. The CGH empirical formulas go as U_{10}^4 , though still give magnitudes comparable to our results. Our results do point to a significantly higher power-law dependence for β_I than the approximately cubic dependence on wind speed suggested by the FL data. A more detailed model for β_I would likely have a wind-speed dependence that varies, perhaps, for example, transitioning to a cubic dependence, as suggested by Kerman (1986) and Wu (1988), at higher wind speeds and for fully developed conditions.

Finally, it is useful to interpret the models for β_I in Table II in terms of resonant scattering and absorption. Our purpose is twofold. One is to evaluate the effect of bubble-mediated changes in sound speed on the interpretation of our backscattering results. The other is to facilitate further comparison with other kinds of *in situ* bubble measurements. Using the resonant approximation, we can relate β_I directly to the depth-integrated distribution of resonant-sized bubbles N_I , where

$$N_I(a_R) = \frac{\beta_I \delta_r}{2\pi^2 a_R^3}. \quad (20)$$

This quantity when multiplied by da (usually taken to be 1 μm) then is equal to the number of resonant-sized bubbles per unit area of sea surface. Equation (20) follows from inspection of our Eq. (19) and Sarkar and Prosperetti's Eq. (A8), who also outline in detail the cautions in using the resonant approximation, the errors from which are mitigated for measurements made within our frequency range (see also Commander and Moritz, 1989). To recover an estimate of the bubble-size distribution function $N(a)|_{a_R}$, we must assume a depth dependence for $N(a)$, as this was not an observable in our measurements. Taking an exponential depth dependence for $N(a)$ with e -folding scale $L \approx 0.1 U_{10}$ gives

$$N(a_R; z) = \frac{N_I(a_R)}{L} e^{-z/L}. \quad (21)$$

With our three frequencies spanning resonant bubble radii between 46 and 121 μm , we can construct an empirical

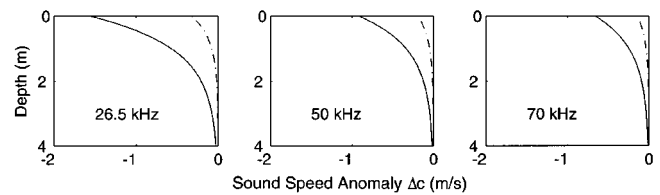


FIG. 7. Representative profiles of the sound-speed anomaly Δc at three frequencies as derived from estimates of β_I . Dashed line is for a wind speed of 7 m/s and solid line for a wind speed of 10 m/s.

size-distribution function $N(a; z) \sim a^{-n}$, where the frequency dependence of $f^{0.8}$ suggests a slope estimate $n \approx 3.8$. With an estimate of $N(a; z)$ we can compute the bubble-mediated sound-speed anomaly (Medwin, 1974; Clay and Medwin, 1977) as a function of depth. (Note that the slope n must apply to all depths, which admittedly is an approximation.) The sound-speed anomaly is defined as $\Delta c = c_1 - c_0$, where c_1 is the sound speed in a bubbly medium and c_0 is the sound speed in a bubble-free medium, and results are shown in Fig. 7 for two representative wind speeds. The example at 10 m/s wind speed (solid line) is quite consistent with Farmer and Vagle's (1989) result for the same wind speed, which was derived from measurements conducted off Vancouver Island using four frequencies that spanned resonant bubble radii between 16 and 116 μm . With $\Delta c < 0$ there is mild upward refraction, but this has negligible impact on the interpretation of our measurements made at a 17° grazing angle. For very low grazing angles, the change in sound speed may have significant impact; its precise quantification, however, is beyond the scope of this paper.

VI. ACOUSTIC NEAR-SURFACE BACKSCATTERING STRENGTH AT MID-INCIDENCE ANGLES

If the acoustic bubble-scattering model and the microwave composite surface model discussed above are indeed the correct interpretations of backscattering at low grazing angles, then we would expect to be able to apply them to predict acoustic backscattering scattering strengths for higher grazing angles than actually measured in our experiment. To this end, we assume that the prediction of composite surface theory for HH backscatter is a reasonable approximation to acoustic rough surface backscatter at grazing angles between 30° and 60° . We then add to this prediction the scattering expected from bubbles in order to predict the total acoustic backscattering from the near-surface region. Before the roughness component is added, its value must first be reduced by a factor of $20(\log_{10} e)\beta_I/\sin \theta$, equivalent to the reduction in intensity caused by two-way travel through the bubble layer. Figure 8 shows the results of our calculations for the acoustic scattering strength expected at 26.5 and 50 kHz and 45° grazing angle. Bubble scattering is shown as dashed lines, surface scattering attenuated by bubbles as dotted lines, and their sum, the total backscattering strength, as solid lines. The data points are measurements obtained during Phase I of SAXON-FPN (Nützel and Herwig, 1995).

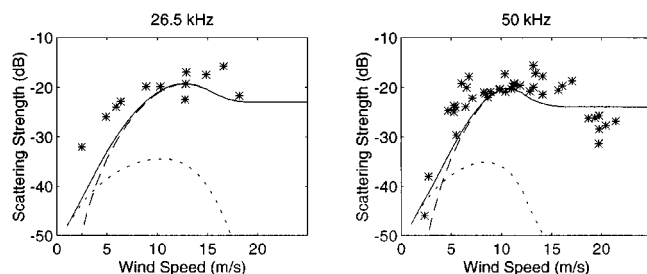


FIG. 8. Comparison of model predictions of total acoustic scattering strength with data from Nützel and Herwig (1995). Dashed line: bubbles only; dotted line: bubble-attenuated surface scatter; solid line: total acoustic backscattering strength.

While the scatter in the data is large, they seem to confirm the model's prediction that at 45° grazing angle surface scattering contributes to the total return only for low wind speeds, say $\lesssim 5$ m/s and for higher wind speeds scattering from near-surface bubbles comprises the main contribution. For low wind speeds, the addition of surface scattering produces a marginal effect, although it is in the correct direction for improving the agreement between the model and the data. The data points are plotted versus wind speed measured at a height of 47 m on the FPN whereas the models refer to neutral-stratification winds at a height of 10 m above the sea surface. Agreement between the model and measurements would be changed somewhat, especially in the low-wind-speed region, if corrections could be made for height and stratification. Unfortunately, the information to make these corrections was not available at the time the data were taken.

VII. SUMMARY AND CONCLUSIONS

Our study has shown that measurements of acoustic backscattering from the sea surface, made at 17° grazing angle and within the frequency range 26–70 kHz, are dominated by scattering from near-surface bubbles. A model for scattering from near-surface bubbles predicts the functional form of the data versus wind speed: the backscatter rises steeply with increasing wind speed and then reaches a constant saturation value, with a slight overshoot just prior to saturation. By inverting the measurements, we obtained an expression for the model's driving parameter β_I versus wind speed. This parameter is the depth-integrated extinction cross section per unit volume, and it succinctly characterizes the density of near-surface bubbles vis-à-vis their effect on acoustic scattering and absorption. Our expression for β_I is consistent with other estimates of β_I for similar wind speeds which were obtained by inverting ambient-noise measurements. We have also interpreted β_I in terms of resonant-bubble scattering and absorption in order to estimate the equivalent sound-speed anomaly Δc versus frequency. We found the average Δc near the surface to be of order 1 m/s (negative) in our frequency range for a wind speed of 10 m/s which is also consistent with other measurements.

The success of our bubble model suggests that most of these bubbles are confined to a layer of order 1 m in thickness. In the Appendix we point out that certain combinations of acoustic pulse width, beamwidth, and range can influence

the interpretation of surface backscattering in terms of scattering from near-surface bubbles. In particular, if the bubble layer exceeds a certain thickness, interpretations will be biased if the geometric considerations discussed in the Appendix are not accounted for.

We have also pointed out some differences in acoustic and microwave terminology which can confuse comparisons of the two types of backscatter from the sea surface. In particular, the standard conventions for reporting acoustic and microwave measurements of sea surface backscattering differ by 11 dB (or a factor of 4π), and 11 dB must therefore be subtracted from microwave measurements in order to compare them with acoustic measurements.

The fit of our bubble model to the measurements at 17° was used to predict backscattering at a higher grazing angle (45°). We used composite surface theory, as applied to HH-polarized microwave backscatter and attenuated by subsurface bubbles, to predict the contribution of surface scattering to the total acoustic scattering strength. While the agreement between the model and data at higher incidence angles is not entirely satisfactory, it lends credence to our primary conclusion, namely, that acoustic backscatter from near the sea surface is dominated by bubbles for moderate to high wind speeds. The wind speed at which bubble scattering becomes as important as surface scattering varies with grazing angle. At a grazing angle of 17° , this wind-speed threshold does not exceed 3 m/s; at 45° , this threshold is increased somewhat, to approximately 5 m/s.

ACKNOWLEDGMENTS

The U.S. participation in this study was supported by the Office of Naval Research through Grant No. N00014-89-J-3224 from the Remote Sensing Program and Grant No. N00014-95-1-0048 from the Ocean Acoustics Program. Much of this work was performed while William Plant was on sabbatical at the Max-Planck-Institut für Meteorologie in Hamburg, Germany, and while Peter Dahl visited Universität Hamburg. These two authors thank Dr. Klaus Hasselmann and Dr. Werner Alpers for their hospitality and assistance. German participation in this study was supported by the German Ministry of Defense and by the Bundesministerium für Forschung und Technologie in Bonn, Germany. This is Woods Hole Oceanographic Institution contribution number 9334.

APPENDIX: GEOMETRICAL CONSIDERATIONS IN ACOUSTIC BACKSCATTERING FROM NEAR-SURFACE BUBBLES

When attempting to compare microwave and acoustic scattering strengths, geometrical configurations in the acoustic scattering must be considered in addition to the terminology issues dealt with in Sec. I. This need arises because we are attempting to compare backscattering cross sections normalized by the ensonified surface area but, away from saturation, the acoustic bubble scattering modeled in Sec. II B is a volume-scattering effect whose level may vary even if the ensonified surface area is held constant. Here we define the assumptions about the bubble profile and the limitations on

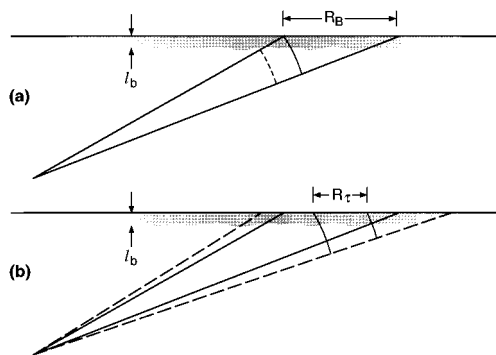


FIG. A1. Geometry of acoustic backscattering from the near-surface region. (a) Beamlimited conditions for two different pulse widths; R_B shows the range extent of ensonified surface A. (b) Pulse-limited conditions for two different beamwidths; R_T shows the range extent of ensonified surface A.

scattering geometries that must be satisfied in order for the comparison of surface and volume scattering to be unique.

Figure A1 defines the problem. Figure A1(a) shows beam-limited conditions for two different pulse widths. Though the ensonified volumes are different for the two pulse widths, the ensonified surface area, A , is constant. Similarly, Fig. A1(b) shows pulse-limited conditions for two different beamwidths. Here the ensonified surface area is not the same owing to different azimuthal extents; nevertheless, the area does not increase in proportion to the ensonified volume. Shown in both configurations is a bubble layer (which can be viewed in an ensemble-average sense) extending down to a nominal depth, l_b . The crux of the matter is how the volume defined by Al_b differs from V_{eff} , the effective volume of the ensonified cone as limited by the pulse width, beamwidth, and bubble depth.

To examine this issue to first order, we assume that the bubble density is constant down to l_b and zero below that. We then calculate the ensonified area for various beam- and pulse-limited geometries along with V_{eff} . As long as the ratio $\rho = V_{\text{eff}}/(Al_b)$ is 1 or nearly so, then measurements of surface scattering strength may be interpreted in terms of the model for scattering from bubbles (σ_0^b) presented in Sec. II B, with-

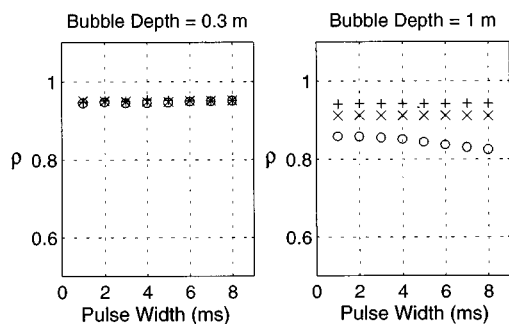


FIG. A2. Comparison of the ratio, ρ , of effective volume (V_{eff}) to ensonified area (A) times bubble depth (l_b) versus pulse width for two bubble depths and three beamwidths ($\circ = 2^\circ$, $\times = 3^\circ$, and $+ = 6^\circ$). The transducer depth (24 m), grazing angle (17°), and beamwidths are representative of the experimental configuration discussed in this paper.

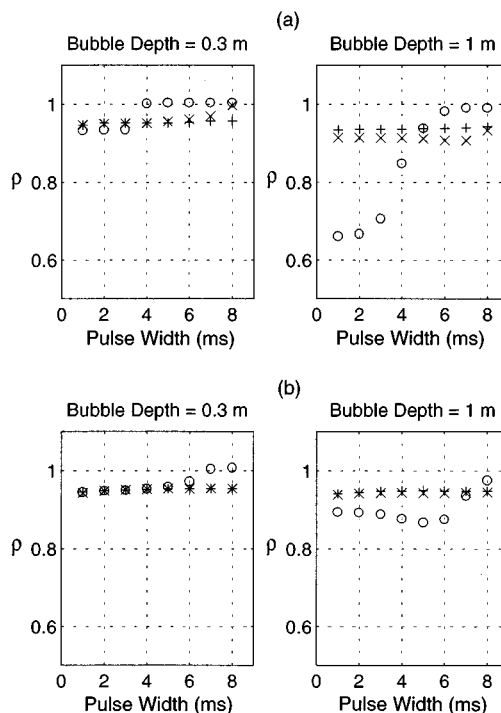


FIG. A3. As in Fig. A2, but for a grazing angle of 45° , beamwidths of $\circ = 2^\circ$, $\times = 5^\circ$, and $+ = 8^\circ$, and transducer depths of (a) 50 m and (b) 100 m.

out modification. If $\rho < 1$, then scattering measurements inverted to obtain bubble concentration estimates via Eq. (17) are biased low by approximately $10 \log_{10} \rho$, in decibels. Figure A2 shows ρ as a function of pulse width, with beamwidth as a parameter, for two characteristic bubble depths and for the geometry under which the acoustic data analyzed in this paper were collected. It is apparent that the geometry affected ρ little as long as the bubble depth was less than about 1.0 m. Even at $l_b = 1$ m, the problem is mitigated at 17° grazing angle, and lower grazing angles in general, because of the onset of saturation, at which point the bubble depth l_b becomes irrelevant.

In general, ρ decreases with decreasing beamwidth, pulse width, and transducer depth, with pulse-limited situations more likely to reduce ρ than beam-limited ones. At higher grazing angles, the saturation point is reached at higher bubble concentrations and therefore values of l_b are likely to be larger. In such cases there is greater possibility of bias, especially in situations involving shallow depths and narrow apertures. Figure A3 shows plots of ρ versus pulse width for a 45° grazing angle and two transducer depths, 50 and 100 m. Obviously the assumption that σ_0^b is independent of pulse width and beamwidth is worse for shallower source/receiver depths, and the possibility of errors on the order of 3 dB is apparent.

Apel, J. R. (1994). "An improved model of the ocean surface wave vector spectrum and its effects on radar backscatter," *J. Geophys. Res.* **99**(C8), 16 269–16 291.
 Bass, F. G., and Fuks, I. M. (1979). *Wave Scattering from Statistically Rough Surfaces* (Pergamon, New York).
 Bass, F. G., Fuks, I. M., Kalmykov, A. I., Ostrovsky, I. E., and Rosenberg, A. D. (1968a). "Very high frequency radiowave scattering by a disturbed

- sea surface. Part I: Scattering from a slightly disturbed boundary," IEEE Trans. Antennas Propag. **AP-16**(5), 554–559.
- Bass, F. G., Fuks, I. M., Kalmykov, A. I., Ostrovsky, I. E., and Rosenberg, A. D. (1968b). "Very high frequency radiowave scattering by a disturbed sea surface. Part II: Scattering from an actual sea surface," IEEE Trans. Antennas Propag. **AP-16**(5), 560–568.
- Blanchard, D. C., and Woodcock, A. H. (1957). "Bubble formation and modification in the sea and its meteorological significance," Tellus **9**, 145–157.
- Commander, K., and Moritz, E. (1989). "Off-resonance contributions to acoustical bubble spectra," J. Acoust. Soc. Am. **85**, 2665–2669.
- Clay, C. S., and Medwin, H. (1964). "High-frequency acoustical reverberation from a rough-sea surface," J. Acoust. Soc. Am. **36**, 2131–2134.
- Clay, C. S., and Medwin, H. (1977). *Acoustical Oceanography: Principles and Applications* (Wiley-Interscience, New York).
- Crowther, P. A. (1980). "Acoustical scattering from near-surface bubble layers," in *Cavitation and Inhomogeneities in Underwater Acoustics*, edited by W. Lauterborn (Springer-Verlag, New York), pp. 194–204.
- Crowther, P. A., Griffiths, H. J. S., and Hansla, A. (1993). "Dependence of sea surface noise in narrow beams on windspeed and vertical angle," in *Natural Physical Sources of Underwater Sound*, edited by B. R. Kerman (Kluwer, Dordrecht), pp. 31–44.
- Dahl, P. H., and Jessup, A. T. (1995). "On bubble clouds produced by breaking waves: An event analysis of ocean acoustic measurements," J. Geophys. Res. **100**(C3), 5007–5020.
- Devin, Jr., C. (1959). "Survey of thermal, radiation, and viscous damping of pulsating air bubbles in water," J. Acoust. Soc. Am. **31**, 1651–1667.
- Donelan, M. A., and Pierson, Jr., W. J. (1987). "Radar scattering and equilibrium ranges in wind-generated waves with application to scatterometry," J. Geophys. Res. **92**(C5), 4971–5029.
- Durden, S. L., and Vesecky, J. F. (1985). "A physical radar cross-section model for a wind-driven sea with swell," IEEE J. Ocean. Eng. **OE-10**(4), 445–457.
- Farmer, D. M., and Lemon, D. D. (1984). "The influence of bubbles on ambient noise in the ocean at high wind speeds," J. Phys. Oceanogr. **14**, 1762–1778.
- Farmer, D. M., and Vagle, S. (1989). "Waveguide propagation of ambient sound in the ocean-surface layer," J. Acoust. Soc. Am. **86**, 1897–1908.
- Foldy, L. L. (1945). "The multiple scattering of waves, I. General theory of isotropic scattering by randomly distributed scatterers," Phys. Rev. **67**, 107–119.
- Goldstein, H. (1951). "Sea echo," in *Propagation of Short Radio Waves*, edited by D. E. Kerr, MIT Radiation Laboratory Series (McGraw-Hill, New York), Vol. 13, p. 481.
- Ishimaru, A. (1978). *Wave Propagation and Scattering in Random Media* (Academic, New York), Vol. II.
- Igarashi, Y., and Stern, R. (1971). "Observations of wind-wave-generated Doppler shifts in surface reverberation," J. Acoust. Soc. Am. **49**, 802–809.
- Jackson, D. R., Winebrenner, D. P., and Ishimaru, A. (1986). "Application of the composite roughness model to high-frequency bottom backscattering," J. Acoust. Soc. Am. **79**, 1410–1422.
- Jones, W. L., Schroeder, L. C., and Mitchell, J. L. (1977). "Aircraft measurements of the microwave scattering signature of the ocean," IEEE J. Ocean. Eng. **OE-2**, 52–61.
- Keller, M. R., Keller, W. C., and Plant, W. J. (1992). "A wave tank study of the dependence of X band cross sections on wind speed and water temperature," J. Geophys. Res. **97**(C4), 5771–5792.
- Kerman, B. R. (1986). "Distribution of bubbles near the ocean surface," Atmos. Ocean **24**(2), 169–188.
- Kolaini, A. R., Crum, L. A., and Roy, R. A. (1994). "Bubble production by capillary-gravity waves," J. Acoust. Soc. Am. **95**, 1913–1921.
- Lilly, J. C., and McConnell, S. O. (1978). "Surface reverberation measurements in Dabob Bay and the open ocean," J. Acoust. Soc. Am. Suppl. **1** **63**, S24.
- McConnell, S. O., and Dahl, P. H. (1991). "Vertical incidence backscatter and surface forward scattering loss from near-surface bubbles," in *Oceans '91 Proceedings* (Institute of Electrical and Electronics Engineers, New York), Vol. 1, pp. 434–441.
- McDaniel, S. T. (1988). "Acoustical estimates of subsurface bubble densities in the open ocean and coastal waters," in *Sea Surface Sound: Natural Mechanisms of Surface Generated Noise in the Ocean*, edited by B. R. Kerman (Kluwer, Norwell, MA), pp. 225–236.
- McDaniel, S. T. (1993). "Sea surface reverberation: A review," J. Acoust. Soc. Am. **94**, 1905–1922.
- McDaniel, S. T., and Gorman, A. D. (1982). "Acoustic and radar sea surface backscatter," J. Geophys. Res. **87**(C6), 4127–4136.
- Medwin, H. (1974). "Acoustic fluctuations due to microbubbles in the near-surface ocean," J. Acoust. Soc. Am. **56**, 1100–1104.
- Medwin, H., and Breitz, N. D. (1989). "Ambient and transient bubble spectral densities in quiescent seas and under spilling breakers," J. Geophys. Res. **94**(C9), 12 751–12 759.
- Nützel, B., and Herwig, H. (1995). "Wind speed dependence of acoustic backscattering," J. Geophys. Res. **100**(C12), 24 885–24 892.
- Nützel, B., Herwig, H., Koenigs, P. D., and Monti, J. M. (1994). "Acoustic backscatter measurements in the North Sea: 3–18 kHz," J. Acoust. Soc. Am. **95**, 2488–2494.
- Plant, W. J. (1986). "A two-scale model of short wind-generated waves and scatterometry," J. Geophys. Res. **91**(C9), 10 735–10 749.
- Plant, W. J. (1987). "Correction to 'A two-scale model of short wind-generated waves and scatterometry' by William J. Plant," J. Geophys. Res. **93**(C2), 1347.
- Plant, W. J., and Alpers, W. (1994). "An introduction to SAXON-FPN," J. Geophys. Res. **99**(C5), 9699–9703.
- Plant, W. J., Keller, W. C., Pettitt, R. A., and Terray, E. A. (1994). "The dependence of microwave backscatter from the sea on illuminated area: Correlation times and lengths," J. Geophys. Res. **99**(C5), 9705–9723.
- Reeves, J., Igarashi, Y., Beck, L., and Stern, R. (1969). "Azimuthal dependence of sound backscattered from the sea surface," J. Acoust. Soc. Am. **46**, 1284–1288.
- Sarkar, K., and Prosperetti, A. (1994). "Coherent and incoherent scattering by oceanic bubbles," J. Acoust. Soc. Am. **96**, 332–341.
- Schroeder, L. C., Boggs, D. H., Dome, G., Halberstam, I. M., Jones, W. L., Pierson, W. J., and Wentz, F. J. (1982). "The relationship between wind vector and normalized radar cross section used to derive SEASAT-A satellite scatterometer winds," J. Geophys. Res. **87**(C5), 3318–3336.
- Skolnik, M. I. (1970). "An introduction to radar," in *Radar Handbook* (McGraw-Hill, New York), pp. 1-1 to 1-16.
- Stoffelen, A., and Anderson, D. (1993). "ERS-1 scatterometer data characteristics and wind retrieval skill," Proc. First ERS-1 Symposium, Cannes, France, 4–6 November 1992, ESA SP-359, 41–47.
- Stolte, S. (1994). "Short-wave measurement by a fixed tower-based and drifting buoy system," IEEE J. Ocean. Eng. **19**(1), 10–22.
- Thorpe, S. A. (1982). "On the clouds of bubbles formed by breaking wind-waves in deep water, and their role in air-sea gas transfer," Philos. Trans. R. Soc. London, Ser. A **304**, 155–210.
- Urick, R. J. (1983). *Principles of Underwater Sound* (McGraw-Hill, New York), 3rd ed.
- Wentz, F. J., Peteherych, S., and Thomas, L. A. (1984). "A model function for ocean radar cross sections at 14.6 GHz," J. Geophys. Res. **89**(3), 3689–3704.
- Wille, P. C., and Geyer, D. (1984). "Measurements on the origin of the wind-dependent ambient noise variability in shallow water," J. Acoust. Soc. Am. **75**, 173–185.
- Wright, J. W. (1966). "Backscattering from capillary waves with application to sea clutter," IEEE Trans. Antennas Propag. **AP-14**, 749–754.
- Wright, J. W. (1968). "A new model for sea clutter," IEEE Trans. Antennas Propag. **AP-16**, 217–223.
- Wu, J. (1988). "Bubbles in the near-surface ocean: A general description," J. Geophys. Res. **93**(C1), 587–590.

The variability of high-frequency acoustic backscatter from the region near the sea surface

Peter H. Dahl and William J. Plant

Applied Physics Laboratory, College of Ocean and Fishery Sciences, University of Washington, Seattle, Washington 98105

(Received 30 May 1996; accepted for publication 17 November 1996)

The temporal variability of acoustic backscattering from the region near the sea surface is examined for frequencies in the 30–70 kHz range. A variance spectrum of the scattering strength exhibits effects associated with three different processes. Below about 0.1 Hz, the spectrum contains a large contribution associated with temporal variations in the advection of bubble clouds through the measurement volume by large-scale processes. At high frequencies, the spectrum asymptotes to a level characteristic of a Gaussian backscattered pressure field from randomly moving bubbles within the scattering volume. The overall variability is treated as a slow modulation of this Gaussian process by larger-scale processes and a probability density function is derived for the scattering strength using Bayes' theorem. Finally, in some cases, the spectrum exhibits a peak at the frequency of the dominant surface waves. Attempts to compute coherence functions between the backscattered acoustic power and surface wave orbital velocities, measured by a microwave system observing the same spot as the acoustic system, resulted in very low values. This leads to the belief that the wave-induced peak in the acoustic backscatter variance spectrum is caused by highly nonlinear processes. A time series of acoustic backscatter from a vertically pointing system confirms the existence of this modulation at the dominant wave frequency and also suggests its nonlinear character. © 1997 Acoustical Society of America. [S0001-4966(97)06105-5]

PACS numbers: 43.30.Gv, 43.30.Ft, 43.60.Cg [JHM]

INTRODUCTION

In a related paper (Dahl *et al.*, 1997), we demonstrated that high-frequency acoustic backscattering from the region near the sea surface is dominated by scattering from bubbles at all but very low wind speeds, for low to moderate grazing angles. Mean values of the backscattering cross section normalized by surface area were used to fit an interpretive model based on scattering from subsurface bubbles. The model adequately described the functional form of scattering versus wind speed and yielded reasonable estimates for the parameter β_I , equal to the depth-integrated extinction cross section per unit volume associated with bubbles. However, a sizable variance about the mean for a given wind speed was also observed, and here we examine this variability. In doing so we discuss measurements of both surface backscattering and volumetric backscattering from subsurface bubbles taken from two different experiments.

In the following section the two experiments from which the surface and volume scattering data originate are briefly described. In Sec. II, variance spectra computed from time series of surface backscattering data (expressed in decibels) are discussed. The variance spectra illustrate how processes of differing time scales contribute to the total variance of surface backscattering strength. Looking ahead, we establish that the variance spectrum can be roughly partitioned into a low-frequency region (less than about 0.1 Hz) and a high-frequency region. A third region, associated with the frequencies of the dominant waves, sometimes exists. Then in Sec. III, a probability density function (PDF) for surface backscattering strength is developed based on partitioning the variance between low- and high-frequency fluctuations.

Statistical goodness-of-fit tests confirm that the model PDF gives a good representation of the data. In Sec. IV the time scale associated with the dominant surface waves is examined in depth using the measurements of volume scattering from subsurface bubbles. It is also shown that the volume scattering data fit the same model PDF as developed in Sec. IV.

I. EXPERIMENTAL DESCRIPTION

The data used in this paper come from two separate experiments. The first data set originates from Phase II of the Synthetic Aperture Radar and X Band Ocean Nonlinearities (SAXON)-Forschungsplattform *Nordsee* (FPN) experiment conducted in December 1991. Measurements relevant to this paper are described by Dahl *et al.* (1997). More information regarding the SAXON-FPN experimental program is given by Plant and Alpers (1994). The acoustic backscattering measurements were made by the Forschungsanstalt der Bundeswehr für Wasserschall-und Geophysik of Kiel, Germany. Here we discuss those made at a grazing angle of 17° and frequencies of 50 and 70 kHz, using two-way beamwidths of 2.9° and 2°, respectively, and a pulse width of 4 ms. The transducer depth was 24 m, and pulses were transmitted every 0.4 s. Backscatter data from a coherent microwave system were obtained from the same area of the sea surface simultaneously with the acoustic data, and we have utilized those from a cw K_u band (14-GHz) system for some of the results described in this paper. The half-power, two-way beamwidth of this system was 6.6° in the E plane and 5.0° in the H plane; it was mounted 26 m above mean sea level and operated at a grazing angle of 17°.

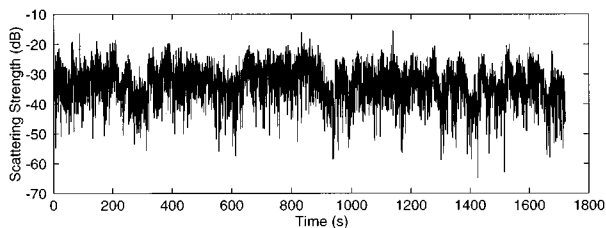


FIG. 1. Time series of 50-kHz acoustic scattering strength at a 17° grazing angle. The wind speed was 5.8 m/s at the time of data collection, 1616 UTC on 3 December 1991.

The second data set originates from an experiment described by Dahl and Jessup (1995) that was conducted from the floating instrument platform FLIP in January 1992. In this case acoustic measurements of subsurface bubbles were made at vertical incidence from a 13-m-long subsurface boom attached to FLIP's hull 28.5 m below the water line. The FLIP measurements discussed here were made at 30 kHz, using a system with a nominal two-way beamwidth of 16° and a pulse width of 1 ms, with pulses being transmitted every 0.25 s.

II. PROCESSES CONTRIBUTING TO THE VARIANCE OF SURFACE BACKSCATTERING STRENGTH

Figure 1 shows a time series of acoustic backscatter from the sea surface that represents a continuous half-hour set of measurements made at 50 kHz. Each value is the decibel equivalent of σ_0^a , the backscattering cross section normalized by surface area; this decibel quantity is referred to as scattering strength. The data were taken during steady wind conditions, with a mean wind speed of 5.8 m/s and a standard deviation of 0.41 m/s, giving a coefficient of variation for wind speed of around 10%. We make the assumption that this represents a single environment vis-à-vis scattering conditions. When we use the logarithmic transformation to scattering strength, $s = 10 \log_{10} \sigma_0^a$, the time series appears to have a slowly varying mean with a roughly constant additive noise component. With the variance now rather more stabilized, the time series data are more appropriate for spectral analysis than had we used linear units. Makris (1995) establishes a more complete mathematical justification for the use of logarithmic units in a variety of engineering applications, and in particular for variables that undergo fluctuations similar to σ_0^a .

Figure 2(a) shows estimates of the variance spectrum for the time series in Fig. 1 (thick, solid curve) and for two other scattering-strength time series made at 70 kHz during which the mean wind speed was 9.7 m/s (dashed curve) and 11.1 m/s (thin, solid curve). Figure 2(b) shows wave-height variance spectra from the FM part of the microwave return; we will discuss these further below. The relative wind-speed variation associated with these data was approximately the same ($\sim 10\%$) as for the data shown in Fig. 1. Note that the spectra in Fig. 2(a) are plotted such that the average value of the spectral function over the Nyquist bandwidth equals the total variance, and thus the ordinate is not in dB^2/Hz but rather in dB^2 . We postulate that the spectral peak at ~ 0.14

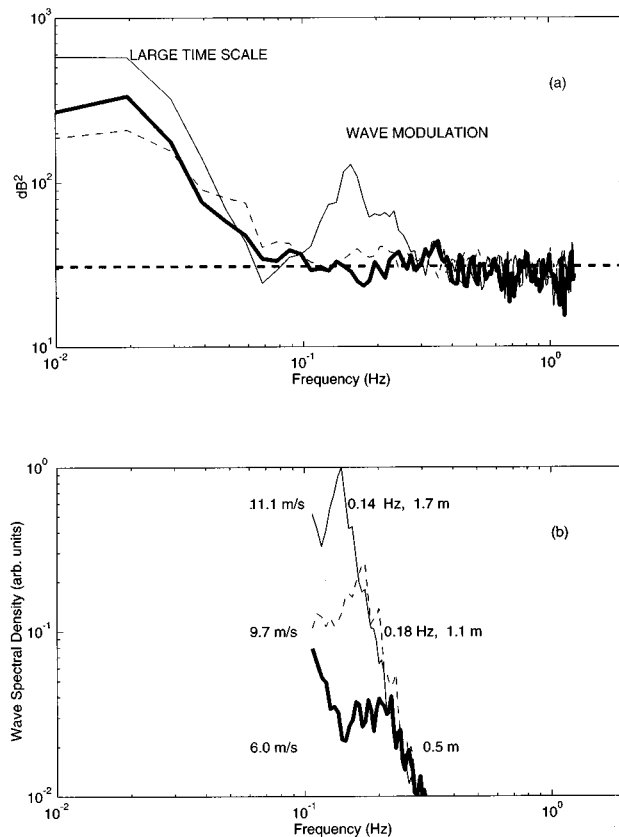


FIG. 2. (a) Variance spectra for the data shown in Fig. 1 (thick bold curve) and for other time series of acoustic scattering strength made at 70 kHz during which the mean wind speed was 9.7 m/s (dashed curve) and 11.1 m/s (thin, solid curve). The dashed horizontal line equals the theoretical spectrum baseline value of 5.57^2 dB^2 . (b) Surface wave-height spectra derived from K_u -band Doppler offset. All spectra are normalized by the peak value of the spectrum taken at a wind speed of 11.1 m/s. The line types correspond to those in (a), to the right of each spectrum is an estimate of the peak wave frequency and significant wave height (with the exception of the bold curve, where the peak in the spectrum is very broad).

Hz for the higher wind-speed data is a result of a modulation caused by surface gravity waves and postpone its discussion for now.

Two features are common to all three spectra in Fig. 2(a). The first is a low-frequency region corresponding to the slow variations shown in Fig. 1. This region is very similar in spectral slope (the spectra fall between 1 and 2 orders of magnitude in the decade 0.01–0.10 Hz) and bandwidth ($\sim 0.1 \text{ Hz}$) to the spectra for low-frequency variation in bubble-mediated sound speed reported by Lamarre and Melville (1994), who attribute this variation to the passage of bubble clouds over their instrument. There is recent speculation (Thorpe, 1992; Melville *et al.*, 1995) that such low-frequency variation is associated with the group frequencies of surface waves. In any case, we would expect a similar variation caused by the passage of bubble clouds under the surface scattering patch and through the effective scattering volume for the acoustic measurements discussed here. Note that the spectral averaging in Fig. 2(a) was carried out using 256-point FFTs with a Hanning window and 50% overlap; analysis using longer processing windows shows that the spectral density continues to increase with decreasing frequency for frequencies $< 0.01 \text{ Hz}$ (though at diminished

rate), reaching a broad peak at approximately 0.005 Hz.

The second feature is the spectrum baseline level of approximately 5.57² (dashed horizontal line), a value equal to the variance of a random variable expressed in decibels, whose probability density function (PDF) for the linear form of the variable is exponential (Dyer, 1970; Dahl and Mathisen, 1983). The implication in our case is that the exponential PDF is a good model for fluctuations in σ_0^a , provided the time scale over which these fluctuations are observed is kept short enough. It is reasonable to assume that each observation of σ_0^a separated by 0.4 s is an independent random variable. Thus, away from the low-frequency variation, the time series data represent a white-noise, or *purely random*, process (e.g., see Percival and Walden, 1993). Note, however, that if the observations are spaced too closely in time they will eventually become correlated. Here the relevant correlation time scale, τ , is the period over which the phase of the complex scattered amplitude has changed by $\pi/2$ owing to random motion within the assembly of bubbles. For example, at 50 kHz the wavelength is 3 cm, and we would expect decorrelation in time when $|v|\tau \approx 0.75$ cm, where $|v|$ is the characteristic speed of bubble motion, which is largely determined by the orbital motion due to gravity waves. We estimate τ is $O(10)$ ms. The exponential PDF arises naturally in many high-frequency acoustic applications where the observed scatter is the vector sum of several randomly phased scattering sources, as would occur in scattering from a random assemblage of bubbles. Were nothing to change other than the random movement of bubbles, the result would be a Gaussian backscattered pressure field with an amplitude following a Rayleigh PDF and its square, or intensity, following an exponential PDF.

III. THE PROBABILITY DENSITY FUNCTION AND VARIANCE OF SURFACE BACKSCATTERING STRENGTH

The bandwidth of approximately 0.1 Hz roughly bounding the low-frequency, high-energy region of the spectra in Fig. 2(a) points to an integral time scale, T , of $O(10)$ s with which to average out the higher-frequency fluctuations. We thus compute a linear average of σ_0^a defined as $\lambda = \langle \sigma_0^a \rangle_T$, where T equals 10 s, or 25 realizations of σ_0^a based on the 0.4-s ping cycle. A histogram (Fig. 3) of $10 \log_{10} \lambda$ for the data in Fig. 1 is reasonably well approximated by a Gaussian curve with a mean, μ_T , equal to -30.63 dB and a standard deviation, σ_T , equal to 3.07 dB. This suggests that λ , the slowly varying mean value of σ_0^a , is distributed lognormally, with a PDF equal to

$$p_\lambda(\lambda) = \frac{e^{-(\log \lambda - \mu_1)^2 / (2\sigma_1^2)}}{\lambda \sigma_1 \sqrt{2\pi}}, \quad (1)$$

where $\sigma_1 = \sigma_T/k$, $\mu_1 = \mu_T/k$, and $k = 10 \log_{10} e$. For reference, we note that the expected value of λ is given by

$$E(\lambda) = e^{\mu_1 + \sigma_1^2/2}. \quad (2)$$

One would use Eq. (2) to relate estimates of μ_T and σ_T back to $\langle \sigma_0^a \rangle_{T^*}$, where T^* is the time over which measurements of σ_0^a were made and during which the environment is

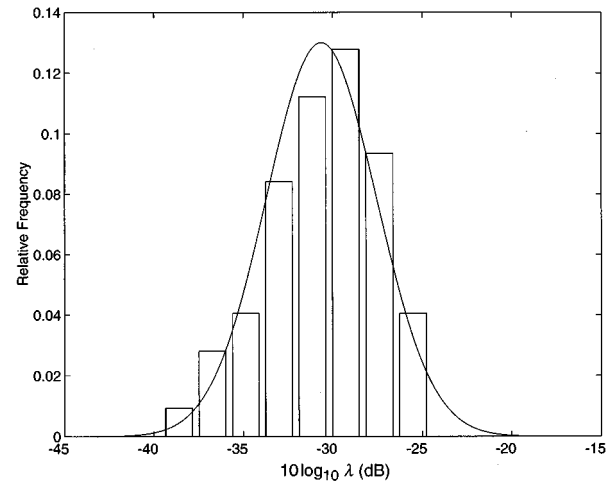


FIG. 3. Histogram of $10 \log_{10} \lambda$ for the data of Fig. 1. Here, λ is the linear mean of σ_0^a over 10 s. The curve represents a Gaussian probability density function with a mean equal to -30.63 dB and a standard deviation equal to 3.07 dB.

viewed as essentially unchanged. Thus, for example, theoretical estimates based on linear ensemble averages would relate to $E(\lambda)$.

Next, we take an approach analogous to that taken by Gotwols and Thompson (1994) in their statistical modeling of microwave backscattering data. Our approach differs from theirs in that we choose to work in decibels or log-transformed variables (as in the spectral estimates). Specifically, we model the PDF $p_s(s)$ for the data in Fig. 1 as an instantaneous PDF $p_{s|\lambda}(s|\lambda)$ governed by a slowly varying mean whose PDF is $p_\lambda(\lambda)$. To remove the condition on λ we multiply by the density for λ and integrate (Bayes' theorem), giving

$$p_s(s) = \int_0^\infty p_{s|\lambda}(s|\lambda) p_\lambda(\lambda) d\lambda, \quad (3)$$

where $p_\lambda(\lambda)$ is Eq. (1) and the conditional PDF is

$$p_{s|\lambda}(s|\lambda) = \frac{e^{-e^{s/k}/\lambda + s/k}}{k\lambda}. \quad (4)$$

Equation (4) represents a PDF that has been used to study decibel quantities such as acoustic transmission loss (Dyer, 1970) or acoustic scattering from fish (Dahl and Mathisen, 1983), where λ (now used as a parameter) is either equal, or proportional, to mean intensity, and where variation in the linear intensity variable is described by an exponential PDF.

Equation (3) is evaluated numerically, and the results for the PDF $p(s)$ [Fig. 4(a)] and its cumulative distribution function (CDF) counterpart [Fig. 4(b)] are shown compared with the data from Fig. 1. Both forms give a good representation of the data, with, as expected, the integrated form looking better. A chi-square (χ^2) goodness-of-fit test (Sachs, 1984) comparing observed with expected relative frequencies was carried out to assess the hypothesis that the data in Fig. 1 are distributed according to (3) when using sample parameters μ_1 and σ_1 . The probability of the χ^2 statistic was 0.20, indicating that (3) is indeed an adequate statistical model for the data.

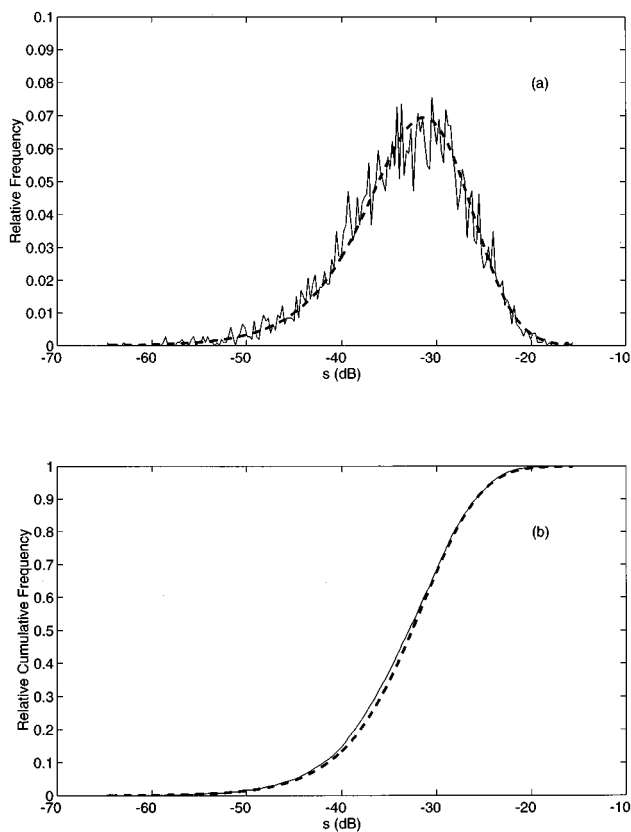


FIG. 4. Probability density function (a) and cumulative distribution function (b) for the SAXON-FPN data of Fig. 1 (thin, solid curves) taken at a wind speed of 5.8 m/s compared with model predictions (dashed curves). The model is computed from Eqs. (1), (3), and (4), using $\mu_T = -30.63$ dB and $\sigma_T = 3.07$ dB. The probability of the χ^2 statistic is 0.20.

The first two moments of $p(s)$ are evaluated via

$$E(s^n) = \int_0^\infty \int_{-\infty}^\infty s^n p_{s|\lambda}(s|\lambda) ds p_\lambda(\lambda) d\lambda, \quad (5)$$

giving

$$E(s) = \mu_T - C_E k, \quad (6)$$

where C_E is Euler's constant (0.577215...), with the variance given by

$$\text{Var}(s) = \frac{k^2 \pi^2}{6} + \sigma_T^2. \quad (7)$$

Thus the effect of the slow variation is simply to add σ_T^2 to an underlying variance of 5.57^2 . The standard deviation of the time series in Fig. 1, for example, is 6.42 dB, a value well approximated by $\sqrt{\sigma_T^2 + 5.57^2}$ when using the sample standard deviation for the slowly varying mean of 3.07 dB.

IV. LONG WAVE MODULATION

We now return to the peak at ~ 0.14 Hz seen in the variance spectrum of the acoustic data taken at the higher wind speed [thin, solid curve in Fig. 2(a)]. Figure 2(b) shows three surface wave-height spectral densities obtained from the microwave Doppler shifts. This is a standard technique for measuring wave spectra (Plant *et al.*, 1983), though typically applied for grazing angles near 45° ; at 17° grazing

angle, the total wave-height variance is not completely recovered for reasons discussed elsewhere (Plant, 1997). In any case, our main purpose for presenting these spectra is to show information on the dominant wave frequency, so we have normalized the spectral densities to the peak value of the highest density. The peak frequency (when resolved) and significant wave heights are noted for each case; the significant wave-height estimates correspond in time but were obtained from a nearby wave buoy. Except for the lower-wind-speed case, these wave-height spectra correspond precisely in time, and space, to the variance spectra shown in Fig. 2(a). For the lower wind speed, the wave-height spectrum was obtained at the same location but during the hour immediately following the acoustic measurements (hence the small difference in wind speed); the significant wave height measured by the wave buoy did not change over the 2-h period.

Referring to the wave-height spectra, the peak frequency in the data taken at a wind speed of 11.1 m/s matches the peak in the acoustic variance spectrum in Fig. 2(a), whereas the peak frequency in the data taken at 9.7 m/s has no acoustic counterpart. Likewise, only a small, broad peak is seen in the wave spectrum obtained at 6.0 m/s, and none is seen in the acoustic data. The frequency coherence between the time series of acoustic backscattering strengths and the microwave Doppler offset (proportional to wave orbital velocity) was < 0.2 for frequencies in the range of the dominant surface waves. This low coherence leads us to postulate that the peak in the acoustic variance spectrum is caused by a process that is related to the dominant surface waves but that is highly nonlinear. The nonlinearity would also explain why clearly resolvable peaks in the wave spectra as measured by the radar show an acoustic counterpart only for large-amplitude waves. At this stage we can only speculate as to the nature of the source of acoustic modulation near the dominant wave frequency. One possibility might be micro-scale breaking occurring preferentially on wave crests; another is discussed below in the context of FLIP data. Beam-pattern effects may also be in part responsible for the observed modulation.

In contrast to the SAXON-FPN data, where measurements of surface scattering are interpreted in terms of volume scattering from near-surface bubbles, the FLIP data represent unambiguous volume scattering from near-surface bubbles. We therefore studied the variability seen in this data set to obtain further insight into the variability observed in the surface-scattering data from SAXON-FPN. Figure 5(a) shows the variance spectrum for a 20-min time series of acoustic volume backscattering made at vertical incidence from FLIP during which the mean wind speed was 9.8 m/s. In this case, each value of the original time series is $s = 10 \log_{10} s_V$, where s_V is the backscattering cross section per unit volume for a given ping cycle. (The decibel quantity here will be referred to as volume backscattering strength.) The scattering center for s_V is 1 m below the time-varying surface elevation determined by the two-way acoustic travel time for the surface return.

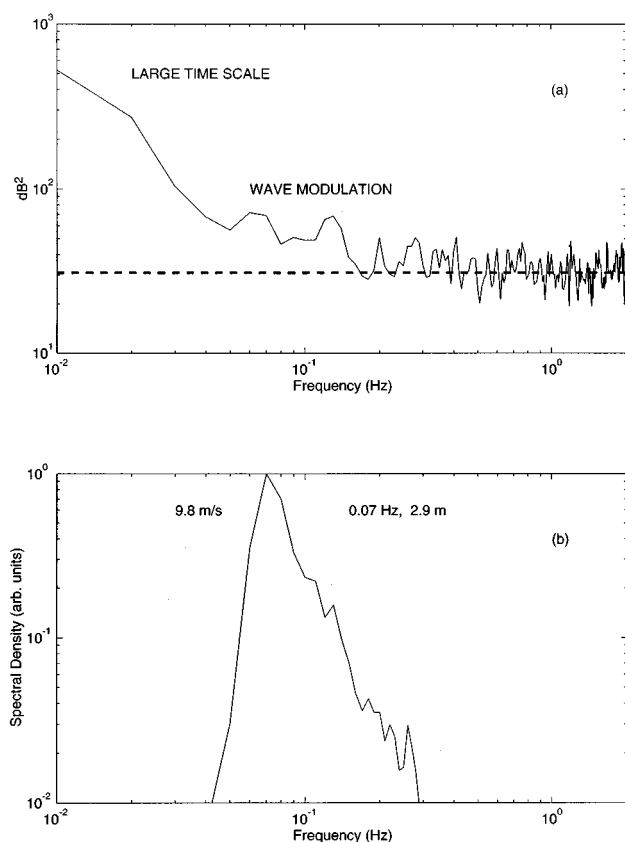


FIG. 5. (a) Variance spectrum of volume backscattering strength measured at vertical incidence from the research platform FLIP. The mean wind speed was 9.8 m/s. The dashed horizontal line equals the theoretical spectrum baseline value of 5.57^2 dB². (b) Surface wave-height spectrum derived from acoustic travel time, for times corresponding to the data shown in (a). The spectrum is normalized to its peak value with peak wave frequency and significant wave height noted on the right.

The variance spectrum for the volume-backscattering-strength measurements from FLIP is remarkably similar to the spectra for the surface-backscattering-strength measurements from SAXON-FPN shown in Fig. 2(a), insofar as having a high-energy region for low frequencies ($f \leq 0.1$ Hz) and coming close to the theoretical background limit of 5.57^2 for higher frequencies out to the Nyquist limit, which in this case extends to 2 Hz. [Figure 5(a) is plotted in the manner of Fig. 2(a) to account for the difference in Nyquist interval.] There is a peak within the low-frequency region near 0.07 Hz, and we postulate that this, too, is associated with the dominant surface waves.

Lamarre and Melville (1994) also discuss a peak in their sound speed anomaly spectra associated with the spectral peak of the dominant waves, in the context of orbital motion of these dominant waves. An alternative (though closely related) explanation is as follows: First, the inhomogeneous, or patchy, bubble field lateral structure is periodically displaced to and fro a distance of order $2A$, where A is amplitude of the dominant surface waves. We would expect to see a signal in the variance spectrum of acoustic measurements at dominant wave frequencies, once $2A$ reaches or exceeds some integral horizontal scale of these patches. To be sure, however, much remains to be learned about the wave-number spectrum of the bubble field near the ocean surface. A start-

ing point should include Gilbert (1993), who discusses this wave-number spectrum in the context of low-frequency acoustic backscatter, and Thorpe (1992). Second, the bubble field is periodically *distorted* while undergoing displacement by the surface wave orbital velocities, which may also add to the acoustic variation observed.

Figure 5(b) shows the wave-height spectrum as determined by the acoustic-travel-time data and normalized by its peak value. As in the spectra derived from the radar measurements, those derived from the acoustic travel time are simultaneous, colocated, and noninvasive with respect to the scattering measurements. The 3-dB surface footprint of the sonar (diameter ~ 8 m) sets the high-frequency spectral cut-off for this method at approximately 0.4 Hz. However, this has little effect on the total wave-height variance, which was dominated by swell, and wave-height spectra derived in this manner agree well with those derived from measurements by a wire wave gauge and a radar altimeter which operated on the opposite side of FLIP. The dominant peak in the wave-height spectrum at 0.07 Hz [Fig. 5(b)] is roughly aligned with a broad peak in the variance spectrum [Fig. 5(a)]. Again, however, the modulation is a nonlinear one, as the frequency coherence between the time series of wave height and $10 \log_{10} s_V$ is < 0.2 for all frequencies.

This kind of modulation, linked to both wave orbital velocity and bubble field patchiness, would also be in effect for the SAXON-FPN data, for which we have argued that bubbles are the primary source of backscatter. It is tempting to associate this modulation with the large peak at 0.14 Hz in Fig. 2(a). The geometry of the SAXON-FPN experiment, however, also requires that we allow for the possibility of effects due simply to the variable location of the air/sea interface in the ensonified volume. The modulation observed at these low grazing angles, therefore, may well result from a combination of bubble-field modulation and purely geometric effects.

Before leaving this topic, we briefly mention two other cases in which acoustic measurements of bubbles show fluctuations related to the surface wave spectrum. Medwin *et al.* (1975) measured sound phase fluctuations near the sea surface and showed that small changes in water pressure (due to surface wave action) at the depth of the measurement probe can modulate the resonance frequency of bubbles and therefore their phase measurements. This change in resonance frequency as a function of pressure can be linearized, and a high coherence was observed between sound phase fluctuations and surface wave height at the dominant wave frequencies. This mechanism is less applicable to the FLIP data because these measurements were effectively wave following, and so crest-to-trough pressure difference is no longer a first-order quantity. Zedel and Farmer (1994) observed a modulation in ambient sound level that was in phase with surface displacement directly above their measurement probe. A number of potential source mechanisms were identified, including small-scale wave-breaking events triggered by long-wave/short-wave interaction. Again, however, as in the observations by Medwin *et al.*, there was high coherence between fluctuations in noise level and surface wave height, which we do not see in our data. Thus the picture remains unclear, and

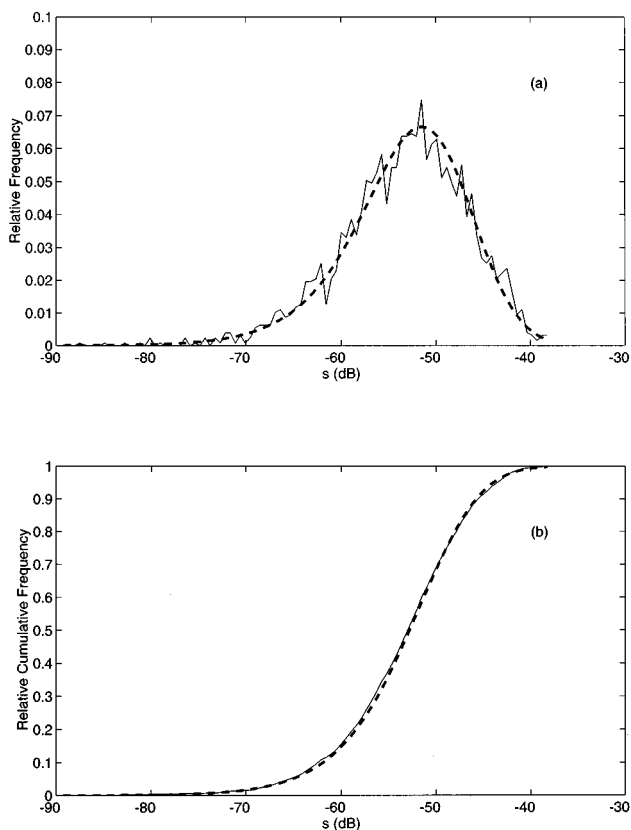


FIG. 6. Probability density function (a) and cumulative distribution function (b) for the FLIP data in Fig. 5 (thin, solid curves) taken at a wind speed of 9.8 m/s compared with model predictions (dashed curves). The model is computed from Eqs. (1), (3), and (4), using $\mu_T = -50.67$ dB and $\sigma_T = 3.45$ dB. The probability of the χ^2 statistic is 0.25.

we will seek to clarify the role of wave modulation in future studies.

Finally, it is useful to view the portion of the variance spectra above the dotted line in Figs. 2(a) and 5(a) as variability in excess of the theoretical 5.57^2 . Let us define the low-frequency excess as the portion above the dashed horizontal line for $f < 0.1$ Hz. This component makes up roughly 20–30% of the total variance for the SAXON-FPN data [Fig. 2(a)] and about 40% of the total variance for the FLIP data [Fig. 5(a)]. The component due to wave modulation in Fig. 2(a), between approximately 0.1 and 0.3 Hz, comprises about 10% of the total variance. For the FLIP data, wave modulation is, in part, added to the low-frequency region as we have defined it, with the region between 0.1 and 0.2 Hz contributing less than 10% to the total variance.

Following the procedure used earlier with the backscattering data, we compute a linear average of the FLIP data, defined in this case as $\lambda = \langle s_V \rangle_T$, where we again use an averaging time scale, T , equal to 10 s, giving $\mu_T = -50.7$ dB and $\sigma_T = 3.4$ dB to use in Eqs. (2) and (3). The results are shown in Fig. 6. The probability of the χ^2 statistic is 0.25, which suggests that our PDF model applies equally well to the s_V measure.

In general, the variance spectrum provides guidance as to the choice of T . In the FLIP data [Fig. 5(a)], for example, 0.1 Hz roughly defines the point at which the excess variability is confined to lower frequencies, and thus $T = 10$ s

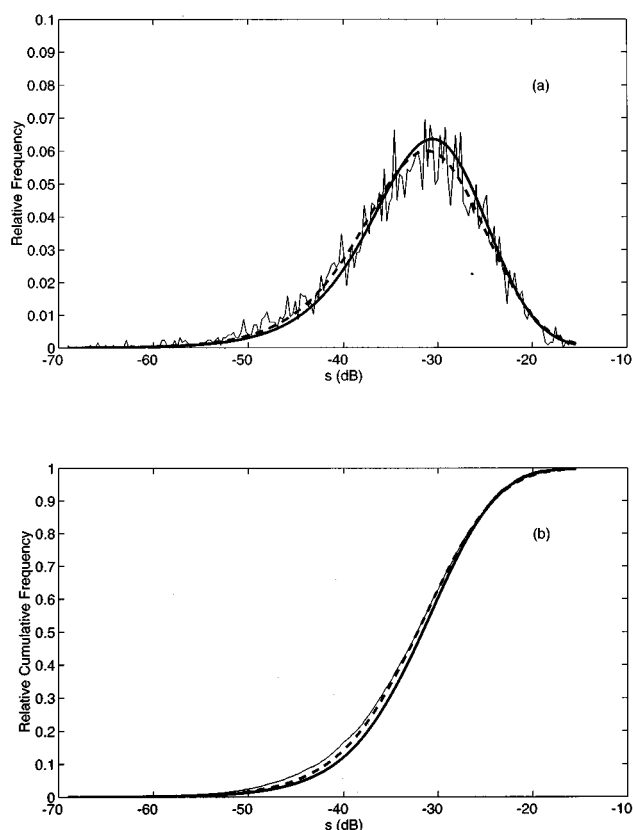


FIG. 7. Probability density function (a) and cumulative distribution function (b) for the SAXON-FPN data (thin, solid curves) taken at a wind speed of 11.1 m/s compared with model predictions. The model is computed from Eqs. (1), (3), and (4). The probability of the χ^2 statistic is < 0.01 when using $\mu_T = -29.59$ dB and $\sigma_T = 3.88$ dB, based on $T = 10$ s (thick, solid curve). The probability of the χ^2 statistic is 0.05 when using $\mu_T = -30.06$ dB and $\sigma_T = 4.41$ dB, based on $T = 3$ s (dashed curve).

should be appropriate, as confirmed in Fig. 6 and the reasonably probable value for the χ^2 statistic. On the other hand, the variance spectrum for the SAXON-FPN data taken at 11.1 m/s [thin, solid curve in Fig. 2(a)], with its large and isolated peak, points to an averaging time scale of 3 s, because ~ 0.3 Hz now roughly defines the point at which the excess variability is confined to lower frequencies. If we, in fact, use $T = 10$ s for these data, we find $\mu_T = -29.59$ dB and $\sigma_T = 3.88$ dB, giving a rather a low probability (< 0.01) for the χ^2 goodness-of-fit parameter. Using these values for μ_T and σ_T , we would reject the hypothesis (at the 5% level of significance) that the data are distributed according to our PDF model. However, if we use $T = 3$ s then μ_T lowers somewhat to -30.06 dB and σ_T increases to 4.41 dB, with the probability of the χ^2 statistic improving to 0.05. This probability is less than in the two earlier trials but sufficiently high that the hypothesis is not rejected.

The results for both T values are shown in Fig. 7. The model/data agreement is better with $T = 3$ s (dashed curve), although using $T = 10$ s (thick, solid curve) still captures the essential features of the empirical PDF [Fig. 7(a)] and CDF [Fig. 7(b)]. Thus, notwithstanding the existence of peaks in the variance spectra associated with the dominant surface waves, an averaging time scale of 10 s seems to be remarkably robust. Accordingly, we examined σ_T , for $T = 10$ s, for

the entire SAXON-FPN Phase II data set spanning 4 days (and including data taken at the third frequency of 26.5 kHz). Contiguous segments of data were defined by having the coefficient of variation for wind speed remain $<10\%$, following our definition of an unchanging environment, with typical segment lengths (such as the one shown in Fig. 1) equaling 30 min. The result, for wind speeds between 3.8 and 11.6 m/s, can be summarized by $\sigma_T \approx -0.18U_{10} + 4.85$, in decibels, where U_{10} is wind speed corrected to 10-m height. The reduction in variance as wind speed increases is consistent with a less patchy bubble field at higher wind speeds and with saturation (e.g., see Dahl *et al.*, 1997), which places an upper bound on the scattering and thereby reduces the variance.

V. SUMMARY AND CONCLUSION

We have studied fluctuations in measurements of acoustic backscattering from the sea surface made during the SAXON-FPN experiment, for which the mean values are shown in another paper to be linked to the concentration of near-surface bubbles. For comparison, our study also included the fluctuations in volumetric backscattering from near-surface bubbles, using data gathered from FLIP. Spectral analysis was carried out on the equivalent decibel form for the data from each experiment, and three features in the variance spectrum were identified. At low frequencies, $f \lesssim 0.1$ Hz, we suggest that the variability is caused by bubble clouds being advected through the ensonified region; the spectral roll-off in this region was found to be between f^{-1} and f^{-2} . At the frequencies of the dominant surface waves, a modulation is induced if the wave height is sufficiently large, and we suggest, based on the low coherence of this modulation with the dominant waves, that it may be highly nonlinear. The FLIP data showed consistency with this modulation being linked to the action of the dominant surface waves. Finally, at high frequencies, the backscattered pressure field is Gaussian, as expected for scattering from a large number of independent scatterers. The standard deviation of this part of the scatter, on a logarithmic scale, is 5.57 dB.

Decibel quantities were used in our analysis to stabilize the variance and to allow comparison of the fluctuations in acoustic measures, such as σ_0^a and s_V , that differ in both dimension and nominal value but otherwise are proportional to backscattered intensity. We modeled the PDF for a decibel quantity, e.g., scattering strength, as the integral over the product of an instantaneous PDF, which corresponds to an exponential PDF for the quantity in linear space conditioned on a value λ , the normalized cross section averaged over a time T and the PDF for λ . The λ value undergoes low-

frequency modulation, and we model its PDF as lognormal. This model reproduces the data well and indicates that the total variance of the scattering strength is simply 5.57^2 plus the variance of the low-frequency modulation, σ_T^2 . A representative estimate of σ_T (for $T = 10$ s) for the entire SAXON-FPN Phase II data set taken at a grazing angle of 17° is approximately 3.7 dB.

ACKNOWLEDGMENTS

This study was supported by the Office of Naval Research through Grant No. N00014-95-1-0048 from the Ocean Acoustics Program. Constructive comments by an anonymous reviewer are much appreciated.

- Dahl, P. H., and Jessup, A. T. (1995). "On bubble clouds produced by breaking waves: An event analysis of ocean acoustic measurements," *J. Geophys. Res.* **100**, 5007–5020.
- Dahl, P. H., and Mathisen, O. A. (1983). "Measurement of fish target strength and associated directivity at high frequencies," *J. Acoust. Soc. Am.* **73**, 1205–1211.
- Dahl, P. H., Plant, W. J., Nützel, B., Schmidt, A., Herwig, H., and Terray, E. A. (1997). "Simultaneous acoustic and microwave backscattering from the sea surface," *J. Acoust. Soc. Am.* **101**, 2583–2595.
- Dyer, I. (1970). "Statistics of sound propagation in the ocean," *J. Acoust. Soc. Am.* **48**, 337–345.
- Gilbert, K. E. (1993). "A stochastic model for scattering from the near-surface oceanic bubble layer," *J. Acoust. Soc. Am.* **94**, 3325–3344.
- Gotwols, B. L., and Thompson, D. R. (1994). "Ocean microwave backscatter distributions," *Geophys. Res.* **99**, 9741–9750.
- Lamarre, E., and Melville, W. K. (1994). "Sound-speed measurements near the ocean," *J. Acoust. Soc. Am.* **96**, 3605–3616.
- Makris, N. C. (1995). "A foundation for logarithmic measures of fluctuating intensity in pattern recognition," *Opt. Lett.* **20**, 2012–2014.
- Medwin, H., Fitzgerald, J., and Rautmann, G. (1975). "Acoustic miniprobing for ocean microstructure and bubbles," *J. Geophys. Res.* **80**, 405–413.
- Melville, W. K., Terrell, E., and Ding, L. (1995). "Field measurements of air entrainment by breaking waves," in *Air-Water Gas Transfer: Selected Papers from the Third International Symposium on Air-Water Gas Transfer*, edited by B. Jähne and E. C. Monahan (AEON Verlag & Studio, Hanau), pp. 286–295.
- Percival, D. B., and Walden, A. T. (1993). *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques* (Cambridge U.P., Cambridge, U.K.).
- Plant, W. J. (1997). "A model for microwave Doppler sea return at high incidence angles: Bragg scattering from bound, tilted waves," *J. Geophys. Res.* (in press).
- Plant, W. J., and Alpers, W. (1994). "An introduction to SAXON-FPN," *J. Geophys. Res.* **99**, 9699–9703.
- Plant, W. J., Keller, W. C., and Cross, A. (1983). "Parametric dependence of ocean wave-radar modulation transfer functions," *J. Geophys. Res.* **88**, 9747–9756.
- Sachs, L. (1984). *Applied Statistics: A Handbook of Techniques* (Springer-Verlag, New York).
- Thorpe, S. A. (1992). "Bubble clouds and the dynamics of the upper ocean," *Q. J. Meteorol. Soc.* **118**, 1–22.
- Zedel, L., and Farmer, D. (1994). "Surface wave period modulations in near surface ambient sound," *J. Geophys. Res.* **99**, 8041–8052.

Acoustic measurements of suspended sediments in turbulent currents and comparison with *in-situ* samples

Peter D. Thorne and Peter J. Hardcastle

Proudman Oceanographic Laboratory, Bidston Observatory, Merseyside L43 7RA, United Kingdom

(Received 4 March 1996; accepted for publication 11 November 1996)

Increasingly in recent years the application of acoustic backscattering to the quantitative measurement of suspended sediment particle size and concentration at sea has gained acceptance. A number of works describing the interaction of sound with suspensions have been published, and the scattering properties of suspended sediments formulated. However, there have been relatively few experiments conducted in the marine environment, which have attempted to assess the accuracy of the acoustic measurements by direct comparison with *in-situ* samples, taken simultaneously with the acoustic observations. The purpose of the present work is to report on such an experiment, and to evaluate the accuracy of the acoustic technique. To this end multifrequency acoustic measurements of suspended sediment profiles were collected in an estuarine environment, subject to strong turbulent tidal currents, which generated high concentrations of suspended sediments. To obtain the sediment parameters from the acoustic data an inversion needs to be applied, and this inversion is examined here in some detail, particularly for the case when sediment attenuation is substantial. To assess the sediment parameters derived from the acoustic inversion, mean acoustic estimates of particle radius and concentration are compared with the benchmark of *in-situ* pumped sampling. In addition to the analysis of the mean data, high-resolution images of the suspension dynamics have been generated, and the validity of these observations appraised by evaluating the internal consistency of the multifrequency results. © 1997 Acoustical Society of America. [S0001-4966(97)06705-2]

PACS numbers: 43.30.Pc, 43.30.Gv, 43.30.Ma [JHM]

INTRODUCTION

The application of sound scattering, coupled with inversion methodologies, for remotely investigating forms and structures, is a well-established field of study in acoustics. Applications range from medical ultrasonics¹ typically operating in the region of 10 MHz and with submillimeter resolution, to ocean acoustic tomography² which uses sources at a few hundred Hertz and probes the marine environment over megameters. Invariably the common theme running through these works is the formulation of an understanding of the interaction of sound with the phenomenon under investigation, and the development of approaches to extract information on the phenomenon from this interaction. The former is a forward problem; knowing the system predicts the outcome of sound interactions. The latter and somewhat more elusive goal is to use inversion techniques on the measured acoustic field after interaction, and to reconstruct the phenomenon. It is in this latter mode that sound is now being used at sea to quantitatively probe the dynamics of suspended sediment transport processes. As the sound propagates through the suspension the backscattered signal is recorded, and an inversion applied to extract profiles of suspended sediment parameters. The outcome from such observations is providing new and valuable insights³⁻⁸ into the fundamental mechanisms of sediment transport.

The typical mode of operation of acoustic backscatter systems used to measure suspended sediments is usually to mount a downward looking transceiver, at $\sim 1-2$ m above the bed, operating in the range $\sim 1-5$ MHz, and transmitting a sound pulse of the order of several microseconds duration.

As the sound propagates from the transceiver to the bed, the envelope of the backscattered signal is recorded. This record is used to obtain information on the suspension. The motivation for applying acoustics is that it provides data on suspended sediment profiles unobtrusively and remotely, and with a temporal and spatial resolution capable of examining within wave and turbulent sediment processes. Also since the method uniquely provides profiles, which can be rapidly updated, the suspension processes can be imaged, and the nearbed dynamics displayed. Further coupled with this resolving capability, records can be collected over long periods, thereby providing the ability to move seamlessly from individual suspension events to processes on the scale of tidal periods. The acoustic approach therefore provides the opportunity to examine processes at varying scales of spatial and temporal detail.

Although the acoustic technique has significant advantages over more conventional methods, the interpretation of the data is somewhat less straightforward if the attenuation due to the suspension is high,⁹ which is at times of high concentration and hence greatest interest. Therefore the purpose of the present work has been to investigate the use of acoustic backscattering in an environment where concentration levels are high and attenuation is significant. To examine the inversion, simulations have been conducted, and marine measurements taken in an estuary. In the present study a triple frequency system operating at 1.0, 2.5, and 5.0 MHz was employed, and the system fully calibrated using a polyvinylidene fluoride (PVDF) membrane hydrophone. This calibration has allowed inversions to be conducted without

reference to *in-situ* sampling, or laboratory calibrations using sediments from the marine site, to give both the suspended concentration and particle size solely from the acoustic backscatter data. To assess the accuracy of the inversion, the acoustic measurements are compared with *in-situ* samples taken with the acoustic data, and these are used as the benchmark test for the inversions. As far as the authors are aware this is the first experiment where the accuracy of the results from a multifrequency backscatter system has been compared with independent *in-situ* suspended sediment sampling.

I. ACOUSTIC FORMULATION AND INVERSION

The theoretical description of the interaction of sound with suspended sediment is now well-documented.¹⁰⁻¹⁴ If multiple scattering can be ignored, though attenuation by the suspension is included, and the backscattered signal is treated as incoherent, then the root-mean-square backscattered pressure, P_{rms} , at range, r , from a piston transceiver, can be written as,

$$P_{\text{rms}} = \frac{P_0 r_0 f_a}{\psi} \left\{ \frac{3M}{16\pi\rho_s \langle a_s \rangle} \right\}^{1/2} \times \left[\int_0^{\pi/2} \int_0^{2\pi} \int_{r-(\tau c/4)}^{r+(\tau c/4)} \left\{ \frac{2J_1(x)}{x} \right\}^4 \times \frac{e^{-4r\alpha}}{r^2} \sin \theta \, dr \, d\phi \, d\theta \right]^{1/2}, \quad (1)$$

$$\alpha = \alpha_w + \frac{1}{r} \int_0^r \zeta M \, dr, \quad (2)$$

where P_0 is the source level at $r_0 = 1$ m, ψ accounts for the departure of the backscattered signal from spherical spreading within the transceiver near-field,¹⁵ and $J_1(x)$ is the first-order Bessel function, where $x = ka_t \sin \theta$; k is the wave number of the sound, a_t is the transceiver radius, and θ is the angle subtended to the acoustic axis. Here, M is the suspended sediment concentration, ρ_s is the density of the sediment particles, $\langle a_s \rangle$ is the mean particle radius, α_w is the absorption due to the water, f_a is the mean backscatter form function given by $f_a = \{ \langle a_s \rangle \langle a_s^2 f^2 \rangle / \langle a_s^3 \rangle \}^{1/2}$, where $\langle \rangle$ denotes average over particle size, and f is the form function for a uniform particle size.¹⁶ $\zeta = 3\chi_a / 4\rho_s \langle a_s \rangle$ and χ_a is the mean normalized total scattering cross section, $\chi_a = \langle a_s \rangle \langle a_s^2 \chi \rangle / \langle a_s^3 \rangle$, and χ is the normalized total scattering cross section for a uniform particle size.¹⁰

Although Eq. (1) has been previously evaluated to varying degrees of approximation the problem is reexamined here, and the solution further developed. The integral over ϕ is simply 2π . To evaluate the beam pattern integral let

$$I(\theta) = \int_0^{\pi/2} \left\{ \frac{2J_1(x)}{x} \right\}^4 \sin \theta \, d\theta. \quad (3)$$

The only significant contribution to the integral is from the main lobe, this allows the bracketed term in Eq. (3) to be written as¹⁷

$$2x^{-1} J_1(x) \approx 1 + \epsilon_1 x^2 + \epsilon_2 x^4 + \epsilon_3 x^6. \quad (4)$$

This order of expansion is adequate for the present work. $\epsilon_1 = -1.125/9$, $\epsilon_2 = 0.4219/81$, and $\epsilon_3 = -0.0791/729$. If the restriction is made that $ka_t \geq 10$, then $\sin \theta \approx \theta$ is a reasonable approximation within the main lobe, therefore

$$I(\theta) \approx \int_0^{\theta_0} \{ 1 + \epsilon_1 (ka_t \theta)^2 + \epsilon_2 (ka_t \theta)^4 + \epsilon_3 (ka_t \theta)^6 \}^4 \theta \, d\theta, \quad (5)$$

where $J_1(ka_t \sin \theta_0) = 0$ and $\theta_0 = \sin^{-1}(3.832/ka_t)$. Owing to the choice of the upper limit of the integral the evaluation of Eq. (5) simply becomes

$$I(\theta) \approx \left\{ \frac{3.832}{ka_t} \right\}^2 \left[\frac{1}{2} - (3.832)^2 \epsilon_1 + \frac{(3.832)^4}{6} (4\epsilon_2 + 6\epsilon_1^2) + \dots \right]. \quad (6)$$

The expression within the brackets is lengthy, and is not shown for brevity, however, it has no functional dependence and is simply a constant. Evaluation of terms inside the bracket gives 0.0629 and therefore

$$\sqrt{I(\theta)} = 0.96/ka_t. \quad (7)$$

This solution is accurate to 0.5% for $ka_t \geq 10$. This condition is invariably fulfilled for most backscatter systems. In the present study the minimum value for ka_t was 41.

The range-dependent integral can be readily evaluated if either the exponential^{10,11} or the spherical spreading¹⁴ components are removed outside the integral. An exact solution can be obtained by expanding the exponential term, letting

$$I(r) = \int_{r-(\tau c/4)}^{r+(\tau c/4)} \frac{e^{-4r\alpha}}{r^2} \, dr, \quad (8)$$

expanding gives

$$I(r) = \int_{r-(\tau c/4)}^{r+(\tau c/4)} \left(\frac{1}{r^2} - \frac{4\alpha}{r} + \frac{(4\alpha)^2}{2!} - \frac{(4\alpha)^3}{3!} r + \dots \right) dr, \quad (9)$$

integrating yields

$$I(r) = \frac{8\tau c}{16r^2 - (\tau c)^2} - 4\alpha \ln \left(\frac{r + \tau c/4}{r - \tau c/4} \right) - \sum_{n=1}^{\infty} \frac{(-1)^n (4\alpha)^{n+1} [(r + \tau c/4)^n - (r - \tau c/4)^n]}{n(n+1)!}. \quad (10)$$

Evaluation of Eq. (10) shows that for all cases a good approximation is given by

$$I(r) = \frac{8\tau c e^{-4\alpha r}}{16r^2 - (\tau c)^2} \cdot \frac{\sinh(\tau c \alpha_r)}{\tau c \alpha_r}, \quad (11)$$

where $\alpha_r = \alpha_w + \zeta M$ at range r . The $\sinh(\tau c \alpha_r)/\tau c \alpha_r$ term has been previously derived.¹⁴ If, as in the present study, the attenuation over the pulse width is not substantial and $r > \tau c$, then, respectively, $\sinh(\tau c \alpha_r)/\tau c \alpha_r \approx 1$, and $8\tau c/[16r^2 - (\tau c)^2] \approx \tau c/2r^2$, Eq. (1) can then be written as

$$P_{\text{rms}} = \frac{K_s K_t}{\psi r} M^{1/2} e^{-2\alpha r}, \quad (12)$$

where

$$K_s = \frac{f_a}{\sqrt{\rho \langle a_s \rangle}}, \quad K_t = \frac{0.96 P_0 r_0}{k a_t} \left\{ \frac{3 \pi c}{16} \right\}^{1/2}.$$

This formulation of P_{rms} effectively deals with the forward problem; if the suspension parameters are known the back-scattered pressure can be predicted, however, the objective of the present work was to obtain suspended sediment measurements from the backscattered pressure. Inverse problems are frequently somewhat more difficult because of the large number of variables which can affect the measured quantity or quantities, and this can lead to the requirement that some assumptions may be necessary to conduct the inversion. In the present study the objective was to obtain both suspended sediment concentrations and particle size. The particle size can be readily expressed¹² using Eq. (12) and gives

$$\frac{f_{ai}}{f_{aj}} = \frac{(\psi_i(r) P_{\text{rms}i} / K_t) e^{2r\alpha_i}}{(\psi_j(r) P_{\text{rms}j} / K_t) e^{2r\alpha_j}}, \quad (13)$$

where i and j refer to pairs of backscatter measurements taken at different frequencies. Similarly the concentration is given by

$$M_i = \left\{ \frac{P_{\text{rms}i}}{K_{si} K_{ti}} \right\}^2 \psi_i^2 r^2 e^{4r\alpha_i}, \quad (14)$$

where i refers to a single frequency.

One of the first problems to arise is that the relationship between the form function ratios and the particle size is not necessarily unique. In the present study the standard deviation of the sediments was expected to be approximately 25% of the mean,¹³ this leads to mean form function ratios for the frequencies used in the present work having the relationships shown in Fig. 1. The 5/1 MHz ratio is monotonic, however, for the 2.5/1 and 5/2.5 there are regions of ambiguity where multiple particle sizes are associated with a single form function value. Also if the inversion yields values near the local maxima or minima, then if the measured value is below the local minimum or above a local maximum there will be discontinuities in particle size, as indicated in Fig. 1. Internal checks within the inversion can be used to improve the estimation of particle size consistency. A simple one would be to check with the monotonic 5/1 result, however, at high concentration, the 5 MHz is highly attenuated, and the inversion of this signal can be problematic. Therefore there is the requirement with varying concentrations and particle sizes to use a mixture of methods to obtain consistent results. Also to evaluate Eqs. (13) and (14) the attenuation is required, which needs a knowledge of particle size and concentration, neither of which are known. This is a problem^{9,18,19} common to both acoustics and electromagnetics when the backscattered signal from the radiation propagating through scatterers is being used to measure parameters of the scattering medium. In the present work the problem is resolved by estimating initial values for M and a_s at the first range bin using $\alpha_s = 0$, and then using these values to estimate α_s . Using this attenua-

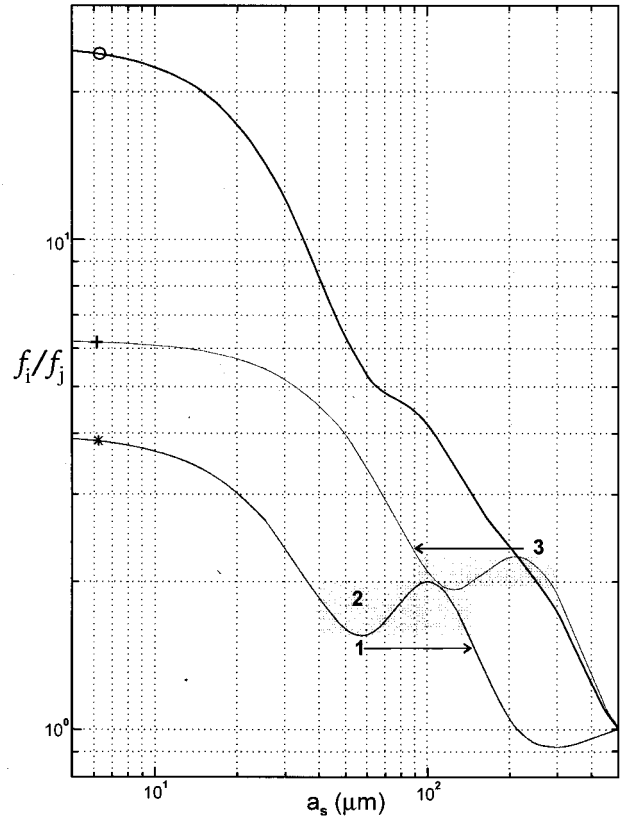


FIG. 1. Variation of the form function ratios with particle radius; \circ : 5/1, $+$: 2.5/1, and $*$: 5/2.5. (1) Particle size discontinuously jump if the measured value of the form function ratio drops below a local minimum. (2) Multiple size output regions. (3) Particle size discontinuity jump if the measured value of the form function ratio increases above a local maximum.

tion value, improved estimates of M and a_s are obtained. The iteration is then repeated until convergence occurs for M and a_s for the first range bin. This process is repeated down through the backscattered profile to obtain $M(r)$ and $a_s(r)$ for each of the frequencies. There is the problem with the inversion that errors can rapidly accumulate and erroneous profiles may be generated.⁹ Further, because the inversions are coupled through the attenuation and signal ratios are used to calculate particle size, erroneous results at one frequency affects the output at the other two frequencies.

To illustrate these inversion problems a synthetic suspension field was generated and sound propagated through the suspension, and the backscattered pressure computed. A number of inversions were conducted on the computed backscattered pressure. The suspension field used is shown in Fig. 2, and is reasonably representative of what was expected in the present marine study. It is possibly worth noting that if iterations are formulated around profiles rather than obtaining convergence at each range bin before calculating the next bin, then convergence to the original fields was generally not possible at high levels of sediment attenuation. As an example of the results obtained an inversion is shown which was conducted with a value for the system calibration which was 10% less than that used to compute the pressure field. Figure 3 shows the ratios of the output field divided by the original field, the inverse of this is used for visualization of

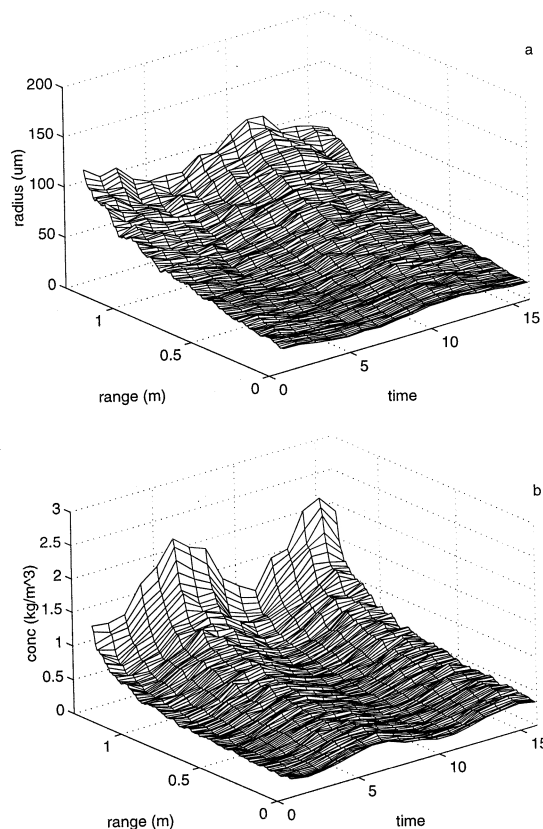


FIG. 2. Synthetic fields used to examine the inversion procedure. (a) Particle radius, (b) concentration.

particle size, Fig. 3(a)–(c). It can be readily seen that the 1.0- and 5.0-MHz inversions show particle sizes which reduce to half the original input size, while the concentration shows values that are five times greater. However, the 2.5-MHz results show significantly less error. This is because the 5.0-MHz signal was used to compute the 1.0- and 5.0-MHz particle sizes using the ratios 5/1 and 5/2.5, while the 2.5-MHz inversion used the 2.5/1 ratio. The 5.0-MHz signal is highly attenuated at the concentrations shown in Fig. 2, and small errors in the system calibration rapidly accumulate in the inversion generating erroneous particle sizes for the 1.0- and 5.0-MHz inversion which directly couple to the concentration values at these frequencies. If the inversion is recomputed, and the 2.5/1 particle size is used to calculate all the concentrations, the outcome shown in Fig. 4 is obtained. These results show a significant improvement, particularly for the 1.0-MHz results. This example indicates some of the problems which can occur at high concentration, in particular the possible problematic inversion at 5.0 MHz which can significantly affect the 1-MHz results. This coupling of inversion errors has implications in the analysis of the marine data.

II. EXPERIMENTAL SITE AND INSTRUMENTATION

The experiment^{20,21} was conducted during spring tides in the River Taw estuary in North Devon, United Kingdom. The estuary was subject to strong rectilinear turbulent tidal

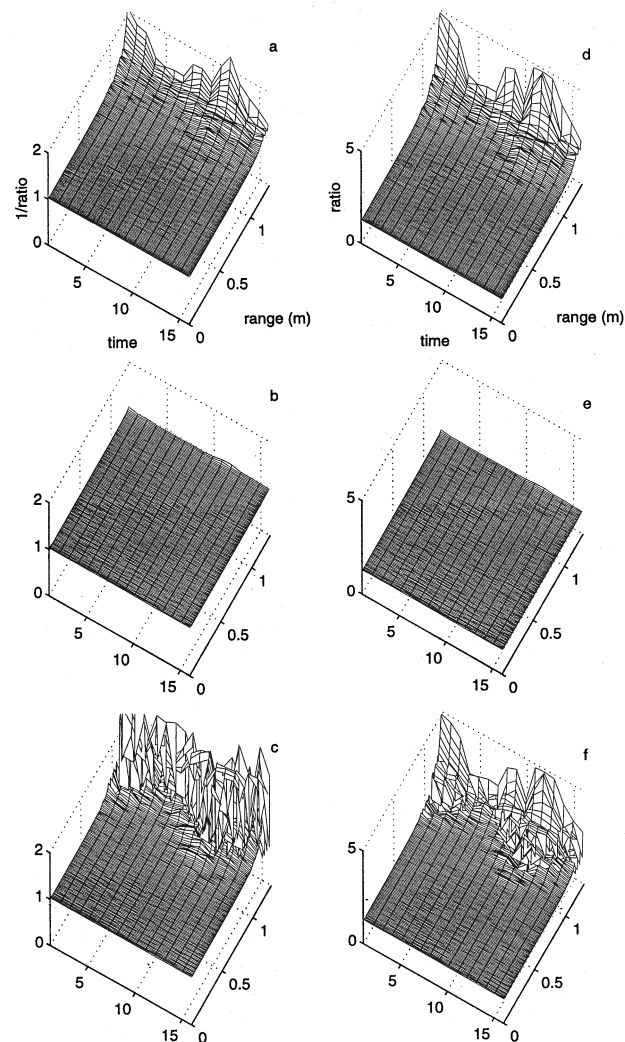


FIG. 3. Acoustic inversion using the backscattered pressure with a 10% underestimate for K_t . (a)–(c) are the calculated particle radius using ratio 5/1, 2.5/1, and 5/2.5, and (d)–(f) the mass at 1.0, 2.5, and 5.0 MHz. The ratio is defined as the output field divided by the input field given in Fig. 2. The time units are arbitrary.

currents which generated high concentrations of suspended sediments. At low water the site was dry and readily accessible for arranging the experiment, while at high water the depth was approximately 6 m. No surface waves were present and the suspension process was driven by turbulent currents. The bed consisted of well-sorted fine to medium sand, with the particle radius having a Gaussian distribution with a mean of $\langle a_s \rangle \approx 80 \mu\text{m}$, and standard deviation of $\sigma(a_s) \approx 20 \mu\text{m}$. Figure 5 shows the location of the estuary, and the size distribution of sediments collected from the bed at the site. The area was dominated by sandwaves which typically had wavelengths of around 20 m and amplitudes of the order of 0.8 m, although there was substantial variability. Measurements were recorded over five flood periods, at different locations along a sandwave. The data presented in this work comes from one of the locations. The position of the instrumentation, and the surveyed sandwave profile for the reported observations is shown in Fig. 6(a).

The instrumentation is shown in Fig. 6(b), and was com-

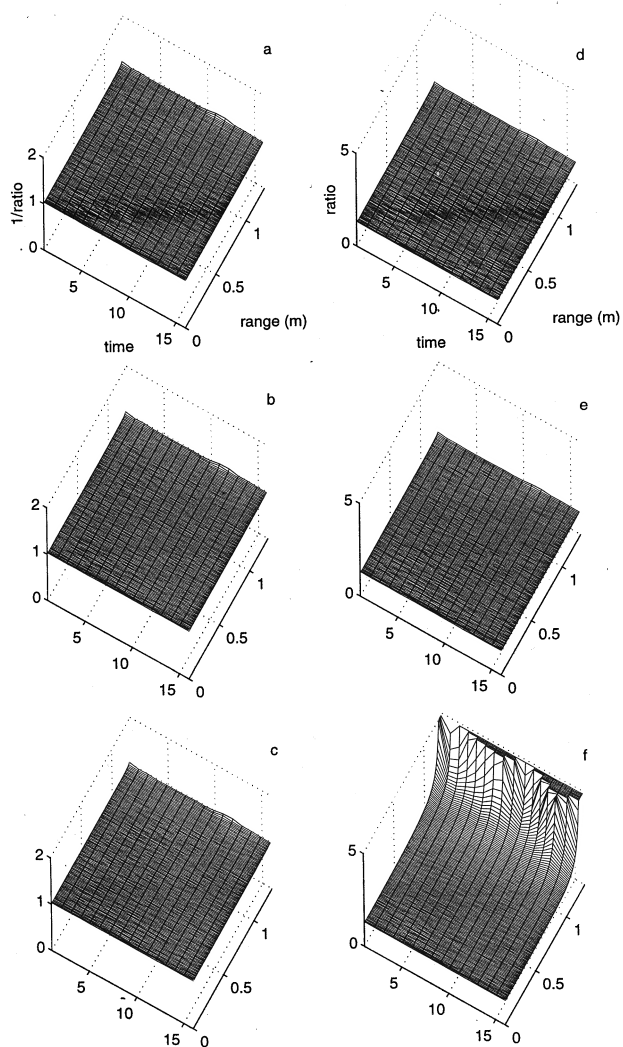


FIG. 4. Acoustic inversion using the backscattered pressure with a 10% underestimate for K_t . (a)–(c) are the calculated particle radius using the 2.5/1 ratio and (d)–(f) the mass at 1.0, 2.5, and 5.0 MHz. The ratio is defined as the output field divided by the input field given in Fig. 2. The time units are arbitrary. The particle radius from the 2.5/1 ratio was used in the calculation for the three concentrations.

posed of a triple frequency backscatter system, pumped sampling (PS) to assess the acoustic inversions, and electromagnetic current metres (EMCM) to monitor the turbulent flow (the flow data is not examined in the present work). The backscatter system used three transceivers operating at 1.0, 2.5, and 5.0 MHz. The backscattered signal was amplified by a logarithmic amplifier capable of operating over a dynamic range of 80 dB, envelope detected and the output data were digitized with 11-bit resolution. The spatial resolution of these devices was 0.01 m, and the operating range was 1.28 m. The setup provided 128 backscattered profiles, at each of the three frequencies each second. Data at each range bin and frequency were averaged using on line processing over 32 records to provide measurements at 0.25-s intervals, and these profiles were recorded. The temporal averaging is necessary to reduce configurational noise, arising from variations in the backscattered signal due to the random-phase distribution of the echoes originating from the varying loca-

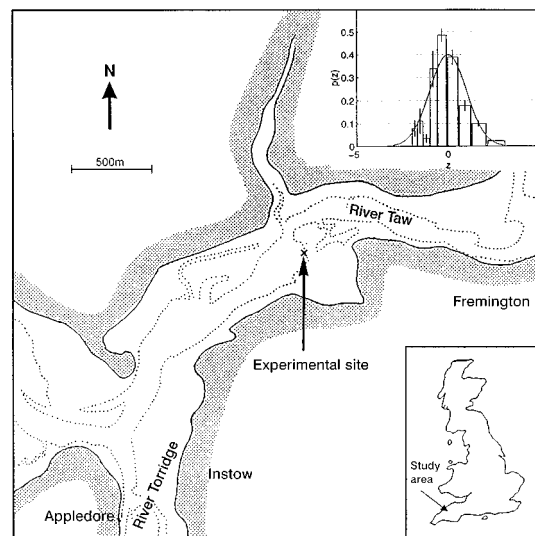


FIG. 5. Location of the experimental site. Insert shows the measured particle size distribution compared with a normalized Gaussian distribution.

tion of the particles within theinsonified volume. Data were collected over a 512-s block, stored, and data collection resumed. The triple frequency system was fully calibrated; the transfer function of the electronics, the transmit and receive sensitivities and beam patterns for the transceivers were all measured. Therefore it was possible to obtain acoustic estimates of particle size and concentration independent of *in-situ* sampling or laboratory calibrations taken using sediments from the site. To assess the accuracy of the acoustic measurements, *in-situ* measurements of the suspended sedi-

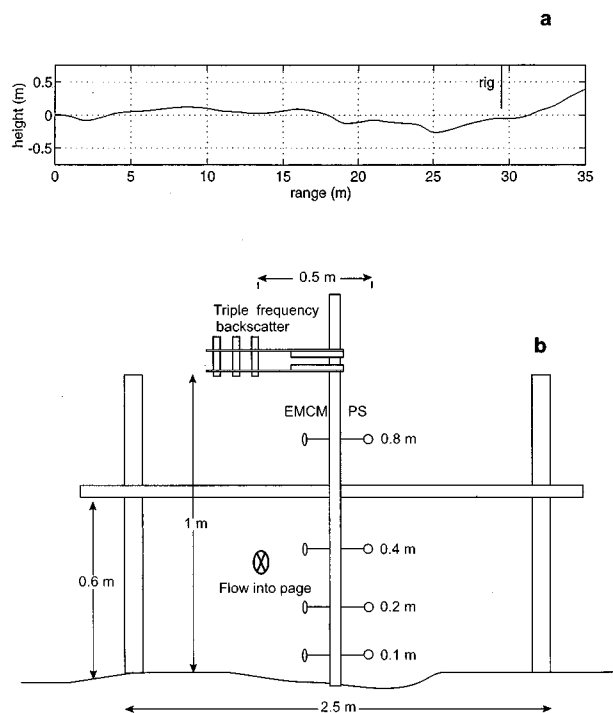


FIG. 6. (a) Position of the rig, and the sandwave profile, flow was from the left to the right. (b) Instrumentation used, PS: pumped sample, EMCM: Electromagnetic current meter, and the triple frequency, 1.0, 2.5, and 5.0 MHz transducers.

ments were taken at 0.1, 0.2, 0.4, and 0.8 m above the bed using pumped sampling. These were obtained at each height simultaneously every 10 min. All the observations were collected on a small Rotork barge anchored within approximately 10 m and to the side of the rig.

III. ANALYSIS OF THE MARINE OBSERVATIONS

A. Mean results

To examine the acoustic measurements of particle size and concentration, the 512-s data blocks were averaged to yield a single profile at each frequency. Using Eqs. (13) and (14) the variation in particle size and concentration with range and time has been computed from the acoustic data for the flood period under examination. The results are shown as 3-D plots in Fig. 7. The discontinuity observed at approximately 0.98 m from the transceiver is due to the bed echo. The particle size results are shown in Fig. 7(a)–(c) and the concentration values are given in Fig. 7(d)–(f). The measurements for the particle radii in Fig. 7(a)–(c) were respectively calculated from the backscattered root-mean-square pressures for frequency ratios of 5/1, 2.5/1, and 5/2.5 MHz. The particle radius figures show the 5/1 measurements decreasing with height above the bed, and the 2.5/1 and the 5/2.5 measurements being relatively uniform. The acoustic measurements of suspended particle size in Fig. 7(a), at around 50 min into the record, and within about 0.5 m of the bed are erroneous. The bed did not contain particles of this size, and the currents were not sufficient to suspend material with $\langle a_s \rangle$ in excess of 300 μm . Also although the 5/1-MHz results, Fig. 7(c), appear reasonable, there is a condition within the inversion which restricts the particle size from changing by more than 50% from one range bin to the next, the plateau regions and discontinuities are symptomatic that there were problems with the 5/1 particle size inversion. The reason for these problems is associated with the 5.0-MHz signal. Inspection of the 5-MHz concentrations in Fig. 7(f) show that at the highest concentrations, around 50 min into the record, the measurements at 5.0 MHz underestimate the concentration relative to the 2.5-MHz measurements, shown in Fig. 7(e). Also in Fig. 7(f) the concentration does not increase as the height above the bed is reduced, which would be the normal expectation. The inversion of the 5.0-MHz signal, at these periods of highest concentration, is somewhat problematic because of the high attenuation by the suspension around this time. The difference between including and excluding the sediment attenuation at the highest concentrations gave a difference in the nearbed concentration at 5.0 MHz of 3000%. As illustrated in the previous section if attenuation is high, relatively small calibration errors can accumulate in the inversion and generate substantial errors in the computed profiles. It appears that in the present case there may be some overestimate of the system response, or some inaccuracy in the description of the scattering properties of the sediment. Given that inversion errors arising from the 5.0-MHz signal couple through to the results at the other frequencies it was considered judicious, in the light of the output from the simulations, to use the 2.5/1 particle size ratio in the inversions.

Using the 2.5/1 calculated particle size, the profiles were recomputed. The results from this inversion are shown in Fig. 8. This shows the particle size, the mean concentration derived from the three frequencies, and the standard error in the mean concentration as a fraction of the mean concentration, i.e., the normalized standard error. The computations for the particle radius, presented in Fig. 8(a), provide a global view of the data set over the flood period. The figure shows a relatively uniform particle size, both with height above the bed, and over the duration of the experiment. The mean and standard deviation for the data shown in Fig. 8(a) are $\langle a_s \rangle = 74 \mu\text{m}$, and $\sigma(a_s) = 10 \mu\text{m}$, respectively. This low standard deviation reflects the temporal and spatial homogeneity of the suspended sediments over the observational period. The variability in particle size towards the end of the observational period is due to system noise, and 1.0-MHz echoes from the rig, both of which became noticeable at low concentrations. Figure 8(b) shows the mean concentration derived from the concentrations obtained at the three frequencies. This shows the variation in suspended sediment concentration over the flood period, extending to just over three orders of magnitude, $0.001\text{--}2 \text{ kg m}^{-3}$. The maximum concentration is seen at approximately 40 min into the record, this was at peak flow, reducing to low levels towards the end of the record at slack high water. The anomalously high values seen towards the end of the record at particular ranges are due to 1.0-MHz echoes from the arrangement of the rig. Since three estimates of concentration were used to compute the results in Fig. 8(b), an estimate for the internal consistency of the data was obtained. The result is shown in Fig. 8(c), which shows the normalized standard error. The values can generally be seen to be less than 0.1, although the first profile reaches 0.2 near the bed, and values are higher where the rig echoes contaminated the 1.0-MHz signal. However, the average of the whole data set is 0.07, and the data is therefore considered to be internally consistent. This level of consistency would not be achieved if the particle size estimates were significantly in error, owing to the different backscattering strengths with size and frequency. Therefore this level of consistency in the concentrations provides some confidence in the accuracy of the particle size measurements.

To assess the accuracy of the acoustic measurements of particle size and suspended sediment concentration, pumped sample data were collected simultaneously on known water volumes at 0.1, 0.2, 0.4, and 0.8 m above the bed, at 10-min intervals. The filtered samples were dry weighed to give suspended concentration, and on a selection of the samples a sieve analysis was conducted and the mean particle radius obtained. The acoustic measurements of particle size are compared with the pumped sample values in Fig. 9(a). This shows a time series of the ratio of the acoustic to the pumped sample measurements of the particle size, at the same heights above the bed as the pumped samples. Ideally this ratio should be unity and constant with time. The results at all four heights have values close to unity, apart from the first and last ratios which overestimate the particle size. The last result is probably associated with errors arising from forming ratios of signals at concentration levels near the minimum operat-

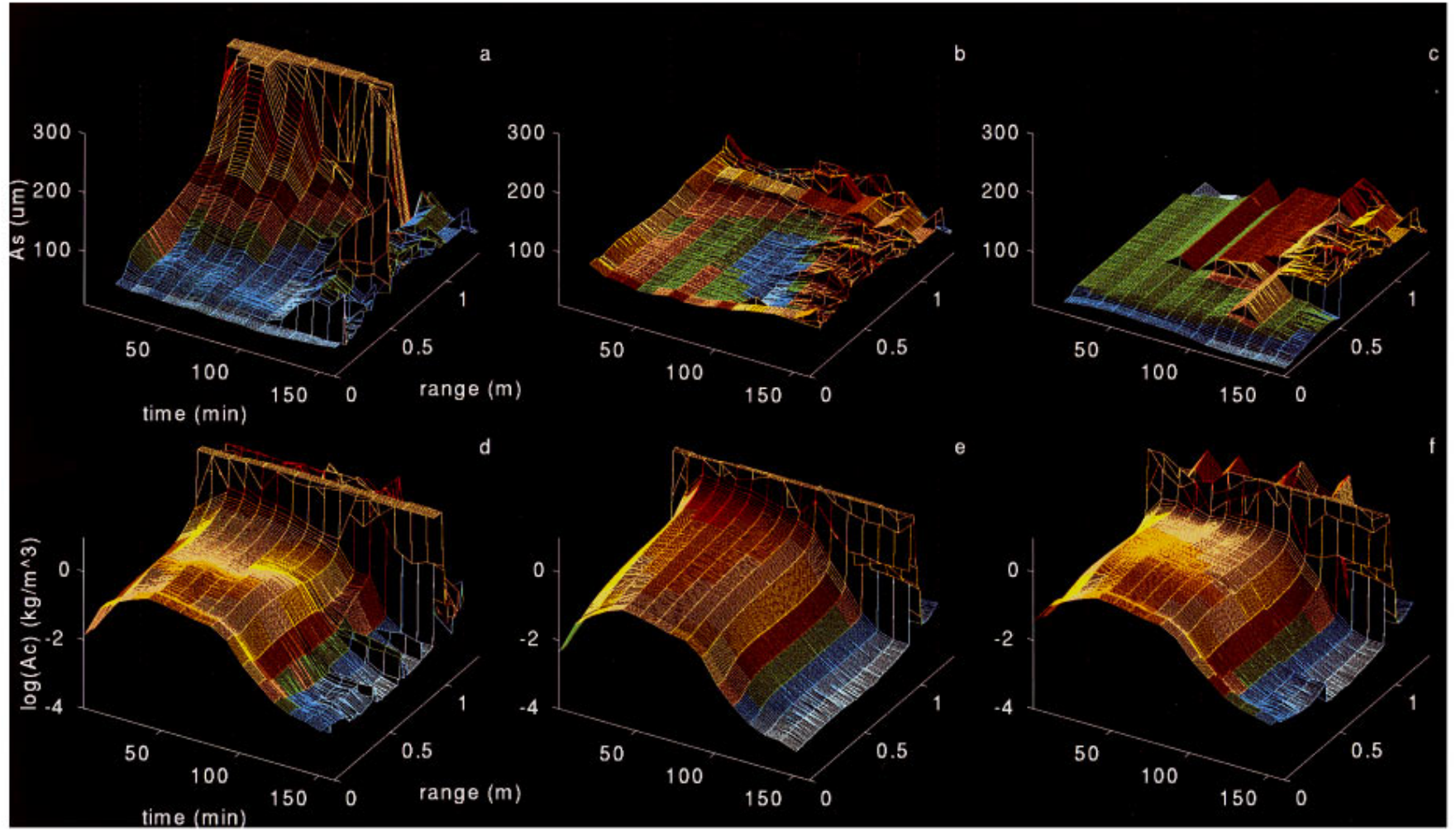


FIG. 7. Acoustic measurements of the suspended particle radius, (a)–(c) obtained using the 5/1, 2.5/1, and 5/2.5 backscattered pressure ratios, and the concentration, (d)–(f) obtained using the 1.0, 2.5, and 5.0-MHz backscattered pressures. The bed echo was at approximately 0.98 m from the transceivers.

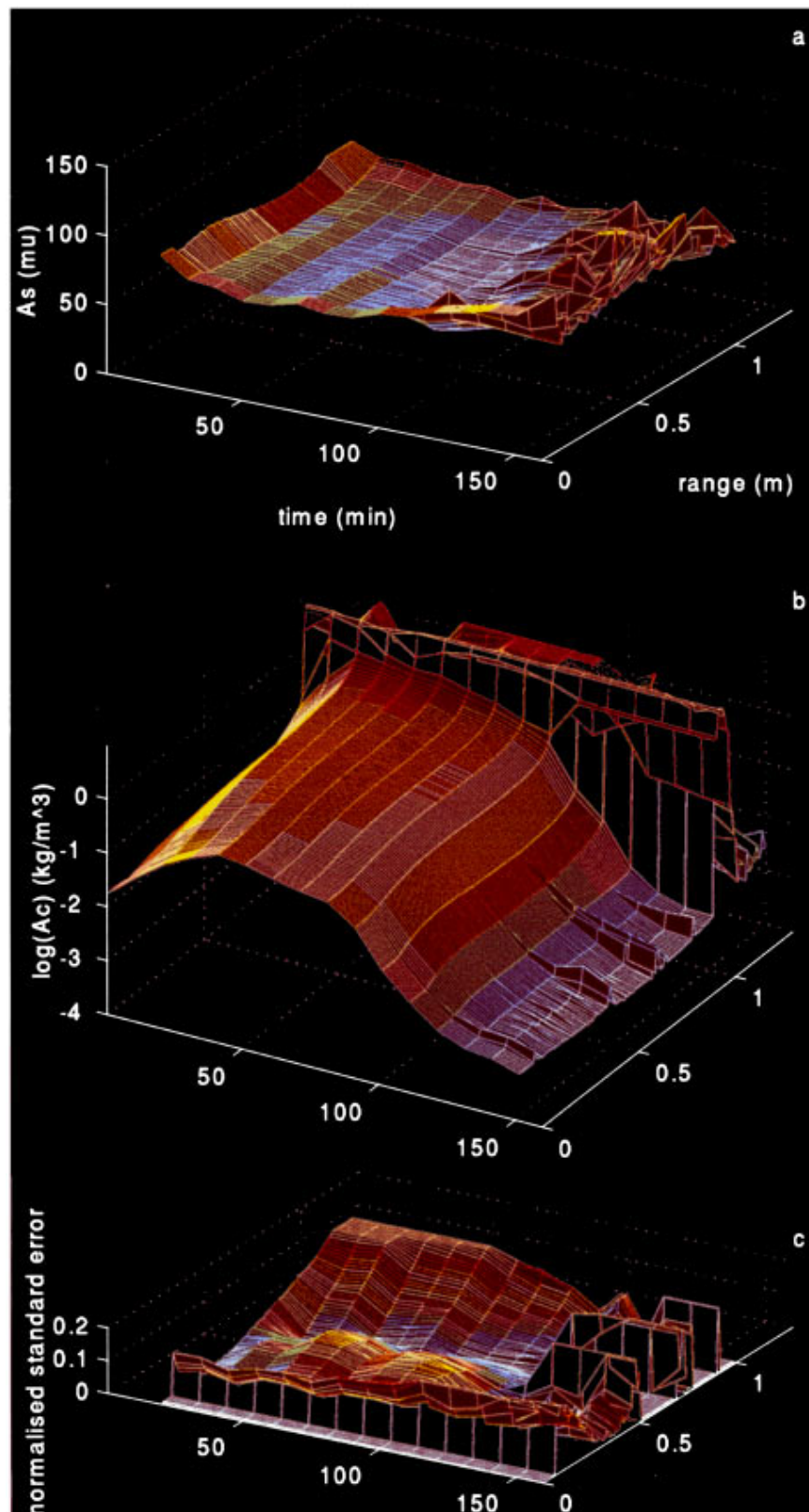


FIG. 8. (a) Acoustic measurements of particle radius from the 2.5/1 backscattered pressure ratio. (b) The mean concentration obtained from the three calculated concentrations. (c) The ratio of the standard error of the concentrations to the mean concentration.

ing range of the system, around $10^{-3} \text{ kg m}^{-3}$. The error in the first measurements is unclear, there is no signal to noise problem. There was the possibility that since the sediments were exposed at low water, there may have been some sedi-

ment degassing at the start of the record (N.B. this was not the beginning of the flood, the water had to be over 1-m deep before the transceivers were submerged) which is enhancing the 1.0-MHz signal, reducing the form function ratio, and

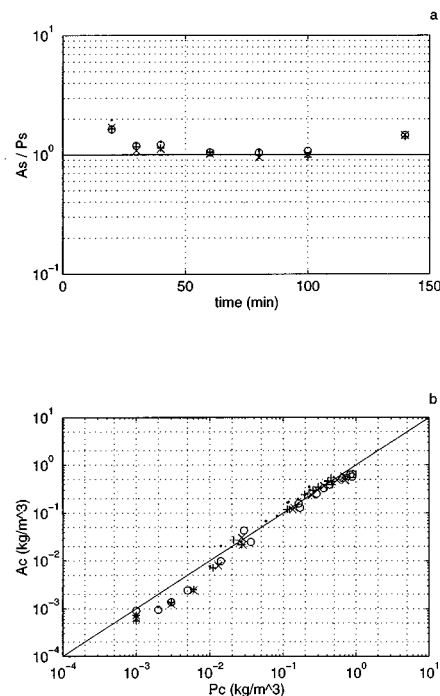


FIG. 9. (a) Time series of the ratio of the acoustic measurements of particle radius A_s to the pumped sample measurements P_s . (b) Comparison of the acoustic A_c and pumped sample P_c measurements of suspended concentration. Comparison conducted at \circ : 0.1, \times : 0.2, $+$: 0.4, and \bullet : 0.8 m above the bed.

leading to a particle size overestimate. However, acknowledging that there is some uncertainty at the beginning and end of the flood, the mean ratio of $A_s/P_s = 1.2 \pm 0.3$ for the whole data set, and this reduces to 1.1 ± 0.1 if the first and last values are excluded. Therefore the mean for over 70% of the acoustic estimates of particle size is within 10% of the pumped sample data, and since the standard deviation of the ratio is 10% there is no significant difference between the majority of the two data sets. The result for the comparison between the acoustic and pumped sample measurements of suspended sediment concentration is shown in Fig. 9(b). The acoustic values are the mean of the results for the concentrations obtained at each of the frequencies. The data cover just over three orders of magnitude, and as can be observed the measurements generally lie close to the line $A_c = P_s$, although there is some divergence at the lower concentrations. This divergence at low concentration is due in part to the overestimate of particle size at lower concentration levels. Linear regression on the data gives a correlation coefficient of 0.97 which is significant²² at the 99.95 level. The mean ratio of $A_s/P_s = 0.9 \pm 0.3$, therefore there is no significant difference from unity. If concentration levels below 0.01 kg m^{-3} are neglected, the result is $A_s/P_s = 1.0 \pm 0.2$, the ratio is unity, and the standard deviation reduces by 30%. Therefore, notwithstanding some divergence at lower concentrations, the acoustic estimates of suspended concentration are in good agreement with the pumped sample concentrations over the three orders of magnitude variation in concentration, and at the different heights above the bed. Taken together, both the acoustic estimates of particle size and suspended concentration levels show no significant dif-

ference from the pumped sample data, and at times of greatest interest, when significant quantities of suspended sediment are being transported, the correspondence is at its highest. This comparison of acoustic measurements of concentration and particle size with *in-situ* samples demonstrates definitively that accurate measurements can be obtained acoustically.

B. Detailed images

The above analysis provides an assessment of the acoustic technique. However, the motivation for developing the acoustic approach is to provide high temporal and spatial resolution measurements of suspended sediments to further our understanding of the fundamental mechanisms of sediment processes. To illustrate the results which can be obtained, a 3-D visualization of 50 s of data is shown in Fig. 10. Owing to the nominal uniformity of the particle size with range and height above the bed [cf. Fig. 8(a)], the high-resolution concentration images were calculated with a fixed particle radius of $74 \mu\text{m}$. Using an input size simplifies and substantially reduces the time to compute the inversion, and as long as the input particle size is accurate the iterative inversion can be obtained with higher precision. Using the concentration fields obtained to evaluate the sediment attenuation, the particle size was computed using the ratio of the 2.5/1 pressure profiles. This particle size field could have been used to recompute the concentration, however, the homogeneity of the calculated particle sizes indicated this was unnecessary. To form the acoustic images a 0.75-s temporal filter, and a 0.03-m spatial filter was applied to the pressure field. This was applied to reduce the ‘‘salt and pepper’’ type noise in the image associated with configurational noise, and real high-frequency fluctuations in the suspended concentration. Figure 10 shows 1.0-, 2.5-, and 5.0-MHz measurements of concentration and a single estimate of particle size using the 2.5/1 ratio. The three concentration images are provided to illustrate the internal consistency of the high-resolution inversions. As a measure of the internal consistency of the data the normalized standard error for the three concentration fields was calculated, this gave a mean value of 0.09. This value, and the coherence of the images provides confidence in the accuracy of the suspended sediment dynamics depicted, and indicates the structures observed are not associated with any acoustic artifacts. Such images are unique to the acoustic technique, and provide a significant advance over more conventional approaches, as they supply the facility to visualize and map out features as the suspension advects below the sensors. It can readily be seen in the concentration figures that the suspended sediment concentration is not homogeneous, but variable, with suspension events interspersed with relatively clearer water. Passing beneath the transceivers is an admixture of background suspended material carried through from suspension events occurring upstream, and local resuspension driven by turbulent events near the instrumentation. The suspension field can be seen to be composed of a number of relatively high-concentration occurrences, with vertical extents of the order of 0.6 m and variable periods of duration. These images show the stochastic character of the suspended sediment concentration in tur-

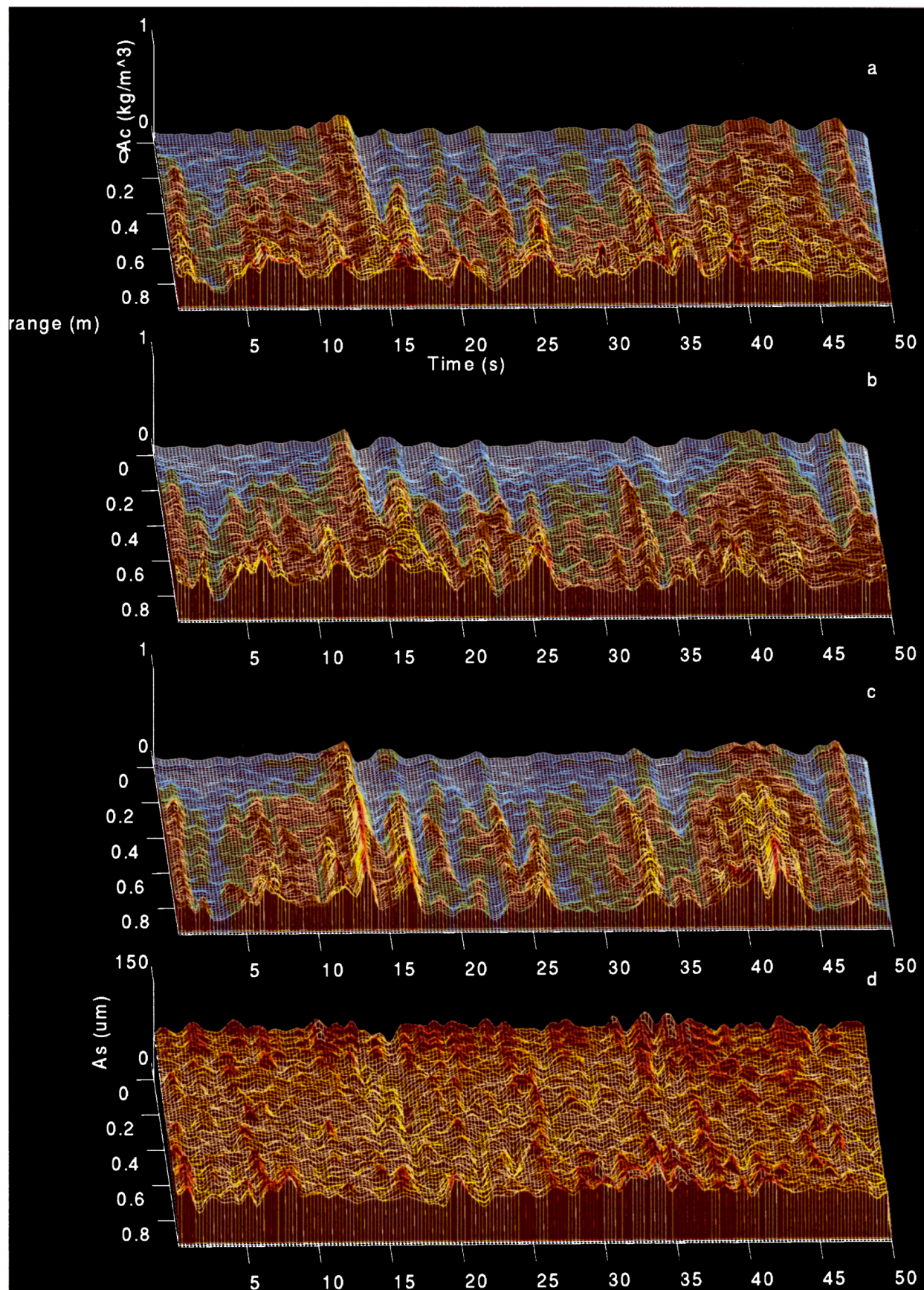


FIG. 10. High-resolution acoustic measurements of concentration at (a) 1.0 MHz, (b) 2.5 MHz, and (c) 5.0 MHz. (d) High-resolution acoustic measurements of the particle radius using the 2.5/1 ratio.

bulent tidal flow. The particle size on the other hand is relatively homogeneous. The mean and standard deviation of the data in Fig. 10(d) are $\langle a_s \rangle = 71 \mu\text{m}$ and $\sigma(a_s) = 14 \mu\text{m}$, this relatively small standard deviation is again indicative of the uniformity in the particle size. This limitation in size variation is both a function of the turbulent mixing process and the narrow size distribution of the source of supply, the bed.

The internal consistency of the detailed multifrequency images shown in Fig. 10 provides evidence for the high-resolution accuracy of the acoustic technique. This gives confidence that the accuracy of the fine resolution imagery of Fig. 10 can be considered a reliable description of the suspension field. It is this type of visualization of the suspended sediments which is enhancing our understanding of the basic mechanisms of sediment transport.

IV. CONCLUSION

The primary object of the present work was to collect multifrequency acoustic backscatter data on suspended sediments in turbulent tidal currents, calculate profiles of the concentration and particle size, and evaluate the accuracy of the results by comparison with *in-situ* samples. The estuarine environment in which the data were collected had relatively high concentrations, with peak nearbed levels of the order of 10 kg m^{-3} . These high concentrations resulted in inversion difficulties particularly at 5.0 MHz, and coupled with the particle size ambiguities led to a degree of uncertainty in the results. To investigate these difficulties simulations were conducted, an example of which was presented. These exercises indicated that the inversion of the 5.0-MHz signal could be problematic at high concentrations, and due to coupling in the inversion this particularly contaminated the 1.0-MHz results. It was therefore considered reasonable to solely use the particle radii derived from the 1.0- and 2.5-MHz backscattered pressure profiles when conducting the comparison with the *in-situ* samples.

The assessment of the acoustic measurements of particle radius and concentration profiles was carried out using pumped samples. Linear regression and statistical analysis on the data showed that for most of the data set there were no significant differences between the acoustic and *in-situ* measurements of particle size and concentration. The deviations noted were at low concentrations where both techniques incurred some inaccuracies, and at the beginning of the record when there was the possibility of the bed degassing. However, the comparison clearly demonstrates that for the majority of the data, accurate measurements of mean profiles of particle size and concentration can be obtained using triple frequency acoustics. Given the validity of the mean observations, the high-resolution output for the triple frequency system was examined. This was primarily assessed by considering the internal consistency of the three concentration measurements. This showed that the normalized standard error had a mean value of 0.09. Given that the devices were not colocated, and therefore some differences were to be expected, this value and the coherence of the visualization of the suspension provided confidence in the integrity of the details observed in the suspension field. These images of the suspension depict a fluctuating field with high-concentration

events interspersed with regions of clearer water. The particle size on the other hand remained nominally uniform with time and height above the bed. This type of visualization is unique to the acoustic approach owing to its unobtrusive high-resolution profiling capability. Although not detailed here, such visualizations are beginning to be coupled with high-resolution flow velocity measurements from a recently developed coherent Doppler system.^{23,24} It is expected that the combination of colocated detailed visualization of the flow and suspended sediments should provide a powerful tool for studying suspended sediment dynamics.

ACKNOWLEDGMENTS

Tom Wylie conducted the acoustic calibrations and Glenn Holdaway carried out some of the initial inversion work, both are thanked for their input. The H. R. Wallingford, UK, field team helped with the collection of the data, and Roy Atkins in particular is thanked. The work was funded by the Natural Environment Research Council, UK, and the Flood Protection Division of the Ministry of Agriculture Fisheries and Food, UK.

- ¹E. M. Lee and L. M. Shapiro, "Echocardiographic assessment of coronary arteries and bypass grafts," *Brit. Med. Ultrasound Soc. J.* **3** (3), 27–31 (1995).
- ²W. H. Munk and P. F. Worcester, "Ocean acoustic tomography," *Oceanography* **1**, 8–10 (1988).
- ³D. M. Hanes, C. E. Vincent, D. A. Huntley, and T. L. Clarke, "Acoustic measurements of suspended sand in the C^2S^2 experiment at Stanhope Lane," *Prince Edward Island Mar. Geol.* **81**, 175–183 (1988).
- ⁴C. Libicki, K. W. Bedford, and J. F. Lynch, "The interpretation and evaluation of a 3-MHz acoustic backscatter device for measuring benthic boundary layer sediment dynamics," *J. Acoust. Soc. Am.* **85**, 1501–1511 (1989).
- ⁵C. E. Vincent and M. Green, "Field measurements of the suspended sand concentration profiles and of the resuspension coefficient γ_0 over a rippled bed," *J. Geophys. Res.* **95**(C7), 11591–11601 (1990).
- ⁶J. F. Lynch, T. F. Gross, B. H. Brumley, and Richard A. Filyo, "Sediment concentration in HEBBLE using a 1-MHz acoustic backscatter system," *Mar. Geol.* **99**, 361–385 (1991).
- ⁷J. F. Lynch, J. D. Irish, C. R. Sherwood, and Y. C. Agrawal, "Determining suspended sediment particle size information from acoustical and optical backscatter measurements," *Continental Shelf Res.* **14** (10/11), 1139–1165 (1994).
- ⁸A. E. Hay and A. J. Bowen, "Coherence scales of wave-induced suspended sand concentration fluctuations," *J. Geophys. Res.* **99**(C6), 12749–12765 (1994).
- ⁹P. D. Thorne, G. P. Holdaway, and P. J. Hardcastle, "Constraining acoustic backscatter estimates of suspended sediment concentration profiles using the bed echo," *J. Acoust. Soc. Am.* **98**, 2280–2288 (1995).
- ¹⁰J. Sheng and A. E. Hay, "An examination of the spherical scatterer approximation in aqueous suspensions of sand," *J. Acoust. Soc. Am.* **83**, 598–610 (1988).
- ¹¹P. D. Thorne and S. C. Campbell, "Backscattering by a suspension of spheres," *J. Acoust. Soc. Am.* **92**, 978–986 (1992).
- ¹²A. E. Hay and J. Sheng, "Vertical profiles of suspended sand concentration and size from multifrequency acoustic backscatter," *J. Geophys. Res.* **97**(C10), 15661–15677 (1992).
- ¹³P. D. Thorne, P. J. Hardcastle, and R. L. Soulsby, "Analysis of acoustic measurements of suspended sediments," *J. Geophys. Res.* **98**(C1), 899–910 (1993).
- ¹⁴A. E. Hay, "Sound scattering from a particle-laden, turbulent jet," *J. Acoust. Soc. Am.* **90**, 2055–2074 (1991).
- ¹⁵A. Downing, P. D. Thorne, and C. E. Vincent, "Backscattering from a suspension in the nearfield of a piston transducer," *J. Acoust. Soc. Am.* **97**, 1614–1620 (1995).
- ¹⁶A. M. Crawford and A. E. Hay, "Determining suspended sand size and

- concentration from multifrequency acoustic backscatter," J. Acoust. Soc. Am. **94**, 3312–3324 (1993).
- ¹⁷ M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972), Chap. 9.
- ¹⁸ K. G. Foote, E. Ona, and R. Toresen, "Determining the extinction cross section of aggregating fish," J. Acoust. Soc. Am. **91**, 1983–1989 (1992).
- ¹⁹ M. Marzoug and P. Amayenc, "Improved range-profiling algorithm of rainfall rate from a spaceborne radar with path-integrated attenuation constraint," IEEE Trans. Geosci. Remote Sens. **29**(4), 584–591 (1991).
- ²⁰ P. D. Thorne and P. J. Hardcastle, "Multifrequency acoustic imaging of suspended sediment processes," in *Proceedings of the Conference Oceans '94, Brest, France* (Clôître, St. Thonan, France, 1994), Vol. III, pp. 263–268.
- ²¹ P. D. Thorne, P. J. Hardcastle, G. P. Holdaway, and A. J. Born, "Analysis of results obtained from a triple frequency acoustic backscatter system for measuring suspended sediments," in *Proceedings of the IEE Conference on Electronic Engineering on Oceanography, Churchill College, Cambridge, UK, 1994* (Omega, UK, 1994), Conference publication number 394, pp. 83–89.
- ²² J. S. Bendat and A. G. Piersol, *Random Data* (Wiley, Chichester, UK, 1986), Chap. 4.
- ²³ P. J. Hardcastle, "A high resolution near bed coherent acoustic Doppler current profiler for measurement of turbulent flow," in *Proceedings of the IEE Conference on Electronic Engineering on Oceanography, Churchill College, Cambridge, UK, 1994* (Omega, UK, 1994), Conference publication number 394, pp. 73–76.
- ²⁴ P. J. Hardcastle, "A high resolution coherent acoustic doppler profiler for the measurement of near bed turbulent flow," Proc. Oceans 95, MTS/IEEE, 1361–1366 (1995).

An investigation of the small slope approximation for scattering from rough surfaces. Part II. Numerical studies

Shira Lynn Broschat

School of Electrical Engineering and Computer Science, Washington State University, Pullman, Washington 99164-2752

Eric I. Thorsos

Applied Physics Laboratory, College of Ocean and Fishery Sciences, University of Washington, Seattle, Washington 98105

(Received 3 May 1996; revised 1 November 1996; accepted 1 December 1996)

The small slope approximation (SSA) of Voronovich [Sov. Phys. JETP **62**, 65–70 (1985)] is a promising method for modeling wave scattering from rough surfaces. The SSA T -matrix series, which can be interpreted as an expansion in a generalized surface slope, satisfies the appropriate reciprocity condition at each order and reduces to the standard perturbation series for small surface roughness. When the SSA T matrix is found to second order in generalized slope, it reduces to that of the Kirchhoff approximation as the frequency is increased. In an earlier paper [E. I. Thorsos and S. L. Broschat, J. Acoust. Soc. Am. **97**, 2082–2093 (1995)] the derivation of the SSA for surfaces subject to the Dirichlet boundary condition was examined in detail. In this paper the accuracy of the SSA for the Dirichlet problem is investigated through comparison with exact results. Expressions for the first three terms of the SSA incoherent bistatic scattering cross-section series are presented, followed by numerical results for one-dimensional surfaces with Gaussian statistics and a Gaussian roughness spectrum. Surfaces with rms slope angles up to 45° are considered. It is found that, for the numerous cases studied, the SSA results agree well with the exact results over a broad range of scattering angles. When the lowest-order results are inaccurate, successive addition of each higher-order term generally yields improvement. The range of scattering angles over which the SSA results are accurate depends on both the rms slope angle and the surface correlation length, as well as on the angle of incidence. A simple rule of thumb, however, is that for an incident angle of 45° , the highest-order SSA scattering cross section examined here is accurate to within ± 1 dB from backscatter to a forward grazing angle of 5° for rms slope angles less than about 30° . When the surface roughness is such that perturbation theory is accurate, the SSA is accurate over the full range of scattering angles for small to moderate slopes. © 1997 Acoustical Society of America. [S0001-4966(97)06405-9]

PACS numbers: 43.30.Hw, 43.20.Fn [JHM]

INTRODUCTION

In the past two decades much research has been done on wave scattering from rough surfaces. A major objective has been to devise an analytic model for bistatic scattering that is accurate when the small perturbation method and Kirchhoff approximation apply but, in addition, is accurate in regions where they do not apply. One of the more promising of the new methods is the small slope approximation (SSA), which was introduced by Voronovich in the mid-1980s (1985, 1986).

In Part I of this two-part series (Thorsos and Broschat, 1995), we presented a derivation of the SSA for the Dirichlet problem and discussed in detail the assumptions made. The SSA takes the form of a systematic expansion that can be interpreted as a series in a generalized surface slope (henceforth, referred to simply as slope), as discussed later. The series is manifestly reciprocal at each order, and reduction to standard perturbation theory is inherent in the derivation. The first three terms in the small slope series T matrix— T_0 , T_1 , and T_2 —are derived in Part I.

Several numerical studies of the small slope approxima-

tion for rough surface scattering are available in the literature. Berman (1991) has compared Monte Carlo SSA results for bistatic scattering with exact results for both the Dirichlet and fluid–solid boundaries for one-dimensional (1-D) surfaces satisfying a Pierson–Moskowitz power-law spectrum (Dirichlet boundary) or a modified power-law spectrum (fluid–solid boundary). Broschat (1993) has compared formally averaged (theoretical) results for the reflection coefficient with exact results for a Pierson–Moskowitz power-law spectrum. These initial comparisons with exact calculations give encouraging support for the accuracy of the SSA. In addition, Yang and Broschat (1992, 1994) have presented results for 2-D surface, bistatic scattering for the Dirichlet problem for a Gaussian spectrum as well as 1-D surface results for acoustic scattering from a fluid–elastic-solid interface for both Gaussian and modified power-law spectra.

In this paper we investigate the validity of the SSA for scalar-wave scattering from 1-D, randomly rough surfaces satisfying the Dirichlet boundary condition and having Gaussian statistics and a Gaussian roughness spectrum. We consider the first three terms of the SSA series for the bistatic scattering cross section, σ , and compare SSA results for the

formally averaged bistatic scattering strength ($= 10 \log_{10} \sigma$) with exact Monte Carlo integral equation results. As discussed in Sec. I, these terms in the SSA series are second, third, and fourth order in slope, respectively.

Surfaces characterized by a Gaussian spectrum are single-scale surfaces, and the surface correlation length determines this scale. The correlation length can be varied to yield, at one extreme, surfaces with only small-scale roughness and, at the other, surfaces with only large-scale roughness. Thus, use of the Gaussian spectrum over the full range of correlation lengths and bistatic scattering angles provides a useful test of the SSA for varying length scales of roughness. A detailed study of low grazing angle scattering with application to power-law-type spectra will be presented in the future. Preliminary results have been reported for the lowest-order SSA cross section using a Pierson–Moskowitz power-law spectrum, and these are encouraging (Broschat and Thorsos, 1995).

In this numerical study we consider several aspects of the SSA for an incident angle of 45° , unless otherwise noted. First we examine the SSA for surfaces with a modest rms slope angle of 10° . The SSA is generally very accurate for these cases, although as the surface correlation length is increased, scattering is underpredicted at low grazing angles of scatter in the back direction and the SSA cross section becomes a little high at forward grazing angles less than about 5° . The underprediction at low grazing angles depends on the surface roughness parameters used and may extend up to 10° , 20° , or even 30° .

Next we compare the SSA with perturbation theory for surfaces with small roughness. The lowest-order term in the series for the SSA scattering cross section is generally accurate when lowest-order perturbation theory is accurate. There is an important difference in the manner in which the two approximations become inaccurate. For moderate slopes, the lowest-order SSA remains accurate near the specular direction for all kh (where k is the wave number and h is the rms surface height). However, for $kh > 1$ the lowest-order perturbation result becomes inaccurate over the entire angular range. Addition of both third- and fourth-order terms in the SSA cross-section series yields accurate results in and significantly beyond the region where higher-order perturbation theory is accurate (Thorsos and Jackson, 1989). Only beyond the perturbation theory regime do the results become low in the back direction and high at forward grazing angles less than 5° as the correlation length is increased.

Comparison of the SSA and perturbation theory is followed by comparison of the small slope and Kirchhoff approximations. In the Kirchhoff regime, the SSA and Kirchhoff approximation results agree well over a broad range of scattering angles. For scattering angles in the backward direction the lowest-order SSA prediction is lower than the Kirchhoff prediction when the Kirchhoff approximation is accurate. [For the Gaussian spectrum, the Kirchhoff approximation becomes inaccurate at grazing angles less than about twice the rms slope angle. See Thorsos (1988) for further discussion.] Successive addition of third- and fourth-order terms improves the result, but the SSA prediction is still slightly low. In the forward direction the lowest-order SSA is

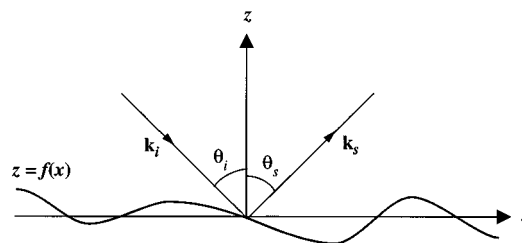


FIG. 1. Scattering geometry.

very accurate at all but the lowest grazing angles ($< 5^\circ$) for moderate surface slopes, unlike the Kirchhoff approximation whose results are typically high over a considerably larger grazing angle region.

As noted in Part I, the sum of the first two terms in the SSA T -matrix series reduces to the Kirchhoff approximation T matrix as the frequency is increased. When the corresponding terms in the SSA cross-section series are retained, reduction to the Kirchhoff cross section occurs as the frequency is increased. We find, however, that this reduction, first to the Kirchhoff approximation result and then to the geometrical optics result, occurs very slowly with increasing frequency for forward-scattering angles near grazing. Comparison with exact calculations shows that the SSA is accurate in this region, except at angles less than about 5° grazing, for rms surface heights at least up to one-and-a-half wavelengths.

In addition to the studies above, we investigate the accuracy of the fourth-order SSA for kh fixed and kl increasing (decreasing slope) and for kl fixed and kh increasing (increasing slope), where l is the surface correlation length. In general, we find that the fourth-order SSA and exact results agree well over a broad range of scattering angles for rms slope angles less than about 30° . In particular, we find that the results become low in the back direction as kl is increased; however, for increasing kh they remain remarkably accurate as the surface slope approaches 1, that is, as the surface slope angle approaches 45° , except at scattering angles near grazing where they become high.

For most of the cases examined the angle of incidence is 45° ; however, for two cases we also consider incident angles of 0° and 80° . The first case uses a moderate correlation length, and the fourth-order SSA results are excellent for both incident angles. The second case uses a longer correlation length, and the range of angles over which the fourth-order SSA is accurate becomes smaller as the incident angle is increased.

In Sec. I the terminology used for the SSA slope orders is described, and expressions for the first three terms of the SSA bistatic scattering cross-section series are given. Numerical results are presented in Sec. II.

I. BISTATIC SCATTERING CROSS SECTIONS

Since our numerical studies are for 1-D surfaces (the 2-D scattering problem), we restrict our derivation to this case; however, extension to 2-D surfaces is straightforward. Our scattering geometry is shown in Fig. 1 with the rough surface denoted by $z = f(x)$. The incident field is a scalar

plane wave $\exp[i\mathbf{k}_i \cdot \mathbf{r}]$ with $\mathbf{r}=(x,z)$ and $\mathbf{k}_i=(k_{ix}, k_{iz})=(k_{ix}, -\kappa_{iz})$, where $\kappa_{iz}=[k^2 - k_{ix}^2]^{1/2} > 0$ and $k=\omega/c$ is the radiation wave number. The small slope series for the T matrix is written as $T=T_0+T_1+T_2+\dots$. The first two terms of this expansion are given by

$$T_0(k_{sx}, k_{ix}) = -\frac{2\kappa_{iz}}{\nu_z} \frac{1}{2\pi} \int dx \exp[i\mathbf{v} \cdot \mathbf{r}]|_{z=f(x)}, \quad (1)$$

$$T_1(k_{sx}, k_{ix}) = \frac{i\kappa_{iz}}{\nu_z} \frac{1}{2\pi} \int dx \exp[i\mathbf{v} \cdot \mathbf{r}]|_{z=f(x)} \times \int dK_1 \exp[iK_1 x] F(K_1) \times [k\beta_{1+i} + k\beta_{s-1} - \nu_z], \quad (2)$$

where $\mathbf{v}=\mathbf{k}_i-\mathbf{k}_s=(\nu_x, -\nu_z)$ so that $\nu_x=k_{ix}-k_{sx}$ and $\nu_z=\kappa_{iz}+k_{sz}$, $k\beta_{1+i}=[k^2-(K_1+k_{ix})^2]^{1/2}$ with $\text{Im}[\beta_{1+i}]>0$, $k\beta_{s-1}=[k^2-(k_{sx}-K_1)^2]^{1/2}$ with $\text{Im}[\beta_{s-1}]>0$, and $F(K)$ is the Fourier transform of $f(x)$. Equations (1) and (2) as well as the expression for the third term of the series, T_2 , are derived in Part I. In general, the T_n term in the T -matrix series includes n factors of $F(K)$; in addition, the surface profile $f(x)$ enters into each term via the factor $\exp[i(\nu_x x - \nu_z f(x))]$. As discussed by Voronovich (1985) and in Part I, it can be shown that T_0 is first order in slope, T_1 is second order in slope, and so on. Actually, when we refer to a term as n th order in slope, we mean n th order in generalized slope. The term generalized slope is used because, as described in Part I, the integrand of the x integral in T_n ($n>0$) can be expressed in a coordinate space series as a product of $\exp[i\mathbf{v} \cdot \mathbf{r}]$ and a series beginning with the term $[f'(x)]^{n+1}$ followed by terms in higher powers of $f'(x)$ as well as by terms involving higher derivatives of $f(x)$. For $n=0$ the coordinate space series begins at $[f'(x)]^0$ and includes all contributions of order $f'(x), f''(x), f'''(x), \dots$. For simplicity, we refer to T_n as order $n+1$ in slope. The entire dependence of T_n on k_{sx} occurs in the factor $\exp[i\mathbf{v} \cdot \mathbf{r}]$ in this coordinate space series.

In terms of the T matrix, the bistatic scattering cross section for scattering from 1-D surfaces is given by (see Appendix A in Berman, 1991, and Appendix B in Thorsos and Jackson, 1989)

$$\sigma \delta(k_{ix} - k'_{ix}) = \frac{k_{sz}^2}{k} - [\langle T(k_{sx}, k_{ix}) T(k_{sx}, k'_{ix})^* \rangle - \langle T(k_{sx}, k_{ix}) \rangle \langle T(k_{sx}, k'_{ix})^* \rangle], \quad (3)$$

where the angle brackets denote an ensemble average, and the asterisk denotes the complex conjugate. The right-hand side of (3) is proportional to $\delta(k_{ix} - k'_{ix})$ which is a result of the assumed translational invariance of the surface statistics. Using the T -matrix series, $T=T_0+T_1+T_2+\dots$, in (3) then gives the scattering cross-section series, $\sigma=\sigma_{00}+\sigma_{01}+\sigma_{11}+\dots$, where

$$\sigma_{00} \delta(k_{ix} - k'_{ix}) = \frac{k_{sz}^2}{k} [\langle T_0(k_{sx}, k_{ix}) T_0(k_{sx}, k'_{ix})^* \rangle - \langle T_0(k_{sx}, k_{ix}) \rangle \langle T_0(k_{sx}, k'_{ix})^* \rangle], \quad (4)$$

$$\sigma_{01} \delta(k_{ix} - k'_{ix}) = \frac{k_{sz}^2}{k} 2 \text{Re}[\langle T_0(k_{sx}, k_{ix}) T_1(k_{sx}, k'_{ix})^* \rangle - \langle T_0(k_{sx}, k_{ix}) \rangle \langle T_1(k_{sx}, k'_{ix})^* \rangle], \quad (5)$$

$$\sigma_{11} \delta(k_{ix} - k'_{ix}) = \frac{k_{sz}^2}{k} [\langle T_1(k_{sx}, k_{ix}) T_1(k_{sx}, k'_{ix})^* \rangle - \langle T_1(k_{sx}, k_{ix}) \rangle \langle T_1(k_{sx}, k'_{ix})^* \rangle]. \quad (6)$$

Since T_0 is first order in slope, it follows that σ_{00} is second order in slope. Similarly, σ_{01} and σ_{11} are third and fourth order in slope, respectively. Thus we write the second-, third-, and fourth-order cross sections as

$$\sigma^{(2)} = \sigma_{00}, \quad (7)$$

$$\sigma^{(3)} = \sigma^{(2)} + \sigma_{01}, \quad (8)$$

$$\sigma^{(4)} = \sigma^{(3)} + \sigma_{11}. \quad (9)$$

Note that (9) does not give the complete fourth-order cross section since it is missing the σ_{02} term associated with T_0 and T_2 ; this omission will be discussed later. For simplicity, in this paper we will refer to (9) as the fourth-order cross section.

Evaluation of the formal averages in (4)–(6) is carried out in a way similar to that used for perturbation theory (e.g., Appendix B in Thorsos and Jackson, 1989). However, the moment evaluations are more complicated because, in addition to the factors of $F(K)$, the random variable $f(x)$ in the argument of the exponential function $\exp[i\mathbf{v} \cdot \mathbf{r}]|_{z=f(x)}$ must be considered. The first moments of T_0 and T_1 are

$$\langle T_0(k_{sx}, k_{ix}) \rangle = -\frac{2\kappa_{iz}}{\nu_z} \frac{1}{2\pi} \int dx \exp[i\nu_x x] \times \langle \exp[-i\nu_z f(x)] \rangle = -\exp[-\chi^2/2] \delta(k_{sx} - k_{ix}), \quad (10)$$

$$\langle T_1(k_{sx}, k_{ix}) \rangle = \frac{i\kappa_{iz}}{\nu_z} \frac{1}{(2\pi)^2} \int dx \exp[i\nu_x x] \times \int dK_1 \exp[iK_1 x] [k\beta_{1+i} + k\beta_{s-1} - \nu_z] \int dx' \exp[-iK_1 x'] \times \langle \exp[-i\nu_z f(x)] f(x') \rangle = \kappa_{iz} \exp[-\chi^2/2] \delta(k_{sx} - k_{ix}) \times \int dK_1 W(K_1) g(K_1), \quad (11)$$

where

$$\chi = \nu_z h,$$

$$g(K_1) = k\beta_{1+i} + k\beta_{s-1} - \nu_z$$

[note that $g(K_1)$ can be complex], and $W(K_1)$ is the surface roughness spectrum given by

$$W(K_1) = \frac{1}{2\pi} \int dx \exp[-iK_1x] \langle f(x_0)f(x+x_0) \rangle.$$

To obtain (11), $F(K)$ in (2) is expressed as the Fourier transform of $f(x)$, giving a moment of the form $\langle \exp[i\alpha_1 X_1] X_2 \rangle$, where X_1 and X_2 are zero-mean Gaussian random variables. This moment is found by use of the identity

$$\langle e^{i\alpha_1 X_1} X_2 \rangle = -i \left. \frac{\partial \langle e^{i\alpha_1 X_1} e^{i\alpha_2 X_2} \rangle}{\partial \alpha_2} \right|_{\alpha_2=0}. \quad (12)$$

The partial derivative in (12) is readily evaluated using (Parzen, 1962)

$$\langle e^{i\alpha_1 X_1} e^{i\alpha_2 X_2} \rangle = \exp[-(1/2)(\alpha_1^2 \langle X_1^2 \rangle + 2\alpha_1 \alpha_2 \langle X_1 X_2 \rangle + \alpha_2^2 \langle X_2^2 \rangle)]. \quad (13)$$

The second moments are obtained similarly as described shortly. When evaluating the right-hand sides of (4)–(6), the result for each contains a factor of $\delta(k_{ix} - k'_{ix})$, which then cancels out to give the cross section terms. The integral in (1) is identical to that of the Kirchhoff approximation (Ishimaru, 1978). Thus, the result of performing the formal averages in (4) differs from that of the Kirchhoff approximation by a coefficient only, and the second-order SSA cross section is given by

$$\sigma^{(2)} = \frac{2k_{sz}^2 \kappa_{iz}^2}{\pi k \nu_z^2} \exp[-\chi^2] \int dx \exp[i\nu_x x] B_C(x), \quad (14)$$

where

$$B_C(x) = \exp[\chi^2 C(x)] - 1,$$

where $C(x)$ is the surface correlation function normalized so that $C(0) = 1$ and is given by

$$C(x)h^2 = \langle f(x_0)f(x+x_0) \rangle = \int dK W(K) \exp[iKx].$$

The relationship between the second-order SSA and Kirchhoff approximation cross sections is given by

$$\sigma^{(2)} = \sigma_{KA}/g^2, \quad (15)$$

where

$$g = \frac{\kappa_{iz} \nu_z - k_{ix} \nu_x}{2k_{sz} \kappa_{iz}}. \quad (16)$$

In the specular direction $g=1$, and (14) reduces to the Kirchhoff approximation cross section.

The terms $\langle T_0 T_1^* \rangle$ and $\langle T_1 T_1^* \rangle$ in (5) and (6) contain ensemble averages of the type $\langle e^{i\alpha_1 X_1} e^{i\alpha_2 X_2} X_3 \rangle$ and $\langle e^{i\alpha_1 X_1} e^{i\alpha_2 X_2} X_3 X_4 \rangle$ where X_1 through X_4 are zero-mean Gaussian random variables. The latter average can be found using

$$\langle e^{i\alpha_1 X_1} e^{i\alpha_2 X_2} X_3 X_4 \rangle = - \left. \frac{\partial^2 \langle e^{i\alpha_1 X_1} e^{i\alpha_2 X_2} e^{i\alpha_3 X_3} e^{i\alpha_4 X_4} \rangle}{\partial \alpha_3 \partial \alpha_4} \right|_{\alpha_3=\alpha_4=0} \quad (17)$$

and the former average using a similar relationship. The partial derivative is readily evaluated using (Parzen, 1962)

$$\begin{aligned} & \langle e^{i\alpha_1 X_1} e^{i\alpha_2 X_2} e^{i\alpha_3 X_3} e^{i\alpha_4 X_4} \rangle \\ &= \exp \left[- \left(\frac{1}{2} \right) \sum_j \sum_k \alpha_j \alpha_k K_{jk} \right], \end{aligned} \quad (18)$$

where $K_{jk} = \langle X_j X_k \rangle$. The final results after some lengthy but straightforward manipulations are given by

$$\begin{aligned} \sigma_{01}(k_{sx}, k_{ix}) &= \frac{2k_{sz}^2 \kappa_{iz}^2}{\pi k \nu_z} \exp[-\chi^2] \\ &\times \text{Re} \left\{ -J^* \int dx \exp[i\nu_x x] B_C(x) \right. \\ &+ \int dx I(x) \int dK \exp[iKx] W(K) \\ &\times g^*(K) \left. \right\}, \end{aligned} \quad (19)$$

$$\begin{aligned} \sigma_{11}(k_{sx}, k_{ix}) &= \frac{k_{sz}^2 \kappa_{iz}^2}{2\pi k} \exp[-\chi^2] \\ &\times \left\{ |J|^2 \int dx \exp[i\nu_x x] B_C(x) \right. \\ &+ \int dx I(x) \int dK \exp[iKx] W(K) \\ &\times \left\{ \frac{1}{\nu_z^2} |g(K)|^2 - 2 \text{Re}[Jg^*(K)] + g(K) \right. \\ &\times \left. \int dK' \exp[iK'x] W(K') g^*(K') \right\} \left. \right\}, \end{aligned} \quad (20)$$

where

$$I(x) = \exp[i\nu_x x] \exp[\chi^2 C(x)],$$

$$J = \int dK W(K) g(K).$$

Since \mathbf{v} and $g(K)$ are both reciprocal, it follows that $\sigma^{(2)}$, σ_{01} , and σ_{11} given by (14), (19), and (20), respectively, all satisfy the reciprocity condition, that is, exchanging k_{ix} and k_{sx} with $-k_{sx}$ and $-k_{ix}$ results in identical expressions for the cross-section terms. For the angle convention of Fig. 1 the reciprocity condition implies that the cross-section terms are unchanged when θ_i and θ_s are replaced by $-\theta_s$ and $-\theta_i$. In addition, because of the assumed translational invariance of the surface statistics, moments remain the same when $f(x)$ is replaced by $f(-x)$ for each surface in the surface ensemble. It follows that the cross section terms are unchanged when θ_i and θ_s are replaced by $-\theta_i$ and $-\theta_s$. This combined with the reciprocity condition implies that the cross-section terms do not change when θ_i and θ_s are replaced by θ_s and θ_i .

As discussed in Part I (Sec. II) and earlier by Voronovich (1985), $T_0 + T_1$ reduces to the Kirchhoff approximation T matrix for large k . It is straightforward to show the corresponding result for the cross section, that is, $\sigma^{(4)}$ given by (9) reduces to σ_{KA} , the Kirchhoff approximation cross sec-

tion, for large k . Thus, in the high-frequency limit as $k \rightarrow \infty$, $\sigma^{(4)} \rightarrow \sigma_{GO}$, the geometrical optics result.

As mentioned earlier in this section, the fourth-order SSA cross section given by (9) is missing the σ_{02} term and, thus, does not contain all possible fourth-order components. This is not consistent with the second- and third-order expressions given by (7) and (8) since they contain all second- and third-order components, respectively. There are several reasons for including σ_{11} and not σ_{02} in (9). As stated in the previous paragraph, $\sigma^{(4)}$ given by (9) reduces to the Kirchhoff approximation cross section for large k . This is strong motivation for including σ_{11} . Excluding the σ_{02} term is motivated by the need to balance numerical accuracy and numerical complexity. From (14), (19), and (20), it can be seen that evaluating σ_{00} involves a single integral, while σ_{01} and σ_{11} involve double integrals. Evaluation of the σ_{02} term, not presented here, involves triple integrals. Thus, inclusion of σ_{02} would generally be considered less practical for most applications. Since σ_{02} and σ_{11} are of the same order of slope, they may in fact be of comparable size. However, comparisons with integral equation results show $\sigma^{(4)}$, without σ_{02} , to be an improvement over $\sigma^{(3)}$. This improvement is sufficient justification for keeping only part of the fourth-order slope contribution to the SSA cross section series.

In standard perturbation theory (PT) a cross-section series in even powers of kh is obtained. The lowest-order term is denoted σ_{PT}^{11} or PT(2) and is of order $(kh)^2$. [The notation follows that of Thorsos and Jackson (1989) with the subscript PT added to distinguish between perturbation theory and the SSA. The superscript 11 indicates the absolute square of the first-order scattered field.] The next order in the perturbation cross-section series is $\sigma_{PT}^{22} + \sigma_{PT}^{13}$, where σ_{PT}^{13} is the contribution from the product of the first- and third-order scattered fields, or equivalently, from the product of the first- and third-order T matrices. Both σ_{PT}^{22} and σ_{PT}^{13} are of order $(kh)^4$. For the SSA, as discussed in Part I, the small height expansion of $T_{SSA} = T_0 + \dots + T_n$ agrees exactly with standard perturbation theory to order $(kh)^{n+1}$. This implies that when $\sigma^{(2)}$ is expanded in powers of kh , the lowest-order term is exactly σ_{PT}^{11} , and additional terms are of higher even powers in kh . Among these terms are reciprocal approximations to σ_{PT}^{22} and σ_{PT}^{13} . When $\sigma^{(3)}$ is expanded in a perturbation series, these approximations are improved. Finally, in the perturbation expansion of $\sigma^{(4)}$, both σ_{PT}^{11} and σ_{PT}^{22} are obtained exactly. To obtain σ_{PT}^{13} exactly, it would be necessary to extend the T_{SSA} series through T_2 in order to include σ_{02} in $\sigma^{(4)}$. The perturbation expansion of $\sigma^{(4)}$ would then yield the standard perturbation theory result to order $(kh)^4$.

II. NUMERICAL RESULTS

In this section, the accuracy of the second-, third-, and fourth-order SSA bistatic scattering cross sections— $\sigma^{(2)}$, $\sigma^{(3)}$, and $\sigma^{(4)}$ —is examined by comparison with Monte Carlo integral equation calculations (Thorsos, 1988). Numerical results are presented for the scattering strength ($= 10 \log_{10} \sigma$) and are restricted to scalar wave scattering from 1-D Dirichlet surfaces satisfying the Gaussian roughness spectrum as given by

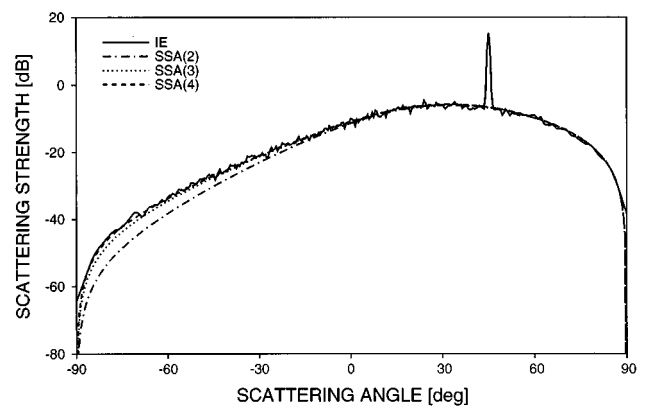


FIG. 2. Comparison of the second-order [SSA(2)], third-order [SSA(3)], and fourth-order [SSA(4)] small slope approximation and Monte Carlo integral equation (IE) scattering strengths ($= 10 \log_{10} \sigma$) for a modest rms surface slope angle. Here $kh=0.5$, $kl=4.0$, where k is the wave number, h is the rms surface height, and l is the correlation length, the rms slope angle γ is 10° , and the incident angle θ_i is 45° .

$$W(K) = \frac{h^2 l}{(2\sqrt{\pi})} e^{-K^2 l^2 / 4},$$

where h is the rms surface height and l is the surface correlation length. For each example we calculate the scattering strength over the complete 180° range of scattering angles for a particular incident angle. When the SSA results agree with exact results to within 1 dB, we consider them to be accurate. However, we limit our investigation to scattering strengths greater than -80 dB.

A. Accuracy of the SSA for modest surface slopes

We begin by examining two cases with rms surface slope angles of 10° . Figure 2 shows the SSA and exact results when the surface roughness parameters kh and kl are 0.5 and 4.0, respectively, the incident angle θ_i is 45° , the rms slope s is 0.177 ($s = \sqrt{2}h/l$ for Gaussian surfaces), and the rms slope angle γ is $\tan^{-1} s = 10^\circ$. The peak in the specular direction of the integral equation result is due to the coherent component of the total cross section and represents specular reflection, which has not been included in the approximate methods. The specular peak can be subtracted out easily, but will be included for most of the cases presented as a qualitative indication of the size of the coherent component. Fluctuations in the integral equation results are due to the finite number of surface realizations used in the ensemble average. Here, 50 realizations have been used. As more surface realizations are added, the fluctuations decrease in magnitude; in the limit of an infinite number of realizations, the curve would be smooth. For details on the integral equation technique, see Thorsos (1988). From Fig. 2 we see that the second-order SSA [SSA(2)] is accurate in the forward direction ($0^\circ < \theta_s \leq 90^\circ$), but it is not accurate in the back direction ($-90^\circ \leq \theta_s < 0^\circ$). The third-order result shows marked improvement, and the fourth-order SSA [SSA(4)] is accurate everywhere.

In Fig. 3 kh and kl have both been increased by a factor of three to 1.5 and 12.0, respectively, yielding a smoother surface. The coherent component at $\theta_s = 45^\circ$ is reduced to nearly the incoherent level because of the greater surface height, and the scattering level in the back direction is re-

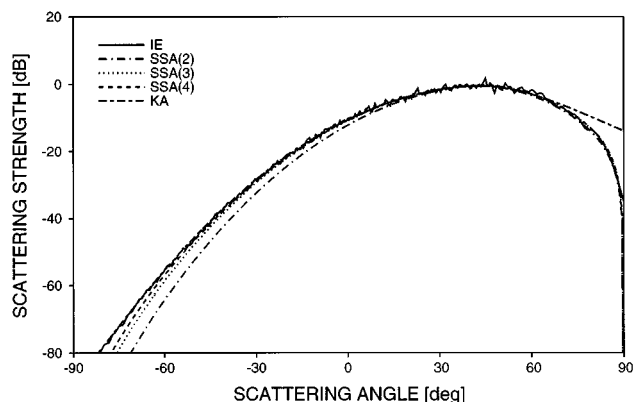


FIG. 3. Comparison of the SSA(2), SSA(3), SSA(4), Kirchhoff approximation (KA), and IE scattering strengths. The rms height and correlation length of Fig. 2 have been increased threefold to $kh=1.5$ and $kl=12.0$, and $\gamma=10^\circ$ and $\theta_i=45^\circ$.

duced because of the increased correlation length. Once again the SSA(2) underpredicts scattering in the back direction, and the third- and fourth-order results show successive improvement. In addition, while not visible in Fig. 3, at very low forward grazing angles ($<5^\circ$) all three orders of the SSA overpredict the scattering strength by slightly more than 1 dB. This behavior is discussed in Sec. II G.

The examples presented in Figs. 2 and 3 illustrate a couple of general points. First, except when the rms slope angle becomes large ($\gamma > 30^\circ$), extending the cross-section series to higher orders increases the region of accuracy to scattering angles further from the specular direction. Second, even the SSA(2) is accurate in the forward-scattering region for modest rms surface slopes except at grazing angles below 5° . (Initial numerical results show that this underprediction of scattering levels in the back direction for smooth surfaces with a Gaussian spectrum does not extend to surfaces with power-law roughness spectra since these surfaces usually contain sufficient small-scale roughness to be rough on the scale of a wavelength. These results will be presented in a future paper.)

B. Comparison with perturbation theory

The series for the scattering cross section obtained from perturbation theory is in even orders of kh , beginning at $(kh)^2$. Therefore, the lowest-order perturbation cross section is second order in kh (commonly referred to as the “first-order” perturbation result since the scattered field is found to first order in kh) and the next order is fourth order in kh . By construction, the SSA(2) cross section must reduce to that of lowest-order perturbation theory [PT(2)] in the limit as kh becomes small. [Note that reference to orders of the SSA is to orders in surface slope; whereas for perturbation theory orders refer to orders in kh . The PT(2) cross section is of order $(kh)^2$.] In Fig. 4 we examine a case when PT(2) is accurate and where the Kirchhoff approximation does not apply. The surface roughness parameters kh and kl are 0.1 and 1.4, respectively, the incident angle θ_i is 45° , and the

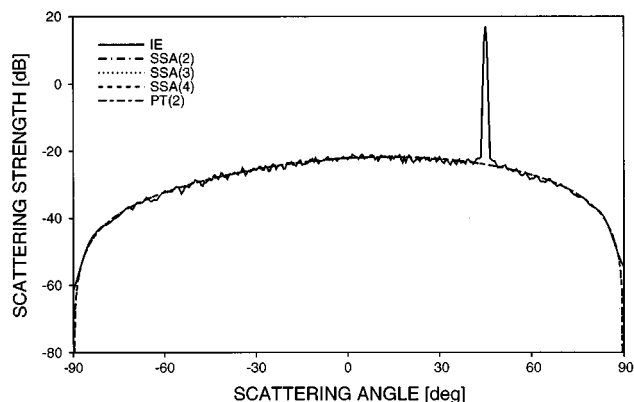


FIG. 4. Comparison of the SSA(2), SSA(3), SSA(4), lowest-order perturbation theory [PT(2)], and IE scattering strengths for a case when PT(2) is accurate. The SSA and PT(2) results are coincident. Here $kh=0.1$, $kl=1.4$, $\gamma=5.8^\circ$, and $\theta_i=45^\circ$.

rms slope angle γ is 5.8° . For this case, we see that the SSA results for the first three orders are the same as the PT(2) result.

A large number of cases were examined to determine the accuracy of the SSA(2) cross section. The accuracy of the SSA(2) results generally extends beyond the region when PT(2) is accurate, but as with perturbation theory, degradation occurs as the correlation length is increased. There is an important difference between the manners in which PT(2) and the SSA(2) become inaccurate for increasing kh . In the region of kh - kl parameter space with moderate slopes (γ less than $\sim 20^\circ$), the SSA(2) remains accurate near the specular direction for all kh . In contrast, PT(2) becomes inaccurate over the entire angular range for $kh > 1$.

We have examined a number of different cases outside the region of validity of PT(2) for which fourth-order PT [PT(4)=PT(2)+ $\sigma_{PT}^{22}+\sigma_{PT}^{13}$] is accurate. For the case shown in Fig. 5 with $kh=0.67$, $kl=2.83$, $\theta_i=45^\circ$, and $\gamma=18.4^\circ$, the SSA(4) cross section is accurate over all scattering angles. In fact, it is more accurate than the PT(4) result which barely meets the ± 1 -dB criterion for accuracy over a broad range of angles.

Finally, we have examined several cases beyond the region of validity of PT(4), where both PT(4) and the Kirchhoff approximation have limited applicability. For example, in Fig. 6 $kh=1$, $kl=4.3$, and $\gamma=18.2^\circ$. Since $kh=1$, perturbation theory is not expected to do well; in fact, the dropout in the PT(4) curve occurs because the cross section becomes negative. The Kirchhoff approximation is better, but for this case, the SSA(4) is clearly the best. It is accurate over almost the entire range of scattering angles.

C. Comparison with the Kirchhoff approximation

The Kirchhoff approximation (KA) is accurate away from low grazing angles for rough surfaces smooth on the scale of a wavelength, assuming small surface slopes. For a Gaussian roughness spectrum, the KA is accurate over a broad range of scattering angles when $kl > 6$, that is when $l/\lambda > 1$ (Thorsos, 1988). In this section we consider the SSA and KA for several examples with $kl > 6$.

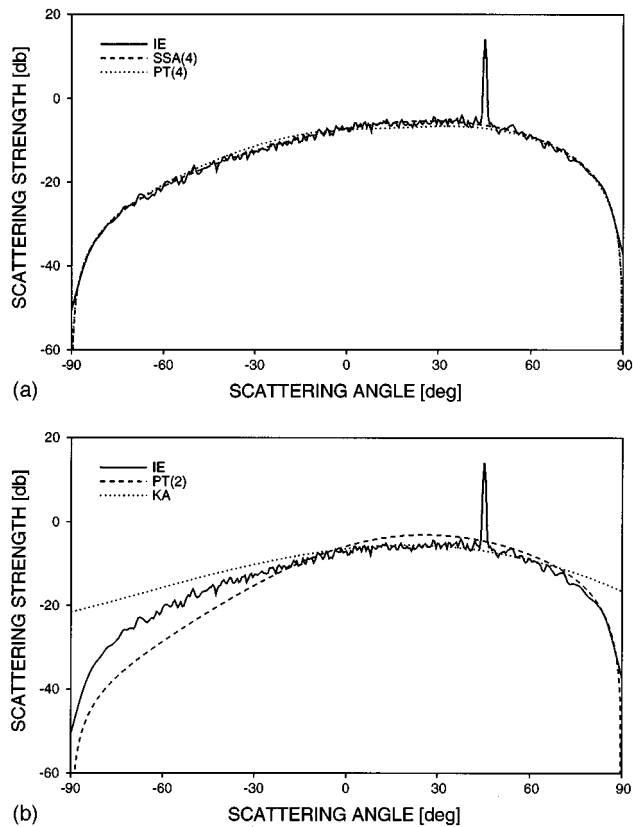


FIG. 5. Comparison of the (a) SSA(4) and fourth-order perturbation theory [PT(4)] and (b) PT(2) and KA scattering strengths with the IE results for a case when neither PT(2) nor the KA is accurate. Here $kh=0.67$, $kl=2.83$, $\gamma=18.5^\circ$, and $\theta_i=45^\circ$.

In Fig. 3 ($kh=1.5$, $kl=12$, $\theta_i=45^\circ$, and $\gamma=10^\circ$) the KA is accurate away from low grazing angles, while perturbation theory does not apply. The SSA(2) result reduces to the KA result in the specular region; the third- and fourth-order results do as well. In the backscatter region the KA is more accurate than the SSA. On the other hand, the KA overpredicts the scattering strength in the forward direction at low grazing angles. In contrast, the SSA scattering strength, even for the lowest order, is remarkably good in this same region, although as explained later the SSA results

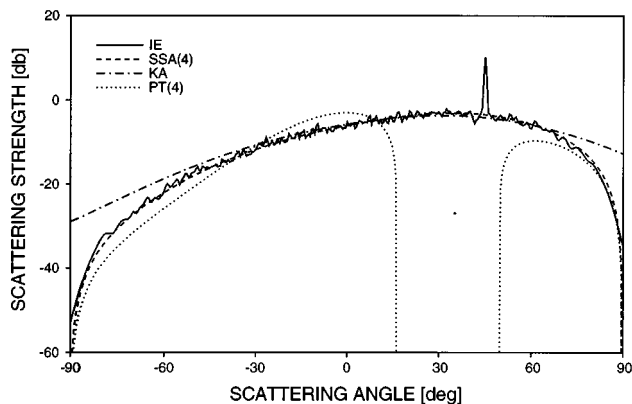


FIG. 6. Comparison of the SSA(4), KA, and PT(4) scattering strengths with the IE results for a case when neither PT(4) nor the KA is accurate. Here $kh=1.0$, $kl=4.3$, $\gamma=18.2^\circ$, and $\theta_i=45^\circ$.

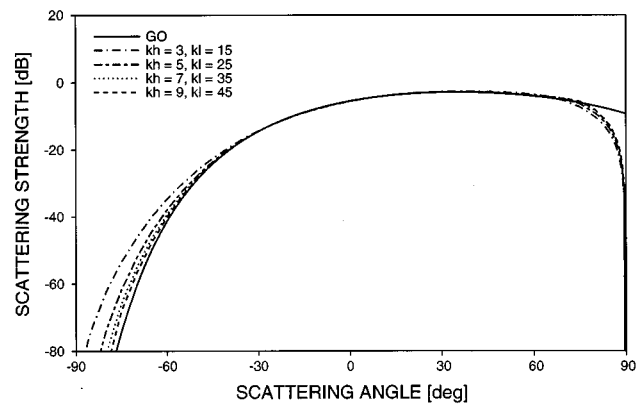


FIG. 7. Reduction of the SSA(4) result to the geometrical optics result (GO) as kh and kl are increased for a fixed rms slope angle $\gamma=15.8^\circ$ and an incident angle of 45° . The SSA(4) scattering strength reduces analytically to the GO result in the high-frequency limit, but the reduction is nonuniform in scattering angle and, at low grazing angles, is very slow.

are somewhat high at the lowest forward grazing angles.

The SSA(4) cross section reduces to the KA cross section as the frequency is increased (Voronovich, 1985; Thorsos and Broschat, 1995), which suggests that the SSA may become significantly inaccurate at low grazing angles for high frequencies. To study the reduction of the SSA(4) to the KA, we considered several cases with fixed surface slope angle and increased the frequency, that is, we increased both kh and kl so that the surface slope angle remained constant. We then compared the SSA and KA results and found that the reduction occurs quickly in the specular region and much more slowly away from this region, especially near grazing. In fact, as will be seen shortly, the SSA remains quite good at low grazing angles as the frequency is increased.

In the high-frequency limit ($k \rightarrow \infty$) both the SSA(4) and KA cross sections reduce to the geometrical optics (GO) expression

$$\sigma_{GO}(\theta_s, \theta_i) = \frac{(1 + \cos \theta_i \cos \theta_s - \sin \theta_i \sin \theta_s)^2}{(\cos \theta_i + \cos \theta_s)^3} \times \exp \left\{ -\frac{(\sin \theta_i - \sin \theta_s)^2 l^2}{4(\cos \theta_i + \cos \theta_s)^2 h^2} \right\}.$$

In Fig. 7, we show a case for $l/h=5$, an rms slope angle γ of 15.8° , and an angle of incidence θ_i of 45° . As we increase both kh and kl , the SSA(4) scattering strength approaches the GO result. This reduction is nonuniform in scattering angle: Over a large range of scattering angles reduction occurs quickly; however, at grazing angles less than about 10° the reduction is quite slow. In the backscatter region, the SSA(4) curves lie close to the KA curves (not shown) for each case, except for $\theta_s < -80^\circ$ as in Fig. 3, as both converge to the GO limit. In contrast, for $\theta_s > 70^\circ$, the KA scattering strengths all reduce to the GO limit, while the SSA(4) results converge very slowly to this limit. In Fig. 8 we compare the SSA(4) prediction with the GO and shadowed GO results (Wagner, 1967) in the forward-scattering region for the highest-frequency case in Fig. 7. We also show the integral equation scattering strength. The two GO approximations differ significantly only for scattering angles greater

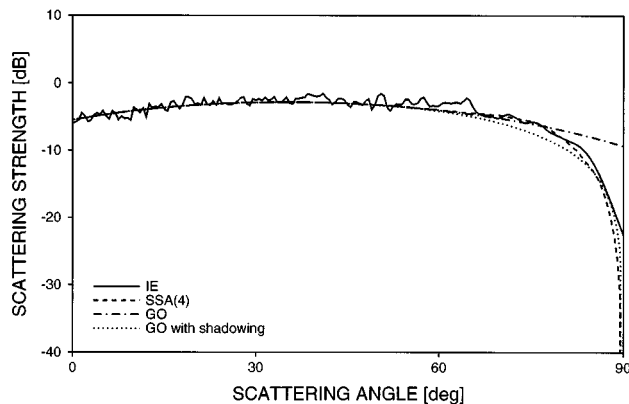


FIG. 8. Comparison of the SSA(4), GO, GO with shadowing, and IE scattering strengths for $kh=9$, $kl=45$, $\gamma=15.8^\circ$, and $\theta_i=45^\circ$. Note that the SSA(4) result is more accurate in the forward-scattering region than the shadowed geometrical optics result.

than approximately 70° . We see in Fig. 8 that the SSA(4) is quite good in this region; in fact, it is superior to both the unshadowed and shadowed GO approximations. Even the SSA(2) cross section gives results superior to those of the Kirchhoff or GO approximations, while it is as readily calculated as the KA cross section.

To explore the relationship between the surface slope and the accuracy of the SSA, we focus on the effects of changing kh for a fixed kl and kl for a fixed kh , that is, increasing and decreasing the rms surface slope, respectively, in Secs. II D and II E. Then in Sec. II F we consider angles of incidence other than 45° . Finally, in Sec. II G we return to the behavior of the SSA at low forward grazing angles when the correlation length is increased.

D. Increasing the rms surface height for a fixed correlation length

The SSA is not restricted to small surface roughness and, as shown previously, can be accurate over a broad range of scattering angles for moderate surface slopes. As explained in Sec. I, the terms of the T -matrix series depend on more than just simple surface slope. However, we might still expect a relationship between the accuracy of the SSA and rms surface slope values. In fact, when we increase the rms slope for a fixed correlation length by increasing the rms surface height, the SSA(4) results become visibly high in the forward-scattering direction. In Fig. 9(a) and (b) we have chosen kl to be 2.83 and have used a kh of 1.33 and 2.0, respectively. Figures 5, 9(a), and 9(b) correspond to increasing rms slope angles from 18.5° to 45° for a fixed kl . The incident angle is 45° . When $kh=1.33$ and $\gamma=33.7^\circ$ [Fig. 9(a)], the accuracy of the SSA(4) scattering strength starts to degrade at low forward grazing angles where the result is too high (the error extends up to about 1.5 dB). This overprediction becomes more pronounced and occurs at higher grazing angles as the rms slope is further increased as shown in Fig. 9(b). In addition, as the rms slope angle is increased up to and beyond 45° , scattering levels become high at low back-scatter grazing angles as well. This overprediction, just becoming evident in Fig. 9(b), also becomes greater and occurs at higher grazing angles as the rms slope is further increased.

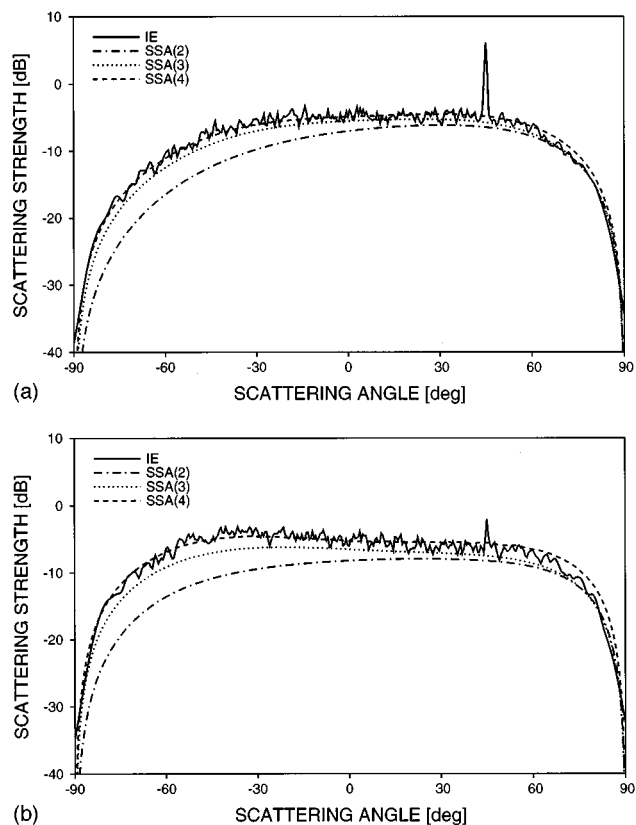


FIG. 9. The effect of increasing the slope by changing kh for a fixed kl and fixed incident angle, where $kl=2.83$ and $\theta_i=45^\circ$, (a) $kh=1.33$, $\gamma=33.6^\circ$, and (b) $kh=2.0$, $\gamma=45.0^\circ$. Also, compare the result in Fig. 5 ($kh=0.67$, $\gamma=18.5^\circ$). As the rms slope is increased beyond 30° , the SSA(4) scattering strength becomes high in the forward direction. When it is increased beyond 45° , the SSA(4) scattering strength becomes high in the low back direction as well.

A similar study was done for a fixed kl of 12.0 and an incident angle of 45° ; the results show the SSA(4) scattering strength becomes inaccurate at a smaller rms surface slope for this higher-frequency case.

One way of observing the decrease in accuracy of the SSA with increasing slope is to compare the third- and fourth-order results. When $\sigma^{(2)}$, $\sigma^{(3)}$, and $\sigma^{(4)}$ are coincident as in Fig. 4 the terms σ_{01} and σ_{11} make negligible contributions to the cross section, and the first three orders of the SSA are all found to be accurate. When the difference between $\sigma^{(3)}$ and $\sigma^{(4)}$ is small, σ_{11} makes a negligible contribution to the scattering strength and the third- and fourth-order results are found to be accurate. When the difference between $\sigma^{(3)}$ and $\sigma^{(4)}$ becomes significant, the fourth-order result may no longer be accurate. To determine how large this difference must be, we examined $\Delta SS = 10 \log_{10}[\sigma^{(4)}/\sigma^{(3)}]$ for the cases given in Figs. 5 and 9 as well as for other cases. We found that for a fixed kl and increasing kh , that is, for increasing slopes, ΔSS shows the largest increase near grazing in both the forward- and backward-scattering directions. This is consistent with our discussion in the preceding paragraph. Furthermore, larger values of kl have larger values of ΔSS in the backward-scattering direction relative to the forward-scattering direction for the same values of slope. Unfortunately, the relation-

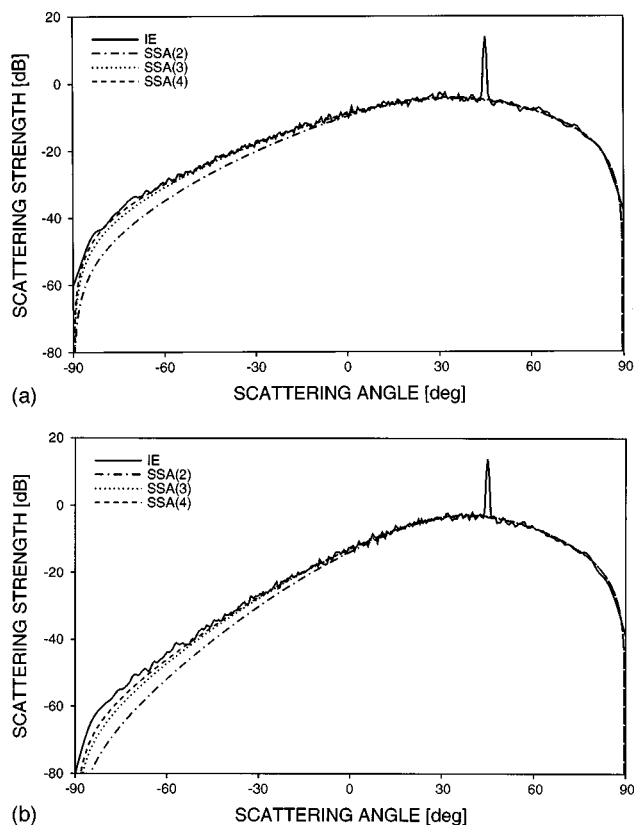


FIG. 10. The effect of decreasing the slope by changing kl for a fixed kh and fixed incident angle, where $kh=0.67$ and $\theta_i=45^\circ$, (a) $kl=4.24$, $\gamma=12.6^\circ$, and (b) $kl=6.28$, $\gamma=8.6^\circ$. Also, compare the result in Fig. 5 ($kl=2.83$, $\gamma=18.5^\circ$). As the correlation length is increased beyond about 5, the SSA(4) scattering strength becomes somewhat low in the back direction.

ship between ΔSS and the accuracy of $\sigma^{(4)}$ cannot be quantified in a simple manner.

E. Increasing the correlation length for a fixed rms surface height

Next it is of interest to examine what occurs as the correlation length is increased for a fixed rms surface height. This is equivalent to decreasing the rms slope for a fixed kh . In Fig. 10(a) and (b) we have chosen kh to be 0.67 and have used a kl of 4.24 and 6.28, respectively. Figures 5, 10(a), and 10(b) correspond to decreasing rms slope angles from 18.5° to 8.6° for a fixed kh . The incident angle is 45° . For kl up to 4.24, there is excellent agreement between the SSA(4) and exact results. However, beginning at a kl of about 5 the SSA(4) underpredicts scattering in the back direction. This underprediction is true in Fig. 3 as well, where $kh=1.5$ and $kl=12$. An additional study was done for a fixed kh of 2.0, an incident angle of 45° , and kl varying between 4 and 15 (equivalent to rms slope angles between 35.3° and 10.7°) with similar results, that is, as the correlation length is increased, backscatter is underpredicted. The discrepancy between the SSA and exact results is greater at smaller backward grazing angles and worsens as the correla-

tion length is increased. In the forward direction the results become slightly high at forward grazing angles smaller than 5° as kl is increased.

Again we examined $\Delta SS = 10 \log_{10}[\sigma^{(4)}/\sigma^{(3)}]$, this time for the cases shown in Figs. 5 and 10 as well as for other cases. We found that for a fixed kh and increasing kl , that is, for decreasing slopes, ΔSS becomes larger in the backward-scattering direction, indicating the decrease in accuracy of the fourth-order result and reflecting the importance of the σ_{11} term in the back direction.

The underprediction of backscatter by the SSA for large kl and small kh can be explained as follows: The SSA T matrix is expanded in powers of the surface profile (or its Fourier transform) and the resulting terms are equated with those of the standard perturbation series T matrix (Thorsos and Broschat, 1995). To find T_0 , terms up to first order in the surface profile are retained; to find T_1 , terms up to second order are retained; and to find T_2 , terms up to third order are retained. In earlier work (Thorsos and Jackson, 1989) it was shown that for a Gaussian roughness spectrum, lowest-order perturbation theory becomes inaccurate as the correlation length is increased even when $kh \ll 1$. In particular, the perturbation prediction for scattering in the back direction becomes low. However, in the forward direction the results are quite accurate, even for very large kl . When the perturbation cross section is computed to fourth order in kh , scattering levels in the back direction are accurately predicted for larger kl . For the SSA the same behavior occurs for small kh since the SSA reduces to PT in this region, but interestingly enough, it occurs for larger kh as well. For general kh the SSA(2) result underpredicts scattering in the back direction as the correlation length is increased. The results improve when the third- and fourth-order cross sections are used, particularly at fourth order which includes two factors of T_1 . Thus, if the SSA were extended to include the T_2 term, we would expect our results to improve even further in the back direction for large correlation lengths.

F. Other angles of incidence

In the previous sections we considered examples with incident angles of 45° . In this section we briefly discuss other angles of incidence and present a few examples. For the cases studied, we noted that when the scattering strength results are accurate over all angles of scatter for an incident angle of 45° , in general they are good for incident angles up to 80° . When the results are low in the backscatter region for an incident angle of 45° , they are better for 0° incidence and worse for 80° incidence. These results are illustrated in Figs. 11 and 12. In Fig. 11(a) and (b) the incident angles are 0° and 80° , respectively, for the same roughness parameters used in Fig. 5, $kh=0.67$, $kl=2.83$, and $\gamma=18.5^\circ$. In Fig. 11(b), and later in Fig. 12(b), the coherent component has been subtracted from the integral equation result (see Thorsos, 1990, for details) so the specular peak, which is wider at lower grazing angles, does not obscure the incoherent component. For the results presented in Fig. 11(a) and (b) the SSA(4) scattering strengths show excellent agreement with the integral equation results over the entire range of scattering angles. However, when the correlation length is in-

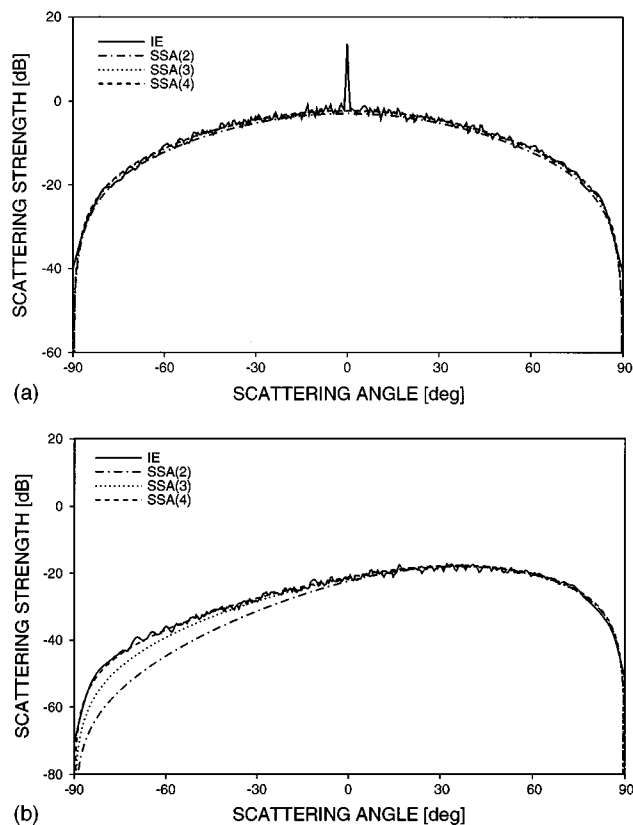


FIG. 11. Comparison of the SSA(2), SSA(3), SSA(4), and IE scattering strengths for the case shown in Fig. 5 but with an incident angle of (a) 0° and (b) 80° . In (b) the coherent component has been subtracted from the IE result to facilitate comparisons of the incoherent scatter.

creased, the accuracy of the SSA is not as good. For example, in Fig. 12(a) and (b) we consider incident angles of 0° and 80° for the same surface parameters used in Fig. 3, $kh = 1.5$, $kl = 12$, and $\gamma = 10^\circ$. For the 0° case, the SSA(4) clearly performs better than the KA at low scattered grazing angles. Also, for the 0° incidence case the SSA is accurate over a larger range of scattering angles than for the 45° incidence case shown in Fig. 3. For the 80° case, the decrease in accuracy of the SSA is evident in comparison to Fig. 3, but still the result is better than that of the KA.

G. Behavior at low forward grazing angles

As mentioned several times in earlier sections, the SSA scattering strength can become somewhat high at low forward grazing angles as the correlation length is increased. However, because this occurs when the curves are sloping steeply downward, it is difficult to detect the difference between them in the plots presented thus far. The region of inaccuracy occurs at very small grazing angles, and the angular resolution of the exact results decreases with grazing angle, which also obscures the difference. To highlight the difference between the integral equation and SSA(4) results, in Fig. 13 we show part of the forward grazing angle region for the case shown in Fig. 3. In addition, to reduce the effects of fluctuations the integral equation result shown is the average of nine different runs, each using 50 surface realizations. Error bars provide an estimate for the uncertainty in

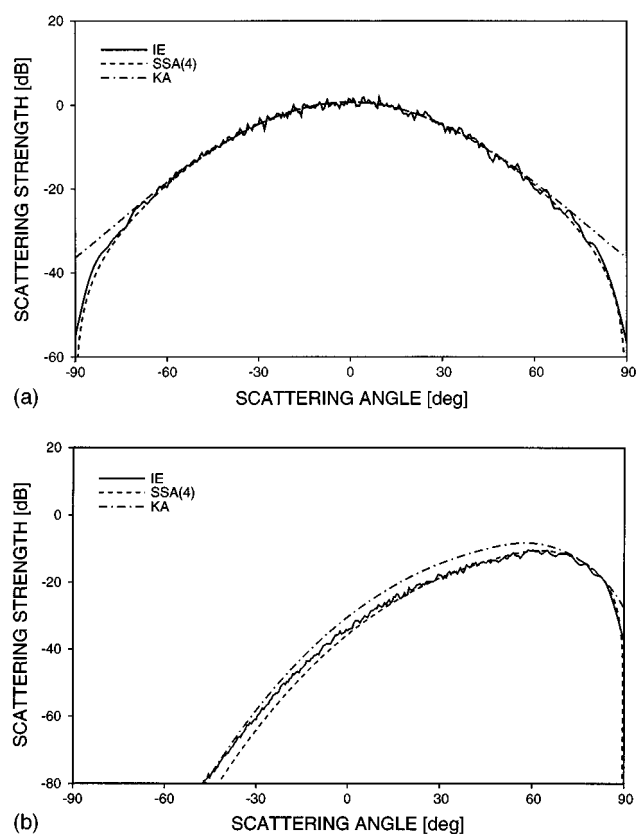


FIG. 12. Comparison of the SSA(4), KA, and IE scattering strengths for the case shown in Fig. 3 but with an incident angle of (a) 0° and (b) 80° . In (b) the coherent component has been subtracted from the IE result to facilitate comparisons of the incoherent scatter.

the Monte Carlo result for the scattering strength based on all 450 surfaces. This estimate is obtained by dividing the standard deviation of the nine 50-surface averages by the square root of nine. In this example, the SSA(4) overpredicts the forward-scattering strength by 1–1.5 dB in this very low grazing region. The crossover, which occurs at about 88.6° in Fig. 13 with the integral equation curve rising above the

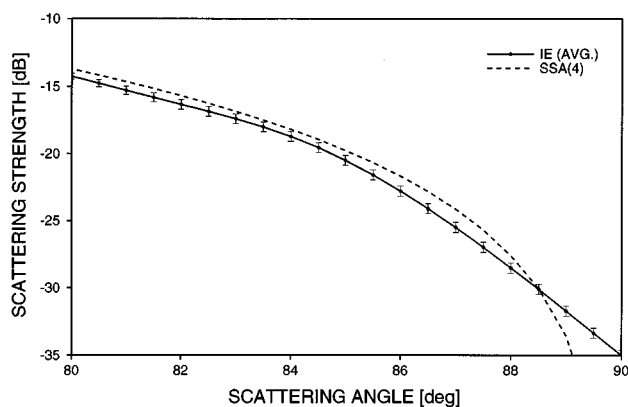


FIG. 13. The forward grazing region for the case shown in Fig. 3. The IE result is the average of nine different runs, each using 50 surface realizations. Error bars provide an estimate for the uncertainty in the Monte Carlo result for the scattering strength based on all 450 surfaces. This estimate is obtained by dividing the standard deviation of the nine 50-surface averages by the square root of nine. The SSA(4) result becomes high at very low forward grazing angles, exceeding 1 dB at approximately 87° .

SSA(4) curve as 90° is approached, is due in part (and perhaps entirely) to the reduced angular resolution of the integral equation result near 90° . Stated another way, the integral equation result is found with surfaces of finite length, while the SSA(4) curve is for infinite length surfaces. Scattering strengths for these two types of surfaces differ significantly only within a few degrees of grazing, with the finite surface lengths yielding higher scattering strengths in this region. Thus, it is consistent with our results for the SSA(4) to have a positive error all the way to a scattering angle of 90° (0° grazing). While the finite angular resolution of the integral equation result affects the comparison with the SSA(4) at very low grazing angles, studies with longer surface lengths which give improved angular resolution indicate that our comparison correctly shows the scattering angle at which the error first exceeds 1 dB as 90° is approached. We have examined this behavior for a large number of cases and find a 1–2 dB overprediction at very low grazing angles to be typical of the SSA(4).

We find that the behavior of the low grazing angle error depends on both the correlation length and the rms slope. In the region when perturbation theory is accurate, the error does not occur. As the correlation length is increased, the inaccuracy occurs earlier for steeper slopes. Furthermore, the inaccuracy begins at higher grazing angles for steeper slopes. For example, for a slope angle of 25° , the inaccuracy begins at a scattering angle of about 85.5° ; whereas for a slope angle of 15° it begins at about 87° . We suspect that this low grazing angle error is caused by nonlocal multiple scattering that is not adequately treated by the SSA.

III. SUMMARY

Expressions for the second-, third-, and fourth-order small slope bistatic scattering cross sections have been presented together with extensive numerical results for scalar wave scattering from 1-D surfaces with Gaussian statistics and a Gaussian roughness spectrum. These results were compared with exact numerical results as well as with the results of other approximate techniques.

The initial appeal of the SSA was its formulation as a systematic series, its reduction to perturbation theory which is inherent in the derivation, and its reduction to the Kirchhoff approximation under appropriate conditions when the T matrix is found to second order in slope. Also, the theory is manifestly reciprocal at each order and, at low orders, the formally averaged expressions for the scattering cross section are numerically tractable. In this numerical study we have shown that the appeal of the SSA extends further. The fourth-order SSA cross section is remarkably accurate for surfaces with rms slope angles up to about 30° . Even with an rms surface slope angle of 45° , the fourth-order result is accurate over a wide range of scattering angles. As the surface correlation length is increased, the SSA results become low in the back direction and slightly high at forward grazing angles less than 5° . Finally, although the fourth-order SSA

cross section analytically reduces to the geometrical optics result in the high-frequency limit, we found in our numerical studies that the reduction is nonuniform in the scattering angle and, at low forward grazing angles, occurs very slowly. Comparison with exact results shows that the SSA gives better results than geometrical optics with shadowing in this region.

Because the SSA reduces to low-order perturbation theory when the latter is accurate and to the Kirchhoff approximation in the specular region, we expect it might be good for multiscale surfaces. In fact, preliminary results show that for a Pierson–Moskowitz spectrum the SSA gives accurate results even in the back direction (Broschat and Thorsos, 1995). This is the subject of further study.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research, Code 321OA. The authors wish to acknowledge helpful comments made by the reviewers.

- Berman, D. H. (1991). "Simulations of rough interface scattering," *J. Acoust. Soc. Am.* **89**, 623–636. Appendix A in this reference shows that $\langle T(k_{sx}, k_{ix}) \rangle$ vanishes unless $k_{sx} = k_{ix}$ and that $\langle T(k_{sx}, k_{ix}) T(k'_{sx}, k'_{ix})^* \rangle$ vanishes unless $(k_{ix} = k'_{ix})$. It does not show that $\langle T(k_{sx}, k_{ix}) \rangle$ and $\langle T(k_{sx}, k_{ix}) T(k'_{sx}, k'_{ix})^* \rangle$ must be proportional to $\delta(k_{sx} - k_{ix})$ and $\delta(k_{ix} - k'_{ix})$, respectively, since other distributions were not ruled out; these properties are needed to obtain (3). An extension of the arguments given in Berman's Appendix A can be used to show these moments of T must be proportional to delta functions, but the analysis is not presented here.
- Broschat, S. L. (1993). "The small slope approximation reflection coefficient for scattering from a 'Pierson-Moskowitz' sea surface," *IEEE Trans. Geosci. Remote Sens.* **31**, 1112–1114.
- Broschat, S. L., and Thorsos, E. I. (1995). "Numerical studies of the small slope approximation for rough surface scattering using a Pierson–Moskowitz spectrum," *J. Acoust. Soc. Am.* **97**, 3404.
- Ishimaru, A. (1978). *Wave Propagation and Scattering in Random Media* (Academic, New York), Vol. II, Chap. 21.
- Parzen, E. (1962). *Stochastic Processes* (Holden-Day, San Francisco), pp. 88–97.
- Thorsos, E. I. (1988). "The validity of the Kirchhoff approximation for rough surface scattering using a Gaussian roughness spectrum," *J. Acoust. Soc. Am.* **83**, 78–92.
- Thorsos, E. I. (1990). "Acoustic scattering from a 'Pierson–Moskowitz' sea surface," *J. Acoust. Soc. Am.* **88**, 335–349.
- Thorsos, E. I., and Broschat, S. L. (1995). "An investigation of the small slope approximation for scattering from rough surfaces. Part I. Theory," *J. Acoust. Soc. Am.* **97**, 2082–2093.
- Thorsos, E. I., and Jackson, D. R. (1989). "The validity of the perturbation approximation for rough surface scattering using a Gaussian roughness spectrum," *J. Acoust. Soc. Am.* **86**, 261–277.
- Voronovich, A. G. (1985). "Small-slope approximation in wave scattering by rough surfaces," *Sov. Phys. JETP* **62**, 65–70.
- Voronovich, A. G. (1986). "A unified description of wave scattering at boundaries with large and small scale roughness," *Progress in Underwater Acoustics*, edited by Harold M. Merklinger (Plenum, New York).
- Wagner, R. J. (1967). "Shadowing of randomly rough surfaces," *J. Acoust. Soc. Am.* **41**, 138–147.
- Yang, T. Q., and Broschat, S. L. (1992). "A comparison of scattering model results for two-dimensional randomly rough surfaces," *IEEE Trans. Antennas Propag.* **40**, 1505–1512.
- Yang, T. Q., and Broschat, S. L. (1994). "Acoustic scattering from a fluid–elastic solid interface using the small slope approximation," *J. Acoust. Soc. Am.* **96**, 1796–1804.

The detection of tethered and rising bubbles using multiple acoustic techniques

Timothy G. Leighton, David G. Ramble, and Andy D. Phelps

Institute of Sound and Vibration Research, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom

(Received 22 January 1996; accepted for publication 10 November 1996)

There exists a range of acoustic techniques for characterizing bubble populations within liquids. Each technique has limitations, and complete characterization of a population requires the sequential or simultaneous use of several, so that the limitations of each find compensation in the others. Here, nine techniques are deployed using one experimental rig, and compared to determine how accurately and rapidly they can characterize given bubble populations. These are, specifically (i) two stationary bubbles attached to a wire; and (ii) injected, rising bubbles. © 1997 Acoustical Society of America. [S0001-4966(97)00905-3]

PACS numbers: 43.30.Jx, 43.30.Pc, 43.25.Yw, 43.35.Ei [JHM]

INTRODUCTION

Bubble detection is required for many industrial,¹ medical,² and environmental³ applications.^{4,5} Throughout the range of acoustic techniques by which this can be achieved, there are inherent limitations. If, for example, a signal^{6,7} is capable of interpretation in terms of assigning homogeneous bulk properties to the “bubbly liquid” as a whole, then such interpretation may be limited to relatively high, relatively spatially uniform, bubble population densities, where the interbubble spacing is very much less than the acoustic wavelength. In contrast, other signals may be practicable only at low number densities.^{8,9} Several are prone to false triggering, in that some other object (e.g., a solid body, or a cluster of small bubbles)¹⁰ may give the same signal as that obtained from a given bubble.

In water with ambient pressure p_0 , an air bubble of radius $R_0 > \sim 10 \mu\text{m}$ has a well-defined resonance frequency $f_0 = \omega_0/2\pi \approx 0.01(\sqrt{p_0})/R_0$, and pulsates as a lightly damped oscillator: On entrainment the pulsations generate an acoustic “signature,” an exponentially decaying sinusoid, the frequency of which indicates the bubble size.^{11,12} A few milliseconds after entrainment these passive emissions have decayed to below the level of the noise. However, the bubble may still be driven, and active acoustic techniques exploit this acoustic resonance^{6,7,13,14} through measurements of sound speed, attenuation, scattering, etc. In such procedures, at a particular frequency the acoustic response of a bubbly liquid is taken to be dominated by bubbles which are resonant with that frequency. The maximum number of different bubble sizes that can be measured at any one time is determined by the number of different frequencies investigated, which historically is usually one, but in notable cases has been four¹³ or around nine.⁷ However, simple linear theory demonstrates that the acoustic scattering cross section of the fundamental frequency is only a local, and not a global, maximum at resonance.¹⁵ Bubbles very much larger than resonance size can geometrically scatter sound to a greater degree than can smaller, resonant bubbles. If an ultrasonic interrogating signal is employed, the frequency of which is very much higher than the resonances of any bubbles in the

sample, geometric scattering can detect bubbles.^{16–18} If MHz sound is, for example, employed to detect mm-sized bubbles, the small wavelengths involved ($\approx 0.4 \text{ mm}$ in water at 3.5 MHz) allow the bubble to be located, but do not accurately give the bubble size. Geometric, nonresonant scattering relies on the acoustic impedance mismatch between the inhomogeneity and the surrounding liquid. It is therefore insensitive to the nature of the inhomogeneity, and in practice may not distinguish between bubbles and solid bodies of a similar size.

A bubble in an acoustic pressure field $P = A \cos \omega_p t$ tends to linear, low-amplitude oscillations if the driving amplitude A is small, or if the bubble is far from resonance. However, as the bubble pulsations become larger (for example, at resonance) the inherent nonlinearities in the motion become more pronounced, and manifest in the scattered acoustic signal as harmonics of the driving frequency. For example, a quadratic nonlinearity (i.e., a system response $\propto P^2$) will generate harmonics at $2\omega_p$; higher-order nonlinearities give commensurate harmonics. This has been used to detect bubbles of specific size, resonant at 0.89 MHz in one experiment¹⁹ and at both this and at 1.64 MHz in another.²⁰ If such systems are to be perfect bubble detectors then the condition must hold that only resonant bubbles can generate the required nonlinearity, and in the presence of only nonresonant bubbles, ω_p alone is detected. However, while the emission of the second harmonic is a global maximum at resonance, the $2\omega_p$ signal can arise through nonbubble sources of nonlinearity, such as signal distortion in the equipment, which must be carefully examined. Such sources do not include solid inhomogeneities. The same condition holds if the applied field contains two frequencies, i.e., $P = A \cos \omega_p t + B \cos \omega_i t$ where $\omega_p \ll \omega_i$. The “imaging” frequency (ω_i) scatters geometrically from a target (the pulsating bubble) whose cross-sectional area varies periodically.⁹ The detected signal consists of ω_i , modulated at frequency ω_p , and the resulting detection of $\omega_i \pm \omega_p$ in the received spectrum has been used to size a bubble spectrum by employing the assumption that, bar the presence of resonant bubbles, only ω_i and ω_p are detected.^{8,21} The as-

TABLE I. The various acoustic techniques available for bubble detection. Numerals in columns 4 and 5 are references.

Scatters	Advantage	Disadvantage	Prior application	Bubble sizes investigated in a single expt.
Geometric	Rapidly obtains images with high-spatial (location) resolution.	Cannot distinguish between bubbles and solid particles.	Laboratory ^{16,18,23}	Distribution (low-radius resolution)
Fundamental	Apparatus simple.	Large bubbles and bubble clouds may falsely register as resonant bubble (geometric scattering). Low-spatial resolution. False triggering and off-resonance scattering may occur. High-number densities only are valid if "bulk properties" are assigned to the liquid.	Resonator ^{7,22} Attenuation ⁶ Backscatter ¹⁴	Four; ¹³ around nine ^{7,22}
Second harmonic	Little contribution from geometric scattering.	Low-spatial resolution. False triggering and off-resonance scattering may occur.	Clinical, detecting $\approx \mu\text{m}$ radius bubbles ^{19,20}	One ¹⁹ or two ²⁰ per trial
$\omega_i \pm \omega_p$	No threshold.	False triggering and off-resonance scattering may occur.	Lab., ^{8,21} field ²⁴	Distribution
$\omega_i \pm \omega_p/2$	Minimal false triggering or, at threshold, off-resonance scattering.	Insonation at the threshold acoustic pressure is required for fine radius resolution.	Laboratory ^{9,25-27}	One at 25 Hz resolution ⁹

sumption fails if the pulsation of nonresonant bubbles, or the presence of a quadratic nonlinearity anywhere in the system (for example, through turbulent water motion), is sufficient to generate an $\omega_i \pm \omega_p$ signal. One advantage of combination-frequency methods is that the bubble resonance generates a signal in the MHz range (close to ω_i), removing it from "masking" signals such as the acoustic input and ambient noise.

All the above techniques for bubble sizing which exploit the bubble resonance suffer in that sources other than resonant bubbles (e.g., turbulence, transducer effects, etc.) can to a greater or lesser extent generate the desired signal, indicating the presence of a resonant bubble when one is not present.⁴ Such "false triggering" has not to date been found when signals at $\omega_i \pm \omega_p/2$ are used for bubble sizing.⁹ These signals are generated when the amplitude component A of the insonating field $P = A \cos \omega_p t + B \cos \omega_i t$ exceeds the threshold value required to generate Faraday waves on the bubble surface. Characteristics of the various acoustic sizing techniques are summarized in Table I.

The less prone a system is to "false triggering," the more complicated in general it is to deploy. It therefore would be desirable to be able to deploy a range of these techniques to interrogate a given liquid sample, either sequentially or concurrently as defined by the problem. This would enable optimization of the process of characterizing the bubble population in the liquid with respect to minimizing the ambiguity of the result and the complexity of the task. The task itself involves first the detection of inhomogeneities in liquids. In certain circumstances it is then necessary to analyze the sample further to distinguish gas bubbles from solid or immiscible liquid-phase inclusions. The final stage of analysis would involve not only the detection, but also the sizing of the gas inclusions, leading to the characterization of the bubble population. This can be summarized in a four-part *ideal objective*:²⁸ (i) Detect inhomogeneities in liquids; (ii) Distinguish gas bubbles from solids; (iii) Measure radii of bubbles present; (iv) Measure number of bubbles in each radius class.

This study introduces a method by which the *ideal objective* might eventually be achieved, using a range of tech-

niques. The limitations of each can be compensated through the deployment of others. Since the ambiguities of each have been studied theoretically and experimentally,¹⁵ the initial emphasis of this study will be how successfully each technique can provide information about simple controlled populations (stationary single and paired bubbles). A rising bubble stream will then be measured. The techniques listed in Table I are used, so that bubble detection is achieved through the geometric scattering of 3.5-MHz ultrasound (using a scanner in both B and M modes simultaneously), and through scattering of signals at ω_p , $2\omega_p$, $\omega_p/2$, $\omega_i \pm \omega_p$, $\omega_i \pm 2\omega_p$, $\omega_i \pm \omega_p/2$, and $\omega_i \pm 3\omega_p/2$. This is done for broadband, and increasing, incremented, tonal "pump" signals. The study was carried out using relatively low-amplitude acoustic fields to drive the bubble, which is desirable to minimize the invasiveness of the technique.

I. METHOD

There exist detection zones, at 15-cm depth, for the various active acoustic sizing systems (including those listed in Table I), comprising the overlap of beam patterns of relevant transducers held in rigid "cage" configuration (Fig. 1). The cage is placed at depth 0.15 m in a $1.8 \times 1.2 \times 1.2$ -m deep vibration isolated glass-reinforced plastic tank. The bubble population is either injected into the tank below these zones, and then rises to pass through them; or consists of bubbles attached to a wire, held within the intersection of the zones. A Gearing and Watson UW60 loudspeaker (having a frequency response flat to within ± 5 dB over the range 500 Hz–10 kHz) is used to generate the required "pump" signal. This signal drives the bubbles into oscillation, and it may be broadband, or a series of tones $P = A \cos \omega_p t$, where ω_p is incremented in 50 Hz (tethered bubbles) or 100 Hz (moving bubbles) steps.

During combination-frequency tests the imaging signal $P = B \cos \omega_i t$ is generated by a Therasonic 1030 (Electro-Medical Supplies) fixed at 1.134 MHz. A Panametrics V302 receiver detects the MHz signal before it is heterodyned with the Therasonic signal. The Bruel & Kjaer 8103 hydrophone ("HP1") is used to detect signals not involving combination

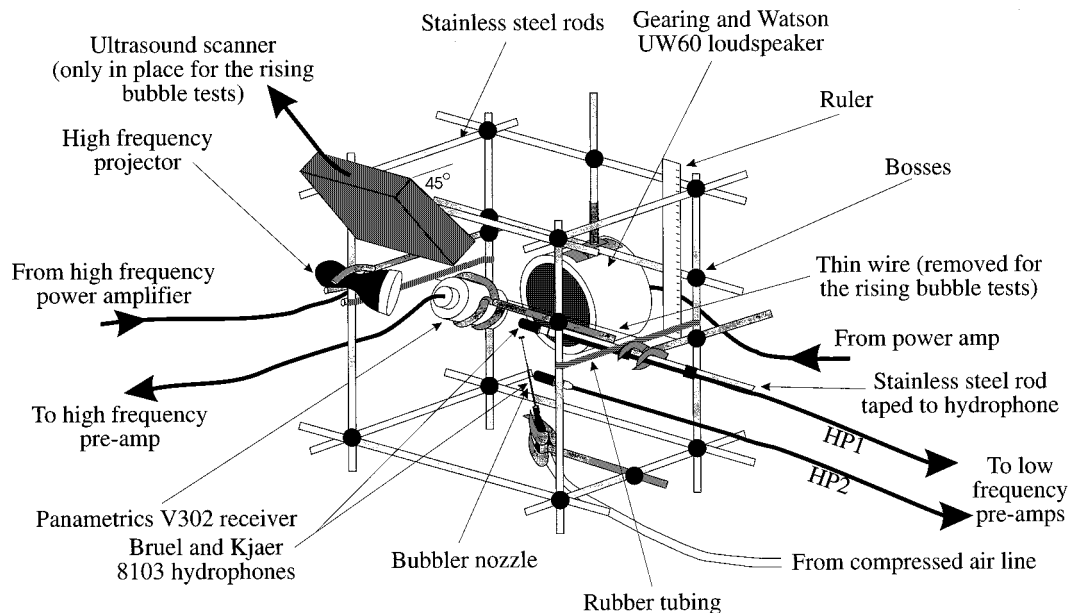


FIG. 1. Schematic of apparatus mounted in cage. For tethered bubble tests the ultrasound scanner is removed. For rising bubble tests the thin wire in the cage center is removed.

frequencies. The heterodyned high-frequency signal and the B&K 8103 signals are acquired to the PC via a general purpose interface bus (GPIB)-controlled Digital Storage Oscilloscope (LeCroy 9314L). Calibration is made with no bubbles present to allow compensation for the acoustic response of the water, apparatus, and tank. This enables the sample to be insonated at equal amplitudes when interrogated by a sequence of tonal pumping signals, each of 0.2-s duration. Data is only collected after a “start-up” time of the first 7.5 ms for tethered bubbles, to allow transients to die away. No such delay can be afforded with rapidly rising mm-sized bubbles, though averaging over the 10^4 samples of each increment reduces the transient effect. Including data collection, the individual incremented tones start 1.6-s apart.

The rising bubbles are injected from a needle attached to a compressed air line. The passive acoustic signal so generated is detected by “HP2,” a hydrophone (Bruel & Kjaer 8103) placed 10 mm from the needle tip at a depth of 29 cm. This signal is analyzed for the exponentially decaying sinusoid “signatures” which indicate the generation of each bubble, the frequency of the sinusoid revealing the bubble size. However, with higher entrainment rates (where signatures overlap) in noisy environments, individual entrainments may not be detected even in time-frequency representation (TFR), where resolution in time and frequency is a compromise determined by the size of the window imposed upon the data. However, a TFR of the Gabor coefficients, rather than the acoustic power invested in each frequency band, will readily identify the bubble signatures.^{29–31} A routine uses thresholds on the value and gradient of the Gabor coefficients, then automatically counts and sizes the bubbles, giving their rate of production before they rise into the active detection zones. A second count is made by placing a greased Petri dish in the rising bubble stream above the detection zones. Photographic measurement of the bubbles adhering to the thin layer of silicone grease were taken. How-

ever, compensation must be made in comparing the bubble population measured in given volumes of liquid by the active techniques, with the captured population on the dish and the rate of bubble generation measured at the needle by the Gabor technique. This is because, for example, the volume of the bubble stream sampled by the Petri dish increases with the bubble rise speed. The volume changes caused by the varying hydrostatic pressure are accounted for in comparing all measurements. The sizes of the two bubbles attached to the wire were checked by detaching them from the wire into small glass flasks, in which they were transferred to a traveling microscope for measurement.²⁵ A Hitachi EUB-26E 3.5-MHz ultrasound scanner, mounted in the cage, gave M- and B-mode images of the rising bubbles. Atmospheric pressure was 0.1003 MPa.

Detection through scattering at ω_p and $\omega_i \pm \omega_p$ requires only linear bubble pulsations, so that the relatively low-energy densities per frequency band afforded by broadband insonation (bandlimited white noise between 1000–8000 Hz) is appropriate. This rapidly allows an estimate of the region wherein the bubble resonances lie, for later application of the nonlinear detection signals ($2\omega_p$, $\omega_p/2$, $\omega_i \pm \omega_p/2$, $\omega_i \pm 2\omega_p$, $\omega_i \pm 3\omega_p/2$). These nonlinear signals require an incremented pure-tone pump signal, rather than broadband insonation, for two reasons. First, it is necessary to drive at a sufficiently high amplitude to generate detectable nonlinearities. Second, the detector frequency emitted by a bubble differs from that which drives it at resonance, which would cause ambiguity if broadband excitation were employed.

II. RESULTS

A. Two tethered bubbles

The first of the results are shown in Fig. 2 for the broadband excitation of two bubbles attached to a wire 10-mm apart. Throughout the paper a dashed line indicates signal

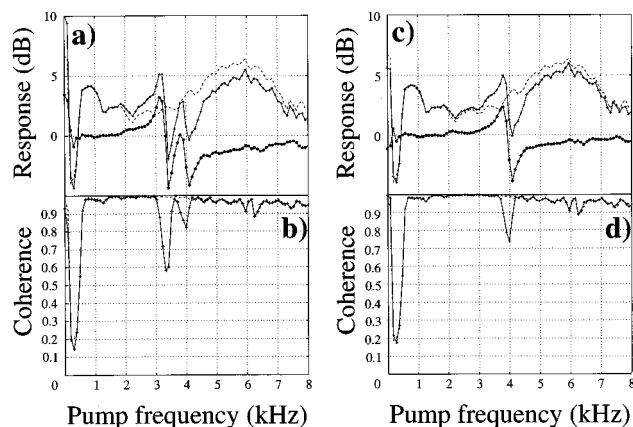


FIG. 2. Response [modulus of voltage transfer function, plots (a) and (c)] and coherence [(b) and (d)] for broadband insonation (band limited 1–8 kHz) of both [(a) and (b)] tethered bubbles, and for just the smaller [(c) and (d)]: dashed line=“in absence of bubbles”; —+—+—=“in presence of bubbles”; ●—●—●=ratio of “bubble present” to “bubbles absent” signals. Resolution: 98 Hz.

with no bubbles present; a solid line with crosses indicates the signal in presence of bubble(s); and a thick solid line with closed circles indicates the ratio of the signal “with” bubbles to that “without” bubbles, i.e., the bubble-mediated amplification. Data points occur at symbols, and at equivalent frequencies for dashed lines. Figure 2(a) illustrates the difference in the modulus of the voltage transfer function (the ratio of output to input) when the bubbles were driven by bandlimited (1–8 kHz) white noise. The response shows peaks at 3.1 and 3.9 kHz (± 0.1 kHz), with a sharp dip ~ 300 Hz above each. This reflects the through-resonance behavior of each bubble: At frequencies just below resonance the sound field and the bubble pulsations (which scatter significantly more than they do away from resonance) are in phase and constructively interfere. However, above resonance the bubble undergoes a π phase shift such that it now pulsates in antiphase with the driving sound field, resulting in destructive interference. This behavior suggests that the change in signal which results from bubble presence does not represent geometric scattering from a large bubble or other body, but is due to the presence of resonant bubbles in that frequency range. Even several kHz above the resonance of the pair, the detected signal is ~ 1 dB less than the levels at low frequencies, and those found in the absence of bubbles. This is a result of the destructive interference caused by the whole population, and it may be that this can be used to characterize a population [compare this reduction with the smaller one seen in Fig. 2(c) for one of these bubbles on its own]. The coherence between the signal input to the source and the returned signal [Fig. 2(b)] shows a definite bubble-mediated reduction in the signal around 3.3 ± 0.15 and 4 ± 0.15 kHz. As these coherence dips appear at frequencies midway between the peaks and troughs in the transfer function [Fig. 2(a)] they appear to indicate a bubble nonlinearity rather than a reduced signal to noise ratio, which would be the case if the dips in Fig. 2(a) and (b) occurred at the same frequency.

Figure 2(c) and (d) show the transfer function and co-

herence resulting from broadband excitation when the bubble resonant at ~ 3.3 kHz is removed after completion of the two-bubble tests. Its peak disappears [Fig. 2(c)]. The other peak remains at 3.9 ± 0.1 kHz, suggesting that to within this resolution the bubbles were far enough apart (approximately 10 bubble radii) for the bigger bubble not to significantly influence the resonance frequency of the other.³² The peak is about 3 dB higher than in the two bubble test even though the same excitation amplitude was used. This is due to the removal of the antiphase bubble pulsation of the larger bubble beyond its resonance, which therefore means that there is no destructively interfering component on the smaller bubble’s pulsation below its resonance. The coherence again shows a similar dip to the relevant one found in the two bubble test [Fig. 2(d)].

Having, through 1 s (5 averages) of broadband insonation of the two bubbles, reduced the range of interest for further investigation from 1–8 kHz to 2.7–4.7 kHz, the bubble pair was excited (with pump amplitude 120 Pa) at 40 discrete increasing frequencies in 50-Hz increments: At 1.6 s per increment, the test took 64 s. The results are given in Fig. 3 for the harmonic [parts (a)–(c)] and sum-and-difference [parts (d)–(f)] signals, monitored simultaneously. The data is displayed as the magnitude in the frequency domain at the location of the signal of interest (i.e., at ω_p , $2\omega_p$, $\omega_p/2$, $\omega_i \pm \omega_p$, $\omega_i \pm 2\omega_p$, and $\omega_i \pm \omega_p/2$) corresponding to each pump frequency. The data was sampled at 50 kHz, and the FFT frequency resolution with 8192 points was 6 Hz. The test was repeated following the removal of the larger bubble (Fig. 4).

The fundamental backscatter [Fig. 3(a)] shows a rippled amplitude response in the absence of a bubble, which is due to the differences in the proximity of each pumping signal tone to an FFT bin center frequency. This effect disappears when the dB difference (“amplification”) between the signal with, and without, bubbles is taken, revealing again the characteristic through-resonance response indicating the presence of resonant bubbles at 3325 ± 70 and 3900 ± 100 Hz. The response of the second harmonic [Fig. 3(b)] is less clear. The height of the signal in the absence of the bubble can be affected for instance by the relative levels of harmonic distortion in the equipment and also the proximity of the signal to a frequency bin. Nevertheless, there still appears to be a clear increase in the signal between 3200–3400 and 3800–4100 Hz. Removal of the larger bubble has negligible effect in the peaks in the first harmonic and second harmonic response for the smaller bubble as shown in Fig. 4(a) and (b). The emissions of $\omega_p/2$ from both two bubbles [Fig. 3(c)] and the smaller one [Fig. 4(c)] are too small to differentiate from the noise floor. The amplitude of the heterodyned returned signal from the high-frequency receiver at $\omega_i \pm \omega_p$, $\omega_i \pm 2\omega_p$, and $\omega_i \pm \omega_p/2$ are shown in Fig. 3(d)–(f) as a function of the incrementing pumping frequency ω_p . Though there are maxima at 3.25 ± 0.05 and 3.9 ± 0.2 kHz, the signal at $\omega_i \pm \omega_p$ [Fig. 3(d)] is present at more than 12 dB above the “no bubble” signal over the entire pumping frequency range. Clearly, the off-resonance contribution to the returned signal limits the resolution of the measurement for the bubble’s resonance frequency. Though the off-resonance contri-

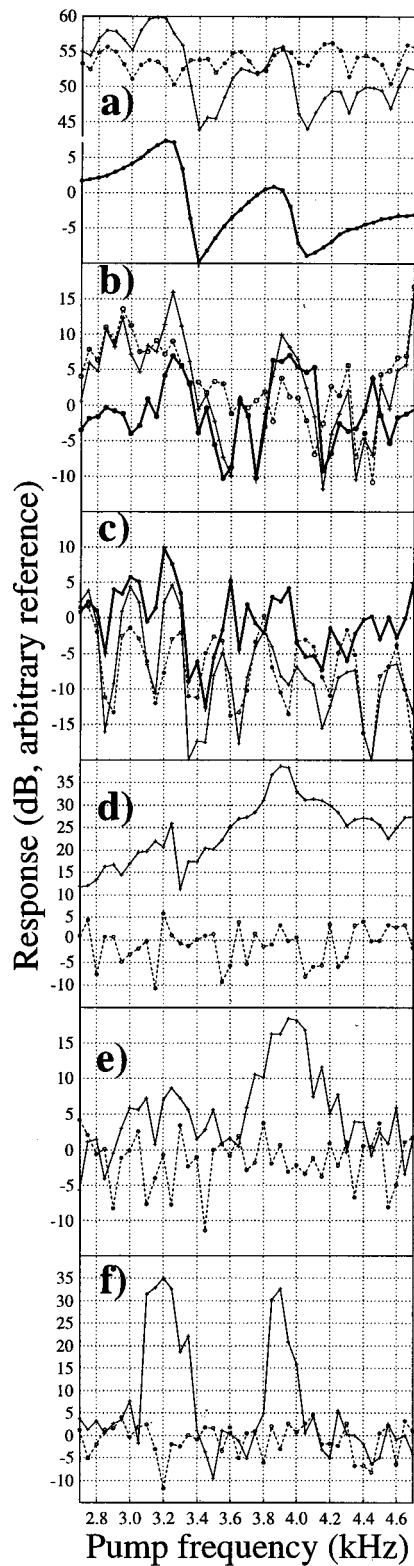


FIG. 3. The HP1 signals for the two-bubble tests (50-Hz increments) showing (a) ω_p , (b) $2\omega_p$, (c) $\omega_i \pm \omega_p/2$, (d) $\omega_i \pm \omega_p$, (e) $\omega_i \pm 2\omega_p$, (f) $\omega_i \pm \omega_p/2$. Key as for Fig. 2, with open circles showing data points on dashed line.

tribution is less for $\omega_i \pm 2\omega_p$ [Fig. 3(e)] the resolution of the high-frequency peak is similarly poor (4 ± 0.2 kHz), and there are spurious maxima. It is clear that the $\omega_i \pm \omega_p/2$ [Fig. 3(f)] signal best shows the presence of two bubbles, resonat-

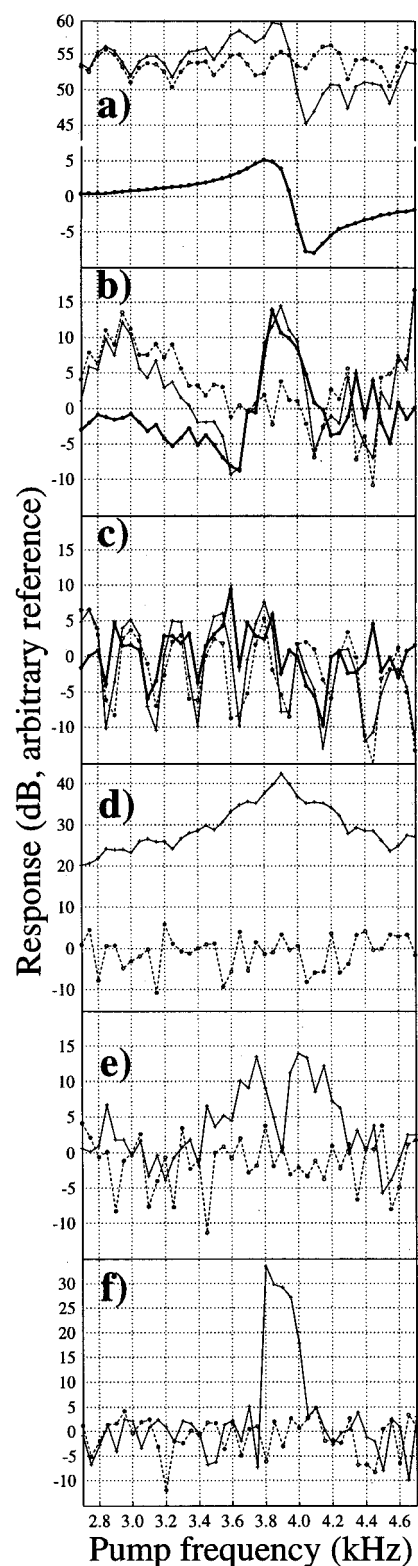


FIG. 4. As for Fig. 3, but for the smaller bubble only. Key as for Fig. 3.

ing at 3.2 ± 0.1 and 3.88 ± 0.05 kHz. The off-resonance contributions are negligible. Removal of the larger bubble demonstrates the same features in the detection of the remaining bubble (Fig. 4) by the (d) $\omega_i \pm \omega_p$, (e) $\omega_i \pm 2\omega_p$, and (f) $\omega_i \pm \omega_p/2$ signals.

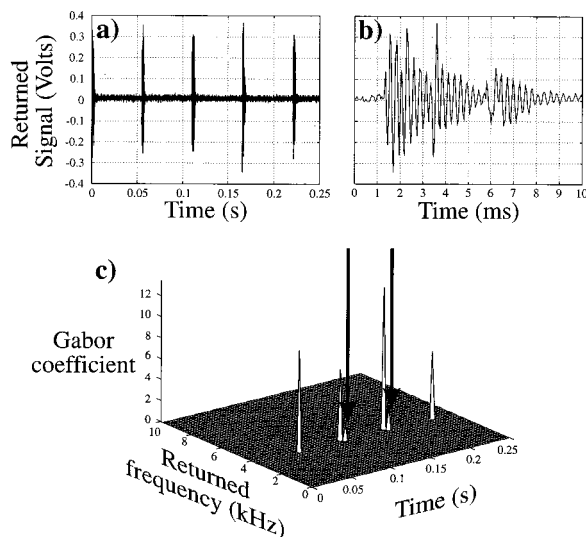


FIG. 5. The HP2 signal during injection. (a) Time series [detail shown in (b)]. (c) Time-frequency representation of Gabor coefficients associated with (a) (first peak removed for clarity). Where multiple coefficients are identified with injection of a single bubble, the later one (arrowed) gives natural frequency.

B. Rising bubbles

Figure 5 shows a portion of the bubble stream as measured through the passive acoustic emissions generated on injection. In Fig. 5(a) a 0.25-s section of the time series recorded by the hydrophone HP2 indicates individual bubbles being repeatably generated every ~ 0.06 s. Each of the bubble signatures has the form, not of a single exponentially decaying transient, but of multiple ones, revealing that the released bubble is excited on three subsequent occasions following the initial release from the needle [Fig. 5(b)]. These excitations arise through contact, and usually coalescence, between the newly released bubble and the successor gas pocket growing at the nozzle tip.³³ As a result, the plot of the Gabor coefficients [Fig. 5(c)] may reveal multiple peaks for a single bubble (which vary each time, showing the nozzle process is not entirely repeatable). Clearly the frequency at which the final peak of each group occurs [arrowed in Fig. 5(c)] is the one which relates to the size of the final bubble after it has escaped clear of the contact/coalescent processes that occur at the nozzle. It is this size which is taken to be a measure of the bubble size upon injection.

In Fig. 6 the results of broadband insonation in the frequency range 1–8 kHz is shown. In Fig. 6(a), the signal picked up by HP1 is shown, both for the situation before the bubble stream began, and for the scattered signal in the presence of the bubble stream. The difference between the two signals is plotted, showing significant changes in the frequency range 3.5–5 kHz, indicating the through-resonance effect described above, centered around 4 ± 0.1 kHz. In Fig. 6(b), the heterodyned signal from the high-frequency receiver transducer shows bubble-mediated change from 3.5 to 4.9 kHz (centered at 4.2 ± 0.3 kHz). An 800-Hz high-pass filter was placed after the heterodyning so that the strong Doppler components of the returned signal did not overload the input channel to the oscilloscope.

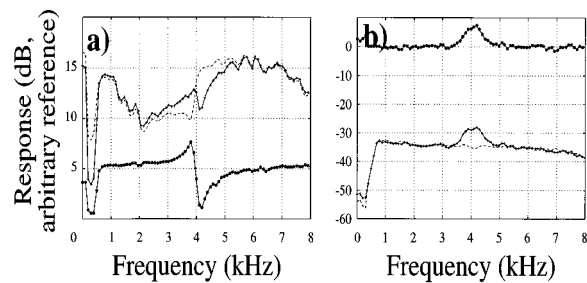


FIG. 6. Response (modulus of voltage transfer) for broadband insonation (bandlimited 1–8 kHz) of rising bubbles, from (a) HP1, and (b) heterodyned high-frequency (from V302) signals. Resolution: 98 Hz. Key as for Fig. 2.

Having rapidly found the region of interest (3.3–4.3 kHz) through the broadband technique, the pump sound field is incremented in this range in steps of 100 Hz, at a pressure amplitude of 240 Pa (0-pk). Figure 7 shows the results of analysis of the signal recorded by hydrophone HP1. In Fig. 7(a), the scattering of the fundamental frequency ω_p gives $f_0 \approx 3850 \pm 20$ Hz. The second harmonic $2\omega_p$ neither immediately indicates a distribution around a single bubble size [Fig. 7(b)], nor accurately indicates what the size might be ($f_0 \approx 3.9 \pm 0.2$ kHz). The $\omega_p/2$ results are similarly unclear [Fig. 7(c)]. During the same single pass from 3.2 to 4.4 kHz as was made for Fig. 7, were taken the results for Fig. 8, a histogram showing the received, heterodyned spectrum as a function of the pump frequency (this, on the horizontal axis,

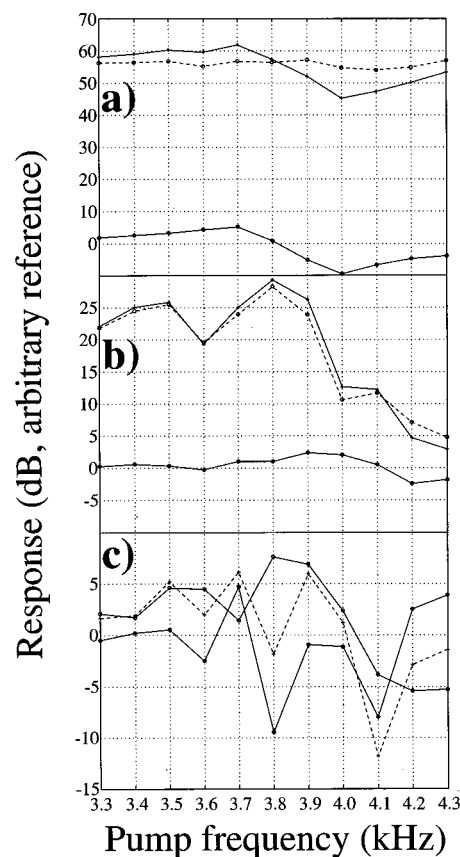


FIG. 7. Response at (a) ω_p , (b) $2\omega_p$, (c) $\omega_p/2$ in the HP1 signal for insonation in 100-Hz increments. Key as for Fig. 3.

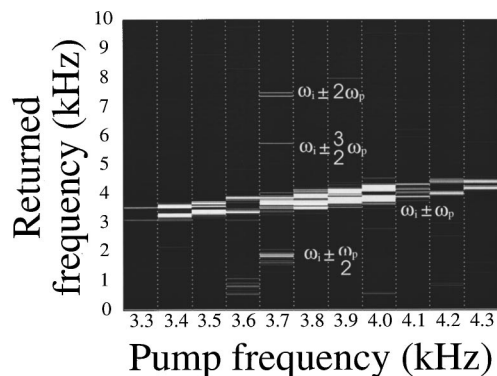


FIG. 8. Greyscale histogram showing heterodyned received signal (from V302) for each discrete setting of the pump frequency (100-Hz increments). Light shades indicate strong signal. Signals at $\omega_i \pm \omega_p/2$, $\omega_i \pm \omega_p$, $\omega_i \pm 3\omega_p/2$, and $\omega_i \pm 2\omega_p$ are indicated.

indicating not a continuum but the 12 settings of the pump frequency, since the latter was incremented in 100-Hz steps).

The clearest indication of resonance is that only for the pump frequency setting of 3.7 kHz does structure in the heterodyned spectrum at frequencies which are multiples of $\omega_p/2$ (corresponding to $\omega_p/2$, ω_p , $3\omega_p/2$, and $2\omega_p$) occur. All other peaks do not correspond to multiples of ω_p . Figure 9 shows both the (a) M- and (b) B-mode images obtained using the Hitachi ultrasound scanner, the section shown being a slice at 45° to vertical (Fig. 1). The bubble (labeled B) can be located in Fig. 9(b) (near field is at top of image), which also images the loudspeaker (S) and part of the cage. The images which intersect the vertical line (L) in 1 s are plotted in Fig. 9(a): Almost 19 bubbles pass through the beam in that time, with rise speed (from the image, within the limits of the rectilinear bubble motion, adjusting for the

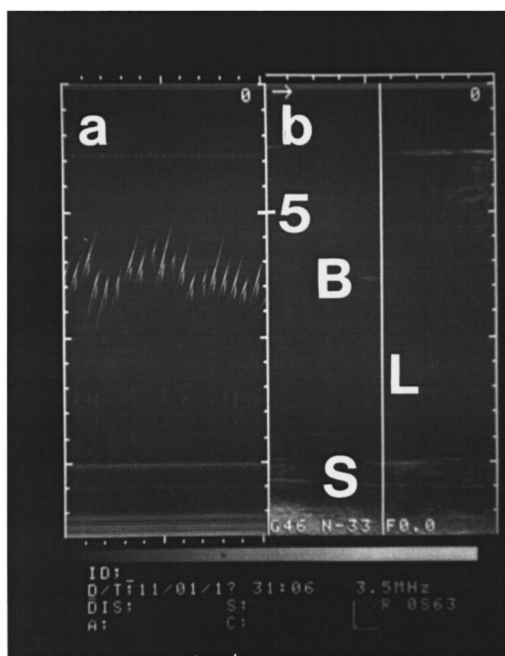


FIG. 9. (a) M-mode (1-s sweep) and (b) B-mode images from Hitachi ultrasound scanner. In (b) a bubble (B), UW60 speaker (S), the 5-cm marker from transducer faceplate (at top of image) and the line (L, occurrence of a target in which defines the M-mode image) are indicated.

45° orientation) of 20 ± 2 cm/s. Comparison of ‘a’ with ‘b’ allows the transient features (e.g., bubbles) to be distinguished from the time-invariant ones (e.g., cage and speaker).

III. DISCUSSION

For the two tethered bubbles, optical measurements gave radius estimates of 1.1 ± 0.1 and 0.8 ± 0.1 mm. Figure 3(f), which plots $\omega_i \pm \omega_p/2$, most clearly indicates the presence of two bubbles. Table II summarizes the information gleaned from each signal type in the two-bubble test. Though no high-resolution technique determines both resonances to the same accuracy, the best overall resolution is obtained from $\omega_i \pm \omega_p/2$ using incremented pump signals. Initial use of broadband first reduced the test time by a factor of 64. The resolution of $\omega_i \pm \omega_p/2$ can be dramatically affected by the acoustic pressure at the bubble: While it could be improved to ± 12 Hz by insonating at the threshold pressure,⁹ there is no guarantee that in the general case this threshold can be accurately delivered.

This is particularly true when considering the results from moving bubbles (Table III), since each bubble is transitory. Also because of this, not only the accuracy but also the population sampling must be considered. In fact, the results in Table III refer to two quite separate populations. First the incremented techniques (while they can be repeated to average a steady-state population)⁹ were here in fact applied in one pass, and so would ideally detect signals only from resonant bubbles which are in the detection zone during the 0.2 s of each tone. Since bubbles are generated at ~ 60 ms intervals, and have a rise time of 20 ± 2 cm/s, all the incremented tests (columns 4–10) sample in each increment the same population of ~ 4 bubbles (different sets of ~ 4 bubbles for each of the 40 increments—‘population 1’). Three minutes later the broadband techniques sample across the entire frequency range for five 0.2-s averages, totalling 1 s: The results in columns 2 and 3 therefore sample a population of ~ 19 bubbles (‘population 2’). Though there are differences in resolution between the broadband and the incremented techniques, the results in Table III indicate that the two populations differed, the one measured first having a lower resonance (3.7 ± 0.05 kHz) than the other (4.0 ± 0.1 kHz). This issue will be discussed later.

Resolution of the ω_p and $\omega_i \pm \omega_p$ signals is roughly constant between broadband and incremented forcing at around 100 and 300–500 Hz, respectively (Table III). The ω_p signal is not pronounced and would readily be confused by a wide range of sizes (see Table II). The resonance is indicated not by the maximum (strong emission almost in phase with driver), but by the in-phase point between the maximum and the minimum (antiphase) point: This has implications for studies where the scattering is assumed to be from resonant bubbles only. Only the simultaneous occurrence of the structure at $\omega_i \pm \omega_p/2$, $\omega_i \pm 3\omega_p/2$, and $\omega_i \pm 2\omega_p$ allows accurate active characterization. It is not surprising that the $\omega_i \pm \omega_p/2$ signal should so clearly indicate the resonance,

TABLE II. Resonances and calculated radii of the two tethered bubbles ($p_0=101\,770$ Pa). References in row 2 are to figures.

	← Broadband pump signal →		← Pump signal (50 Hz increments, frequency increasing) →					
	Amplitude ω_p	Coherence ω_p	ω_p	$2\omega_p$	$\omega_p/2$	$\omega_i \pm \omega_p$	$\omega_i \pm 2\omega_p$	$\omega_i \pm \omega_p/2$
Distribution indicated	(2a) Two bubbles	(2b) Two bubbles	(3a) Two bubbles	Spurious peaks (3b)	Spurious peaks (3c)	(3d) Two bubbles	Unclear peaks (3e)	(3f) Two bubbles
Resonance freq. f_0 /kHz	3.3 ± 0.1 4.0 ± 0.2	3.3 ± 0.15 4.0 ± 0.15	3.33 ± 0.07 3.9 ± 0.1	3.3 ± 0.1 3.95 ± 0.15	3.2 ± 0.1 ...	3.25 ± 0.05 3.9 ± 0.2	... 4 ± 0.2	3.2 ± 0.1 3.88 ± 0.05
$\frac{R_0}{\mu\text{m}} \approx \frac{p_0^{1/2}}{100 f_0}$	970 ± 30 800 ± 40	970 ± 44 800 ± 30	960 ± 20 820 ± 20	970 ± 30 810 ± 30	1000 ± 30 ...	980 ± 15 820 ± 42	... 800 ± 40	1000 ± 30 820 ± 10

whereas the $\omega_p/2$ signal does not, since the surface activity that generates the subharmonic emission cannot itself propagate to distance as it does not involve any bubble volume changes. However, as these Faraday waves change the effective area presented to the imaging beam, they can cause a modulation in the scattered signal, and this signal will propagate to distance.

The question of whether the two populations, measured by broadband and incremented techniques, could possess the distribution difference suggested above must be addressed by reference to the other techniques used for determining the bubble size some minutes after the conclusion of the broadband tests. The ± 2 -cm/s standard deviation on the 20-cm/s rise time translates³⁴ to estimated lower and upper limits for radius in this water of 0.87 and 1.13 mm. Clearly this is not sufficiently discerning. The distribution of rising bubbles from four Petri dish photographs (taken 10 minutes after the end of the passive Gabor tests and corrected for hydrostatic head) gives for the size at 15-cm depth: 790 ± 60 μm (28 bubbles collected in 1.5 s); 790 ± 120 μm (24 bubbles in 1.3 s); 830 ± 80 μm (27 bubbles in 1.4 s); 820 ± 130 μm (32 bubbles in 1.7 s). There is some indication of occasional larger bubbles in a more uniform distribution.

The actual stability of the population is best determined by the Gabor tests. Three of these were performed at one-minute intervals after the broadband tests, and before the ultrasonic images were taken. In each test 0.25 s of passive emissions, comprising the injection emissions of five consecutive bubbles, were taken [Fig. 5(a) represents test 2]. The natural frequencies so found are shown in Table IV, with the average for each test, and the calculated bubble size distribution

at the needle (29-cm depth) and at the zone of the active detector (15-cm depth). Clearly, variation in the size of the generated bubbles can occur. This is not unexpected when compressed air, supplied from a line, is bubbled at rates high enough for interbubble contact/coalescence to occur. Table IV suggests that the variation found during the 1 s of the broadband test, and the 40×1.6 s of the incremented test, is of the same order as the standard deviations quoted in Table III. Clearly for all but the technique with the highest resolution in each population, the standard deviation must represent the resolution limitations of the techniques. For the highest resolution (columns 8–10 for population 1; column 2 for population 2) the uncertainties in Table III are similar to those quoted for these techniques during the two-bubble test (Table II), when the population was stable. This suggests that, here too, the standard deviations reflect limits in resolution. It seems that in fact the best resolution limits in each case are very similar to the variability one might expect in the population. Though by no way conclusive, it is suggestive that the large standard deviations in tests 1 and 3 result from single outlying values. These values could well escape detection in the 0.2-s duration of each incremented tone, and if the item 3190 Hz is removed from test 1 the average becomes 3686 ± 90 Hz (871 ± 21 and 875 ± 21 μm at 29- and 15-cm depth, respectively), and if the item 3219 Hz is eliminated from test 3, the average becomes 4004 ± 30 Hz (giving 802 ± 6 and 806 ± 6 μm at 29- and 15-cm depth, respectively). This variation is less than the resolution limits of Table II and the uncertainties quoted in Table III, for the $\omega_i \pm \omega_p/2$ and related tests.

The Gabor technique for sizing bubbles from their pas-

TABLE III. Resonances and calculated radii of rising bubbles for populations 2 (broadband pump) and 1 (incremented pump).

Column 1	← Broadband →		← Incremental pump, 100-Hz steps, frequency increasing →						
	Population 2		Population 1						
Signal:	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8	Column 9	Column 10
	ω_p	$\omega_i \pm \omega_p$	ω_p	$2\omega_p$	$\omega_p/2$	$\omega_i \pm \omega_p$	$\omega_i \pm \omega_p/2$	$\omega_i \pm 2\omega_p$	$\omega_i \pm 3\omega_p/2$
Distribution indicated	narrow [Fig. 6(a)]	broad [Fig. 6(b)]	narrow [Fig. 7(a)]	bimodal [Fig. 7(b)]	bimodal [Fig. 7(c)]	broad (Fig. 8)	narrow (Fig. 8)	narrow (Fig. 8)	narrow (Fig. 8)
Resonance freq. f_0 /kHz	4 ± 0.1	4.2 ± 0.3	3.85 ± 0.1	3.9 ± 0.2	3.9 ± 0.2	3.8 ± 0.5	3.7 ± 0.05	3.7 ± 0.05	3.7 ± 0.05
$\frac{R_0}{\mu\text{m}} \approx \frac{p_0^{1/2}}{100 f_0}$	800 ± 20	760 ± 50	830 ± 22	840 ± 33	820 ± 42	840 ± 110	862 ± 12	862 ± 12	862 ± 12

TABLE IV. Natural frequencies and calculated average radii from Gabor tests at 29, and 15-cm depths.

Trial	Test 1	Test 2	Test 3
Natural frequencies/Hz	3722	3751	4018
	3737	3699	4015
	3190	3642	3219
	3550	3758	3965
	3736	3835	4021
Average freq./Hz	3580	3740	3850
	± 240	± 70	± 350
$R_0/\mu\text{m}$ at 29 cm	897	859	834
	± 60	± 15	± 76
$R_0/\mu\text{m}$ at 15 cm	901	863	838
	± 60	± 16	± 75

sive ringing upon formation is not only the most simple and accurate but also samples the entire population, being capable of logging the natural frequency of each and every bubble that is generated in near real time to 1 Hz accuracy (even giving details of nozzle processes). However, the Gabor signal must be interpreted carefully. It reflects the natural frequency of a damped system, given by $\omega_0(1 - \delta^2)$, where δ is the dimensionless damping coefficient³⁵ and ω_0 the undamped natural frequency: Active techniques in general measure the maximum of the amplitude response, which occurs at frequency $\omega_0(1 - 2\delta^2)$. The two major limitations of the Gabor technique are, first, that the signal becomes increasingly difficult to interpret as the entrainment rate increases. Second, passive emissions usually give information only about the bubbles being entrained during the measurement interval, the excitation that is strong enough to make adequate emissions usually requiring the closure of a liquid surface:¹⁵ Older, "silent" bubbles would have to be excited by impulse to ring, and a sufficiently strong impulse would alter the bubble population by inducing more closures (i.e., fragmentation).

IV. CONCLUSIONS

Broadband insonation rapidly indicates the range over which bubble resonances may occur, reducing the time required for tonal incrementation. Best resolution and population sampling was achieved using the Gabor technique, though this operates only on entrainment. The best active indicator of the bubble population in these tests, where a relatively low-amplitude pump signal was employed to minimize the invasiveness of the technique,³⁶ was the $\omega_i \pm \omega_p/2$ signal. However, it must be remembered that this signal is not simple to implement: For best resolution the acoustic pressure amplitude at the bubble must be close to the threshold,⁹ and a delay (after insonation at a given frequency commences) is recommended, to allow the transients to decay before data is acquired.

ACKNOWLEDGMENTS

Our thanks to EPSRC (Reference No. GR/H 79815) and NERC (GR3/9992) for funding, and to P. R. White for advice.

- ¹R. M. Detsch and R. N. Sharma, "The critical angle for gas bubble entrainment by plunging liquid jet," *Chem. Eng. J.* **44**, 157–66 (1990).
- ²E. G. Tickner, "Precision microbubbles for right side intercardiac pressure and flow measurements," in *Contrast Echocardiography*, edited by R. S. Meltzer and J. Roeland (Nijhoff, London, 1982).
- ³D. K. Woolf, "Bubbles and the air-sea transfer velocity of gases," *Atmos.-Ocean* **31**, 451–474 (1993).
- ⁴T. G. Leighton, "Acoustic Bubble Detection. I. The detection of stable gas bodies," *Environ. Eng.* **7**, 9–16 (1994).
- ⁵T. G. Leighton, "Acoustic Bubble Detection. II. The detection of transient cavitation," *Environ. Eng.* **8**, 16–25 (1995).
- ⁶H. Medwin, "In situ acoustic measurements of microbubbles at sea," *J. Geophys. Res.* **82**, 971–976 (1977).
- ⁷H. Medwin and N. D. Breitz, "Ambient and transient bubble spectral densities in quiescent seas and under spilling breakers," *J. Geophys. Res.* **94**, 12751–12759 (1989).
- ⁸V. L. Newhouse and P. M. Shankar, "Bubble size measurement using the nonlinear mixing of two frequencies," *J. Acoust. Soc. Am.* **75**, 1473–1477 (1984).
- ⁹A. D. Phelps and T. G. Leighton, "High resolution bubble sizing through detection of the subharmonic response with a two frequency excitation technique," *J. Acoust. Soc. Am.* **99**, 1985–1992 (1996).
- ¹⁰R. Y. Nishi, "Ultrasonic detection of bubbles with Doppler flow transducers," *Ultrasonics* **10**, 173–179 (1972).
- ¹¹M. Strasberg, "Gas bubbles as sources of sound in water," *J. Acoust. Soc. Am.* **28**, 20–26 (1956).
- ¹²T. G. Leighton and A. J. Walton, "An experimental study of the sound emitted from gas bubbles in a liquid," *Eur. J. Phys.* **8**, 98–104 (1987).
- ¹³D. M. Farmer and S. Vagle, "Waveguide propagation of ambient sound in the ocean-surface bubble layer," *J. Acoust. Soc. Am.* **86**, 1897–1908 (1989).
- ¹⁴S. A. Thorpe, "Measurements with an Automatically Recording Inverted Echo Sounder; ARIES and the Bubble Clouds," *J. Phys. Oceanogr.* **16**, 1462–1478 (1986).
- ¹⁵T. G. Leighton, *The Acoustic Bubble* (Academic, London, 1994), pp. 234–243, 295–298, 439–464.
- ¹⁶S. L. Morris and A. D. Hill, "Ultrasonic imaging and velocimetry in two-phase pipe flow," *Trans. ASME, J. Heat Transf.* **115**, 108–116 (1993).
- ¹⁷R. Van Der Welle, "Void fraction, bubble velocity and bubble size in two phase flow," *Int. J. Multiphase Flow* **11**, 317–45 (1985).
- ¹⁸W. F. Kolbe, B. T. Turko, and B. Leskovar, "Fast ultrasonic imaging in a liquid filled pipe," *IEEE Trans. Plasma Sci.* **33**, 715–722 (1986).
- ¹⁹D. L. Miller, "Ultrasonic detection of resonant cavitation bubbles in a flow tube by their second harmonic emissions," *Ultrasonics* **19**, 217–24 (1981).
- ²⁰D. L. Miller, A. R. Williams, and D. R. Gross, "Characterization of cavitation in a flow-through exposure chamber by means of a resonant bubble detector," *Ultrasonics* **22**, 224–230 (1984).
- ²¹R. M. Schmitt, H. J. Schmidt, B. Grohs, H. P. Schwarz, and M. Biebing, "Bubble sizing in the lower micron range based on the 2-frequency in-sonification method," *Ultrason. Imaging* **9**, 63–64 (1987).
- ²²N. Breitz and H. Medwin, "Instrumentation for in-situ acoustical measurements of bubble spectra under breaking waves," *J. Acoust. Soc. Am.* **86**, 739–43 (1989).
- ²³J. Wolf, "Investigation of bubbly flow by ultrasonic tomography," *Part. Syst. Charact.* **5**, 170–173 (1988).
- ²⁴D. Koller, Y. Li, P. M. Shankar, and V. L. Newhouse, "High-speed bubble sizing using the double frequency technique for oceanographic applications," *IEEE J. Oceanic Eng.* **17**, 288–291 (1992).
- ²⁵T. G. Leighton, R. J. Lingard, A. J. Walton, and J. E. Field, "Acoustic bubble sizing by the combination of subharmonic emissions with an imaging frequency," *Ultrasonics* **29**, 319–323 (1991).
- ²⁶A. 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, Proceedings of an IUTAM Symposium*, edited by J. R. Blake, J. M. Boulton-Stone, and N. H. Thomas (Kluwer Academic, Dordrecht, The Netherlands, 1994), pp. 475–483.
- ²⁷A. D. Phelps and T. G. Leighton, "Acoustic bubble sizing using two frequency excitation techniques," *Proceedings of the 2nd European Conference on Underwater Acoustics*, Copenhagen, 1994, edited by L. Bjorno (European Commission, Luxembourg, 1994), pp. 201–206.
- ²⁸T. G. Leighton, A. D. Phelps, D. G. Ramble, and D. A. Sharpe, "Com-

- parison of the abilities of eight acoustic techniques to detect and size a single bubble," *Ultrasonics* **34**, 661–667 (1996).
- ²⁹T. G. Leighton, M. F. Schneider, and P. R. White, "Study of dimensions of bubble fragmentation using optical and acoustic techniques," *Proceedings of the Sea Surface Sound*, Lake Arrowhead, California, 1994, edited by M. J. Buckingham and J. Potter (World Scientific, Singapore, 1995), pp. 414–428.
- ³⁰A. D. Phelps, T. G. Leighton, M. F. Schneider, and P. R. White, "Acoustic bubble sizing, using active and passive techniques to compare ambient and entrained populations," ISVR Technical Report No. 229, University of Southampton, UK, 1994.
- ³¹A. D. Phelps, T. G. Leighton, M. F. Schneider, and P. R. White, "Active and passive acoustic bubble sizing," ISVR Technical Report No. 237, University of Southampton, UK, 1994.
- ³²M. Strasberg, "The pulsation frequency of nonspherical gas bubbles in liquids," *J. Acoust. Soc. Am.* **25**, 536–537 (1953).
- ³³T. G. Leighton, K. J. Fagan, and J. E. Field, "Acoustic and photographic studies of injected bubbles," *Eur. J. Phys.* **12**, 77–85 (1991).
- ³⁴R. Clift, J. R. Grace, and M. E. Weber, *Bubbles, Drops and Particles* (Academic, New York, 1978).
- ³⁵C. Devin, Jr., "Survey of Thermal, Radiation, and Viscous Damping of Pulsating Air Bubbles in Water," *J. Acoust. Soc. Am.* **31**, 1654 (1959).
- ³⁶T. G. Leighton, A. D. Phelps, and D. G. Ramble, "Bubble detection using low amplitude multiple acoustic techniques," *Proceedings of the 3rd European Conference on Underwater Acoustics, Heraklion, 1996*, edited by J. S. Papadakis (Crete University, Heraklion, Crete, Greece), pp. 1143–1147.

Numerical simulations with horizontal-refraction-modal tomography. Part I. Adiabatic propagation

A. G. Voronovich and E. C. Shang

NOAA/Environmental Technology Laboratory, 325 Broadway, Boulder, Colorado 80303 and Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado, Boulder, Colorado 80303

(Received 30 July 1996; accepted for publication 25 January 1997)

Numerical simulations on retrieving the 3-D structure of the sound-speed inhomogeneities based on data on the horizontal refraction angle (HRA) of acoustic modes have been performed. The inversion proceeds in two stages. At the first stage the values of propagation constants of different acoustic modes at the nodes of a horizontal grid are calculated. The corresponding procedure represents a linear 2-D problem which is solved for different modes independently. At the second stage the coefficients of the expansion of the variation of the sound-speed profile with respect to some set of empirical orthogonal functions are calculated based on already determined values of propagation constants. This stage represents a 1-D nonlinear problem which is solved for different nodes of the horizontal grid independently. When error-free HRA data were used in the numerical simulations, the 3-D sound-speed field was retrieved perfectly. Then the HRA data were artificially distorted with random errors. It was found that the relative error of the retrieved values of the propagation constant of modes is proportional to the error in HRA (expressed in radians). Corresponding relative error in retrieved values of sound speed is in its turn proportional to the relative error in propagation constants. When the level of HRA error has been chosen according to the level arising from distortions caused by internal waves, the rms error in the retrieved sound-speed values for the situation considered appeared to be about 0.7 m/s. © 1997 Acoustical Society of America. [S0001-4966(97)06205-X]

PACS numbers: 43.30.Pc [JHM]

INTRODUCTION

Horizontal-refraction-modal tomography (HRMT) is an approach for retrieving a 3-D structure of the sound-speed field $c(x, y, z)$ in the ocean based on measurements of horizontal refraction of horizontal rays associated with different acoustic modes. This approach has some advantages as compared to "classical" time-of-flight inversion schemes.¹ The estimates performed in Ref. 2 demonstrated principal feasibility of the HRMT. This paper describes some numerical simulations with the HRMT based on the assumption of adiabatic propagation of the acoustic modes. The other preliminary numerical experiments showed that exact inversions that take into account mode interactions can be done iteratively based on adiabatic propagation as a good zero-order approximation.³ The results of numerical simulations with full-wave inversions will be reported in a separate publication.

A general description of the inversion scheme is given in Secs. I–III; numerical results are described in Sec. IV; and Sec. V presents some conclusions.

I. DESCRIPTION OF THE INVERSION SCHEME

Let us assume that a tonal acoustic source is towed by a vessel around some area of interest. Horizontal refraction angles (HRA) α_n related to different acoustic modes for different vessel positions can be measured with the help of an acoustic interferometer. It represents two line vertical mode-resolving arrays separated about 10 km apart. The phase dif-

ference φ_n between the complex amplitudes $A_n^{(1)}$ and $A_n^{(2)}$ of the n th acoustic mode measured at the two arrays is related to the arrival angle β of the horizontal ray by the obvious relation

$$\varphi_n = b_n L \cos \beta,$$

where L is the distance between the arrays and b_n is the propagation constant of the n th mode. Hence, HRA α_n arising due to ocean inhomogeneities in the horizontal plane is

$$\alpha_n = \Delta\beta = -\frac{\Delta\varphi_n}{b_n L \sin \beta_0},$$

where β_0 is the direction to the source. When $b_n L \gg 1$, a small value of HRA results in the phase shift $\Delta\varphi_n$ of the order of unity. The phase shift $\Delta\varphi_n$ can be measured by correlating complex amplitudes $A_n^{(1)}$ and $A_n^{(2)}$ as follows:

$$\exp(i\Delta\varphi_n + ib_n L \cos \beta_0) = A_n^{(1)}(A_n^{(2)})^* |A_n^{(1)} A_n^{(2)}|^{-1}. \quad (1)$$

The experiments on modal decomposition of the tonal acoustic field with the help of line vertical arrays (determination of A_n 's) were described, in particular, in Refs. 4 and 5. The ways of eliminating $2\pi N_{\text{intr}}$ -type ambiguity in determining $\Delta\varphi_n$ from Eq. (1) (multilobes of the interferometer) are discussed in Ref. 2. It is important that the correct lobe of the interferometer (i.e., the value of the integer N_{intr}) should be determined in the course of the experiment only once because the vessel changes its position continuously. Note that the accuracy in locating the source provided by conventional

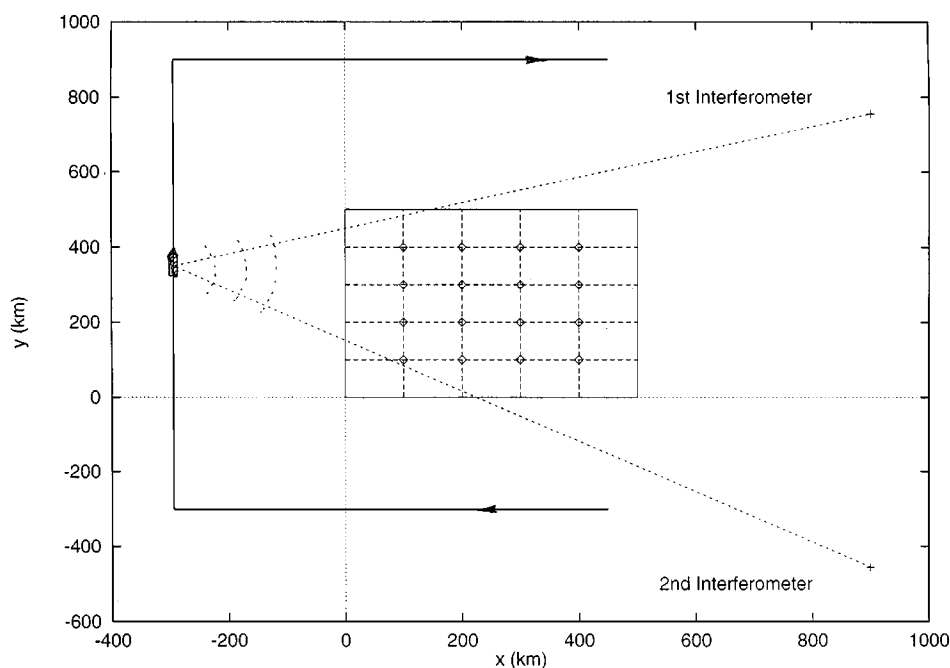


FIG. 1. The geometry of the numerical experiment. The tonal source is towed by the vessel traveling along the solid line marked by arrows. Positions of two acoustic interferometers (each consisting of two vertical line mode-resolving arrays located about 10 km apart) are shown by crosses. Sound-speed profiles are retrieved in 16 nodes of horizontal grid; these nodes are marked by diamonds.

satellite navigation (which is in any case less than a few hundred meters) appears to be sufficient, and no additional systems for the source positioning are necessary.²

In what follows we will assume that the acoustic mode interactions can be neglected and mode propagation is adiabatic, an important concept developed first in Ref. 6. In this case HRA can be expressed with great accuracy as in Ref. 2:

$$\alpha_n = \frac{1}{b_n^{(0)}} \int_0^R \frac{x}{R} \frac{\partial b_n}{\partial y} dx. \quad (2)$$

Here, $b_n^{(0)}$ is the value of the propagation constant in the vicinity of the interferometer; $\partial b_n / \partial y$ is the cross gradient of the propagation constant that changes along the acoustic path connecting the source located at the point $x=0, y=0$ and the receiver which is located at the point $x=R, y=0$. Since the horizontal rays are straight lines with a great accuracy, the integral in Eq. (2) is actually calculated along x axes. Thus, measurement of the HRA gives information about the cross gradients of the propagation constant b_n averaged along the straight line connecting the source and the receiver. By exposing the area of interest from different points along the vessel route, one can hope to get enough information to retrieve the 3-D structure of the sound speed.

The solution of the inverse problem can be performed in two steps. At the first step, based on a set of equations generated by the relation in Eq. (2) for different positions of the vessel, one can retrieve the values of propagation constants b_n as a function of horizontal coordinates x, y . As the acoustic path represents a straight line in the horizontal plane with great accuracy, this task represents a linear 2-D problem. This problem is solved for different acoustic modes independently. Then, one can proceed to the second step and retrieve the sound-speed profile at a given point in the horizontal

plane based on a set of already determined values of propagation constants b_n for different n . At this stage different points in the horizontal plane are considered independently. Now, the inverse problem becomes nonlinear; however, it is in one dimension, and can be easily tackled numerically.

II. THE FIRST STAGE: 2-D LINEAR INVERSION

Let us assume that inhomogeneities of the sound-speed field are located with respect to the horizontal plane inside some rectangular area (see Fig. 1), and outside this area the ocean is horizontally homogeneous. Suppose also for simplicity that both the source and the vertical arrays are located outside the inhomogeneous area. It is assumed also in what follows that the mode interaction with the ocean bottom is negligible, and horizontal refraction is totally due to water-mass inhomogeneities.

Let us introduce in the inhomogeneous area a horizontal grid with m_x and m_y lines along x and y axis, correspondingly. Let us introduce deviations of the propagation constants at the nodes of this grid from the background values \tilde{b}_n :

$$\Delta b_n^{(k)} = b_n - \tilde{b}_n, \quad k=1,2,\dots,M. \quad (3)$$

The values $\Delta b_n^{(k)}$ will be the unknowns to be determined at the first step of the inversion; here, $M = m_x \times m_y$ is the total number of unknowns. Between the nodes of the horizontal grid some interpolation for Δb_n should be used. We will assume that the linear interpolation for Δb_n is applied along the x lines of the grid between the nodes first, and then the linear interpolation is applied in the y direction for all x (continuous set) inside the inhomogeneous area. Thus, the values of Δb_n can now be expressed through the final num-

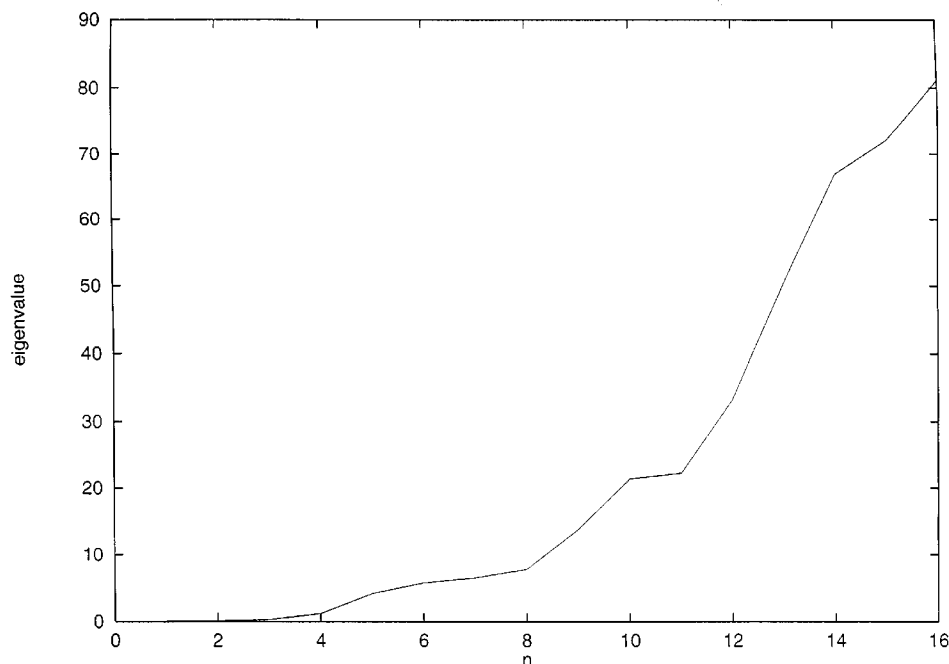


FIG. 2. Spectrum of the matrix T used for retrieving propagation constants at 2-D linear inversion stage [see Eqs. (9) and (10)]. Abscissa n corresponds to the eigenvalue number ordered according to their value, and ordinate shows corresponding eigenvalue.

ber of unknowns $\Delta b_n^{(k)}$ throughout the inhomogeneous area as a function of horizontal coordinates x, y , and the integral in Eq. (2) can be calculated. As a result Eq. (2) can be represented as:

$$\sum_{k=1}^M B_k^l \cdot \Delta b_n^{(k)} = \alpha_n^l, \quad l=1, 2, \dots, L. \quad (4)$$

Here, the upper index l corresponds to different acoustic paths crossing the inhomogeneous area; it is assumed that there are L such paths altogether. The entry of the matrix of the linear set Eq. (4) B_k^l depends on the way the acoustic path crosses the area and can be easily calculated taking into account the method of approximating propagation constants. It is obvious that the matrix B_k^l has generally many zero entries. Note that this matrix does not depend on mode number n . In the linear set Eq. (4) and subsequent equations related to the first stage of inversion, mode number n is a parameter. Thus, different acoustic modes are considered at this stage independently.

Linear equation (2) cannot be solved immediately, because in general $M \neq L$, and the acoustical data α_n^l inevitably contain some errors. For this reason Eq. (4) will be solved in the least square sense, i.e., we will try to minimize the Euclidean norm of the mismatch of the solution Q :

$$Q_{\min} = \min \sum_l \left(\sum_{k=1}^M B_k^l \cdot \Delta b_n^{(k)} - \alpha_n^l \right)^2. \quad (5)$$

From Eq. (5) one easily finds the following linear set:

$$\sum_{k=1}^M T_{mk} \cdot \Delta b_n^{(k)} = A_{mn}, \quad (6)$$

where

$$T_{mk} = \sum_{l=1}^L B_m^l B_k^l \quad (7)$$

and

$$A_{mn} = \sum_{l=1}^L B_m^l \alpha_n^l. \quad (8)$$

Now matrix T_{mk} is a positively determined symmetric $M \times M$ square matrix, and Eq. (6) can be uniquely solved. The matrix T_{mk} can be represented in a diagonal form:

$$T_{mk} = \sum_{i=1}^M t_{mi} \lambda_i t_{ki}, \quad (9)$$

where t_{mi} with index i fixed is an eigenvector of the matrix T related to the real eigenvalue $\lambda_i > 0$: $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_M$. Now the solution of Eq. (6) can be expressed as follows:

$$\Delta b_n^{(k)} = \sum_{i=i_0}^M \frac{1}{\lambda_i} t_{ki} \sum_{j=1}^M t_{ji} A_{jn}. \quad (10)$$

The exact solution of Eq. (6) corresponds to the summation over all i in Eq. (10) starting from $i = i_0 = 1$. However, some of the eigenvalues λ_i can be very small (see Fig. 2 as an example), and when the acoustic data contain errors it is advisable to discard a few of the smallest ones (singular value decomposition). This distorts the solution somewhat, but also makes it more robust. The number of discarded eigenvalues $i_0 - 1$ depends on the level of noise; in the present instance the proper choice of i_0 was made on the basis of numerical simulations.

III. THE SECOND STAGE: 1-D NONLINEAR INVERSION

After the values $\Delta b_n^{(k)}$ are determined for a number of low-order modes $n=1,2,\dots,N$, one can proceed to the second stage of inversion. Let us now consider a set of deviations of propagation constants $\Delta b_n^{(k)}$ corresponding to different acoustic modes n for a given node of the horizontal grid (i.e., for k fixed). The problem now is to retrieve the sound-speed profile (SSP) based on this data. To solve this problem we first parametrize the possible set of SSP with the help of some empirical orthogonal functions (EOF) $\text{EOF}_d(z)$:

$$\frac{1}{c^2(z)} = \frac{1}{c_0^2(z)} + \Delta \left(\frac{1}{c^2(z)} \right), \quad (11)$$

$$\Delta \left(\frac{1}{c^2(z)} \right) = \sum_{d=1}^D C_d \cdot \text{EOF}_d(z). \quad (12)$$

Here, $c_0(z)$ is an initial SSP, and D is a total number of EOFs (the methods of calculating EOFs can be found elsewhere). The problem now is to determine the values of coefficients C_d based on the known initial SSP $c_0(z)$ and the values of variations of propagation constants Δb_n , $n=1,2,\dots,N$ (for brevity, the fixed index k is omitted here and in what follows).

We will solve this task using the method related to the invariant imbedding approach. Let us add to SSP $c_0(z)$ some small addition according to Eq. (12) with C_d sufficiently small. Resulting variations of propagation constants are in this case linearly related to C_d :

$$\delta b_n^2 = \omega^2 \int \Delta \left(\frac{1}{c^2(z)} \right) u_n^2 dz = \sum_{d=1}^D R_{nd} C_d, \quad (13)$$

where

$$R_{nd} = \int \text{EOF}_d(z) \cdot u_n^2 dz.$$

Here, u_n is the eigenfunction of the standard boundary problem

$$\frac{d^2 u_n}{dz^2} + \left(\frac{\omega^2}{c_0^2(z)} - b_n^2 \right) u_n = 0, \quad u_n(0) = u_n(-\infty) = 0 \quad (14)$$

associated with propagation constant b_n .

Let us choose the values of C_d such that the modified propagation constants $\tilde{b}_n + \delta b_n$ approach the desired values $\tilde{b}_n + \Delta b_n$. To achieve this let us set

$$\delta b_n^2 = a \cdot \Delta b_n^2, \quad (15)$$

where a is a sufficiently small constant. The resulting linear set

$$a \cdot \Delta b_n^2 = \sum_{d=1}^D R_{nd} C_d$$

generally cannot be solved. For this reason we determine C_d in the same way as the solution of Eq. (4) minimizing the following mismatch:

$$q_{\min} = \min \sum_{n=1}^N \left(a \cdot \Delta b_n^2 - \sum_{d=1}^D R_{nd} C_d \right)^2. \quad (16)$$

This leads to the linear set

$$\sum_{d=1}^D U_{bd} C_d = f_b, \quad b = 1, \dots, D, \quad (17)$$

where

$$U_{bd} = \sum_{n=1}^N R_{nb} R_{nd}$$

and

$$f_b = a \sum_{n=1}^N \Delta b_n^2 R_{nb}.$$

The value of constant a has been chosen according to the relation

$$a = \epsilon \cdot \frac{\min_n (\tilde{b}_{n-1}^2 - \tilde{b}_n^2)}{\max_n [(\tilde{b}_n + \Delta b_n)^2 - \tilde{b}_n^2]}$$

with $\epsilon \sim 0.1$. After the values of C_d have been determined by solving Eqs. (17) we redefine background SSP $c_0(z)$ incorporating into it the determined variation of SSP:

$$\frac{1}{c_0^2(z)} + \sum_{d=1}^D C_d \cdot \text{EOF}_d(z) \rightarrow \frac{1}{c_0^2(z)}.$$

Accordingly, propagation constants corresponding to the new background SSP will be closer to the desired values $\tilde{b}_n + \Delta b_n$. Now, we can repeat the procedure from the very beginning. Iterations are stopped after propagation constants cease to change or become close enough to the desired values. An example of the resulting evolution of the propagation constants for the numerical example considered in the next section is shown in Fig. 3. The actual procedure applied was somewhat optimized by making use in a new node of the information concerning SSP in the neighboring nodes, which have already been considered. Thus, the initial SSP $c_0(z)$ for a given node in fact represented some interpolation of the SSPs related to neighboring nodes of the horizontal grid.

Note that in many cases 1-D nonlinear inversion can also be accomplished using brute force approach. Namely, one can initially calculate the sets of propagation constants for a sufficiently dense net of points within an appropriate region in the linear D -dimensional space C formed by parameters C_d . In many practical situations D will not exceed 3, and choosing for each of D parameters $N_c = 10-20$ points, one should solve boundary problem Eq. (14) about $(N_c)^D = 10^3-10^4$ times, which would require no more than a day of calculations on a PC. Then, one can select appropriate points in the C -space by comparing calculated sets of propagation constants with those retrieved at the first stage of inversion and choosing the closest point using suitable metrics.

IV. A NUMERICAL EXAMPLE

Let us assume that the inhomogeneities to be reconstructed are located within the square 500×500 km. We will

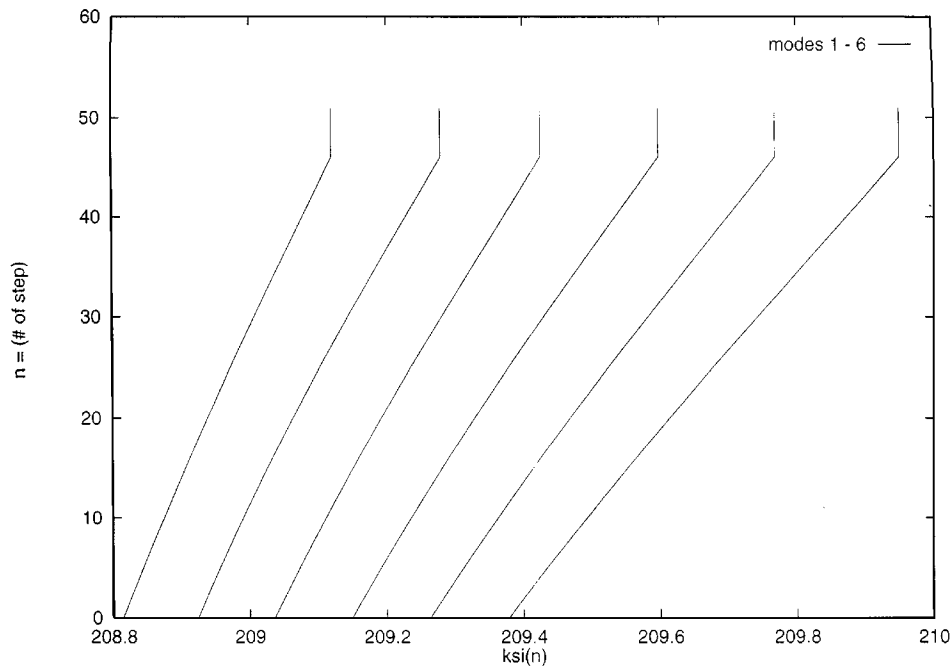


FIG. 3. Evolution of propagation constants when being retrieved according to invariant embedding approach described in Sec. III. The node is located at the point $x=100$ km, $y=100$ km, and error-free data are used in computations. The short vertical cuts at the upper ends of the lines describe desired final values of propagation constants which are perfectly retrieved in the error-free case.

introduce a horizontal grid with 4 lines along both x and y directions: $m_x=m_y=4$. Those lines are uniformly distributed within the square (see Fig. 1). Thus, the SSP will be retrieved at the 16 nodes located 100 km apart from each other. It is assumed that the inhomogeneous region is surrounded by layered ocean with Munk's SSP:

$$c^{(0)}(z) = c_{00}[1 + 0.0057(e^t - t - 1)].$$

Here, $t = 2(z - z_1)/B$ and $c_{00} = 1.5$ km/s. The depth of the

sound channel axis z_1 and its width B in what follows are chosen to be $z_1 = 1$ km and $B = 1.3$ km. Let us assume that the inhomogeneities are described with the help of the two following EOFs (see Fig. 4):

$$\text{EOF}_1(z) = 0.001 \exp(-t^2) \cos t,$$

$$\text{EOF}_2(z) = 0.002522 \exp(-t^2) \sin t.$$

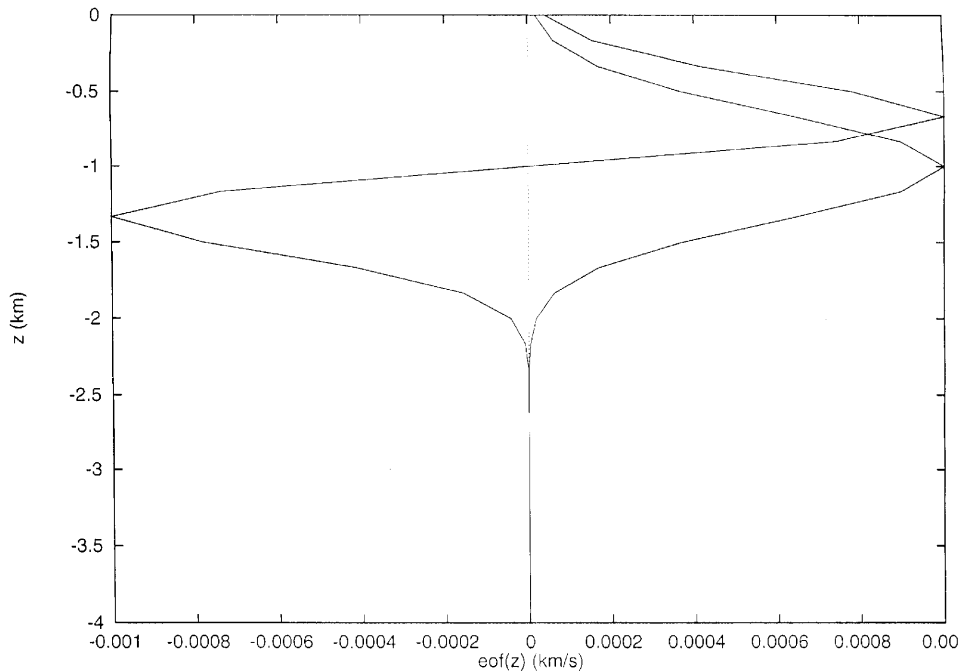


FIG. 4. Vertical profiles of $\text{EOF}_1(z)$ and $\text{EOF}_2(z)$ used in the numerical simulations.

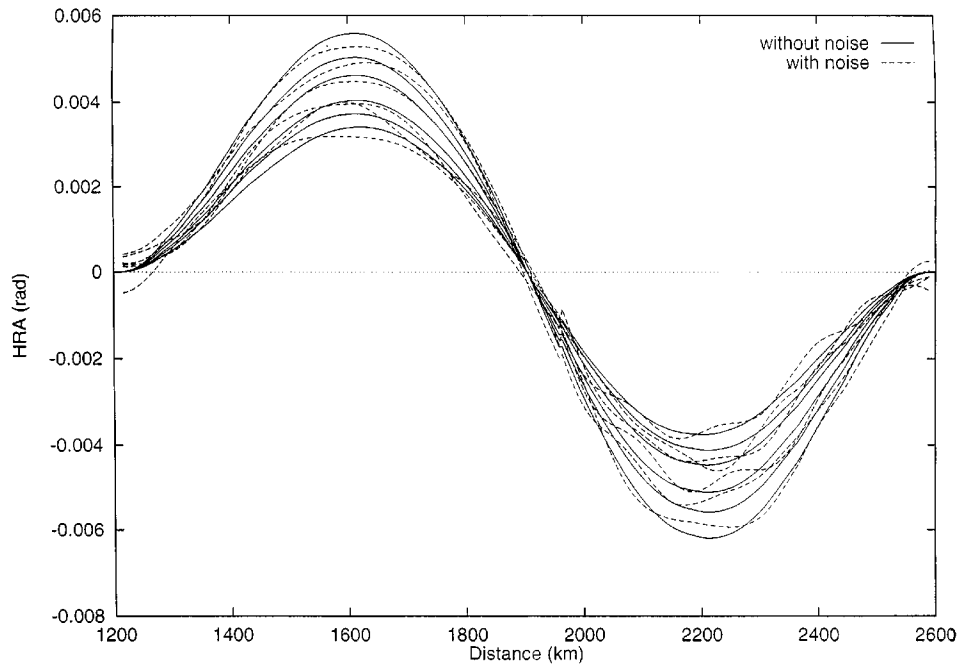


FIG. 5. Horizontal refraction angles (HRA) of the first six modes measured by the first interferometer. Abscissa represents distance along the vessel route counted along the track from the starting point (see Fig. 1). Solid lines correspond to error-free HRA data, and dashed lines incorporate random errors.

Thus in our case $D=2$. The dimension of the EOFs is $(\text{km/s})^{-2}$ and the expansion coefficients C_d are in this case dimensionless. The EOFs were normalized according to the condition:

$$\max_z \text{EOF}_d(z) = 0.001(\text{km/s})^{-2}.$$

It is assumed that the horizontal structure of the inhomogeneities represents a Gaussian eddy:

$$C_1(x,y) = C_{\max} \cdot \exp \left[- \left(\frac{x-x_0}{l_x} \right)^2 - \left(\frac{y-y_0}{l_y} \right)^2 \right],$$

$$C_2(x,y) = 0.$$

In the numerical simulations we have chosen $C_{\max}=8$ with the center of the eddy coinciding with the center of the square: $x_0=y_0=250$ km. It easily follows from Eq. (12) that the maximum of sound-speed variation with respect to

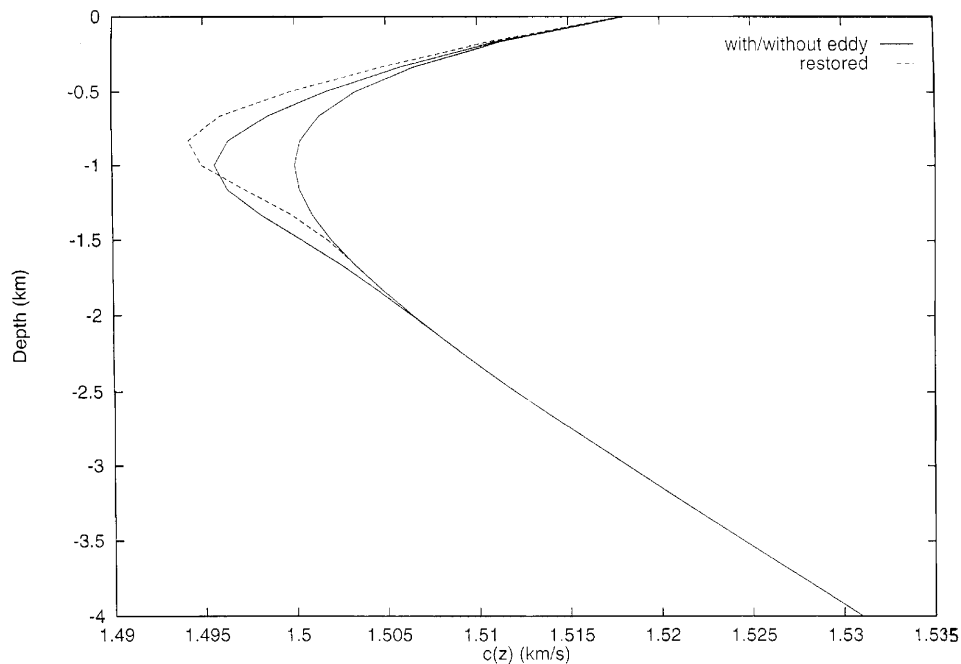


FIG. 6. Sound-speed profiles retrieved in the node with coordinates $x=400$ km, $y=400$ km (the worst case among all nodes). The solid line with deeper sound-speed minimum value represents exact SSP, and the dashed line corresponds to the result of retrieving (note that HRA data used for inversion incorporated errors).

depth can be estimated with good accuracy as Δc (m/s) $\approx -1.5^3/2 \cdot C \approx -1.7C$. In the center of our cold eddy we obtain accordingly $\Delta c \approx -13.6$ m/s. The horizontal scales of the eddy were chosen to be $l_x = l_y = 200$ km.

It is important that in spite of the actual absence in the SSPs of the second EOF, its appearance at the second stage of inversion is allowed. Retrieving the SSP with the only one EOF ($D=1$) is a much more robust procedure. Thus the structure of inhomogeneities within a 500×500 km square in our case is described with the help of $m_x \times m_y \times D = 32$ independent parameters.

The frequency of the tonal signal was taken at $f = 50$ Hz. The route of the vessel towing the source and the positions of two acoustic interferometers are also shown in Fig. 1. When the vessel's route proceeds along the perimeter of the square (which would be more natural), the spectrum of matrix T [see Eq. (7)] deteriorates (the number of small eigenvalues increases), and a stable solution of the inverse problem requires additional interferometers. A similar problem with respect to time-of-flight tomography was addressed at length in Ref. 7; regarding the tomography scheme under consideration it will be discussed in a separate publication. In the present instance we avoided this problem by moving the route and interferometers from the immediate vicinity of the inhomogeneous region.

If no errors are introduced into the numerically simulated data concerning HRA, the inversion scheme described above works perfectly, and SSPs are retrieved to within an accuracy of millimeters per second. In particular, evolution of the propagation constants b_n from the initial values (corresponding to Munk's profiles) to the actual values retrieved at the first (linear) stage of inversion is shown in Fig. 3. One can see that the second stage of inversion represents essentially a nonlinear problem (as the variations in propagation constants exceed the difference between neighboring ones significantly).

However, "noiseless" simulations are obviously of little practical use. Possible sources of errors in measuring HRA are analyzed in Ref. 2. The main source appears to be scattering of the acoustic modes at long internal waves. The rms value of the corresponding error is estimated to be of the order of $\sigma_\alpha \sim 10^{-3}$ rad.² This value can be reduced significantly by averaging. Let us assume that the coherence time of the acoustic signal associated with scattering at internal waves is about 500 s. Let us also assume, that averaging the HRAs proceeds over a 25×500 s ≈ 3.5 h period of time. If the towing speed is about 2 m/s, the vessel will move during this period for 25 km, which should additionally contribute to the decorrelation of the different readings of HRA. As a result one can hope to reduce the rms value of the HRA error to the level of $\sigma_\alpha \sim 10^{-3}/\sqrt{25} = 2 \times 10^{-4}$ rad. To simulate real data the random values ensuring such a level of rms error were added to the exact calculated values of HRA, and the inversion in fact proceeded with "noisy" data. The dependencies of the HRAs for different modes on the distance passed by the vessel are demonstrated in Fig. 5, where both exact and "noisy" cases are shown.

It was assumed that the data concerning HRA of the first $N_m = 6$ acoustic modes were available. The spectrum of ma-

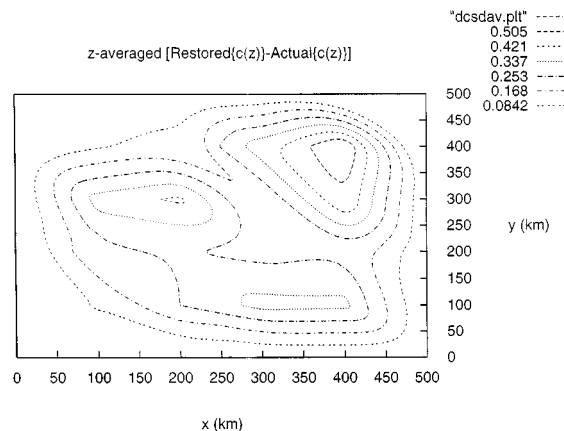


FIG. 7. Isolines of the errors in the retrieved values of the sound speed in m/s averaged over the ocean depth ($H=4$ km).

trix T [see Eqs. (8) and (9)] corresponding to the geometry of the experiment is plotted in Fig. 2. It was determined by numerical simulation that the best results for the given level of HRA errors $\sigma_\alpha = 2 \times 10^{-4}$ rad are obtained for $i_0 = 5$ [i.e., four of the least eigenvalues were discarded, see Eq. (10)].

Ten inversions of the sound-speed inhomogeneities corresponding to different realization of the "noise" have been performed. The error $\overline{\delta b}$ in the retrieved values of the propagation constants averaged over all six modes, 16 nodes, and ten realizations appeared to be

$$\frac{\overline{\delta b}}{b} = 4.90 \times 10^{-4},$$

where $b = 2\pi f/c_{00} \approx 209.3 \text{ km}^{-1}$ is a typical value of the propagation constant. It is reasonable to assume that in a general case one will have

$$\frac{\overline{\delta b}}{b} = g_1 \sigma_\alpha. \quad (18)$$

For a given numerical experiment we thus obtained $g_1 \sim 2.45$.

The results of retrieving the 3-D inhomogeneities for one typical realization are demonstrated in Figs. 6–9. The result of SSP inversions (second stage) appeared to be the

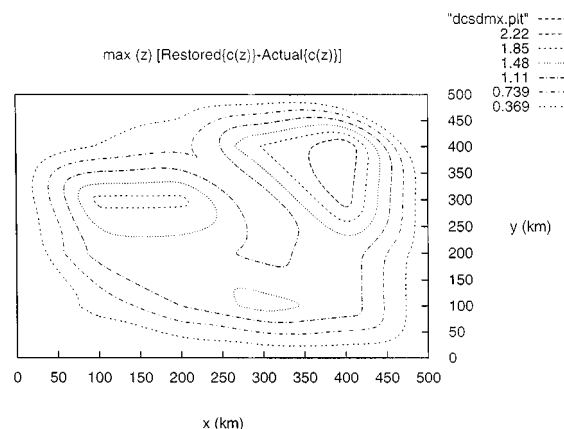


FIG. 8. The same as in Fig. 7, however, with respect to maximum values of errors.

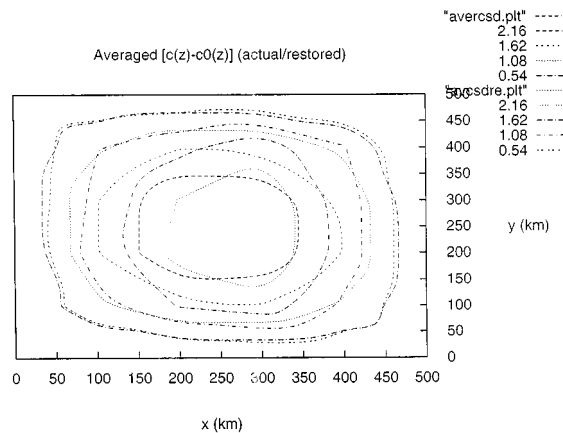


FIG. 9. Isolines of the averaged-over-depth absolute values of the variations of the sound speed (in m/s) associated with an eddy. Both actual and retrieved eddies are shown ("noisy" data were used for retrieving; for "noisy-free" data retrieving is perfect).

worst for the node at $x = y = 400$ km. Both Munk's canonical profile, the real SSP calculated according to Eqs. (11) and (12), and the retrieved one are shown in Fig. 6. The error in the retrieved values of the sound speed (in m/s) averaged over the depth of the ocean $H = 4$ km is shown by contour lines in Fig. 7. The maximum value of the error (with respect to depth) is plotted by contour lines in Fig. 8. The contour lines describing depth-averaged variations of the sound speed for both actual and retrieved eddies are shown in Fig. 9.

The rms error in retrieving coefficients C_d averaged over all $D = 2$ EOFs, $M = 16$ nodes and ten realizations appeared to be $\delta C = 0.41$. Corresponding error in terms of sound-speed value is $\delta c = 1.7 \delta C = 0.7$ m/s, or in the dimensionless form

$$\frac{\delta c}{c_{00}} = 4.6 \times 10^{-4}.$$

Note that in the general case one could assume that the following relation holds:

$$\frac{\delta c}{c_{00}} = g_2 \frac{\delta b}{b} = g_1 g_2 \sigma_\alpha = g \sigma_\alpha. \quad (19)$$

In the case considered we found $g_2 \approx 0.94$ and $g \approx 2.30$. Thus coefficients in Eqs. (18) and (19) can be considered as generally having the order of unity.

V. DISCUSSION

Numerical simulations on retrieving the 3-D structure of the sound-speed inhomogeneities based on data on horizontal refraction of acoustic modes have been performed. The inversion procedure proceeds in two stages. At the first stage the values of propagation constants of different acoustic modes at the nodes of a horizontal grid are calculated. Corresponding procedure represents a linear 2-D problem that is solved for different modes independently. At the second stage the coefficients of the expansion of the variation of SSP with respect to some set of empirical orthogonal functions are calculated based on already determined values of propagation constants. At this time inversion represents a 1-D nonlinear problem that is solved for different nodes of the horizontal grid independently.

This inversion scheme works perfectly for an error-free situation. When some errors are introduced into calculated values of horizontal refraction angles, the resulting relative error in the values of propagation constants can be estimated according to the relation in Eq. (18) with coefficient g_1 being of the order of unity. The relative error in the value of sound speed appears to be close to the relative error in propagation constants.

The results of numerical simulation demonstrate feasibility of the suggested inversion scheme for practical situations.

ACKNOWLEDGMENT

This work was supported by the Office of Naval Research.

- ¹W. Munk and C. Wunsch, "Ocean acoustic tomography: a scheme for large scale monitoring," *Deep-Sea Res.* **26A**, 123–161 (1979).
- ²A. G. Voronovich and E. C. Shang, "A note on horizontal-refraction-modal tomography," *J. Acoust. Soc. Am.* **98**, 2708–2716 (1995).
- ³A. G. Voronovich and E. C. Shang, "Horizontal refraction tomography for nonadiabatic propagation," *J. Acoust. Soc. Am.* **98**, 2913(A) (1995).
- ⁴T. C. Yang, "A Method of range and depth estimation by model decomposition," *J. Acoust. Soc. Am.* **82**, 1736–1745 (1987).
- ⁵A. G. Voronovich, V. V. Goncharov, and Yu. A. Chepurin, "Measurements of modal spectrum of a low-frequency sound field in the deep sea," *Dokl. Akad. Nauk SSSR* **317**, 723–727 (1991).
- ⁶H. Weinberg and R. R. Burridge, "Horizontal ray theory for ocean acoustics," *J. Acoust. Soc. Am.* **55**, 63–79 (1974).
- ⁷B. Cornuelle, W. Munk, and P. Worcester, "Ocean acoustic tomography from ships," *J. Geophys. Res.* **94** (C5), 6232–6250 (1989).

Analytical expressions for high-frequency acoustic fluctuations from suspended particles in the ocean

R. A. Thuraisingham

Aeronautical and Materials Research Laboratory, DSTO, P.O. Box 44, Pyrmont, NSW 2009, Australia

(Received 1 February 1996; accepted for publication 17 November 1996)

Simple analytical expressions for amplitude and phase fluctuations of high-frequency sound from suspended particles in the ocean are derived. The fluctuations depend on the wave number, path length, particle radius, the relative volume density of scatterers (σ), and the ratio of sound speed in the medium to that of the scatterer. The expressions are valid in the Rytov region. © 1997 Acoustical Society of America. [S0001-4966(97)05404-0]

PACS numbers: 43.30.Re, 43.30.Ft, 43.30.Ma [JHM]

LIST OF SYMBOLS

f	frequency	a	radius of particle
\mathbf{r}	position vector	σ	relative volume density
ω	sample	$K_\beta(r)$	spatial correlation function in the β field
Ω	probability sample space	$N(r)$	spatial correlation function of sound-speed fluctuations
$c(\mathbf{r}, \omega)$	sound speed at position \mathbf{r} , in the sample ω	A, A_0	amplitude of received and direct signals
c_0	mean sound speed	ϕ, ϕ_0	phase of received and direct signals
$\delta c/c(\mathbf{r}, \omega)$	fractional sound-speed fluctuation at position \mathbf{r} , in sample ω	$\frac{\phi}{am^2}$	$\frac{(\ln(A/A_0))^2}{(\phi - \phi_0)^2}$
c_p, c_m	sound speed in particle and medium	$\Delta \phi^2$	$\frac{(\phi - \phi_0)^2}{(x' - x''), (y' - y''), (z' - z'')}$
k_0	wave number ($2\pi f/c_0$)	u_1, u_2, u_3	
λ	wavelength	ρ	mean density

INTRODUCTION

When sound is transmitted in the ocean it has been observed that the amplitude and phase of the acoustic signal fluctuate under what appears to be constant transmission conditions. An optical phenomenon similar to this is the twinkling of stars. The acoustic fluctuations observed in the ocean have been attributed to the scattering of sound due to variations in temperature, salinity, current, and to the number density of bubbles and suspended particulate matter. These are usually referred to as random inhomogeneities of the medium. Thuraisingham (1996) investigated the effect of temperature variations on high-frequency acoustic transmission. This paper extends this study to the effect of suspended particles in the ocean on megahertz acoustic transmission. The interest in megahertz sound stems from our current innovation program to develop an underwater acoustic camera of the type described in Jones (1994).

Generally in the ocean two populations of particles have been distinguished: a common population of particles that have radii of less than $2.5 \mu\text{m}$ and a rarer population of larger particles of radii greater than $25 \mu\text{m}$ (Lal, 1977). The fine particle concentration is of the order of 0.02 to 0.05 mg/l with a settling velocity of less than 1 m per day and so they undergo large-scale horizontal transport. The large particles undergo vertical settling with speeds of 50–200 m per day. (Eisma, 1993). Assuming fine particles of radius of $1 \mu\text{m}$ to be silty clay with an approximate density of $2.5 \times 10^6 \text{ g/m}^3$ (Hamilton and Bachman, 1982), the number of par-

ticles corresponding to a concentration of 0.02 mg/l is approximately $1.9 \times 10^9 \text{ per m}^3$. The effect of a large number of small particles on the amplitude and phase of sound in the 1-MHz to 10-MHz frequency range is the focus of this paper.

The effect of inhomogeneities on acoustic transmission depends strongly on the wavelength of the signal. If the inhomogeneity is small compared to the acoustic wavelength then scattering is small. The inhomogeneities of the ocean change slowly compared with the acoustic frequencies so the medium can be assumed to be essentially fixed during a measurement. However, the variations in the inhomogeneities between measurements result in changes in the amplitude and phase of the signals for repeated transmissions. The statistics of the amplitude and phase fluctuations of the signal subject to random inhomogeneities are useful in the study of the temporal coherence of the transmitted signal and in the determination of the coherence length for a receiver array.

The problem that is looked into is referred to as the line of sight problem (Ishimaru, 1978, Vol. 1, p. 116). In this problem the signal at the receiver consists of a coherent and incoherent contribution. In the absence of the inhomogeneities it reduces to the incident direct signal. In the scattering problem the signal is totally incoherent. Scattering and absorption of sound in the megahertz region by suspended particles have already been dealt by Thuraisingham (1994).

This study of the effect of suspended particle scatterers on the amplitude and phase fluctuation of high-frequency transmitted sound is carried out using a form of the correla-

tion function developed by Liu (1991a, 1991b). Liu obtained excellent agreement with experimental ultrasound scattering from blood cells using his correlation function. Our study extends the use of it to model a cloud of suspended particles in the sea and obtains expressions for amplitude and phase fluctuations of high-frequency plane wave.

In this paper theoretical expressions for the mean-square amplitude and phase fluctuations are obtained using the Rytov method for a plane wave (Chernov, 1960, pp. 61–64). For suspended particles in the ocean, at acoustic frequencies of 1–10 MHz and a range of a few meters the ray approximation cannot be used. This is discussed in Sec. II A. However the Rytov method may be appropriate. Nevertheless the variance of the phase fluctuations in the ray model is derived in this paper to provide an independent check on the expressions using the Rytov method (Ishimaru, 1978, Vol. 2, p. 365).

Section I describes the correlation function used while Sec. II gives expressions for amplitude and phase fluctuations using Rytov and ray theory. Section III discusses the results and Sec. IV is the conclusion.

I. CORRELATION FUNCTION

This section describes the correlation function of the sound-speed fluctuations that is used to obtain the variance of the acoustic fluctuations. Liu (1991a, 1991b) considered a group of spherical scatterers, made up of the same material randomly distributed in a region D , assumed to be homogeneous. The speed of sound in the surrounding medium is taken as c_m while that in the particle as c_p . Suppose S is the region occupied by the scatterers and included in D . The speed of sound $c(\mathbf{r}, \omega)$, where $\mathbf{r} \in D$ and $\omega \in \Omega$ is given by

$$c(\mathbf{r}, \omega) = c_p, \quad \mathbf{r} \in S, \quad (1)$$

$$c(\mathbf{r}, \omega) = c_m, \quad \mathbf{r} \notin S, \quad (2)$$

where Ω is the probability sample space. Liu (1991a) then defined a stochastic field $\beta(\mathbf{r}, \omega)$ where

$$\beta(\mathbf{r}, \omega) = \left(\frac{c_0^2(\mathbf{r})}{c^2(\mathbf{r}, \omega)} - 1 \right), \quad \mathbf{r} \in D, \quad (3)$$

$$c_0(\mathbf{r}) = c_0 = c_m c_p / \sqrt{\sigma c_m^2 + (1 - \sigma) c_p^2}, \quad \mathbf{r} \in D. \quad (4)$$

Thus

$$\beta(\mathbf{r}, \omega) = \beta_1, \quad \mathbf{r} \in S, \quad (5)$$

$$\beta(\mathbf{r}, \omega) = \frac{-\sigma}{(1 - \sigma)} \beta_1, \quad \mathbf{r} \notin S, \quad (6)$$

where

$$\beta_1 = \left(\frac{0.5\sigma}{1 - \sigma} \right)^{1/2} \frac{(1 - \sigma)(c_m^2 - c_p^2)}{[\sigma c_m^2 + (1 - \sigma) c_p^2]}. \quad (7)$$

The spatial correlation function in the β field for a sample realization is given by

$$K_\beta(r) = \lim_{v(D) \rightarrow \infty} \frac{1}{V(D)} \int_D \beta(r_1, \omega) \beta(r_1 - r) d\mathbf{r}_1, \quad (8)$$

where $V(D)$ is the volume of D , and r the separation between the particles. Assuming that the distance from any sphere to its nearest neighbor is a random variable that has an exponential probability density, and using Eqs. (5)–(7) along with geometry, Liu (1991b) obtained the following expression for $K_\beta(r)$.

$$K_\beta(r) = \beta_1^2 e^{-br} (1 + \cos sr), \quad (9)$$

where

$$b = \frac{3}{4(1 - \sigma)a}, \quad s = \frac{\pi^3 \sqrt{\sigma}}{a}. \quad (10)$$

In the above expressions a denotes radius of the particle, and σ the relative volume density of the particles, this being volume of the scatterers divided by the volume of the region D .

In what follows the correlation function of the sound-speed fluctuations, $N(r) = \langle (\delta c/c)(\mathbf{r}_1, \omega) (\delta c/c)(\mathbf{r}_2, \omega) \rangle$ is required, where $r = |\mathbf{r}_2 - \mathbf{r}_1|$. Duykers (1967) from acoustic measurements in the upper layers (0–200 m) inferred that the relative volume density of particles is 6×10^{-6} . In general it has been observed that the relative volume density of suspended matter in the ocean varies from 5×10^{-9} to 10^{-6} (Eisma, 1993, p. 38). In coastal waters the relative volume density of particles can be of the order of 10^{-4} (Eisma, 1993, p. 38), while near the sea bottom it can vary from about 10^{-6} at about 1 m from the bottom to about 10^{-3} close to the boundary (Adam and Weatherly, 1981, p. 6). In all cases it is thus possible to make the approximation that the relative volume density of scatterers $\sigma \ll 1$. Under these conditions $c_0(\mathbf{r}) \approx c_m$ and

$$\begin{aligned} \langle \beta(\mathbf{r}_1, \omega) \beta(\mathbf{r}_2, \omega) \rangle &\approx \left\langle \frac{\delta c}{c}(\mathbf{r}_1, \omega) \frac{\delta c}{c}(\mathbf{r}_2, \omega) \right\rangle \\ &\times \left(1 + \frac{c_m}{c_p} \right)^2. \end{aligned} \quad (11)$$

Equation (11) is obtained using Eq. (3). Since $\langle \beta(\mathbf{r}_1, \omega) \beta(\mathbf{r}_2, \omega) \rangle = K_\beta(r)$, it follows therefore

$$N(r) \approx K_\beta(r) / (1 + c_m/c_p)^2. \quad (12)$$

The correlation function for sound-speed fluctuations $N(r)$ can therefore be written as,

$$N(r) = \gamma e^{-br} (1 + \cos sr), \quad \gamma = (\sigma/2) (1 - c_m/c_p)^2. \quad (13)$$

In the study of acoustic fluctuations due to thermal inhomogeneities, Chernov (1960, p. 75) used the Gaussian correlation function

$$N(r) = \overline{\mu^2} e^{-(r^2/a^2)}, \quad \overline{\mu^2} = \overline{(\delta c/c)^2}. \quad (14)$$

The choice of the Gaussian correlation function for thermal inhomogeneities was based on the experimental temperature observations made in the ocean (Liebermann, 1951). In this paper the focus is on acoustic fluctuations caused by suspended particles in the ocean. This has similarities to the problem studied by Liu (1991b) who investigated ultrasound scattering from blood cells. Both acoustic fluctuations from suspended particles in the ocean and ultrasound scattering from blood cells involves discrete particles where the solu-

tion requires the specification of a correlation function for discrete particles. The correlation function developed by Liu (1991b) provided theoretical estimates of backscattering that agreed very well with the experimental ultrasound scattering from blood cells. Further the choice of the Gaussian function would require $(\delta c/c)^2$ and the correlation length to be specified. These prompted the use of Eq. (13) instead of Eq. (14) in the study of acoustic fluctuations due to suspended particles.

II. ANALYTICAL EXPRESSIONS FOR FLUCTUATIONS

A. Expressions using the Rytov method

Consider a continuous monochromatic plane wave traveling along the X direction and incident upon an inhomogeneous medium at $x=0$. The medium is inhomogeneous beyond $x > 0$, and the receiver is at $(L, 0, 0)$. Let us denote the mean-square logarithmic relative amplitude fluctuations $[(\ln(A/A_0))^2]$ by $\overline{am^2}$ where A and A_0 are the amplitude of the received and direct signals. Also let us denote the mean-square phase fluctuations by $(\Delta\phi^2)$ where $\Delta\phi = \phi - \phi_0$, ϕ and ϕ_0 being the phases of received and direct signals. The procedure adopted in this paper to obtain the variance of the acoustic fluctuations follows that given by Chernov (1960, pp. 61–75). The equations are similar except for the correlation function. Following Chernov (1960, p. 70), $\overline{am^2}$, $\overline{\Delta\phi^2}$ are given by

$$\overline{am^2} = 0.5k_0^2(I_1 - I_2), \quad (15)$$

$$\overline{\Delta\phi^2} = 0.5k_0^2(I_1 + I_2), \quad (16)$$

where $k_0 = 2\pi f/c_0$ is the wave number, f being the frequency.

The integrals I_1 and I_2 are given by

$$I_1 = \frac{k_0}{2\pi} \int_0^L dx' \int_0^L dx'' \int_{-\infty}^{+\infty} du_2 \int_{-\infty}^{+\infty} du_3 \\ \times \sin\left(\frac{k_0(u_2^2 + u_3^2)}{2(x' - x'')}\right) \frac{N(u_1, u_2, u_3)}{(x' - x'')}, \quad (17)$$

$$I_2 = \frac{k_0}{2\pi} \int_0^L dx' \int_0^L dx'' \int_{-\infty}^{+\infty} du_2 \int_{-\infty}^{+\infty} du_3 \\ \times \sin\left(\frac{k_0(u_2^2 + u_3^2)}{2[2L - (x' + x'')]} \right) \frac{N(u_1, u_2, u_3)}{[2L - (x' + x'')]}, \quad (18)$$

where $u_1 = (x' - x'')$, $u_2 = (y' - y'')$, $u_3 = (z' - z'')$.

The above expressions for amplitude and phase fluctuations were obtained using the Rytov and Fresnel approximation (Chernov, 1960, pp. 61–66). It also assumes that the acoustic fluctuations are as a result of velocity variations of the suspended particles and not due to density fluctuations. The latter assumption arises as a result of the acoustic pressure in the Rytov formulation being based on the wave equation $(1/c^2)(\partial^2 p/\partial t^2) - \nabla^2 p = 0$. This form of the wave equation neglects density fluctuations. The expression for acoustic fluctuations that will be derived in this section is

therefore for the contribution that arises from the velocity variations due to suspended particles. Let us now look at the validity of the other two approximations.

Flatte (1979, p. 93) presented different sound transmission regions in the $\Lambda - \Phi$ space, where Λ and Φ relate to the spatial extent (size) and strength of the inhomogeneities, respectively. The size of the inhomogeneities Λ is referred to by Flatte as the diffraction parameter. Flatte (1979, p. 91) provided an expression for the diffraction parameter Λ , where

$$\Lambda = L/6a^2k_0. \quad (19)$$

The strength Φ is defined to be the root-mean-square (rms) phase fluctuation in the ray model. An expression for Φ will be derived in Sec. II B. In the region where $\Lambda\Phi^2 \leq 1$ and $\Lambda \leq 1$, geometric acoustics or ray theory as is commonly known is valid. For suspended particles in the ocean, at megahertz frequencies and range of a few metres, $L/a \gg 1$ and $\lambda/a \geq 1$, where $\lambda = (2\pi/k_0)$ is the wavelength of sound. Hence $\Lambda > 1$. The ray approximation is clearly invalid. The Rytov method uses the acoustic wave equation and obtains estimates of amplitude and phase fluctuations assuming that the relative amplitude and phase changes over a wavelength are small (Chernov, 1960, pp. 61–65). It is superior to first-order multiple scattering theory (Ishimaru, 1978, Vol. 1, pp. 134–136). Its validity goes beyond the geometric acoustic region, to the region where $\Lambda > 1$ but $\Phi \leq 1$. Since the Rytov method is applicable when $\Lambda > 1$, it could be used for $L/a \gg 1$ and $\lambda/a \geq 1$. There are no stringent conditions such as the necessity to have large-scale inhomogeneities, that is $k_0a \gg 1$. Using the expression for Φ^2 in the ray approximation that will be derived in Sec. II B one can determine the upper limits on the phase fluctuations under which the Rytov method can be used to provide estimates for acoustic fluctuations.

The other approximation that was used to obtain Eqs. (15)–(18) is the Fresnel approximation. It assumes the horizontal range L to be much greater than the wavelength and the wave number k_0 to be large. This is clearly seen if one looks at the rigorous derivation of Chernov's equations (99) and (100) from Eq. (89) given in Appendix 1 of Chernov (1960, pp. 155–159). It does not depend on the condition $k_0a \gg 1$, that is, that the inhomogeneities must be large compared to the acoustic wavelength. Equations (15)–(18) will certainly apply for large inhomogeneities but its validity does not depend on this condition. The condition $k_0a \gg 1$ is not imposed by the Fresnel approximation.

The radius a of a typical particle varies in range from less than a micrometer to a few hundreds of a micrometer, the latter being mainly present in coastal waters or close to the sea bottom. Although the upper limit size of suspended particles are not fixed, large heavy particles such as sand or gravel rapidly sink to the bottom. Model calculations of suspended sand concentration profiles near bottom boundary (Adam and Weatherly, 1981, p. 6) indicate that at 30 cm from the boundary it consists completely of fine sand whose radius is $88 \pm 30 \mu\text{m}$. Thorne *et al.* (1993) in their measurements of suspended sediments at 10 and 80 cm above the bed of an estuary obtained particle radii of $85 \pm 20 \mu\text{m}$ and

$78 \pm 20 \mu\text{m}$, respectively. In the upper and mid ocean the majority of particles have a radius less than $2.5 \mu\text{m}$ (McCave, 1975).

Since the range $L \gg a$, the above two integrals [Eqs. (17) and (18)] can be simplified. Following a procedure similar to Chernov (1960, pp. 70–73) who has also considered the case $L \gg a$, the integrals I_1 and I_2 can be written as

$$I_1 = 2\gamma L \int_0^L du_1 \int_0^\infty dy \sin y e^{-b\sqrt{u_1^2 + (2u_1 y/k_0)}} \times \left(1 + \cos s \sqrt{u_1^2 + \frac{2u_1 y}{k_0}}\right), \quad (20)$$

$$I_2 = -k_0 \gamma \int_0^L du_1 \int_0^\infty \rho d\rho \operatorname{Si}\left(\frac{k_0 \rho^2}{4L}\right) e^{-b\sqrt{\rho^2 + u_1^2}} \times (1 + \cos s \sqrt{\rho^2 + u_1^2}), \quad (21)$$

where

$$\int_t^\infty \frac{\sin z}{z} dz = -\operatorname{Si}(t), \quad (22)$$

$\operatorname{Si}(t) = \operatorname{Si}(t) - 0.5\pi$, Si being the sine integral, Abramowitz and Stegun (1970, p. 231).

For particle radii generally present in the ocean, the value of $b > 1000$. Let us therefore estimate these integrals for large b . Following the analysis of Chernov (1960, pp. 72, 73), the integral containing y in I_1 is integrated by parts twice. This gives

$$I_1 = 2\gamma L \int_0^L du_1 \left[N_2(u_1, 0) - \int_0^\infty dy \sin y \frac{\partial^2 N_2(u_1, y)}{\partial y^2} \right], \quad (23)$$

where

$$N_2(u_1, y) = e^{-\sqrt{b^2 u_1^2 + (2u_1 y b^2/k_0)}} \times (1 + \cos s \sqrt{u_1^2 + 2u_1 y/k_0}), \quad (24)$$

$b^2/k_0 \approx (\lambda/a)b/8$. For large b and $\lambda/a \geq 1$, the second term in I_1 which involves the integral is negligible compared to the first term. Thus

$$I_1 \approx 2\gamma L \int_0^L du_1 N_2(u_1, 0). \quad (25)$$

For $L/a \gg 1$, $\lambda/a \geq 1$, the expression $k_0 a^2/4L \ll 1$. Under these conditions and for large b , $I_2 \ll I_1$ (Chernov, 1960, p. 75). Thus

$$\overline{am^2} = \overline{\Delta\phi^2} = k_0^2 \gamma L \int_0^L du_1 N_2(u_1, 0). \quad (26)$$

Evaluating this integral one obtains

$$\overline{am^2} = \overline{\Delta\phi^2} = k_0^2 \gamma L \left[\left(\frac{1 - e^{-bL}}{b} \right) + \left(\frac{b + e^{-bL}(-b \cos sL + s \sin sL)}{b^2 + s^2} \right) \right]. \quad (27)$$

For large b and for relative volume densities of particles found in the ocean, $s^2/b^2 \ll 1$ and Eq. (27) can therefore be approximated as

$$\overline{am^2} = \overline{\Delta\phi^2} = 1.333 k_0^2 a L \sigma (1 - c_m/c_p)^2. \quad (28)$$

B. Mean-square phase fluctuations in the ray approximation

In Sec. II A it was shown that the ray approximation cannot be used to estimate the variance of the acoustic fluctuations from suspended particles in the ocean at megahertz frequencies. However when $L \ll k_0 a^2$, that is in the geometric acoustic region, Ishimaru (1978, Vol. 2, p. 365) noted that the expression for mean-square phase fluctuation is twice that obtained in the Rytov method. This prompts one to obtain the ray approximation result independently to verify the derived Rytov result for acoustic fluctuations.

In the ray approximation,

$$\overline{\Delta\phi^2} = k_0^2 \int_0^L dx' \int_0^L dx'' N(x' - x'', y' - y'', z' - z''). \quad (29)$$

Here the sound travels a distance L in the inhomogeneous medium with the receiver at (L, O, O) (Chernov, 1960, p. 28). For $L \gg 1/b$, following Chernov (1960, p. 29) Eq. (29) can be transformed to,

$$\overline{\Delta\phi^2} = 2k_0^2 \gamma L \int_0^L du_1 e^{-bu_1} (1 + \cos su_1). \quad (30)$$

This can be evaluated to give,

$$\overline{\Delta\phi^2} = 2k_0^2 \gamma L \left[\left(\frac{1 - e^{-bL}}{b} \right) + \left(\frac{b - e^{-bL}(b \cos sL + s \sin sL)}{b^2 + s^2} \right) \right]. \quad (31)$$

This is twice that obtained in Eq. (27). For large b , and for relative volume densities of particles found in the ocean, Eq. (31) can be simplified to

$$\overline{\Delta\phi^2} = 2.666 k_0^2 a L \sigma \left(1 - \frac{c_m}{c_p} \right)^2. \quad (32)$$

Again this is twice that of Eq. (28).

III. DISCUSSION

The simple result [Eq. (28)] shows that both amplitude and phase fluctuations increase with particle radius, path length, the wave number of sound, and the relative volume density of scatterers. It also depends on the ratio of sound speed in the medium to that of the scatterer. Ishimaru (1978, Vol. 2, p. 365) has described general formulas for the amplitude and phase fluctuations, but no analytical estimate is provided. To provide a computable result requires specifying the correlation function for sound-speed fluctuation due to particle scatterers, which he has not done. However, he noted certain characteristics common to all wave fluctuation phenomena irrespective of the form of the correlation function. He noted that if the correlation distance is much smaller than $\sqrt{\lambda L}$, where λ is the wavelength of sound used and L the range, then the log-amplitude and phase fluctuations are equal and are proportional to $k_0^2 L$. This is satisfied by Eq. (28).

The Rytov method is valid in the region $\Phi \leq 1$, where Φ is the root-mean-square phase fluctuation in the ray approximation. Since the variance using the Rytov method is half that obtained using the ray method, Eq. (28) will be valid for $\sigma \leq 1$, if the root-mean-square phase fluctuations obtained is less than 0.7 rad. Equation (28) is therefore valid for large k_0 , $L/a \gg 1$, $L/\lambda > 1$, $\lambda/a \geq 1$ and when the root-mean-square phase fluctuations obtained is less than 0.7 rad. The conditions $L/a \gg 1$ and $\lambda/a \geq 1$ can be written as $L \gg a^2/\lambda$.

An expression of the mean-square log-amplitude and phase fluctuations due to thermal inhomogeneities was derived by Chernov (1960, p. 75), using a Gaussian function with a correlation length a for the case $4L \gg k_0 a^2$. This is given by

$$\overline{am^2} = \overline{\Delta\phi^2} = (\sqrt{\pi}/2) k_0^2 a L \overline{\mu^2}, \quad (33)$$

where $\overline{\mu^2}$ is the mean-square fractional sound-speed fluctuation $[(\delta c/c)^2]$. Equation (33) has been written to show the similarity of the functional dependence on k_0 , a and L of Eqs. (28) and (33). In spite of the different correlation functions used for thermal inhomogeneities and velocity fluctuations due to suspended particles, the acoustic fluctuations are proportional to $k_0^2 L$ (Ishimaru, 1978, Vol. 2, p. 365).

Using the single-particle scattered field, and when the particle size is much less than $\sqrt{\lambda L}$, Ishimaru has provided an approximate solution for the variance of the log-amplitude fluctuation (Ishimaru, 1978, Vol. 1, Eq. 6-106). To evaluate this expression a knowledge of the scattering cross section is required. If the scattering cross section for the case where the particle size is much smaller than the wavelength is used (Ishimaru, Vol. 1, p. 68), a k_0^4 dependence is obtained for the wave number. This is contrary to the expected k_0^2 dependence on wave number.

The expression (28) for the variance of the acoustic fluctuations is due to the velocity variations that arises from suspended particles. If $c_m = c_p$ the variance is zero. If one looks at the expression for the scattered pressure due to density and velocity variations in the method of small perturbations, [Chernov, 1960, p. 43, Eq. (28)], the contribution to scattering by velocity fluctuations has a k^2 dependence while the term due to density fluctuations a factor (k/ρ) . One would therefore expect that at megahertz frequencies the predominant contribution to the variance of the acoustic fluctuation is from the velocity variations of the suspended particles.

IV. CONCLUSION

The paper provides a simple analytical expression for the mean-square log-amplitude and phase fluctuations of high-frequency sound due to velocity variations caused by

suspended particles in the ocean. The expressions are applicable for relative volume density of suspended particles much less than one; for large frequency; for a range much greater than the wavelength of sound (λ) and a^2/λ , where a is the particle radius; and in a region where the root-mean-square phase fluctuations are less than 0.7 rad. The latter condition implies that the region is not strongly fluctuating. In the ocean the condition that the relative volume densities of particles are less than one is usually satisfied. Also the conditions on frequency, range, and root-mean-square phase fluctuations are satisfied for most of the suspended particles found in the ocean at frequencies in the 1–10 MHz region and at ranges greater than a meter.

ACKNOWLEDGMENT

The author wishes to thank Dr. I. S. F. Jones for his valuable comments and assistance throughout the course of this work.

- Abramowitz, M., and Stegun, I. A. (1970). *Handbook of Mathematical Functions* (Dover, New York), p. 231.
- Adam, E. A., and Weatherly, G. L. (1981). *Sedimentary Dynamics of Continental Shelves*, edited by C. A. Nittrouer (Elsevier, New York), p. 6.
- Chernov, L. A. (1960). *Wave Propagation in Random Medium* (McGraw-Hill, New York).
- Duykers, L. R. B. (1967). "Sound attenuation in liquid-solid mixtures," *J. Acoust. Soc. Am.* **41**, 1330–1335.
- Eisma, D. (1993). *Suspended Matter in the Aquatic Environment* (Springer-Verlag, New York), p. 38, 160.
- Flatte, S. M., Dashen, R., Munk, W. H., Watson, K. M., and Zachariasen, F. (1979). *Sound Transmission Through a Fluctuating Ocean* (Cambridge U.P., Cambridge, England), pp. 91, 93.
- Hamilton, E. L., and Bachman, R. T. (1982). "Sound velocity and related properties of marine sediments," *J. Acoust. Soc. Am.* **72**, 1891–1904.
- Ishimaru, A. (1978). *Wave Propagation and Scattering in Random Media* (Academic, New York), Vols. 1 and 2.
- Jones, I. S. F. (1994). *High Resolution Underwater Imaging*, Proceedings of the International Conference in Underwater Acoustics, Sydney (Australian Acoustic Society), p. 40.
- Lal, D. (1977). "Ocean microcosm of particles," *Science* **198**, 997–1009.
- Liebermann, L. J. (1951). "The effect of temperature inhomogeneities in the ocean on the propagation of sound," *J. Acoust. Soc. Am.* **23**, 563–570.
- Liu, K. C. (1991a). "Wave scattering in discrete random media by the discontinuous stochastic field method, I. Basic method and general theory," *J. Sound Vib.* **147**, 301–311.
- Liu, K. C. (1991b). "Wave scattering in discrete random media by the discontinuous stochastic field method, II. Contributions of the second order moment of the β -field," *J. Sound Vib.* **147**, 313–321.
- McCave, I. N. (1975). "Vertical flux of particles in the ocean," *Deep-Sea Res.* **22**, 491–502.
- Thorne, P. D., Hardcastle, P. J., and Soulsby, R. L. (1993). "Analysis of acoustic measurements of suspended sediments," *J. Geophys. Res.* **98**, 899–910.
- Thuraisingham, R. A. (1996). "New theoretical expressions for acoustic fluctuations due to temperature structure of the ocean," *J. Acoust. Soc. Am.* **99**, 1380–1386.
- Thuraisingham, R. A. (1994). "Theoretical estimates of high frequency acoustic attenuation and backscattering from suspended sand particles in the ocean and in an estuary," DSTO Technical Report No. 78.

Excess attenuation of leaky Lamb waves due to viscous fluid loading

Adnan H. Nayfeh and Peter B. Nagy

Department of Aerospace Engineering and Engineering Mechanics, University of Cincinnati, Cincinnati, Ohio 45221-0070

(Received 8 July 1995; accepted for publication 7 January 1997)

In two recent papers [J. Acoust. Soc. Am. **97**, 3191–3193 (1995) and **98**, 1057–1064 (1995)], Zhu and Wu presented an analytical technique to assess the effect of viscous fluid loading on the propagation properties of Rayleigh and Lamb waves in fluid-loaded solids. They modeled the viscous fluid as a hypothetical isotropic solid having rigidity $c_{55} = -i\omega\eta$, where η denotes the viscosity of the fluid and ω is the angular frequency. In this way, the vorticity mode associated with the viscosity of the fluid is formally described as the shear-mode in the fictitious solid. In this paper this technique is further developed by removing certain inconsistencies that unnecessarily reduce the accuracy and the range of validity of Zhu and Wu's results. By properly accounting for viscous effects on the bulk compressional wave in the fluid and applying a rigorous treatment of the field equations and boundary conditions, the exact dispersion equations that are not limited to low frequencies and viscosities are derived. Examples of these results are presented to illustrate the effect of fluid viscosity on the lowest-order symmetric and antisymmetric Lamb modes. One interesting feature revealed by these calculations is the presence of a sharp minimum in the viscosity induced attenuation of the lowest-order symmetric mode of thin plates either immersed in or coated with a viscous fluid. This minimum occurs at a particular frequency where the otherwise elliptical polarization of the surface vibration becomes linearly polarized in the normal direction. © 1997 Acoustical Society of America. [S0001-4966(97)04305-1]

PACS numbers: 43.35.Pt [HEB]

INTRODUCTION

The study of the interaction of elastic waves with fluid-loaded solids has been recognized as a viable means for non-destructive evaluation of solid structures. The reflected acoustic field from a fluid–solid interface has a wealth of information which, if exploited, reveals details of many characteristics of the solid. These include its properties, the possible existence of internal defects, and also the quality of the interface, to name a few. Specifically, the reflected field of a finite width incident acoustic beam shows unusual behavior in that it is nonspecular, it displaces (shifts) along the interface, and contains a trailing edge that attenuates away from the center of reflection. This behavior is widely known as the leaky wave phenomenon. These unusual effects are exaggerated at angles associated with the poles of the reflection and transmission coefficients and also, to a lesser extent, at the branch points of the partial waves. These poles also correspond to the zeros of the characteristic equations for the propagation of free waves along the interfaces.

Theoretical and experimental verifications of these phenomena have been conducted for a wide variety of solids extending from the simple isotropic semi-space to the much more complicated systems of multilayered anisotropic media. A review of the available literature on this subject can be found in a recent book.¹ Most of the published literature dealt with ideal fluids and hence did not account for the possible excess attenuation of leaky waves due to viscous fluid loading. Recently, Qi studied the influence of viscous fluid loading on the propagation of leaky Rayleigh wave in the presence of heat conduction effects.² Subsequently, Wu

and Zhu³ introduced an alternative approach to the treatment of Qi when heat conduction is neglected, they presented solutions for the dispersion relations of leaky Rayleigh waves. More recently, Zhu and Wu⁴ adopted the same method for Lamb waves in submerged and fluid-coated plates.

In a recent paper,⁵ we studied the propagation of longitudinal waves in a multilayered transversely isotropic coaxial fiber system totally immersed in a nonviscous fluid. In a subsequent paper,⁶ we proceeded to investigate the excess attenuation due to viscous fluid loading. As in the case of Wu and Zhu, we accounted for the fluid viscosity by modeling the fluid as a hypothetical solid having shear rigidity $c_{55} = -i\omega\eta$. In this way, the vorticity mode associated with the fluid viscosity is formally described as the shear mode in the hypothetical solid. Recognizing that, for high frequencies, the guided modes in cylindrical fibers asymptotically approach the Rayleigh mode of a flat surface, we attempted to compare our numerical results with those of Wu and Zhu; we found a discrepancy.

In this paper, we derive exact characteristic equations for leaky waves propagating along interfaces of several systems involving isotropic solids loaded with viscous fluids. These include semi-spaces and finite-thickness plates totally immersed in fluids or coated on one or on both sides by finite-thickness fluid layers. The validity of our solutions depends on the appropriateness of the suggested model for fluid viscosity. However, within the limits of this model, our results are general and exact for arbitrary values of viscosity and frequency. These results will be compared with those of Refs. 3 and 4.

I. ANALYSIS

Since we are proceeding under the premise that we can model viscous fluids as hypothetical isotropic solids, to facilitate our subsequent analysis and discussions, we shall first review the relevant field equations pertaining to isotropic solids and then proceed to develop their formal solutions. In this way, formal solutions for viscous fluids can be easily constructed by identification of exchangeable parameters.

A. Field equations

To this end, the two-dimensional motion (in the x - z plane) of an isotropic solid is described by the momentum equations

$$\frac{\partial \sigma_x}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} = \rho \ddot{u}, \quad (1)$$

$$\frac{\partial \sigma_z}{\partial z} + \frac{\partial \sigma_{xz}}{\partial x} = \rho \ddot{w}, \quad (2)$$

and the associated constitutive relations

$$\sigma_x = c_{11} \frac{\partial u}{\partial x} + c_{13} \frac{\partial w}{\partial z}, \quad (3)$$

$$\sigma_z = c_{11} \frac{\partial w}{\partial z} + c_{13} \frac{\partial u}{\partial x}, \quad (4)$$

$$\sigma_{xz} = c_{55} \left(\frac{\partial w}{\partial x} + \frac{\partial u}{\partial z} \right), \quad (5)$$

where isotropy requires that $c_{11} - c_{13} = 2c_{55}$. For conventional solids, $c_{11} = \lambda + 2\mu$ and $c_{13} = \lambda$, and hence $c_{55} = \mu$. As shall be apparent later on, Eqs. (3)–(5) are conveniently written in the conventional format of anisotropic media in order to facilitate description of the role of the fluid viscosity. Equations (1)–(5), when combined, result in two coupled partial differential equations in the displacement components u and w . By introducing the potentials φ and ψ in accordance with

$$u = \frac{\partial \varphi}{\partial x} - \frac{\partial \psi}{\partial z} \quad (6)$$

and

$$w = \frac{\partial \varphi}{\partial z} + \frac{\partial \psi}{\partial x}, \quad (7)$$

these equations uncouple into two systems as follows:

$$c_{11} \left(\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial z^2} \right) = \rho \ddot{\varphi} \quad (8)$$

and

$$c_{55} \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial z^2} \right) = \rho \ddot{\psi}. \quad (9)$$

Using the stress-strain relation (3)–(5) yields the following expressions for the stress components in terms of the potentials:

$$\sigma_z = c_{11} \left(\frac{\partial^2 \varphi}{\partial z^2} + \frac{\partial^2 \psi}{\partial x \partial z} \right) + c_{13} \left(\frac{\partial^2 \varphi}{\partial x^2} - \frac{\partial^2 \psi}{\partial x \partial z} \right), \quad (10)$$

$$\sigma_{xz} = c_{55} \left(2 \frac{\partial^2 \varphi}{\partial x \partial z} + \frac{\partial^2 \psi}{\partial x^2} - \frac{\partial^2 \psi}{\partial z^2} \right). \quad (11)$$

B. Formal solutions

For harmonic waves propagating along the x direction, the potentials admit the formal solutions

$$\varphi = (A e^{-qz} + \bar{A} e^{qz}) e^{ikx - i\omega t}, \quad (12)$$

$$\psi = (B e^{-sz} + \bar{B} e^{sz}) e^{ikx - i\omega t}, \quad (13)$$

each consisting of two pairs of partial waves propagating along the positive and negative z directions, respectively, with A , \bar{A} , B , and \bar{B} being unknown wave amplitudes. Here,

$$q^2 = k^2 - k_l^2, \quad s^2 = k^2 - k_t^2, \quad k_l^2 = \frac{\omega^2 \rho}{c_{11}}, \quad \text{and} \quad k_t^2 = \frac{\omega^2 \rho}{c_{55}}. \quad (14)$$

These forms of formal solutions are appropriate for the description of propagation in infinite media.

For propagation in semi-spaces, we choose the partial waves that satisfy the radiation condition that all field variables be bounded for locations deep within the solid. Thus, for propagation in the $z \geq 0$ semi-space, we choose the formal solutions

$$\varphi = A e^{-qz} e^{ikx - i\omega t} \quad (15)$$

and

$$\psi = B e^{-sz} e^{ikx - i\omega t}. \quad (16)$$

An alternative choice of the total solutions (12) and (13) are

$$\varphi = (A_1 \cosh qz + A_2 \sinh qz) e^{ikx - i\omega t}, \quad (17)$$

$$\psi = (B_1 \cosh sz + B_2 \sinh sz) e^{ikx - i\omega t}. \quad (18)$$

These are more appropriate for finite thickness media. Here, once again, A_1 , A_2 , B_1 , and B_2 are unknown constants.

For the motion of a semi-space substitution from Eqs. (15) and (16) into Eqs. (5) and (6) leads to

$$u = ikA e^{-qz} + sB e^{-sz}, \quad (19)$$

$$w = -qA e^{-qz} + ikB e^{-sz}. \quad (20)$$

In a similar manner we obtain

$$\sigma_z = c_{55} [(k^2 + s^2)A e^{-qz} - 2iskB e^{-sz}], \quad (21)$$

$$\sigma_{xz} = -c_{55} [2iqkA e^{-qz} + (k^2 + s^2)B e^{-sz}]. \quad (22)$$

Here and in subsequent definitions, we note and suppress the common factor $e^{ikx - i\omega t}$. For the bounded media, on the other hand, Eqs. (17) and (18), when used in Eqs. (6)–(10), lead to the displacement and stress components

$$u = ik(A_1 \cosh qz + A_2 \sinh qz) - s(B_1 \sinh sz + B_2 \cosh sz), \quad (23)$$

$$w = q(A_1 \sinh qz + A_2 \cosh qz) + ik(B_1 \cosh sz + B_2 \sinh sz), \quad (24)$$

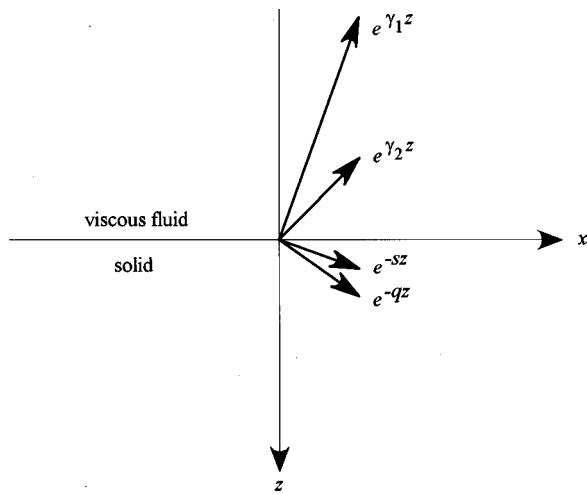


FIG. 1. Coordinate system for leaky Rayleigh wave propagation on an infinite solid/fluid boundary.

$$\sigma_z = c_{55}[(k^2 + s^2)(A_1 \cosh qz + A_2 \sinh qz) + 2isk(B_1 \sinh sz + B_2 \cosh sz)], \quad (25)$$

$$\sigma_{xz} = c_{55}[2ikq(A_1 \sinh qz + A_2 \cosh qz) - (k^2 + s^2)(B_1 \cosh sz + B_2 \sinh sz)]. \quad (26)$$

C. Modeling of viscous fluids

In order to differentiate between properties belonging to the fluid and to the solid, we mark the properties of the fluid by a prime. In modeling the viscous fluid as a hypothetical solid whose shear rigidity equals $-i\omega\eta$, we essentially set $c'_{55} = -i\omega\eta$. For the description of the isotropic fluid, we need another property. This will be λ_f , the compressibility of the viscosity-free fluid. It now remains to identify the values of c'_{11} and c'_{13} in terms of λ_f and η . In the absence of physical or experimental specifications, isotropy of the fluid allows several possible situations. One possibility is to model the viscous fluid, not only as a hypothetical solid but, as a conventional one. This implies that $c'_{11} = \lambda_f - 2i\omega\eta$ and $c'_{13} = \lambda_f$, i.e., longitudinal waves in the viscous fluid become dispersive and attenuated. This is in contrast to the choice suggested by Wu and Zhu in which longitudinal waves are neither dispersive or attenuative and propagate with the speed $c_f = \sqrt{\lambda_f/\rho_f}$ of longitudinal waves in the nonviscous fluid. Wu and Zhu's choice represents a second possible model in which $c'_{11} = \lambda_f$ as λ_f and $c'_{13} = \lambda_f + 2i\omega\eta$. As a third possibility, we suggest a splitting of the viscosity parameter between c'_{11} and c'_{13} . For this purpose, we adopt the so-called Stokes model for viscous fluids by choosing $c'_{11} = \lambda_f - (4/3)i\omega\eta$ and $c'_{13} = \lambda_f + (2/3)i\omega\eta$ which also lead to dispersive and attenuative effects in the longitudinal wave behavior in accordance with the classical approximation⁶

$$\alpha_L = \text{Im} \left\{ \omega \sqrt{\frac{\rho_f}{\lambda_f - 4i\omega\eta/3}} \right\} \approx \frac{2\omega^2\eta}{3c_f^3\rho_f}. \quad (27)$$

We should however emphasize that, regardless of the specific choice of the viscosity model, as long as the fluid is isotropic, all of the formal solutions obtained above can be

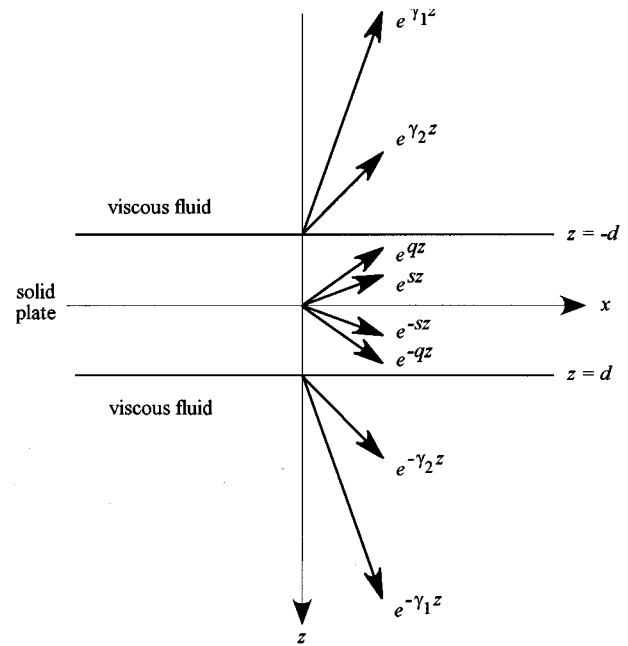


FIG. 2. Coordinate system for leaky Lamb wave propagation in an immersed plate.

also used in their entirety for the fictitious solid modeling the viscous fluid. However, their numerical results can slightly vary depending upon the specific model of viscosity chosen.

II. DISPERSION EQUATIONS

In the following section we shall apply the above formal solutions to the study of specific situations in which the fluid is viscous. These include the derivation of the characteristic dispersion relations for leaky Rayleigh and leaky Lamb waves propagating on fully immersed solid half-spaces and plates, respectively, as well as for nonleaky Lamb waves propagating in a solid plate coated on one or on both sides by finite-thickness viscous fluid layers.

A. Leaky Rayleigh waves

Consider the situation illustrated schematically in Fig. 1. It consists of two semi-spaces; the upper one is a viscous fluid and the lower one is a solid. Since the solid extends in the positive z direction, its partial waves are represented by the two components shown on the sketch and are consistent with the formal solutions given in Eqs. (15) and (16). According to this partial wave identification criterion, in contrast, the two partial waves representing the fluid are then those shown on the same sketch subject to the definitions

$$\gamma_1 = \sqrt{k^2 - k_f^2} = \sqrt{k^2 - \frac{\omega^2\rho_f}{c'_{11}}}, \quad \gamma_2 = \sqrt{k^2 - \frac{i\omega\rho_f}{\eta}}. \quad (28)$$

It is a trivial matter then to write explicit expressions for the displacement and stress components of the fluid in terms of these partial waves. This can be done by applying Eqs. (19)–(22) subject to identifying γ_1 with $-q$ and γ_2 with $-s$ and introducing the two fluid wave amplitudes C and D . By applying the required interface conditions that the stress and

displacement components be continuous across the interface and subsequently requiring nontrivial values for the partial wave amplitudes A , B , C , and D , we arrive at the full characteristic dispersion relation

$$\begin{vmatrix} k^2 + s^2 & -2iks & -r(2k^2 + \gamma_2^2) & -2irk\gamma_2 \\ -2ikq & -(k^2 + s^2) & -2irk\gamma_1 & r(k^2 + \gamma_2^2) \\ -q & ik & -\gamma_1 & -ik \\ ik & s & -ik & \gamma_2 \end{vmatrix} = 0, \quad (29)$$

where $r = c'_{55}/c_{55} = \mu'/\mu = -i\omega\eta/\mu$.

This equation can now be compared with Eq. (8) of Ref. 3 as $h \rightarrow \infty$. Close examination reveals that the results of Wu and Zhu are consistent with our exact solution only for very low viscosity and frequency. This can also explain the unusual choice of their formal solutions in Eq. (7). In particular, their choice of the formal solution of ψ_l does not satisfy their potential Eq. (6). Also, their choice of the solution for φ_l suggests that the potential itself vanishes at the free surface $z=h$ rather than the vanishing of both the normal and shear stress components. However, it seems that all of these discrepancies can be attributed to the low viscosity assumption,

although the authors nowhere made clear this was what they were investigating.

B. Leaky Lamb waves

Next we consider an isotropic solid plate completely immersed in a viscous fluid. The situation is illustrated schematically in Fig. 2. The plate thickness is $2d$ and thus the lower and upper portions of the fluid extend from $z=d$ to ∞ and from $z=-d$ to $-\infty$, respectively. Also shown in this figure are the partial waves in both the plate and the fluid. The appropriate formal solutions for the plate are those given by Eqs. (23)–(26). For the upper fluid, the required formal solutions are those used earlier while investigating the Rayleigh problem, with the exception of replacing z with $z-d$. Due to the symmetry of the loading, we do not need to list formal solutions for the lower portion of the fluid.

Having identified the formal solutions in both the plate and the fluid, we now proceed to treat symmetric and antisymmetric motions of the system separately. For symmetric motion, the displacement and stress components w and σ_{xz} must vanish at the center of the plate. From these conditions, A_2 and B_1 vanish. If this is followed by satisfying the interface conditions at $z=-d$ we finally obtain the dispersion characteristic equation as

$$\begin{vmatrix} (k^2 + s^2)\cosh qd & 2isk \cosh sd & -r(k^2 + \gamma_2^2) & -2rik\gamma_2 \\ 2ikq \sinh qd & -(k^2 + s^2)\sinh sd & -2rik\gamma_1 & r(k^2 + \gamma_2^2) \\ q \sinh qd & ik \sinh sd & -\gamma_1 & -ik \\ ik \cosh qd & -s \cosh sd & -ik & \gamma_2 \end{vmatrix} = 0. \quad (30)$$

For the antisymmetric motion, the displacement and stress components u and σ_z vanish at the center of the plate. These require the vanishing of the wave amplitudes A_1 and B_2 . By inspection, we see that the characteristic equation for this case can be obtained by replacing $\cosh qd$ and $\cosh sd$ with $\sinh qd$ and $\sinh sd$, respectively, resulting in

$$\begin{vmatrix} (k^2 + s^2)\sinh qd & 2isk \sinh sd & -r(k^2 + \gamma_2^2) & -2rik\gamma_2 \\ 2ikq \cosh qd & -(k^2 + s^2)\cosh sd & -2rik\gamma_1 & r(k^2 + \gamma_2^2) \\ q \cosh qd & ik \cosh sd & -\gamma_1 & -ik \\ ik \sinh qd & -s \sinh sd & -ik & \gamma_2 \end{vmatrix} = 0. \quad (31)$$

C. Layered systems

We now proceed to derive characteristic dispersion relations for layered systems; these consist of a solid core layer coated on one or on both sides by viscous fluid layers of finite thicknesses. These define bilayered and trilayered systems, respectively. The derivation procedure of their dispersion relations proceeds as follows: In accordance with Eqs. (23)–(26), formal solutions in each layer consist of the superposition of four partial waves written in terms of four unknown potential wave amplitudes. The characteristic dispersion relation results from satisfying the appropriate interface and boundary conditions. These include the continuity of the displacement and stress components at the interfaces and the vanishing of the stresses at the outer boundaries. For

the bilayered system, this will be in the form of the vanishing of an 8×8 determinant relation. Then it follows that the vanishing of a 12×12 determinant relation is required for the trilayered system. The resulting expressions, although tractable, nevertheless may be less attractive for numerical illustrations.

As an alternative, we use the elegant matrix transfer method, originally introduced by Thomson⁷ and since then widely used in many applications (see the recent review of its application in Ref. 1). According to this method, the four wave amplitudes of each layer are first eliminated by specializing the formal solutions to its two outer faces. As a consequence, the field variables on one side of the layer are written in terms of those at the other side via a 4×4 local transfer

matrix. A 4×4 global transfer matrix for the total system of layers is then constructed by extending the solution from one layer to the next while satisfying the appropriate interfacial continuity conditions. In so doing, we ultimately can construct the characteristic dispersion relation for any number of layers in terms of the vanishing 2×2 determinant equation.

To facilitate our discussion, we consider the geometrical arrangement of Fig. 3 as a typical example of a layered system. For each layer, we introduce a local coordinate system with origin at its upper surface. Thus, layer one extends from $z_1=0$ to h , layer two from $z_2=0$ to $2d$, and layer three from $z_3=0$ to h . Field variables at the upper and lower surfaces of each layer are also identified with $+$ superscripts, respectively. We can take advantage and introduce the method for any number of layers. In general, we then let layer k extend from $z_k=0$ to t_k , where t_k is its thickness.

Next, we adopt the formal solutions (12) and (13) for

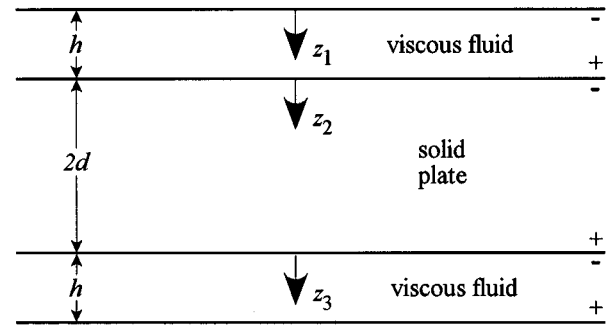


FIG. 3. Coordinate system for the nonleaky Lamb wave propagation in a fluid-coated plate.

each layer with the four unknown amplitudes A , \bar{A} , B , and \bar{B} . Using these, we construct the displacements and stresses in accordance with Eqs. (6) and (7) and (10) and (11), respectively, and get for each layer k

$$\begin{Bmatrix} u \\ w \\ \sigma_z \\ \sigma_{xz} \end{Bmatrix}_k = \begin{pmatrix} ik & ik & s & -s \\ -q & q & ik & ik \\ \mu(k^2+s^2) & \mu(k^2+s^2) & -2i\mu ks & 2i\mu ks \\ -2i\mu kq & 2i\mu kq & -\mu(k^2+s^2) & -\mu(k^2+s^2) \end{pmatrix}_k \begin{Bmatrix} Ae^{-qz} \\ \bar{A}e^{qz} \\ Be^{-sz} \\ \bar{B}e^{sz} \end{Bmatrix}_k, \quad (32)$$

where we substituted μ for c_{55} . Equation (32) is supplemented with the interfacial continuity conditions

$$\begin{Bmatrix} u \\ w \\ \sigma_z \\ \sigma_{xz} \end{Bmatrix}_k = \begin{Bmatrix} u \\ w \\ \sigma_z \\ \sigma_{xz} \end{Bmatrix}_{k+1} \quad (33)$$

at the interface between layers k and $k+1$.

To facilitate the subsequent analysis, Eq. (32) is rewritten in the compact matrix form

$$P_k = X_k D_k U_k, \quad (34)$$

where P_k is the column matrix of the displacement and stress components, X_k is the 4×4 matrix in Eq. (32), D_k is the 4×4 diagonal matrix

$$D_k = \begin{pmatrix} e^{-qz} & 0 & 0 & 0 \\ 0 & e^{qz} & 0 & 0 \\ 0 & 0 & e^{-sz} & 0 \\ 0 & 0 & 0 & e^{sz} \end{pmatrix}, \quad (35)$$

and U_k is the column matrix of the amplitudes A , \bar{A} , B , and \bar{B} . Specializing Eq. (34) to $z_k=0$ and $z_k=t_k$, noting that $D_k^- = I$, and that X_k and U_k are independent of z , we get

$$P_k^- = X_k U_k, \quad (36)$$

$$P_k^+ = X_k D_k^+ U_k. \quad (37)$$

By eliminating the common column amplitude U_k we write

$$P_k^+ = A_k P_k^-, \quad (38)$$

where

$$A_k = X_k D_k^+ X_k^{-1}, \quad (39)$$

defining matrix A_k as the local transfer matrix.

Invoking the continuity conditions (33) across each of the interfaces, we relate the displacement and stress components at one surface of the total system to the other via the global transfer matrix relation

$$P_n^+ = A P_1^-, \quad (40)$$

where $A = A_n A_{n-1} \cdots A_1$.

For future reference, Eq. (40) is now written in the expanded matrix form

$$\begin{Bmatrix} u \\ w \\ \sigma_z \\ \sigma_{xz} \end{Bmatrix}_n^+ = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix} \begin{Bmatrix} u \\ w \\ \sigma_z \\ \sigma_{xz} \end{Bmatrix}_1^-. \quad (41)$$

The dispersion relation is obtained from Eq. (41) by requiring the vanishing of the stress components σ_z and σ_{xz} on both faces, leading to

$$A_{31}A_{42} - A_{41}A_{32} = 0. \quad (42)$$

For the two-layer system (namely, the case of solid plate coated on one side only) we have

$$P_2^+ = A_2 A_1 P_1^- \quad (43)$$

and thus identify A with $A_1 A_2$. Similarly, for the three-layer system

$$P_3^+ = A_3 A_2 A_1 P_1^- \quad (44)$$

and A is now identified as $A_1 A_2 A_3$. For either of these two systems, since layer 2 defines the solid, we recognize that X_2 is the 4×4 matrix in Eq. (32) with

$$X_3 = X_1 = \begin{pmatrix} ik & ik & \gamma_2 & \gamma_2 \\ -\gamma_1 & \gamma_1 & ik & ik \\ \mu'(k^2 + \gamma_2^2) & \mu'(k^2 + \gamma_2^2) & -2i\mu'k\gamma_2 & 2i\mu'k\gamma_2 \\ -2i\mu'k\gamma_1 & 2i\mu'k\gamma_1 & -\mu'(k^2 + \gamma_2^2) & \mu'(k^2 + \gamma_2^2) \end{pmatrix} \quad (46)$$

and

$$D_1^+ = D_3^+ = \begin{pmatrix} e^{-\gamma_1 h} & 0 & 0 & 0 \\ 0 & e^{\gamma_1 h} & 0 & 0 \\ 0 & 0 & e^{-\gamma_2 h} & 0 \\ 0 & 0 & 0 & e^{\gamma_2 h} \end{pmatrix}. \quad (47)$$

For the present trilayered model, the characteristic Eq. (42) contains results pertaining to both symmetric and anti-symmetric modes. We can however identify the conditions which allow us to derive the dispersion relations for symmetric and antisymmetric modes separately. These correspond to the vanishing at the center of the plate of w and σ_{xz} for the symmetric modes and of u and σ_z for the antisymmetric ones. This suggests constructing the global transfer matrix for half of the system which consists of layer 1 and half of layer 2 in accordance with

$$P_2^m = A_2^m A_1 P_1^- = B P_1^-, \quad (48)$$

with $B = A_2^m A_1$, where $A_2^m = X_2 D_2^m X_2^{-1}$ and

$$D_2^m = \begin{pmatrix} e^{-qd} & 0 & 0 & 0 \\ 0 & e^{qd} & 0 & 0 \\ 0 & 0 & e^{-sd} & 0 \\ 0 & 0 & 0 & e^{sd} \end{pmatrix}, \quad (49)$$

where the superscript m designates quantities defined at the middle of the plate. Equation (48) can be expanded as

$$\begin{Bmatrix} u \\ w \\ \sigma_z \\ \sigma_{xz} \end{Bmatrix}_2^m = \begin{pmatrix} B_{11} & B_{12} & B_{13} & B_{14} \\ B_{21} & B_{22} & B_{23} & B_{24} \\ B_{31} & B_{32} & B_{33} & B_{34} \\ B_{41} & B_{42} & B_{43} & B_{44} \end{pmatrix} \begin{Bmatrix} u \\ w \\ \sigma_z \\ \sigma_{xz} \end{Bmatrix}_1^- \quad (50)$$

For antisymmetric modes, $u_2^m = \sigma_{z2}^m = \sigma_{z1}^- = \sigma_{xz1}^- = 0$ yielding

$$B_{11} B_{32} - B_{12} B_{31} = 0. \quad (51)$$

Then, for symmetric modes it follows with $w_2^m = \sigma_{xz2}^m = 0$ rather than u_2^m and σ_{z2}^m that

$$B_{21} B_{42} - B_{22} B_{41} = 0. \quad (52)$$

$$D_2^+ = \begin{pmatrix} e^{-2qd} & 0 & 0 & 0 \\ 0 & e^{2qd} & 0 & 0 \\ 0 & 0 & e^{-2sd} & 0 \\ 0 & 0 & 0 & e^{2sd} \end{pmatrix}. \quad (45)$$

For the fluid, we have

D. Low-frequency approximation

Our analytical model allows us to numerically determine the frequency-dependent velocity and attenuation of Lamb waves propagating in plates either immersed in or coated with a viscous fluid. In spite of all the advantages of numerically determined but otherwise ‘‘exact’’ solutions, wave propagation phenomena can be often better understood from simpler approximations that capture the dominating physical mechanism in a given region. In order to gain better insight into the viscosity induced attenuation phenomenon, we shall shortly consider the lowest-order symmetric mode at low frequencies. Separate asymptotic approximations can be obtained for infinitely thick (immersed) and very thin fluid coatings. It should be mentioned that at high frequencies, both symmetric and antisymmetric fundamental modes asymptotically approach the leaky Rayleigh mode discussed elsewhere.⁶

First, let us consider the lowest-order symmetric mode in an immersed plate. In the low-frequency approximation, the vibration amplitude produced by the fundamental symmetric mode is constant throughout the cross section of the plate. Physically, this mode is analogous to the lowest-order axi-symmetric mode in a thin rod, which can be exploited to calculate the viscosity-induced attenuation of a thin plate from that of a thin rod. For an immersed rod, the viscosity-induced attenuation is⁸

$$\alpha_v = \frac{1}{a} \sqrt{\frac{\omega \eta \rho_f}{2E\rho}}. \quad (53)$$

Here, a is the radius of the rod, ω is the angular frequency, η and ρ_f denote the viscosity and density of the fluid, respectively, and E and ρ denote Young’s modulus and the density in the solid, respectively. For rods coated with a thin fluid layer, the viscosity induced attenuation is⁶

$$\alpha_v = \frac{\omega^2 \rho_f^2 h^3}{3 \eta a \sqrt{E\rho}}, \quad (54)$$

where h denotes the thickness of the coating.

There are only two significant differences between the lowest-order symmetric modes in a thin plate and rod. First,

the effective stiffness $E = c_{11} - (c_{11} - 2c_{55})^2 / (c_{11} - c_{55})$ of a rod is somewhat lower than the effective stiffness $E_p = c_{11} - (c_{11} - 2c_{55})^2 / c_{11}$ of a plate that exhibits Poisson's effect only in one of the lateral dimensions. Second, the crucial ratio between the cross-sectional area and the circumference in contact with the viscous fluid is $a/2$ (i.e., half the radius) for rods and d (i.e., half the thickness) for plates. Due to these differences, the corresponding equations for thin plates can be directly obtained from the above equations by $E_p \rightarrow E$ and $2d \rightarrow a$ substitutions. For immersed plates, the low-frequency asymptote of the viscosity-induced attenuation is

$$\alpha_v = \frac{1}{2d} \sqrt{\frac{\omega \eta \rho_f}{2E_p \rho}}. \quad (55)$$

Similarly, for a coated plate, the low-frequency asymptote of the viscosity-induced attenuation is

$$\alpha_v = \frac{\omega^2 \rho_f^2 h^3}{6 \eta d \sqrt{E_p \rho}}. \quad (56)$$

III. NUMERICAL RESULTS

In the following, we present numerical examples to illustrate the effect of fluid viscosity on the lowest-order symmetric and antisymmetric Lamb modes only, but similar results can be easily obtained by our analytical technique for any of the higher-order modes. The material parameters used in these calculations are $c_{11} = 4.258 \times 10^{11}$ N/m², $c_{55} = 1.675 \times 10^{11}$ N/m², and $\rho = 3.1 \times 10^3$ kg/m³ for silicon carbide and $\lambda_f = 2.25 \times 10^9$ N/m² and $\rho_f = 10^3$ kg/m³ for water. In order to demonstrate important but otherwise rather weak effects, we assumed that the viscosity of the fictitious fluid referred to as "viscous" water changes over four orders of magnitude between 10^1 and 10^{-3} kg/ms (the viscosity of ordinary distilled water is $\eta = 10^{-3}$ kg/ms at room temperature).

The effect of viscosity on the phase velocity of the Lamb modes is usually negligible with respect to the effect on attenuation. Therefore, in the following examples we show only the attenuation effect which can be interpreted in two different ways depending on the physical nature of the system. In the case of plates coated with a finite-thickness fluid layer the Lamb modes are not leaky and viscosity is the sole source of attenuation. The viscosity-induced attenuation then can be directly evaluated from the imaginary part of the complex wave number calculated from the dispersion equation at different frequencies. In contrast, in the case of plates immersed in an infinite fluid bath, the Lamb modes become leaky and the attenuation caused by fluid loading is due to the combined effects of radiation loss due to energy leakage into the fluid and dissipative loss due to viscous friction at the interfaces. Since we cannot separately calculate the viscosity-induced part of the attenuation, we shall define it as the difference between the respective attenuation coefficients caused by immersion in otherwise identical viscous and viscosity-free fluids. However, physically, the leaky and viscous attenuation contributions cannot be simply added together to obtain the total attenuation as strong viscosity can

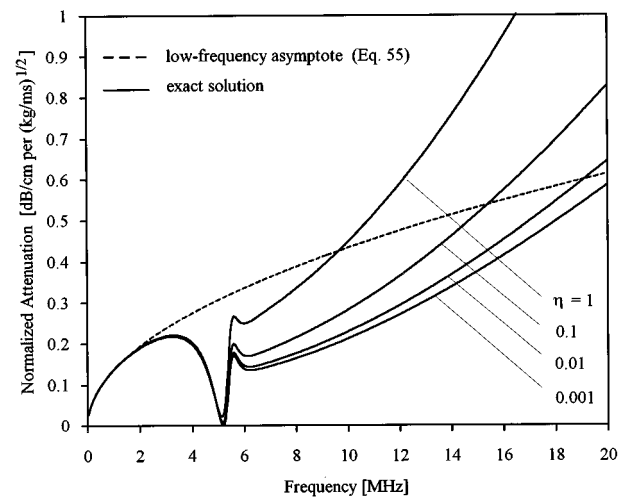


FIG. 4. Normalized "viscosity-induced" attenuation ($\alpha_v/\eta^{1/2}$) versus frequency for the lowest-order symmetric mode in a 1-mm-thick immersed SiC plate.

significantly affect the phase velocity and vibration distribution of a certain mode thereby slightly changing the radiation loss itself. In some cases, the resulting relatively small drop in the very strong radiation loss can overshadow the viscous dissipation and the total attenuation is actually lower for the viscous fluid than for the viscosity-free one. In these rare cases, the relative attenuation caused by viscosity is negative, which was earlier noticed by *Qi*.² In order to distinguish between the real viscosity-induced attenuation of fluid-coated plates from the somewhat artificially defined relative attenuation of immersed plates, in the latter case we shall use the "viscosity-induced" term with quotation marks.

As the first example, Fig. 4 shows the normalized "viscosity-induced" attenuation ($\alpha_v/\eta^{1/2}$) as a function of frequency for the lowest-order symmetric mode in a 1-mm-thick immersed SiC plate. In order to bring out the universal low-frequency asymptotic behavior of the attenuation curves [see Eq. (55)], the attenuation was normalized to the square root of viscosity. Up to about 2 MHz, the "viscosity-induced" attenuation is proportional to the square root of frequency; then it starts to decrease and goes through a sharp minimum at around 5.2 MHz. The position of this minimum is essentially independent of viscosity and depends only on the properties of the solid plate.

Similar behavior can be observed in the viscosity-induced attenuation of the same mode in fluid-coated plates. Figure 5 shows the frequency dependence of the viscosity-induced attenuation for the lowest-order symmetric mode in a 1-mm-thick SiC plate coated with fluid layers of different thicknesses between $3 \mu\text{m}$ and $20 \mu\text{m}$ for a given viscosity of $\eta = 0.1$ kg/ms. An interesting change can be observed at very low frequencies as the coating thickness decreases. At $f = 2$ MHz, the viscous skin depth in the fluid $\delta = \sqrt{2 \eta / \omega \rho_f} \approx 8 \mu\text{m}$ is considerably less than the depth of the thickest coating ($h = 20 \mu\text{m}$), but larger than that of the thinnest one ($h = 3 \mu\text{m}$). As a result, the frequency dependence below 2 MHz smoothly changes from the square-root behavior [Eq. (55)] of immersed plates to the square behavior [Eq. (56)] of fluid-coated plates. The sharp minimum at around 5.2 MHz

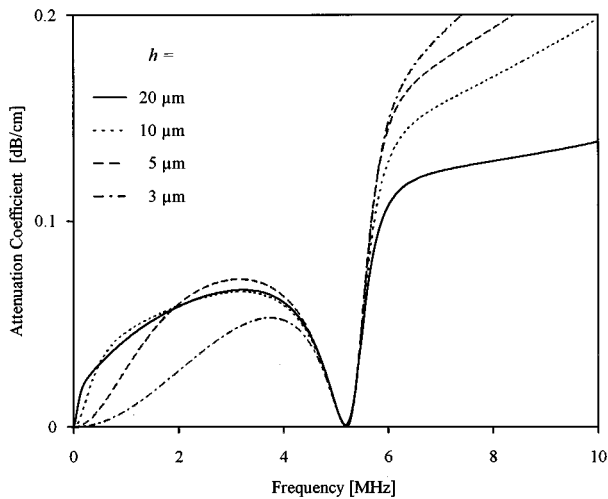


FIG. 5. Viscosity-induced attenuation versus frequency for the lowest-order symmetric mode in a 1-mm-thick SiC plate coated with fluid layers ($\eta = 0.1$ kg/ms) of different thicknesses.

is apparently unaffected by the changing fluid-thickness. This interesting feature is retained even for thin coatings of very high viscosity. Figure 6 shows the normalized viscosity-induced attenuation ($\alpha_v \eta$) versus frequency for the lowest-order symmetric mode in a 1-mm-thick SiC plate coated with 10- μ m-thick fluid layers of different viscosities. Again, in order to bring out the universal low-frequency asymptotic behavior of the attenuation curves [see Eq. (56)], the attenuation was normalized, but this time by multiplying it with the viscosity. It is somewhat surprising that, in contrast with immersion in an infinite fluid bath, for a thin coating the viscosity-induced attenuation decreases with increasing viscosity. The reason for this is that, in the case of a thin coating, increasing viscosity assures that the whole fluid layer moves in union with the plate. In this way, the fluid coating represents a significant mass loading but essentially no friction losses occur.

The above results suggest that the observed minimum in

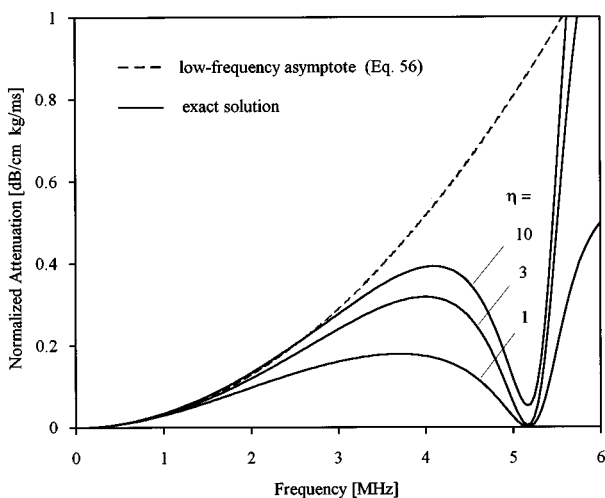


FIG. 6. Normalized viscosity-induced attenuation ($\alpha_v \eta$) versus frequency for the lowest-order symmetric mode in a 1-mm-thick SiC plate coated with 10- μ m-thick fluid layers of different viscosities.

the viscosity-induced attenuation of the fundamental symmetric Lamb mode must be associated with the changing vibration pattern of the lowest-order longitudinal mode in that particular frequency region. At very low frequencies, the surface vibration of this mode is essentially linearly polarized in the propagation direction, i.e., there is but a negligible normal vibration component at the surface with respect to the tangential one. As the frequency increases, the so-called Poisson effect produces more and more transverse motion (i.e., perpendicular to the propagation direction) and the surface vibration becomes elliptically polarized with clockwise rotation. At very high frequencies, the same mode asymptotically approaches a Rayleigh-type surface mode that also exhibits elliptical surface vibration polarization but with counterclockwise rotation. Obviously, there must be a point in the transition region where the surface vibration polarization becomes purely linearly polarized in the transverse direction. At that particular frequency, there is only normal surface vibration as the tangential component completely vanishes. This disappearance of the tangential vibration component and the associated minimum in the viscosity-induced attenuation coefficient has been previously observed in the fundamental axi-symmetric mode of a cylindrical rod.⁶

The ratio of the tangential and normal vibration amplitudes at the surface can be written as follows

$$\frac{u(d)}{w(d)} = \frac{Q_{41}Q_{12} - Q_{42}Q_{11}}{Q_{31}Q_{12} - Q_{32}Q_{11}}, \quad (57)$$

where Q denotes the characteristic matrix previously given in Eq. (30). Pure normal polarization occurs when the numerator $sk_t^2 \cosh sa \cosh qa$ vanishes. Since $k > k_t$, q is always real and $\cosh qa$ never vanishes. On the other hand, $k < k_s$ except for very high frequencies therefore s is pure imaginary. As a result, $\cosh sa$ vanishes when $sa = i\pi/2$, i.e., when the wave number reaches a given value of

$$k_0 = \sqrt{k_t^2 - \left(\frac{\pi}{2d}\right)^2}. \quad (58)$$

Combined with the dispersion relation $k(\omega)$, Eq. (58) represents an implicit equation for the frequency where the viscosity-induced attenuation exhibits a sharp minimum due to the disappearance of the tangential surface vibration component. Figure 7 shows the ratio of the tangential and normal surface vibration amplitudes as a function of frequency for a 1-mm-thick free SiC plate. This ratio is pure imaginary at all frequencies, which results in the previously mentioned elliptical polarization. The tangential surface vibration vanishes exactly at the location of the previously observed minimum of the viscosity induced attenuation, i.e., at 5.2 MHz.

It should be mentioned that Zhu and Wu recently reported the presence of apparently similar sharp minima in the viscosity-induced attenuation spectra of fluid-coated plates,⁴ but those minima are associated with resonance vibrations of the fluid layer and the resonance frequencies are determined by the thickness and material properties of the fluid rather than by those of the solid. We have also observed such resonance behavior when the thickness of the fluid coating is

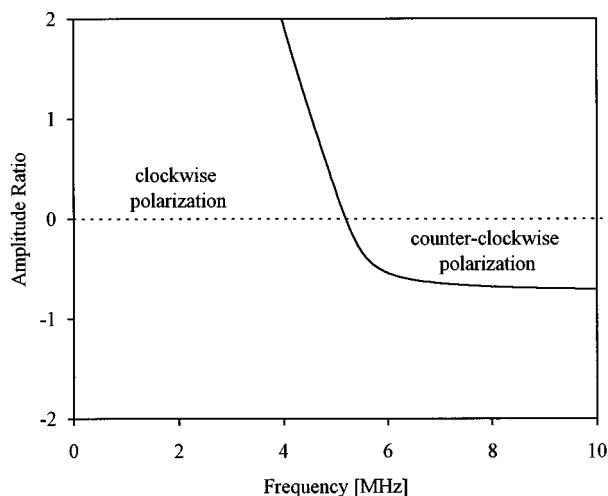


FIG. 7. The ratio of the tangential and normal surface vibration amplitudes of the lowest-order symmetric-Lamb mode as a function of frequency for a 1-mm-thick free SiC plate. Pure normal surface vibration occurs at 52 MHz.

neither very large nor very small with respect to the plate thickness, however, the discussion of these coupled resonances is beyond the scope of this paper.

Figure 8 shows the normalized “viscosity-induced” attenuation ($\alpha_v/\eta^{1/2}$) as a function of frequency for the lowest order antisymmetric mode in a 1-mm-thick immersed SiC plate. Again, in order to bring out the universal low-frequency asymptotic behavior of the attenuation curves, the attenuation was normalized to the square root of viscosity. Due to the large transverse vibration produced by this fundamental flexural mode, there is a large radiation loss that is actually slightly reduced by the viscous friction between the plate and the fluid. The reduced radiation loss more than compensates for the dissipative loss so that the “viscosity-induced” attenuation is negative below approximately 2 MHz. At higher frequencies, the dissipative nature of the viscosity starts to dominate and the “viscosity-induced” attenuation becomes positive. Figure 9 shows the viscosity-

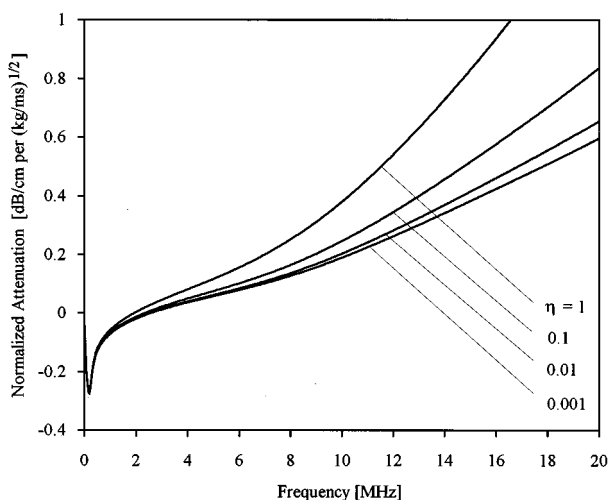


FIG. 8. Normalized “viscosity-induced” attenuation ($\alpha_v/\eta^{1/2}$) versus frequency for the lowest-order antisymmetric mode in a 1-mm-thick immersed SiC plate.

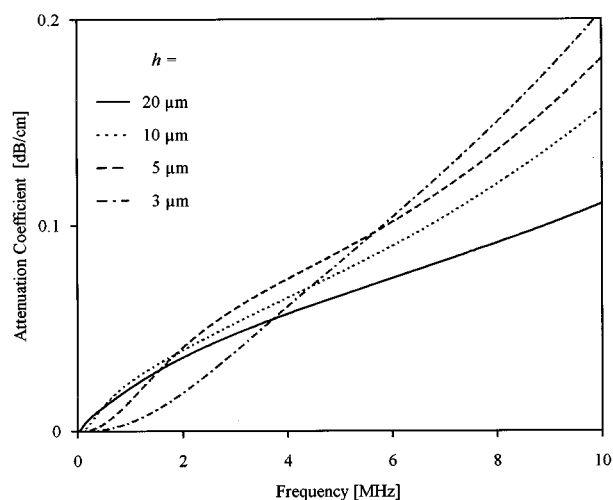


FIG. 9. Viscosity induced attenuation versus frequency for the lowest-order antisymmetric mode in a 1-mm-thick SiC plate coated with fluid layers ($\eta=0.1$ kg/ms) of different thicknesses.

induced attenuation as a function of frequency for the lowest-order antisymmetric mode in a 1-mm-thick SiC plate coated with fluid layers of different thicknesses between 3 μm and 20 μm for a given viscosity of $\eta=0.1$ kg/ms. As we pointed out in connection with the fundamental symmetric mode, the frequency dependence under 2 MHz smoothly changes from the square-root behavior of immersed plates to the square behavior of fluid-coated plates as the coating thickness decreases. Clearly, the viscosity-induced attenuation is a rather complex but not very sensitive function of the coating thickness.

IV. CONCLUSIONS

In this paper we described an analytical technique to estimate the effect of viscous fluid loading on the propagation properties of Lamb waves in submerged and fluid-coated plates. To this goal, we further developed Zhu and Wu’s recently published method by removing certain inconsistencies that unnecessarily reduced the accuracy and range of validity of their results. According to their approach, we modeled the viscous fluid as a hypothetical isotropic solid having rigidity $c_{55}=-i\omega\eta$, where η denotes the viscosity of the fluid and ω is the angular frequency. In addition, we adopted the so-called Stokes model for viscous fluids by choosing $c'_{11}=\lambda_f-(4/3)i\omega\eta$ and $c'_{13}=\lambda_f+(2/3)i\omega\eta$; this also leads to dispersive and attenuative effects in the longitudinal bulk wave behavior in accordance with classical theories. By properly accounting for viscous effects on the bulk compressional wave in the fluid and applying a rigorous treatment of the field equations and boundary conditions, we derived exact dispersion equations that are not limited to low frequencies and viscosities. We presented numerous examples of our results to illustrate the effect of fluid viscosity on the lowest-order symmetric and antisymmetric Lamb modes.

We found that the effect of viscosity on the phase velocity of the Lamb modes is usually negligible with respect to the effect on attenuation. In the case of plates coated with

a finite-thickness fluid layer the Lamb modes are not leaky and viscosity is the sole source of attenuation. In the case of plates immersed in an infinite fluid bath the Lamb modes become leaky and the attenuation caused by fluid loading is due to the combined effects of radiation losses due to energy leakage into the fluid and dissipative losses due to viscous friction at the interfaces. Since we could not separately calculate the viscosity-induced part of the attenuation, we defined it as the difference between the respective attenuation coefficients caused by immersion in otherwise identical viscous and viscosity free fluids. However, strong viscosity can significantly affect the phase velocity and vibration distribution of a certain mode thereby slightly changes the radiation loss itself. We found that in some cases the resulting relatively small drop in the very strong radiation loss can overshadow the additional viscous dissipation and the total attenuation is actually lower for the viscous fluid than for the viscosity free one.² One interesting feature revealed by our calculations is the presence of a sharp minimum in the viscosity-induced attenuation of the lowest-order symmetric

mode of thin plates either immersed in or coated with a viscous fluid. We showed that this minimum occurs at a particular frequency, where the otherwise elliptical polarization of the surface vibration becomes linearly polarized in the normal direction.

¹ A. H. Nayfeh, *Wave Propagation in Layered Anisotropic Media* (North-Holland, Amsterdam, 1995).

² Q. Qi, "Attenuated leaky Rayleigh waves," *J. Acoust. Soc. Am.* **95**, 3222–3231 (1994).

³ J. Wu and Z. Zhu, "An alternative approach for solving attenuated leaky Rayleigh waves," *J. Acoust. Soc. Am.* **97**, 3191–3193 (1995).

⁴ Z. Zhu and J. Wu, "The propagation of Lamb waves in a plate bordered with a viscous fluid," *J. Acoust. Soc. Am.* **98**, 1057–1064 (1995).

⁵ A. H. Nayfeh and P. B. Nagy, "General study of axisymmetric waves in layered anisotropic fibers and their composites," *J. Acoust. Soc. Am.* **99**, 931–941 (1996).

⁶ P. B. Nagy and A. H. Nayfeh, "Viscosity induced attenuation of longitudinal guided waves in fluid-loaded rods," *J. Acoust. Soc. Am.* **100**, 1501–1508 (1996).

⁷ W. T. Thomson, "Transmission of elastic waves through a stratified solid medium," *J. Appl. Phys.* **21**, 89–93 (1950).

⁸ P. B. Nagy and R. M. Kent, "Ultrasonic assessment of Poisson's ratio in thin rods," *J. Acoust. Soc. Am.* **98**, 2694–2701 (1995).

Acoustic scattering of a plane wave by two spherical elastic shells above the coincidence frequency

H. Huang^{a)} and G. C. Gaunaud^{b)}

Naval Surface Warfare Center, White Oak, Silver Spring, Maryland 20903-5640

(Received 13 June 1996; accepted for publication 14 November 1996)

The acoustic scattering by two fluid-filled spherical elastic shells in close proximity to each other and insonified by plane waves at arbitrary angles of incidence is analyzed exactly in the frequency range that includes the midfrequency or coincidence enhancement region of the backscattered echoes. The incident and scattered wave fields are expanded in terms of the classical modal series and the addition theorem for the spherical wave functions is used to determine the exact expression for the sound fields scattered by each spherical elastic shell in the presence of the other, referred to coordinate systems at the centers of either spherical shell. The solution to the scattering problem is found by simultaneously solving the Helmholtz equations governing the wave motion in the fluid medium in which the two shells are submerged as well as in the fluid media contained in the shells, together with the two sets of equations of motion of the two elastic shells obtained from the complete three-dimensional elasticity theory after satisfying the boundary conditions at all fluid-shell interfaces as well as the far-field radiation condition. Again, the numerical computation of the scattered pressure wave involves the solution of a truncated ill-conditioned complex matrix system the size of which depends on how many terms of the modal series are required for convergence. This in turn depends on the value of the frequency, and on the proximity of the two spherical elastic shells. The ill-conditioned matrix equation is solved using the Gauss-Seidel iteration method. Backscattered and bistatic echoes from two identical spherical elastic shells are extensively calculated. The result also exhibits the large enhancement present in the backscattered echoes for the endfire situation after the midfrequency or coincidence enhancement has taken place. This can be attributed to the effects of focusing by the front elastic shell and to the reflection and refocusing by the back elastic shell of the a_0 Lamb wave reradiation in the observer's direction. [S0001-4966(97)01804-3]

PACS numbers: 43.40.Ey, 43.20.Gp [CBB]

INTRODUCTION

The three-dimensional scattering of two strongly interacting elastic spherical shells in a boundless acoustic medium has been previously studied using the classical modal series expansion method in conjunction with the bending theory for the spherical elastic shells.¹ The solution was evaluated exactly and displayed in a nondimensional frequency band of $ka = 0-25$. The numerical results have demonstrated that the low-resonance frequencies of a vibrating submerged elastic structure shift downward with the proximity of a neighboring submerged structure due to the increase in added mass of the two vibrating objects. The shell theory we used earlier was the popular bending theory, described in standard textbooks,² and used in most current finite element method (FEM) codes. The backscattering cross-sectional results of the bending shell theory agree well with those of the general three-dimensional elasticity theory for the shell thickness, material, and the frequency band of $ka = 0-25$ used in the numerical examples shown in Ref. 1. Beyond this frequency band, the validity of this shell theory, for a 3% shell, begins to deteriorate; the scattered field cannot be accurately calculated and many of its significant features fail to

be revealed. The conclusions obtained for this frequency range with the bending theory may not be valid for higher frequencies due to the more complex features of the response of elastic spherical shells at higher frequencies. An approximate solution has been obtained³ based on the assumption that the spheres are in each other's far field using the elasticity equations. Under such an approximation, terms representing multiple scattering effects are neglected, and therefore, this approximate solution would not be suitable for accurately calculating the scattered field. We also found that the numerical results of Ref. 1 for the endfire conditions (shown in Figs. 9 and 10 of Ref. 1) have errors due to an erroneous FORTRAN statement for the phase factor of the second spherical shell. This has been corrected, in the end-fire plots shown here, and the results have been rigorously validated against a Helmholtz integral equation boundary element/finite element approach. Further details of this validation will be reported in a forthcoming paper. In the present paper, the solution in Ref. 1 is extended to the case of acoustic scattering by two fluid-filled spherical elastic shells, employing the three-dimensional elasticity equations of motion to govern the dynamic responses of the shells. The practical frequency range for the validity of the present solution is only limited by the capabilities of the available computational hardware and software. Here, the backscattered form functions for two identical spherical elastic shells are exten-

^{a)}Indian Head Division.

^{b)}Carderock Division.

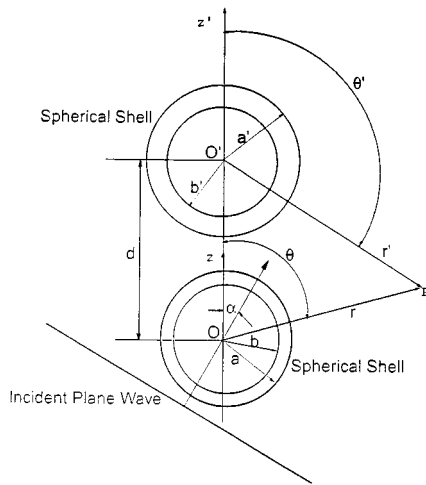


FIG. 1. Two elastic spherical shells insonified by an incident plane wave at arbitrary angle of incidence. Geometry of the problem.

sively calculated in the frequency band of $ka = 0-60$. The result clearly displays the large enhancement of the backscattered echoes for the end-fire situation at the higher end of this frequency range. The magnitude of the enhancement is a function of the frequency, the separation of the two shells, and the material properties of the scatterers. In electromagnetic scattering, this has been attributed⁴ to the effects of focusing by the front sphere and the reflection and refocusing by the back sphere towards the observer, a situation analogous to placing a mirror behind a lens near its focal point. For the present structural acoustics problem, it is specifically the focusing and refocusing of the antisymmetric strong flexural a_0 mode reradiations that produce the significant enhancement of the echoes.

I. DESCRIPTION OF THE PROBLEM

Figure 1 sketches the geometry of the problem. The centers of the two spherical elastic shells are separated by a distance d . The outer and inner radii of the lower and upper shells are denoted by (a, b) and (a', b') , respectively. Their mass densities, dilatation, and shear-wave velocities are, respectively, (ρ_s, c_d, c_s) and (ρ'_s, c'_d, c'_s) . The mass densities and sound speeds of their internal fluids are, respectively, (ρ_1, c_1) and (ρ'_1, c'_1) . The origins O and O' of two spherical coordinate systems (r, θ, ϕ) and (r', θ', ϕ') coincide with the centers of the lower and upper spherical shells, respectively. The primed and the unprimed coordinate systems have the same azimuthal coordinate ϕ which is not shown in the figure. The mass density and sound speed of the infinite fluid medium in which the shells are submerged are denoted by ρ and c , respectively. The propagation vector of the incident plane wave is parallel to the plane of the paper and makes an angle α with the line OO' .

As the incident pressure wave impinges upon the shells, it is reflected, deflected by, and diffracted into and out of the shells while it excites their vibratory dynamic responses. The scattered pressure emanating from one shell also acts as a second incident wave on the other, i.e., the two shells are interacting with each other elastoacoustically. This repeated

rereflection and rescattering between the two as well as all incoming and outgoing waves are completely taken into account by our formulation.

The problem is analyzed in the frequency domain using the classical techniques of separation of variables and modal series expansions. To completely determine the scattered pressure field and the backscattered echo, it is necessary to simultaneously solve the Helmholtz equations governing the wave motion in all fluid media and the two sets of equations of motion of the two elastic shells satisfying the boundary conditions at all shell-fluid interfaces and the far-field radiation condition. Thus, all the complicated reflection, deflection, and diffraction events mentioned in the previous paragraph are treated exactly. When the interaction effect between shells becomes negligible as the separation becomes sufficiently large, the backscattered echo can be obtained as the sum of the contribution from each shell assumed individually isolated, but with an appropriate phase factor affecting the second shell contribution.

II. MATHEMATICAL FORMULATION

Except for the use of the general three-dimensional elasticity equations and the associated boundary conditions, the formulation steps are similar to those in Ref. 1. The governing equations for the scattering by a single fluid-filled spherical elastic shell are also well-known and have been well-publicized.⁵⁻⁸ Portions of the formulation of present interest are only briefly recapitulated here. For clarity, the following dimensionless quantities are used

$$R = \frac{r}{a}, \quad R' = \frac{r'}{a'}, \quad D = \frac{d}{a},$$

$$\Omega = \frac{\omega a}{c} = ka, \quad \Omega_1 = \frac{c}{c_1} \Omega, \quad \Omega'_1 = \frac{c}{c'_1} \Omega, \quad (1)$$

$$\Pi = \frac{p}{\rho c^2}, \quad \Pi_1 = \frac{p_1}{\rho c^2}, \quad \Pi'_1 = \frac{p'_1}{\rho c^2}.$$

In the above equations, p , p_1 , and p'_1 are, respectively, the total pressure field of the fluid medium in which the shells are submerged and of the fluid media inside the lower and upper shells, and ω is the angular frequency of the incident plane wave. Similarly, Π , Π_1 , and Π'_1 are nondimensional pressures that satisfy the Helmholtz equation in their respective domains. The incident plane wave sketched in Fig. 1 can be represented in both coordinate systems as¹

$$\begin{aligned} \Pi^{\text{inc}} &= e^{i\Omega R(\cos \theta \cos \alpha + \sin \theta \sin \alpha \cos \phi)} \\ &= e^{i\Omega R'(\cos \theta' \cos \alpha + \sin \theta' \sin \alpha \cos \phi)} e^{i\Omega D \cos \alpha}, \end{aligned} \quad (2)$$

where $i = \sqrt{-1}$. The modal series expansion of the above expression in terms of spherical wave functions in both the primed and unprimed spherical coordinates can be written as¹

$$\begin{aligned}\Pi^{\text{inc}} &= 2 \sum_{n=0}^{\infty} \sum_{m=0}^n i^n \frac{2-\delta_{0m}}{N_{nm}} J_n(\Omega R) P_n^m(\cos \theta) \\ &\quad \times P_n^m(\cos \alpha) e^{im\phi} \\ &= e^{i\Omega D \cos \alpha} 2 \sum_{n=0}^{\infty} \sum_{m=0}^n J_n^2 \frac{2-\delta_{0m}}{N_{nm}} j_n(\Omega R') \\ &\quad \times P_n^m(\cos \theta') P_n^m(\cos \alpha) e^{im\phi},\end{aligned}\quad (3)$$

where J_n and P_n^m are, respectively, the spherical Bessel function of the first kind and the associated Legendre function,

$$N_{mn} = \frac{2(n+m)!}{(2n+1)(n-m)!} \quad (4)$$

and

$$\delta_{0m} = \begin{cases} 1, & \text{for } m=0, \\ 0, & \text{for } m \neq 0. \end{cases} \quad (5)$$

The pressure scattered (in the exterior fluid medium) by the two shells in the presence of each other, and thus, accounting for all multiple scattering interactions, is

$$\begin{aligned}\Pi^{\text{sca}} &= \sum_{n=0}^{\infty} \sum_{m=0}^n [B_{mn} h_n(\Omega R) P_n^m(\cos \theta) \\ &\quad + b_{mn} h_n(\Omega R') P_n^m(\cos \theta')] e^{im\phi},\end{aligned}\quad (6)$$

where the coefficients B_{mn} and b_{mn} are coupled and are to be determined from the boundary conditions and h_n is the spherical Hankel function of the first kind. The total pressure field in the exterior fluid medium expressed in the (R, θ, ϕ) system is

$$\begin{aligned}\Pi &= \Pi^{\text{inc}} + \Pi^{\text{sca}} = 2 \sum_{n=0}^{\infty} \sum_{m=0}^n i^n \frac{2-\delta_{0m}}{N_{mn}} J_n(\Omega R) P_n^m(\cos \theta) P_n^m(\cos \alpha) e^{im\phi} \\ &\quad + \sum_{n=0}^{\infty} \sum_{m \neq 0}^q b_{qm} \sum_{n=m}^{\infty} Q_{mnmq}(D, \pi) j_n(\Omega R) P_n^m(\cos \theta) e^{im\phi}.\end{aligned}\quad (7)$$

This total pressure field expressed in the (R', θ', ϕ) system is

$$\begin{aligned}\Pi &= 2 e^{i\Omega D \cos \alpha} \sum_{n=0}^{\infty} \sum_{m=0}^n i^n \frac{2-\delta_{0m}}{N_{mn}} J_n(\Omega R') P_n^m(\cos \theta') P_n^m(\cos \alpha) e^{im\phi} \\ &\quad + \sum_{q=0}^{\infty} \sum_{m=0}^q B_{mq} \sum_{n=m}^{\infty} Q_{mnmq}(D, 0) j_n(\Omega R') P_n^m(\cos \theta') e^{im\phi}.\end{aligned}\quad (8)$$

The addition theorem of spherical wave functions⁹ for the geometry described in Fig. 1 has been applied to represent quantities from Eqs. (3) and (6) in the uniformly unprimed and primed spherical coordinates in Eqs. (7) and (8), respectively, and

$$Q_{mnmq}(D, \pi) = \frac{2i^{n-q}}{N_{mn}} \sum_{\sigma=n-q}^{n+q} l^{\sigma} (-1)^{\sigma} b_{\sigma}^{qmn} h_{\sigma}(\Omega D), \quad (9)$$

$$Q_{mnmq}(D, 0) = \frac{2i^{n-q}}{N_{mn}} \sum_{\sigma=n-q}^{n+q} i^{\sigma} b_{\sigma}^{qmn} h_{\sigma}(\Omega D), \quad (10)$$

where

$$\begin{aligned}b_{\sigma}^{qmn} &= (-1)^m \sqrt{\frac{(q+m)!(n+m)!}{(q-m)!(n-m)!}} (q, n, 0, 0/\sigma, 0) \\ &\quad \times (q, n, m, -m/\sigma, 0).\end{aligned}\quad (11)$$

The quantity $(j_1, j_2, m_1, m_2/j, m)$ in the above equation is a Clebsch–Gordan coefficient.^{9,11} The pressure fields in the internal fluids of the lower shell in the unprimed and the upper shell in the primed systems are, respectively,

$$\Pi_i = \sum_{n=0}^{\infty} \sum_{m=0}^n C_{mn} j_n(\Omega_i R) P_n^m(\cos \theta) e^{im\phi}, \quad (12)$$

and

$$\Pi'_i = \sum_{n=0}^{\infty} \sum_{m=0}^n c_{mn} j_n(\Omega'_i R') P_n^m(\cos \theta') e^{im\phi}, \quad (13)$$

where the coefficients C_{mn} and c_{mn} are also to be determined through the boundary conditions.

The three-dimensional elasticity theory equations of motion for the spherical elastic shell^{5–8} are used here. The displacement in a spherical shell can be expressed in terms of a scalar and a vector potential. These equations are well-known and need not be repeated here. The boundary conditions of the problem are that at the boundaries of the shells the displacement and normal stress must be continuous and the tangential stress must be zero and that the scattered pressure must satisfy the radiation condition in the far field. Applying these boundary conditions and using the elasticity equations of motion for the shells yields

$$B_{mn} + \left[\sum_{q=m}^{\infty} b_{mq} Q_{mnmq}(D, \pi) \right] X_n(\Omega) = -2i^n \frac{2 - \delta_{0m}}{N_{mn}} X_n(\Omega) P_n^m(\cos \alpha), \quad (14)$$

$$b_{mn} + \left[\sum_{q=m}^{\infty} B_{mq} Q_{mnmq}(D, 0) \right] X'_n(\Omega) = -2i^n \frac{2 - \delta_{0m}}{N_{mn}} X'_n(\Omega) P_n^m(\cos \alpha) e^{i\Omega D \cos \alpha},$$

where

$$X_n(\Omega) = - \frac{\begin{vmatrix} A_1^* & d_{12} & d_{13} & d_{14} & d_{15} & 0 \\ A_2^* & d_{22} & d_{23} & d_{24} & d_{25} & 0 \\ 0 & d_{32} & d_{33} & d_{34} & d_{35} & 0 \\ 0 & d_{42} & d_{43} & d_{44} & d_{45} & d_{46} \\ 0 & d_{52} & d_{53} & d_{54} & d_{55} & d_{56} \\ 0 & d_{62} & d_{63} & d_{64} & d_{65} & 0 \end{vmatrix}}{\begin{vmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & 0 \\ d_{12} & d_{22} & d_{23} & d_{24} & d_{25} & 0 \\ 0 & d_{32} & d_{33} & d_{34} & d_{35} & 0 \\ 0 & d_{42} & d_{43} & d_{44} & d_{45} & d_{46} \\ 0 & d_{52} & d_{53} & d_{54} & d_{55} & d_{56} \\ 0 & d_{62} & d_{63} & d_{64} & d_{65} & 0 \end{vmatrix}}, \quad (15)$$

where the 30 nonvanishing elements d_{11} and A_1^* , A_2^* were given in Refs. 5–7 and minor typographical corrections were given in Ref. 8. They are functions of n and the properties a/b , ρ , ρ_s , ρ_1 , c_d , c_s , and c_1 of the lower shell. The form function for a single shell based on the above expression for the scattering coefficients X_n has been confirmed and benchmarked by various methods.^{7,8,10} Again, for uniform complete spherical shells, $X_n(\Omega)$ is not a function of m . The analogous quantity $X'_n(\Omega)$ has exactly the same form as Eq. (15) but requires the outer to inner radius ratio, shell density, dilatation, and shear wave velocities, internal fluid density, and sound speed of the upper shell. All the three-dimensional elasticity effects in the shells are contained in the quantities $X_n(\Omega)$ and $X'_n(\Omega)$. This generalization allows us to carry on accurate calculations in arbitrarily broad frequency bands. For each m , Eq. (14) can be arranged as the following matrix equations if N terms are used in the n , q , and m sums,

$$\{B\} + [\Lambda]\{b\} = \{S\}, \quad (16)$$

$$\{b\} + [A]\{B\} = \{S'\} e^{i\Omega D \cos \alpha},$$

where the vectors B and b represent B_{mn} and b_{mn} , respectively, and the other matrix elements are

$$\begin{aligned} \Lambda_{kl} &= Q_{mkml}(D, \pi) X_l(\Omega), \\ A_{kl} &= Q_{mkml}(D, 0) X'_l(\Omega), \\ S_l &= -2i^l \frac{2 - \delta_{0m}}{N_{ml}} X_l(\Omega) P_l^m(\cos \alpha), \\ S'_l &= -2i^l \frac{2 - \delta_{0m}}{N_{ml}} X'_l(\Omega) P_l^m(\cos \alpha), \end{aligned} \quad (17)$$

where k and l range from m to N . Here, Λ and A are $(N + 1 - m) \times (N + 1 - m)$ complex matrices, and B and S are $(N + 1 - m)$ complex vectors for each m . From the matrix system, Eq. (16), the complex coefficients B_{mn} and b_{mn} can be solved for, and then the total or the scattered pressure field can be computed. For very large R , we use the asymptotic relation.

$$h_n(\Omega R) \rightarrow i^{-(n+1)} \frac{e^{i\Omega R}}{\Omega R}. \quad (18)$$

The far-field scattered pressure expression for the two shells can be simplified as

$$\begin{aligned} \Pi^{\text{sca}}(R, \theta, \phi, \Omega) &= \frac{e^{i\Omega R}}{\Omega R} \sum_{n=0}^{\infty} \sum_{m=0}^n i^{-(n+1)} \left[B_{mn} P_n^m(\cos \theta) \right. \\ &\quad \left. + b_{mn} \frac{R}{R'} e^{i\Omega(R' - R)} P_n^m(\cos \theta') \right] e^{im\phi}. \end{aligned} \quad (19)$$

The primed coordinates are given by

$$\begin{aligned} R' &= \sqrt{R^2 + D^2 - 2RD \cos \theta}, \\ \theta' &= \sin^{-1} \left(\frac{R}{R'} \sin \theta \right), \quad 0 \leq \theta \leq \frac{\pi}{2}, \\ &= \pi - \sin^{-1} \left(\frac{R}{R'} \sin \theta \right), \quad \frac{\pi}{2} \leq \theta \leq \frac{3\pi}{2}, \\ &= 2\pi - \sin^{-1} \left(\frac{R}{R'} \sin \theta \right), \quad \frac{3\pi}{2} \leq \theta \leq 2\pi. \end{aligned} \quad (20)$$

The form function is defined as the absolute value of

$$f_{\sigma}(R, \theta, \phi, \Omega) = 2R [\Pi^{\text{sca}}(R, \theta, \phi, \Omega)] e^{-i\Omega R}. \quad (21)$$

III. NUMERICAL RESULTS AND DISCUSSION

The complex coupling coefficients B_{mn} and b_{nm} in Eq. (6) are found by solving the truncated complex linear system. Equation (16), by the Gauss–Seidel iteration technique. The details have been discussed in Ref. 1. The truncation, i.e., the number of terms N used for the sums from $n = 0$ to $n = N$ and from $q = 0$ to $q = N$ in Eqs. (6) and (14), depends on the degree of accuracy required, the value of Ω , and the complexity of the wave interaction picture. For $\Omega = ka = 60.0$, it is required that $N = 80$ to obtain an 8-digit accuracy. For each frequency, it is necessary to solve the complex equation, Eq. (16), N times to find B_{mn} and b_{mn} where m and n range from 0 to N . The sizes of the matrices are $(N + 1 - m) \times (N + 1 - m)$. The present converged numerical results are computed to 8-digit accuracy.

TABLE I. Material parameters of the two identical shells and the fluids.

Material	Density (g/cm ³)	Dilatational wave speed (c_d) $\times 10^5$ (cm/s)	Shear wave speed (c_s) $\times 10^5$ (cm/s)
Water	1.0	1.5	
Steel	7.8	5.88	3.14
Shell internal fluid	0.00001	0.344	

All subsequent numerical results are presented for the case of two *identical* steel shells submerged in water with a thickness to radius ratio $h/a = 0.03$, and for various angles of incidence α . The material parameters used are listed in Table I. The shells are assumed lossless. The value of the density of the internal fluid is chosen so that the internal fluid effects are negligible. The observation point is at $R = 1000.0$ in the backscattering direction.

There are two physical reasons why the present backscattering form function is different from just the sum of the form functions of the two individual shells, namely, the multiple interaction between the two shells and, depending on the position of the observation point, the fact that the signals from the two shells have different phase factors as indicated by the expressions $e^{i\Omega D \cos \alpha}$ in Eq. (14) and $e^{i\Omega(R'-R)}$ in Eq. (19) (in the time domain the two signals reach the observer at different arrival times). For the backscattering of the broadside incidence case (i.e., $\alpha = \pi/2$), the two signals from the

two shells have a zero phase difference and if there is negligible multiple scattering interaction effects between the two shells, the magnitude of the backscattering echo should equal twice that of a single shell. Broadside incidence cases are again studied first to isolate the two shells' interaction effects. The difference between the true form function from twice that of a single shell reveals the features of this interaction. If the backscattering echo has an amplitude exceeding twice that of a single shell, this interaction manifests enhancement effects. The broadside backscattering form functions are calculated at $R = 1000.0$, $\theta = 1.5\pi$, and $\phi = \pi$. Figure 2(a) presents such a form function for the two shells (light curve) when the crown of the lower shell is half a radius from the keel of the upper shell ($D = 2.5$) and compares it to that of the single shell (heavy curve) for a frequency range of $\Omega = 0-60$. This broad range can be displayed only because we are now using the general three-dimensional elasticity formulation for the shells motions. This range encompasses the so-called midfrequency or coincidence enhancement phenomenon for the returning echo of a submerged 3% steel shell.⁷ The significant interaction effect at low frequencies ($\Omega < 2.0$) is that the resonance frequencies of the large amplitude echoes shift downward, and their magnitudes decrease as the two shells are placed closer to each other. This has previously been discussed in detail.¹ The magnitudes of the low-frequency resonance peaks of the pair of shells are clearly lower than twice those of the single shell. It has been previously found¹⁰ that for the form function of a single sub-

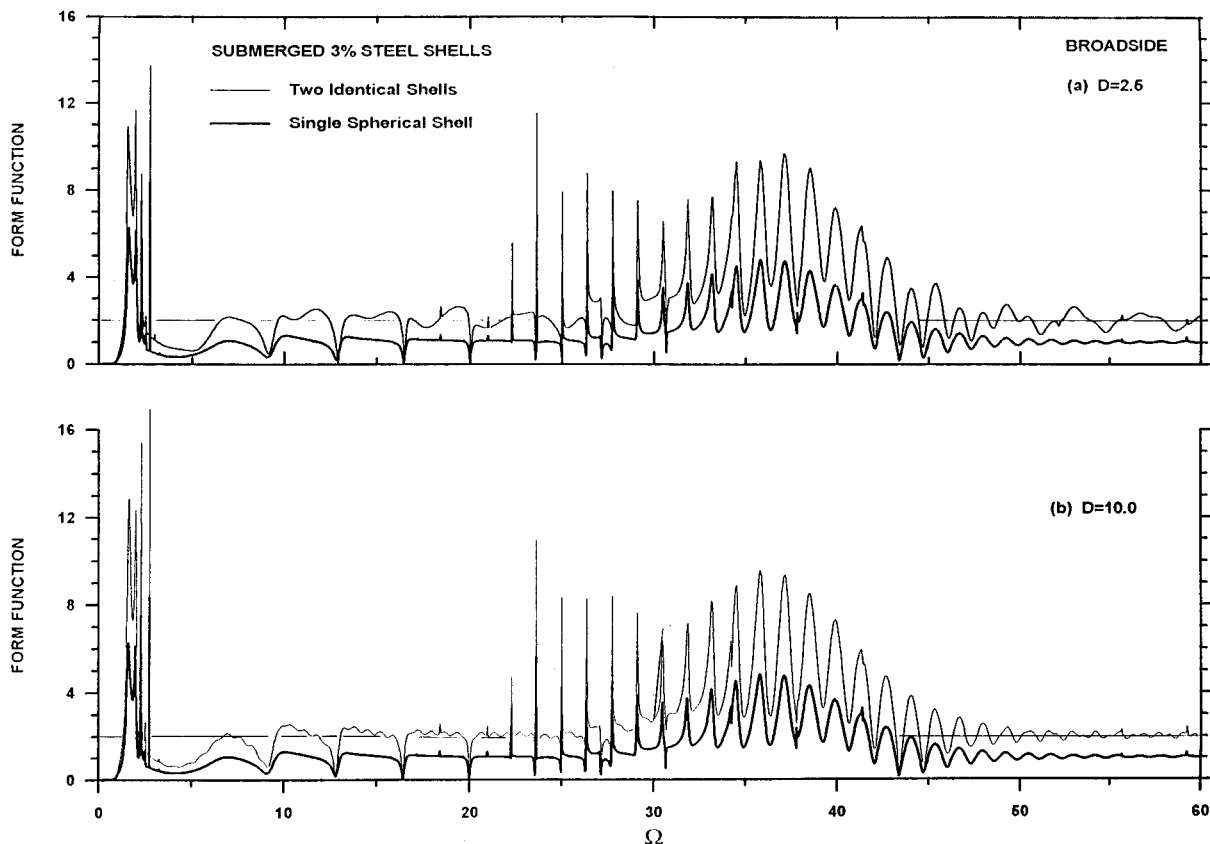


FIG. 2. Backscattering form functions of two interacting 3% steel shells compared to that of a single shell. Here, the incidence is in the broadside ($\alpha = \pi/2$) direction, and (a) $D = 2.5$ and (b) $D = 10.0$.

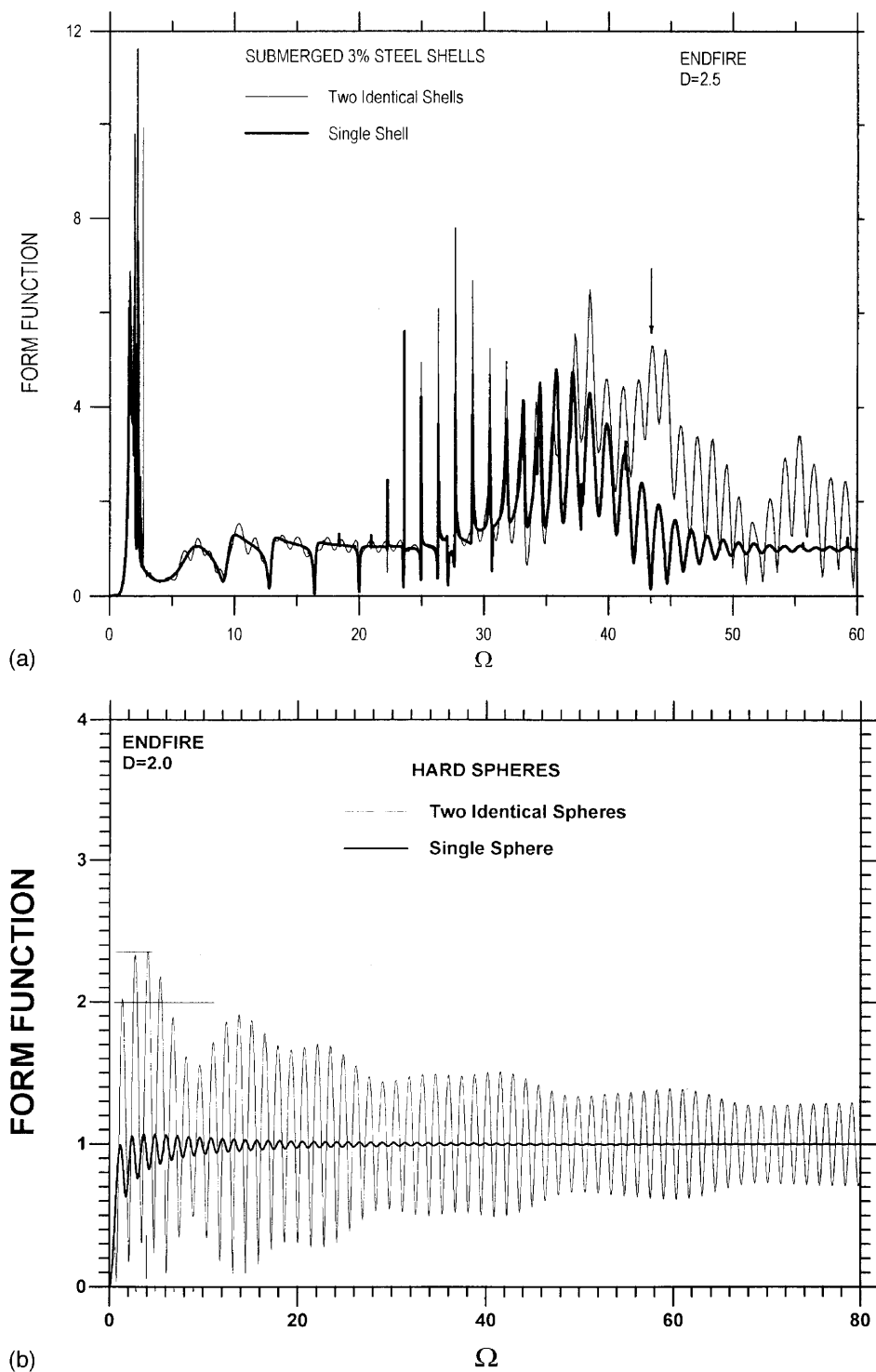


FIG. 3. (a) Backscattering form functions of the two interacting 3% steel shells compared to that of a single shell. The incidence is in the end-fire ($\alpha=0$) direction and $D = 2.5$. (b) Same calculation (thin line) for a pair of rigid spheres, also at end fire, and $D = 2.0$. The thicker curve is for one hard sphere.

merged elastic shell, a coincidence “bump” or enhancement band occurs centered at the coincidence frequency $\Omega_c \cong [\zeta(h/a)]^{-1}$ and extends roughly an amount $0.5\Omega_c$ on both sides of Ω_c . The constant $\zeta \cong 0.87$ for steel and hence for the present submerged 3% steel shell, $\Omega_c \cong 38$ and the coincidence enhancement band extends throughout $19 \leq \Omega \leq 57$. It is within this coincidence band that the first generalized antisymmetric Lamb wave a_0 is “activated” and plays a dominant role in determining the features of the form

function. The first symmetric Lamb wave s_0 is activated throughout the entire frequency band displayed, but its effect is weaker within the coincidence band. In Fig. 2(a), for $\Omega > 2$, the difference between the *interacting* pair and the $2 \times$ single case persists and shows some minor enhancement of the echo at high frequency, i.e., the magnitude of the form function is greater than twice that of a single shell, and this enhancement is more discernible for $\Omega > 45$, i.e., after the Ω_c of the single shell. This can be attributed to the interac-

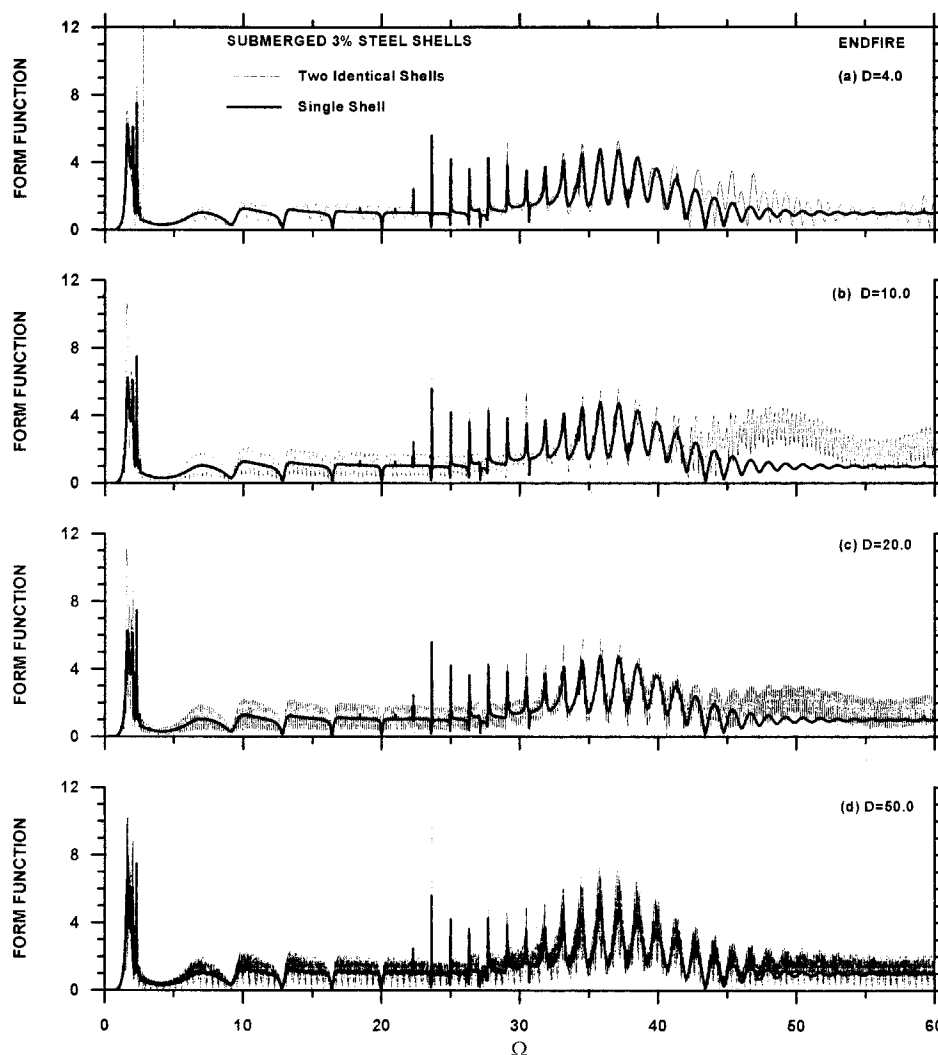


FIG. 4. Backscattering form functions of the two interacting 3% steel shells compared to that of a single shell. Here, the incidence is in the end-fire ($\alpha=0$) direction, and (a) $D = 4.0$, (b) $D = 10.0$, (c) $D = 20.0$, and (d) $D = 50.0$.

tion effect of the a_0 Lamb wave. The corresponding two curves for the case where the crown of the lower spherical shell is 8 radii from the keel of the upper shell ($D = 10.0$) is presented in Fig. 2(b). Here, comparing the form function of the pair of shells to twice that of the single shell, the effect of the interaction between the two shells constitutes a very minor perturbation of the magnitude of twice the form function of a single shell, and the enhancement of the echo is no longer noticeable for this shell separation. Nevertheless, the effect induces a still quite clearly visible oscillation pattern which has a period of oscillation $2\pi/D$ and is therefore significant from the acoustic signal point of view. The magnitude of this oscillation will be further reduced as the separation of the two shells is increased.

The end-fire incidence ($\alpha=0$) cases reveal more interesting features of the multiple scattering interaction. Figures 3(a) and 4 plot the form functions of the pair of shells with separations of $D = 2.5, 4.0, 10.0, 20.0$, and 50.0 , and compares them to that of the single shell for a frequency range of $\Omega=0-60$. At very low frequencies corresponding to incident wavelengths comparable to $D + 2$ or longer, the results correctly show no shielding effect of the frontal shell first im-

pinged upon by the incident wave. In Fig. 3(a), the two shells have a small separation $D = 2.5$, and it can be seen that from $\Omega \approx 3$ to Ω_c the form function of the pair differs from that of the single only by a small perturbation. In other words, the frontal shell shields or eclipses to a large extent the back shell in this spectral region (i.e., the back shell's presence is not clearly noticeable acoustically). From Ω_c onward, this shielding becomes negligible and there are even significant enhancements of the returned echoes (i.e., the magnitude of the form function of the pair is substantially greater than twice that of the single). It is expected that the frontal shell could completely hide the back shell at frequencies much higher than those shown in the presently displayed band. In the electromagnetic end-fire scattering cases, the enhanced echoes have been interpreted⁴ as due to rays focused by the frontal sphere and refocused by the second sphere back to the observer, analogous to placing a mirror behind a lens near its focal point. For the present situation, it is specifically the focusing and refocusing of the Lamb a_0 mode reradiations that result in the significant enhancement of the echoes after Ω_c . The features of the corresponding echoes for the case of two hard spheres are intrinsically different from the above

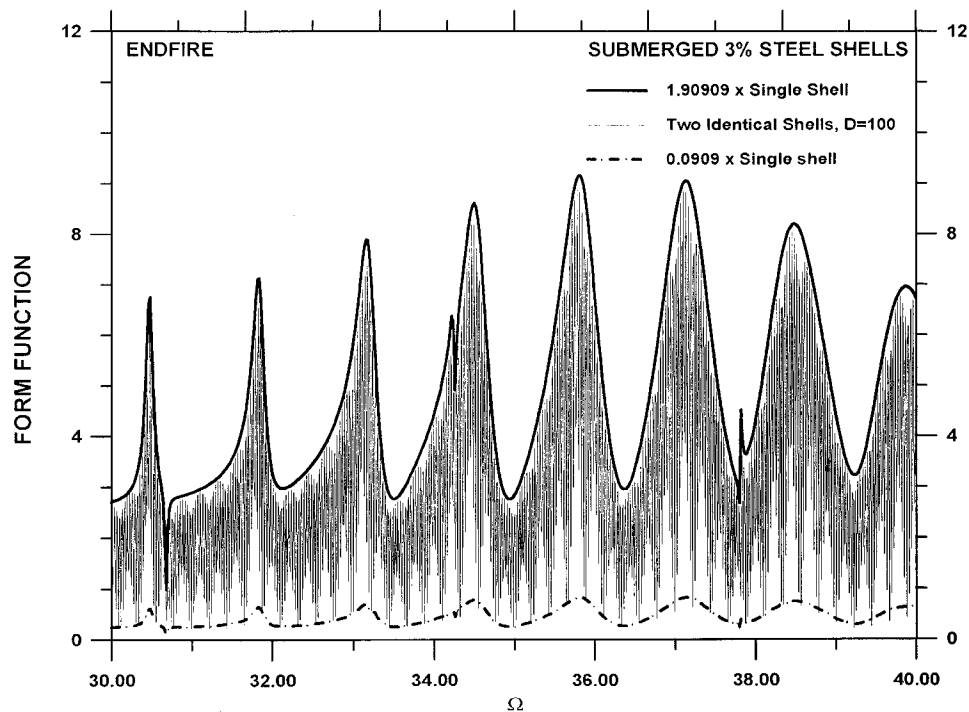


FIG. 5. Backscattering form function of the two interacting 3% steel shells with incidence in the endfire ($\alpha=0$) direction at large shell separations (i.e., $D = 100.0$) in the "coincidence band." There are hardly any observable interaction effects.

descriptions as can be witnessed from Fig. 3(b) where the form function of two identical hard spheres, obtained by an analogous formulation,¹¹ is compared to that of a single hard sphere. In Fig. 3(b) there are some enhancements of the echoes of the two hard spheres for $\Omega < 6$, but from then on, the magnitude of the form function of the pair is always less than

twice that of the single (i.e., no enhancement) and its deviation from the single hard sphere form function decreases as Ω increases. In other words, as Ω increases, the form function of the pair slowly approaches that of the single. The trend shown indicates that the frontal hard sphere will eventually hide the back hard sphere at sufficiently high Ω as one

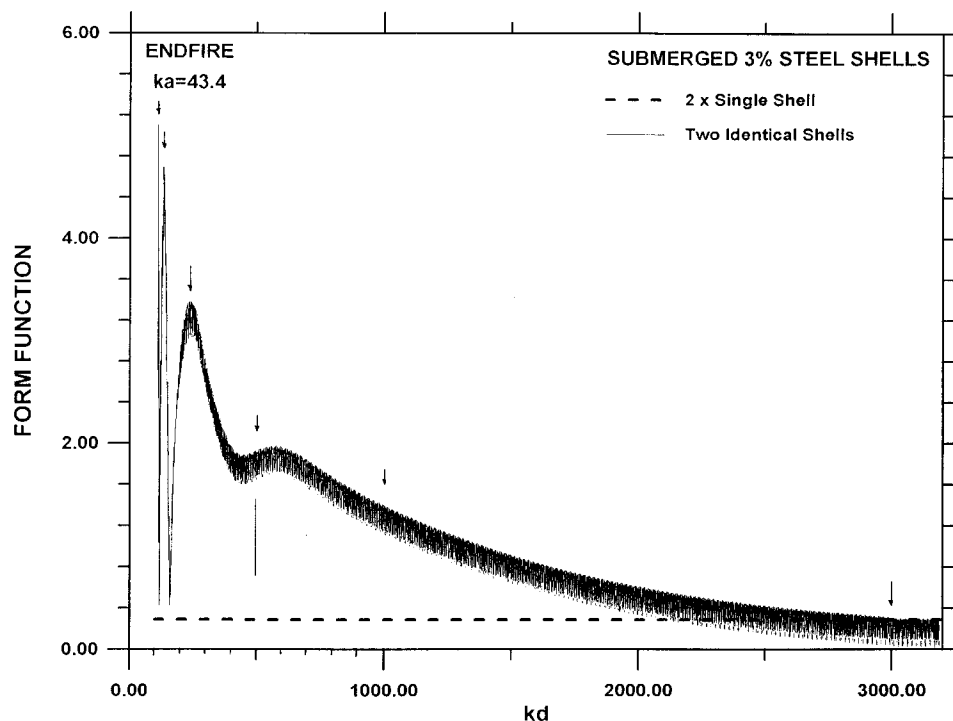


FIG. 6. Backscattering form function of the two interacting 3% steel shells at end-fire ($\alpha=0$) incidence, plotted versus kd , for $ka = 43.4$. The dashed curve shows twice the single shell value. This shows the variation of the focusing enhancement effect with shell separation.

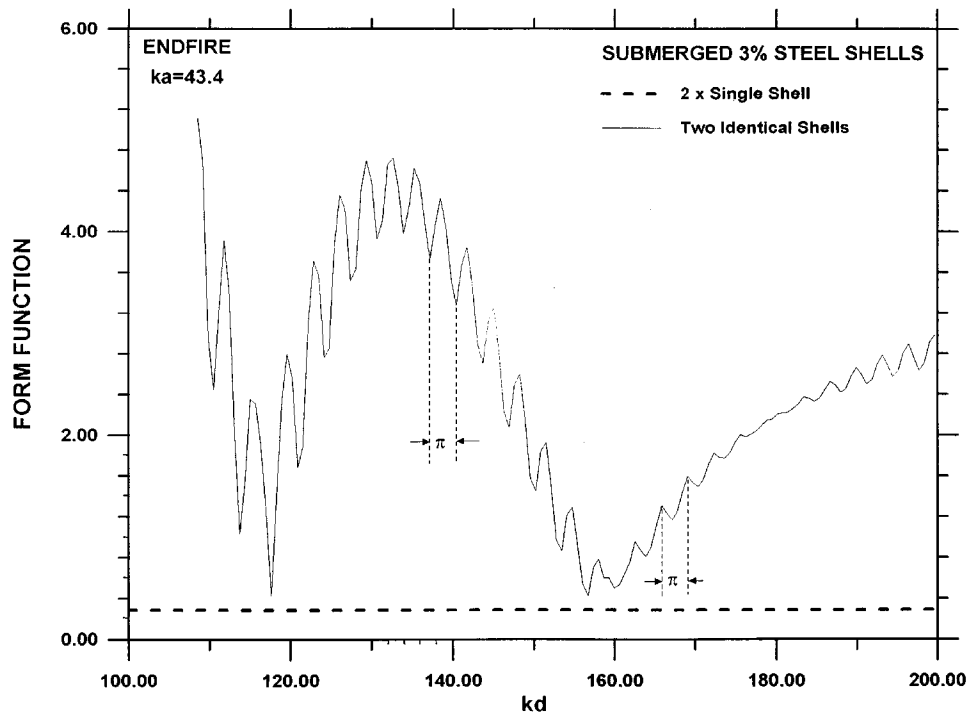


FIG. 7. Enlarged view to show detail of Fig. 6 in the band: $kd = 100-200$.

would expect, above the band shown in Fig. 3(b). From this comparison, it is evident that the present enhancements after Ω_c seen in Figs. 3(a) and 4 are primarily an *elastoacoustical* effect. The form function of the pair with slightly larger separation, $D = 4.0$, is plotted in Fig. 4(a). With this increase in separation, the shielding of the frontal shell for frequencies below the coincidence enhancement band has a larger leakage since the form function has a larger deviation from that of the single shell as compared to Fig. 3(a). Nevertheless, the shielding for the core of the coincidence band ($\Omega \approx 30-40$) is still very effective. If the interaction among the two spherical shells has a negligibly small effect, the end-fire form function of the pair can be deduced from Eqs. (14) and (19) and it is simply given by

$$|f_\infty| = |f_\infty|_s \left(1 + \frac{R}{R+D} e^{2i\Omega D} \right), \quad (22)$$

where $|f_\infty|$ is the form function of the single spherical elastic shell. This type of formula also appears in the radar literature.¹² Therefore, in the absence of interaction, the end-fire form function of the pair oscillates between two envelopes given by $|f_\infty|_s [1 \pm R/(R+D)]$. The form function for the case in which the two shells have a wider separation, $D = 10.0$, is plotted in Fig. 4(b). Here, below the coincidence band, there is very little interaction among the two shells since the form function appears to be oscillating between the above-mentioned envelopes. Even for this separation of 10 shell outside radii, the shielding effect of the frontal shell is still strong in the core of the coincidence band ($\Omega \approx 30-40$) and the focusing enhancement is still significant in the higher spectral region ($\Omega \approx 45-60$). The form function of the pair for $D = 20.0$ is plotted in Fig. 4(c) and it can be seen to have similar features as that for $D = 10.0$ except that the focusing

enhancement is weakened but still persists in the higher spectral region ($\Omega \approx 45-60$). The form function of the pair with still wider separation, $D = 50.0$, is plotted in Fig. 4(d). Now, except in the spectral region of the core of the coincidence band ($\Omega \approx 30-40$), there appears to be negligible interaction among the two spherical shells in the rest of the frequency band displayed in view of the criterion of Eq. (22). In the core of the coincidence band ($\Omega \approx 30-40$) the shielding effect of the frontal shell still persists. This shielding effect should be absent for pairs with sufficiently large separations. Figure 5 shows that the form function in the core of the

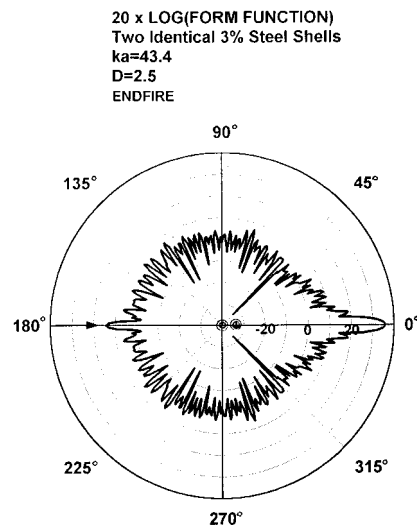


FIG. 8. Bistatic form function (in dB) plotted versus observation angle θ for the two 3% steel shells separated by $D = 2.5$, at the fixed frequency of $\Omega = 43.4$. This value corresponds to a large enhancement peak in Fig. 3(a). The arrow denotes the incidence direction.

coincidence band ($\Omega \approx 30\text{--}40$), of the pair separated by 100 shell outer radii ($D = 100.0$) oscillates between the two envelopes provided by Eq. (22) and, therefore, there is no longer shielding of the back shell and any other interaction effects. We may also add that for the single shell case, the core of the coincidence band ($\Omega \approx 30\text{--}40$) is the one showing the most significant enhancement of the form function. But this is for *backscattering*. In the opposite direction, the forward scattering cross section should exhibit a “coincidence depression” in that same frequency band since less energy is then continuing forward. If now there is a second shell behind the front one, it will receive the least amount of energy from the front one within that band. This may provide a further intuitive reason why there is shielding instead of enhancement for the two shells in $\Omega \approx 30\text{--}40$. The old term “midfrequency enhancement” has lost its meaning for two shells in the end-fire situation. The elastoacoustical effect then takes over and it makes its influence felt above coincidence.

It is more informative to examine the variation of this focusing enhancement effect with the separation of the two spherical shells by plotting the form function versus $\Omega D = kd$. Figure 6 displays such a form function plot for $\Omega \equiv ka = 43.4$ comparing it to twice the single shell echo amplitude (i.e., $2 \times 0.13617 = 0.27232$). The form function of the pair oscillates with a period of π due to the combined phase factor $e^{2i\Omega D}$ and the curve is again being modulated by an envelope whose period increases rapidly with kd . Figure 7 is an enlargement of Fig. 6 for $kd = 100\text{--}200$ to show these features. By comparing to twice the magnitude of the single shell echo, the end-fire echo of the pair is enhanced 17.93 times at $D = 2.5$ ($kd = 108.5$), 16.57 times at $D = 3.55$ ($kd = 132.567$), 11.81 times at $D = 5.25$ ($kd = 228.284$), 5.75 times at $D = 11.52$ ($kd = 499.968$) and 3.95 times at $D = 23.055$ ($kd = 1000.587$). This enhancement stops at a very large separation $D \approx 70$ ($kd \approx 3000$). Finally, the bistatic pattern around the center O of the end-fire echo of the pair for $\Omega = 43.4$ and $D = 2.5$ is displayed in Fig. 8. The arrow denotes the direction of wave incidence. Notice the large amplitudes in the backscattering and forward scattering directions and the extremely low magnitudes at 45° and 315° .

In summary, the present paper has augmented our previous analytical solution¹ to the problem of acoustic scattering by two fluid-filled elastic shells, employing the three-dimensional elasticity equations of motion to govern the dynamic responses of the elastic shells. The practical frequency range for the validity of the present solution is only

limited by the capabilities of the available computational hardware and software. Here, the backscattered form functions for two identical submerged 3% spherical elastic shells are extensively calculated in the frequency band of $ka = 0\text{--}60$ which includes the coincidence enhancement bands of each of the 3% steel shells. The results also exhibit the large enhancement of the backscattered echoes for the end-fire situation at the higher end of this frequency range above Ω_c . The magnitude of the enhancement is a function of the frequency, the shell separation, and the material properties of the scatterers. This enhancement of the echoes is specifically caused by the focusing and refocusing by the spherical shells of the antisymmetric strong flexural Lamb a_0 mode reradiations and is a special *elastoacoustical* effect.

ACKNOWLEDGMENT

The authors acknowledge the partial support of the Independent Research Program of the NSWC Carderock Division.

- ¹H. Huang and G. C. Gaunard, “Acoustic Scattering of a Plane Wave by Two Spherical Elastic Shells,” *J. Acoust. Soc. Am.* **98**, 2149–2156 (1995).
- ²M. C. Junger and D. Feit, *Sound, Structures, and Their Interaction* (MIT, Cambridge, MA, 1972), Chap. 9.
- ³J. P. Sessarego and J. Sageloli, “Étude théorique et expérimentale de la diffusion acoustique par deux coques sphériques élastiques,” *Acustica* **79**, 14–21 (1993).
- ⁴J. Brunning and Y. Lo, “Multiple Scattering of EM Waves by Spheres. Part II—Numerical and Experimental Results,” *IEEE Trans. Antennas Propag.* **AP-19**, 378–390 (1971).
- ⁵R. Goodman and R. Stern, “Reflection and Transmission of Sound by an Elastic Spherical Shell,” *J. Acoust. Soc. Am.* **34**, 338–344 (1962).
- ⁶R. Hickling, “Analysis of Echoes from a Hollow Metallic Sphere in Water,” *J. Acoust. Soc. Am.* **36**, 1124–1137 (1964).
- ⁷G. C. Gaunard and M. F. Werby, “Lamb and Creeping Waves Around Submerged Spherical Shells Resonantly Excited by Sound Scattering,” *J. Acoust. Soc. Am.* **82**, 2021–2033 (1987).
- ⁸H. C. Strifors and G. C. Gaunard, “Multipole Character of the Large-Amplitude, Low Frequency Resonances in the Sonar Echoes of Submerged Spherical Shells,” *Int. J. Solids Struct.* **29**, 121–130 (1992).
- ⁹Ye. Ivanov, *Diffraction of Electromagnetic Waves by Two Bodies* (Nauka i Tekhnika, Minsk, 1968) [NASA Technical Translation #F-587 (1970)], pp. 121–127.
- ¹⁰G. C. Gaunard and M. F. Werby, “Sound Scattering by Resonantly Excited, Fluid-Loaded, Elastic Spherical Shells,” *J. Acoust. Soc. Am.* **90**, 2536–2550 (1991).
- ¹¹G. C. Gaunard, H. Huang, and H. Strifors, “Acoustic Scattering by a Pair of Spheres,” *J. Acoust. Soc. Am.* **97**, 495–507 (1995); “Acoustical scattering by a pair of spheres: Addenda and corrigenda,” *J. Acoust. Soc. Am.* **101**, 2983–2985 (1997).
- ¹²J. Crispin and K. Siegel, *Methods of Radar Cross-Section Analysis* (Academic, New York, 1968), pp. 251–255.

Noise reduction by a barrier having a random edge profile

Steve S. T. Ho, Ilene J. Busch-Vishniac, and David T. Blackstock

Department of Mechanical Engineering, The University of Texas at Austin, Austin, Texas 78712

(Received 14 February 1996; revised 21 November 1996; accepted 21 January 1997)

Sound generally reaches the shadow zone behind a noise barrier by diffraction, or scattering, from the edge at the top of the barrier. By redirecting the incident sound into the shadow zone, the edge acts as a line source. For the traditional straight-edge barrier, the line source is coherent. Since a crooked line source is less coherent, we propose to improve barrier performance by making the edge randomly jagged. Laboratory model experiments to compare insertion loss of straight- and jagged-edge barriers are reported here. A spark was used as a point source, the barriers were thin (compared to a wavelength), and ground and meteorological effects were not important. After preliminary measurements showed that a jagged edge can produce significantly more insertion loss at high frequency, a three level full factorial experiment was done. The results led to an empirical equation for insertion loss of a jagged-edge barrier. Improvement over the straight barrier was found to increase with jaggedness. An unexplained result was the poorer performance of the jagged-edge barrier at low frequency. © 1997 Acoustical Society of America. [S0001-4966(97)05505-7]

PACS numbers: 43.50.Gf, 43.20.Fn [GD]

INTRODUCTION

Sound barriers have been the subject of a large number of experimental, theoretical, and numerical studies. Recent experimental studies have included consideration of the effects of various geometries for barrier caps. For instance, Hothersall *et al.*¹ review work done on barriers with caps having T , Y , and arrow profiles. They conclude that wide (2–3 m) T profiles are the best of these options, and that the ability to increase a barrier insertion loss using a T profile can be enhanced by using strongly absorbent material on the top surface of the T . Fujiwara and Furuta² demonstrate a barrier with a cylinder at the edge, building on theory developed by Keller and Magiros.³ Their results with a practical absorbent cylinder demonstrate an increased insertion loss of 2–3 dB, A-weighted, compared to a conventional thin screen barrier of the same height. Experimental studies of the application of active noise cancellation to barriers have been done. For instance, Ise *et al.*⁴ showed that active cancellation of the virtual source at a barrier edge is possible using edge-mounted secondary sources spaced closer than one-half wavelength. However, they noted that active control including ground reflections would require secondary sources that are not located along the barrier edge.

A number of detailed models of sound scattering at barrier edges have been developed, for a variety of barrier geometries. For instance, Pierce^{5,6} and Kawai and Terai⁷ describe general models of barrier edge diffraction based on wedge theory. Isei *et al.*⁸ and L'Esperance *et al.*⁹ have presented barrier diffraction models which include ground effects. Hayek¹⁰ has developed a model including both ground effects and absorbing surfaces. However, these detailed acoustical models generally are too complicated for the standard computations required by various state agencies involved in design and construction of highway barriers. Simpler models based on geometrical acoustics and empirical inclusion of diffraction are usually employed. The best known simpler models are those of Kurze,¹¹ Kurze and

Anderson,¹² and Maekawa.¹³ Their models find use in virtually all numerical algorithms currently employed by state highway agencies. Maekawa's purely empirical model is based on diffraction measurements made in the laboratory with a thin rigid screen and a pulsed tone.

In our work we start with the model of Kurze and Kurze and Anderson, which is based on a phenomenological fit of full scale data on highway noise barriers. A relationship between the measured insertion loss and the Fresnel number for the given barrier geometry is developed. Here the Fresnel number N_0 is defined as the extra distance the sound must travel with the barrier present, normalized by the half-wavelength $\lambda/2$ of the sound. For the barrier shown in Fig. 1, the Fresnel number is given by

$$N_0 = \frac{2(a+b-d)}{\lambda}, \quad (1)$$

where a , b , and d are the distances defined in Fig. 1. For diffraction of a point source at the edge of an infinitely long barrier, Kurze and Anderson¹² give the following expression for insertion loss IL:

$$\text{IL} = 5 \text{ dB}$$

$$+ 20 \log_{10} \left(\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})} \right), \quad \text{for } -0.2 < N < 12.5, \\ = 24 \text{ dB}, \quad \text{for } N > 12.5, \quad (2)$$

where

$$N = N_0 \cos(\theta), \quad (3)$$

and θ is the angle between the normal to the barrier and the line connecting the source and the receiver (Fig. 1). In this model, as is clear from Eq. (2), insertion loss increases with Fresnel number, up to a point of saturation, 24 dB. Saturation is set by ground reflections and atmospheric effects, which include wind and temperature gradients. However, because the measurements reported here were done in quiet labora-

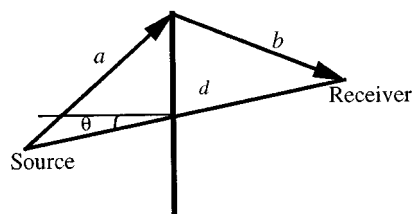


FIG. 1. Barrier geometry.

tory air, under conditions in which ground reflections play no part, saturation was not expected to be important. For our study, therefore, the appropriate version of Eq. (2) is the first line, with no restriction on the value of N .

Consider sound radiating from a stationary point source and striking the edge of a noise barrier. Because the sound scatters at the barrier edge, one can model the edge as a set of infinitesimal sound sources. The strength and phase of each edge source depends on its distance from the real point source. For a conventional barrier with a straight edge, the strength varies slowly, the phase more rapidly, but both monotonically as one moves along the edge away from the point of closest approach. The straight edge of the barrier thus resembles an acoustic line source, that is, phase and amplitude shaded in a deterministic way. Sound observed in the shadow zone is the result of coherent addition of signals from the effective sources along the line. At any location in the shadow zone, the coherent addition results in interference, which is constructive at some frequencies, destructive at others. Because highway noise tends to be broadband, and because we use broadband logarithmic figures of merit for noise barrier effectiveness, the constructive interference dominates the noise metric. This argument leads us to hypothesize that a barrier's performance could be improved by upsetting or preventing the constructive interference. For this purpose we propose that the edge of the barrier be made randomly jagged instead of straight. The aim is to interrupt the monotonic phase variation (and to a lesser extent strength variation) of the edge sources. In other words we propose a crooked line source. As the remainder of this article will show, changing the barrier geometry in this fashion can produce a significant benefit.

The idea of using a random edge to improve noise barrier performance grew out of the work of Bailey *et al.*,¹⁴⁻¹⁶ who used a jagged edge to affect diffraction by circular apertures, disks, and rectangular barrier plates. Underwater measurements with very short pulses from a spark source showed that the scattered signal from an aperture (or disk) can practically disappear when the aperture (or disk) edge is made sufficiently jagged. Measurements with a rectangular barrier also showed degradation of the scattered signal when the barrier edge was made jagged.

Wirt^{17,18} has also considered deviating from the straight edge geometry of conventional barriers in order to improve barrier performance. However, his motivation was to vary barrier properties deterministically, not randomly. His design, a picket-fence-type barrier, was intended to alter the transmission characteristics, or acoustical transparency, of the barrier in a regular way. His scale model experiments

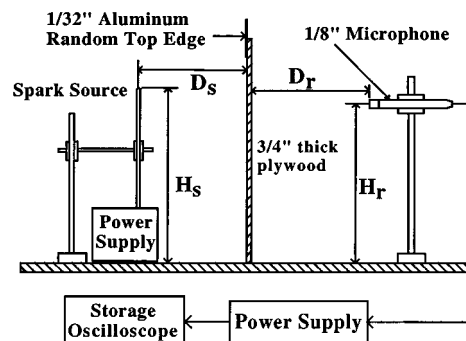


FIG. 2. Schematic of random-edge noise barrier experiment.

with picket-fence-edge barriers demonstrated 1–4 dB improvements in insertion loss, compared to a barrier with the same average height and a straight edge.

The next section of this article describes preliminary experiments performed to find out whether making the barrier edge jagged produces a significant impact on insertion loss. Since the results of the preliminary experiments were positive, we then undertook a more detailed set of experiments (Sec. II). The aim was to develop a phenomenological model of the insertion loss of a jagged edge barrier. Section III is given to a discussion of our results and suggested avenues for further research.

I. PRELIMINARY EXPERIMENTS

As mentioned above, our initial experiments had a single goal: to determine whether introduction of a random fluctuation of the barrier height significantly improves the insertion loss of the barrier. Our approach to investigate this question was to use physical scale models of barriers. The experimental procedure is described below, followed by a presentation of results.

A. Experimental procedure

Figure 2 shows the geometry used in both the preliminary experiments and the later ones. The apparatus was set up on a large table in an indoor laboratory. An electrical spark (Grozier apparatus) was used to generate a short acoustical pulse. The time waveform of the pulse was a sharp positive phase followed by a negative phase. The duration, 40–50 μ s, implies a spatial length of about 1.5 cm. The spark was a point source in the sense that the length of the electrical arc was very short compared with other lengths in the experiment. The height of the spark source above the “ground” (table top) is designated H_s . A $\frac{1}{8}$ -in. B&K microphone at height H_r was the receiver. Located between the spark source and microphone was the barrier, a sheet of plywood 1.5 m long, 0.9 m high, and 1.9 cm ($\frac{3}{4}$ in.) thick. The distance from the barrier to the source is denoted D_s ; that from the barrier to the microphone D_r .

In order to make it possible to vary the top edge of the barrier without completely replacing the whole barrier for each experimental configuration, we capped the plywood with one of two aluminum sheet metal strips. The strips had thickness 0.08 cm ($\frac{1}{32}$ in.), length 1.5 m, and nominal height about 8 cm. The height of one of the barrier caps was uni-

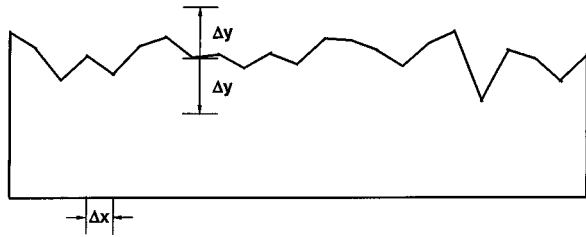


FIG. 3. Example of a random-edge profile used in the preliminary experiments. Here Δx is the horizontal spacing between points of random height within defined bounds of $\pm \Delta y$.

form. The other was cut so as to have a random edge pattern, but average height the same as that of the uniform strip. The pattern was generated in two steps. First, a horizontal spacing δx and the maximum of the vertical fluctuation δy were chosen. Both were scaled to the wavelength at the spectral peak, about 20 kHz, given by the manufacturer for the spark source. The horizontal spacing was chosen to be a fixed value, $\delta x_i = 9$ mm, about half the wavelength at 20 kHz. Randomness was achieved by the height fluctuation δy_i , which was limited to a maximum of ± 18 mm. Second, a random number generator was used to determine the value of δy_i at the end of the i th horizontal interval. Straight lines were used to connect the resulting points, and the sheet metal was cut accordingly. A sample edge profile is shown in Fig. 3. Although the edge is random only vertically, and not horizontally, the design was felt to be sufficient for this study.

The microphone signal was amplified, narrow-band filtered, and displayed on an oscilloscope. In order to maximize the signal-to-noise ratio in the experiments, we chose to use the maximum peak-to-trough excursion in the microphone signal as the measure of the received sound pulse. Within the pulse, specific location of the maximum excursion was not important. Use of an electromagnetic trigger from the spark source allowed us to gate out all signals except those traveling over the barrier. Thus, pulses reflected from the table and other laboratory surfaces, as well as those diffracted from the end edges of the barrier, did not interfere with the measurement because they arrived too late. Indeed it was for this reason that we chose to use a spark rather than a continuous sound source.

The experimental procedure just described differs from most conventional tests of barrier diffraction in two ways. First, impulses were used rather than steady-state sounds. Pulse methods have of course been employed by others, for example, Papadopoulos and Din,¹⁹ who also used acoustic impulses to study sound barrier attenuation. Second, the barrier we used is quite short. While a number of studies have been made of the insertion loss produced by finite length barriers (for example, Lam and Roberts,²⁰ Pirinchieva,²¹ and Takagi *et al.*²²), the focus has been on the issue of sound diffracted by the barrier side edges as well as the top edge. As explained above, however, in our experiments any side scattered sound pulse arrived too late to interfere with measurement of the peak-to-trough pressure in the pulse scattered from the top edge. Our barrier therefore resembles an infinitely long one.

With $D_s = 0.6$ m and $D_r = 0.2, 0.3, 0.4, 0.5, 0.6$, and 0.7

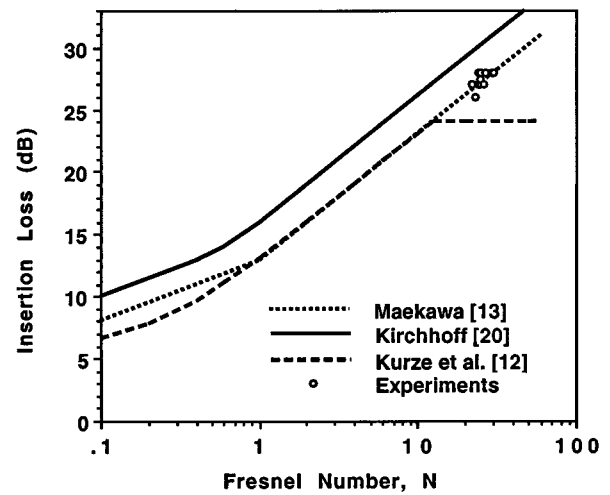


FIG. 4. Measured insertion loss versus predictions by existing models.

m, the following test configurations were examined for the straight edge and the jagged edge barrier: $H_s = H_r = 0.7$ m (source and receiver at the same height), $H_s = 0.7$ m, $H_r = 0.5$ m (source higher than receiver), and $H_s = 0.5$ m, $H_r = 0.7$ m (receiver higher than source). In all cases the microphone was in the shadow zone of the barrier. For each test configuration, results were obtained in the following bands: 2–4 kHz, 4–8 kHz, 8–16 kHz, and 16–25 kHz. Five trials were made for each test and the average result obtained. Variations within each set of five trials were below one decibel. Results were converted to insertion loss through comparison with measured results obtained with the barrier removed.

B. Results

Three types of results were obtained from the preliminary experiments: (1) comparison of the measured and predicted insertion loss for a barrier with a straight edge; (2) comparison of time signals from a straight and a random edge barrier; and (3) comparison of the insertion loss produced by a straight barrier with that by a jagged edge barrier. Each of these is discussed below.

Figure 4 shows the measured insertion loss for the straight-edge barrier compared to predictions of various theoretical models. As the figure shows, although our measurements are concentrated in a very narrow range of Fresnel numbers, the results agree well with the models of Maekawa¹³ and Kurze and Anderson.¹² The prediction based on Kirchhoff's theory taken from optics is well known to be about 3 dB higher the predictions from the other two models.

Figure 5 shows typical microphone time signals captured on the oscilloscope for the straight-edge barrier and a random-edge barrier after low- and high-pass filtering the signal at the indicated frequency (i.e., producing a filter centered at the indicated frequency with 20 dB per decade skirts). The difference between the straight-edge and random-edge waveforms is substantial. The straight edge produces a clean signal in which the leading peak is always the highest. The pulse from the jagged-edge barrier is irregular and diffuse; the second peak is typically higher than the

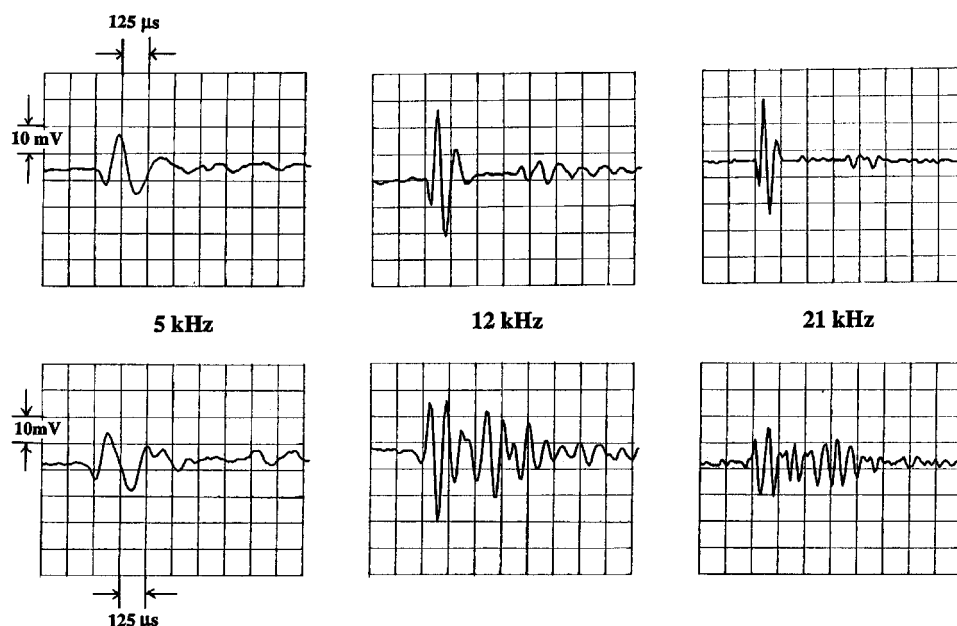


FIG. 5. Microphone time signals of straight-edge barrier (top) and random-edge barrier (bottom). Time signals are high-pass and low-pass filtered at the specified frequency, with 20 dB per decade skirts.

first. The waveforms show the difficulty of selecting an appropriate metric by which to measure the jagged-edge pulse. The peak-to-trough excursion was selected because of its simplicity (only an oscilloscope was required to measure it). Better metrics are indicated for future studies.

Figure 6 shows a typical result for the insertion loss measurements. The difference between $IL(R)$, insertion loss for a random barrier, and $IL(S)$, insertion loss for a straight barrier, is plotted against frequency, where the result at any given frequency was obtained using a filter identical to that employed in Fig. 5. A value greater than zero indicates that the random-edge barrier produces more insertion loss, while a negative value shows better performance by the straight-

edge barrier. For Fig. 6 the jagged-edge barrier is better above about 5 kHz. The improvement generally grows as the Fresnel number increases, or as the distance from the barrier to the receiver decreases, all other parameters being held constant. In the frequency range in which the jagged barrier has the higher insertion loss, the improvement is 3–7 dB, a significant difference.

For the data shown in Fig. 6 the source and receiver are at the same height ($H_r = H_s = 0.7$ m). The effect of varying the two heights was found to be small. For instance, Fig. 7 shows data obtained in the same way but for the receiver lower than the source ($H_r = 0.5$ m, $H_s = 0.7$ m). The largest variation in the results for the two cases is 2 dB.

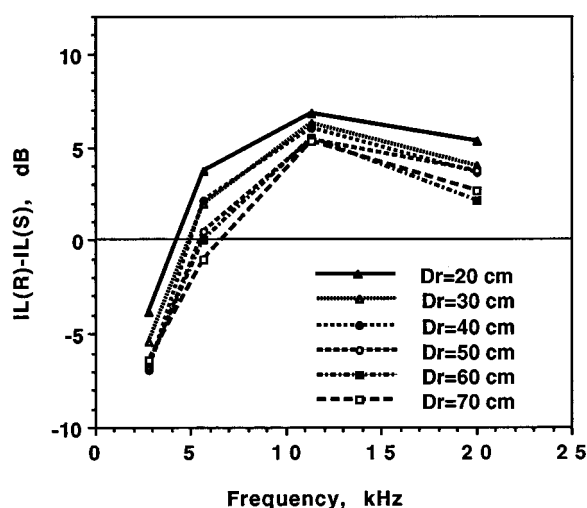


FIG. 6. Comparison of insertion loss for a random-edge barrier, $IL(R)$, and a straight-edge barrier, $IL(S)$, with the source and receiver at heights of 0.7 m, and the distance to the receiver, D_r , varying. Time signals are high-pass and low-pass filtered at the specified frequency, with 20 dB per decade skirts.

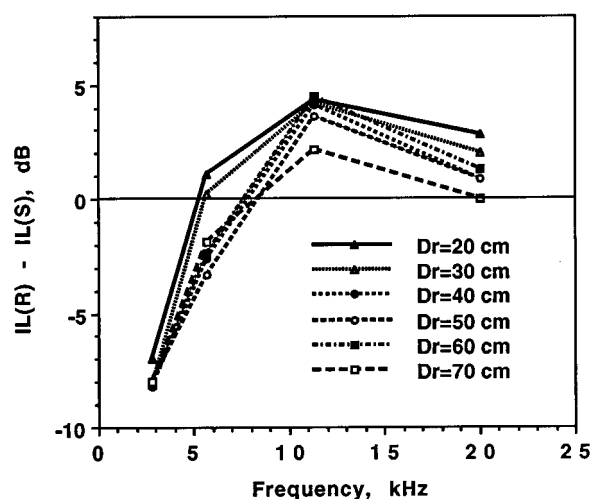


FIG. 7. Comparison of insertion loss for a random-edge barrier, $IL(R)$, and a straight-edge barrier, $IL(S)$, with the source height at 0.7 m, the receiver height at 0.5 m, and the distance to the receiver, D_r , varying. Time signals are high-pass and low-pass filtered at the specified frequency, with 20 dB per decade skirts.

TABLE I. Parameter limits for 3 level full factorial experiment.

Variables	Lower	Intermediate	Upper
Source frequency, f	5 kHz	12 kHz	21 kHz
Fresnel number, N	2	7	20
Δx	9 mm	25 mm	37 mm
Δy	± 5 mm	± 11 mm	± 18 mm

The results of the preliminary study showed that the jagged-edge barrier deserves further attention. Its insertion loss can be significantly greater than that of a straight-edge barrier of the same average height. In the next section a second set of experiments is described. The purpose was to obtain more detailed experimental data and to develop a phenomenological model similar to that used by Kurze and Anderson for the straight-edge barrier.

II. FULL FACTORIAL EXPERIMENTS

The preliminary experiments had a modest goal of simply determining whether a random edge makes a significant difference in a barrier's insertion loss. Armed with a positive answer to this question, we next sought to determine what parameters influence the insertion loss for a random-edge barrier. We designed a set of experiments which would permit us to vary parameters and develop a model to describe the insertion loss observed in terms of the parameters. In the following section the experimental design to obtain an empirical formula for insertion loss is discussed. Next, the method is tested by applying it to a straight-edge barrier. Finally, measurements are used to obtain an empirical formula for a jagged-edge barrier.

A. Design of experiments

While statistical analysis once was considered useful only in psychological experiments, the rise of statistical process control (SPC) in industrial production²³⁻²⁵ has led to the wide acceptance of statistical experimental methods throughout science and engineering. Factorial design is a particularly useful, and widely employed, method of experimental design. The experiment is designed by identifying the parameters that will be varied and the values each will assume. For a full factorial design of the sort we chose to apply, each possible combination of parameter values is tested. For instance, a full factorial experiment with three factors, each having two values (or levels) means that there are 2^3 test configurations. Using regression analysis or analysis of variance methods, it is possible to examine the results of a factorial experiment and develop a phenomenological model of significant parameters, with an estimate of the confidence in the model's fit of the data. (For further information on factorial experiments see Refs. 26 and 27.)

Based on the preliminary experimental results, we chose to vary four factors: the source frequency, the Fresnel number, the horizontal spacing Δx used in generating the random-edge pattern, and the maximum height fluctuation Δy about the mean. We chose to begin with these four factors even though the Fresnel number is not independent of frequency. (Later we reduced the number of factors to three

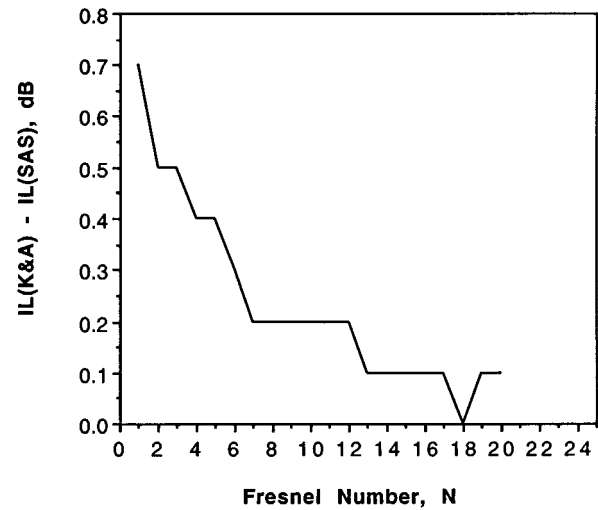


FIG. 8. Comparison of Kurze and Anderson prediction of insertion loss, $IL(K\&A)$, and new phenomenological model, $IL(SAS)$.

by scaling Δx and Δy by the wavelength and removing frequency as a separate parameter.) For each factor we chose three levels, i.e., three values it could assume. Table I shows low, intermediate, and high values for each of the four factors with which we began. Note that our choice of a *full* factorial experimental design meant that nine different random edges had to be constructed.

Given a four factor, three level full factorial design, there were 3^4 or 81 test configurations for the irregular edge barrier. In addition, we obtained data for a straight-edge barrier with the first two factors in Table I identical to those for the random-edge tests, and for the case of no barrier present (only frequency varying). This adds an additional 12 test configurations. For each of the 93 total configurations, we repeated the experiment 10 times, leading to a total of 930 data on peak-to-trough values. The data thus obtained were analyzed using a commercially available statistical package running standard ANOVA and linear regression analyses to determine phenomenological models from data.

The experimental procedure used for the full factorial experiments was the same as that for the preliminary experiments except that the filter was set to low- and high-pass filter at the selected frequency rather than in a relatively wide band.

B. Results—Barrier with a straight edge

Given our prior knowledge of the Kurze and Anderson curve fit given in Eq. (2), we chose to search for a model for the straight edge having the same general form. In other words, rather than simply determine a curve fit using Fresnel number and frequency, we defined the model variables as frequency and the logarithmic function of Fresnel number given in the equation. The result obtained from the regression analysis was a model which follows:

$$IL = 3.88 + 21 \log_{10} \left[\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})} \right]. \quad (4)$$

TABLE II. Variable sets investigated for random-edge barriers.

No.	Var. 1	Var. 2	Var. 3	Var. 4	Var. 5	Quad.	Int.	R^2
1	$\log_{10}\left(\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})}\right)$	Δx	Δy	f		no	no	0.750
2	$\log_{10}\left(\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})}\right)$	Δx	Δy	f		no	full	0.774
3	$\log_{10}\left(\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})}\right)$	Δx	Δy	f		yes	linear	0.832
4	$\log_{10}\left(\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})}\right)$	$\Delta x/\lambda$	$\Delta y/\lambda$			no	no	0.750
5	$\log_{10}\left(\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})}\right)$	$\Delta x/\lambda$	$\Delta y/\lambda$			no	full	0.762
6	$\log_{10}\left(\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})}\right)$	$\Delta x/\lambda$	$\Delta y/\lambda$			yes	linear	0.778
7	$\log_{10}\left(\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})}\right)$	$\log_{10}(\Delta x/\lambda)$	$\log_{10}(\Delta y/\lambda)$			no	no	0.738
8	$\log_{10}\left(\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})}\right)$	$\log_{10}(\Delta x/\lambda)$	$\log_{10}(\Delta y/\lambda)$			no	full	0.755
9	$\log_{10}\left(\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})}\right)$	$\log_{10}(\Delta x/\lambda)$	$\log_{10}(\Delta y/\lambda)$			yes	linear	0.779
10	$\log_{10}(\sqrt{2\pi N})$	$\Delta x/\lambda$	$\Delta y/\lambda$			no	no	0.75
11	$\log_{10}(\sqrt{2\pi N})$	$\Delta x/\lambda$	$\Delta y/\lambda$			no	full	0.762
12	$\log_{10}(\sqrt{2\pi N})$	$\Delta x/\lambda$	$\Delta y/\lambda$			yes	linear	0.778
13	$(\sqrt{2\pi N})$	$\Delta x/\lambda$	$\Delta y/\lambda$			no	no	0.75
14	$(\sqrt{2\pi N})$	$\Delta x/\lambda$	$\Delta y/\lambda$			no	full	0.762
15	$(\sqrt{2\pi N})$	$\Delta x/\lambda$	$\Delta y/\lambda$			yes	linear	0.778
16	$\log_{10}\left(\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})}\right)$	$\Delta x/\lambda$	$\Delta y/\lambda$	$\log_{10}\left(\frac{2\pi N}{\tanh(2\pi N)}\right)$		no	no	0.754
17	$\log_{10}\left(\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})}\right)$	$\Delta x/\lambda$	$\Delta y/\lambda$	$(\sqrt{2\pi N})$	$(2\pi N)$	no	no	0.764
18	$\log_{10}\left(\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})}\right)$	$\Delta x/\lambda$	$\Delta y/\lambda$	$\log_{10}\left(\frac{2\pi N}{\tanh^2(\sqrt{2\pi N})}\right)$		no	no	0.754

The r^2 value for this regression, which is a measure of the experimental deviation from the generated model, is 0.95. As 1.00 is the maximum r^2 value obtainable, our results indicate that almost all of the variation in the data between the various straight-edge test configurations is explained by Eq. (4).

Comparison of Eq. (4) with Eq. (2) (first line only; saturation not appropriate) shows that our results correspond well to those previously published. Figure 8 shows the difference between Eq. (4) and Eq. (2) as a function of Fresnel number. Here IL(K&A) is the insertion loss predicted using Kurze and Anderson's model with no saturation, and IL-(SAS) is the measured result for the straight-edge barrier. The largest discrepancy over the range tested is 0.7 dB. This result gave us confidence in our experimental design and measurements.

C. Results—Random-edge barrier

When the statistical model was tested by applying it to the straight-edge barrier, a clear form of the curve fit from which to start was available. For the random-edge barrier, the problem was more difficult because of the need to assume a form for the important terms in the model, and to do

the regression analysis for the given form. The *best* model is the one that describes the greatest amount of experimental variation that results from changing the values of the test parameters, i.e., that produces the highest value of r^2 .

A great many models were considered in order to find a good fit to the experimental data for the jagged-edge barrier. Table II lists the variables for each of the 18 models and the resulting values for r^2 . As Table II shows, we considered five different approaches in searching for a good fit to the data: (1) inclusion of interaction terms; (2) use of quadratic terms; (3) normalization of the spatial parameters; (4) variation of the form of terms involving Fresnel number or spatial parameters; and (5) use of a power series approach.

For any given set of variables, it is possible to specify how complicated the model should be. The standard linear regression produces a curve fit with linear terms of each variable. However, with full factorial experiments, one may include interaction terms, which are products of more than one of the variables. Further, one may include quadratic terms of the variables, and progressively more sophisticated interaction terms. Indeed, one could certainly get progressively better fit to data by simply adding more and more

terms to the model. However, such an approach tends to produce progressively less physical insight about the problem. Hence, although we tried models with interaction terms and with quadratic terms, we eliminated these because they offered little or no significant improvement in the fit (r^2 nearly the same). As Table II also shows, most of the variable sets we investigated used normalized parameters, and we generally found these to be better than those involving unnormalized spatial parameters. We chose to normalize the spatial parameters defining the random edge with the wavelength, λ .

Of the models shown in Table II, two are simple and as good in terms of r^2 , in fitting the data, as all the other models tested. These two models are identical except for the form assumed for the Fresnel number variable. In one case, number 4 in Table II, we used the logarithmic form used in the curve fit for the straight-edge barrier. In the other, number 10, we simplified this expression to $\log_{10}(\sqrt{2\pi N})$. While one would normally lean toward the least complicated expression, and thus the latter model, two mitigating factors prompted our choice of the former model: ease of comparison with the results for the straight-edge barrier, and applicability over a broader range of Fresnel numbers than tested in our experiments. Indeed, for the range of Fresnel numbers used in our tests, there is almost no distinction between the two forms of the variables used for the Fresnel number influence. Hence, we chose model number 4 for the irregular edge barrier, and with it obtained the following expression:

$$IL = 5.45 + 19.64 \log_{10} \left[\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})} \right] - 1.29 \frac{\Delta x}{\lambda} + 4.11 \frac{\Delta y}{\lambda}. \quad (5)$$

The r^2 value for this model is 0.75, indicating that while it is the best of the models examined, a significant amount of the variation in the insertion loss data still is not explained. From Table I, it can be seen that the ranges of $\Delta x/\lambda$ and $\Delta y/\lambda$ used were 0.134–2.131 and 0.07–1.125, respectively. Equation (5) is valid only within these limits.

Equation (5) permits us to gain valuable physical insight about the behavior of jagged-edge barriers. First, consider the limit of a straight-barrier edge, with Δx and Δy set to zero. The result is remarkably close to both our experiments and the model of Kurze and Anderson, once again conveying faith in the experimental procedures.

Second, the form of the expressions suggest that the insertion loss is the greatest for the minimum possible (non-zero) value for Δx , and the maximum value for Δy . This corresponds to making the edge as jagged or rapidly fluctuating as possible. Figure 9 shows the insertion loss improvement obtained for the most jagged edge in our tests as a function of Fresnel number, along with the prediction of Eq. (4). For our experiments, the improvement is modest, varying from 2.5 dB to 5.0 dB. However, to the extent that the model developed here is applicable beyond the range of parameters used in the test, one can anticipate substantial improvement in insertion loss with a suitably jagged edge.

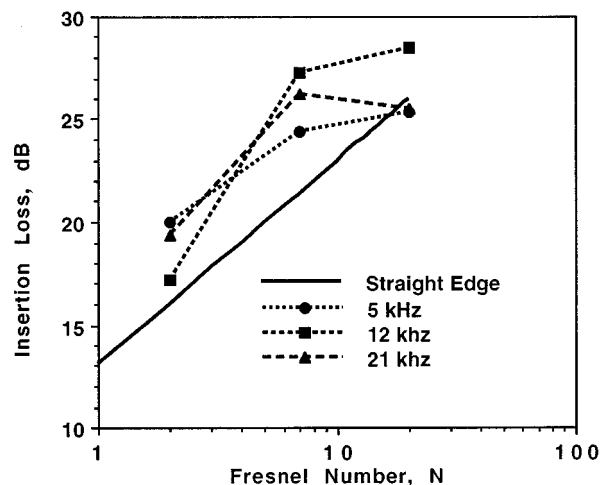


FIG. 9. Measured insertion loss of the most jagged-edge barrier at 5 kHz, 12 kHz, and 21 kHz compared to the measured insertion loss of a straight-edge barrier.

Third, it is clear from the data and Eq. (5) that it is possible to choose a random-edge fluctuation that results in a performance worse than that of a straight-edge barrier of same average height. Indeed, one can calculate the Fresnel number at which the transition from worse to better occurs by solving for the value of $\Delta y/\lambda$ for which the insertion loss predictions of Eqs. (4) and (5) match. This happens when

$$\frac{\Delta y}{\lambda} = 0.331 \log_{10} \left[\frac{\sqrt{2\pi N}}{\tanh(\sqrt{2\pi N})} \right] + 0.314 \frac{\Delta x}{\lambda} - 0.382. \quad (6)$$

Figure 10 shows where the crossover occurs as a function of $\Delta x/\lambda$ for different values of Fresnel number. Above each line in the figure the insertion loss of the random edge barrier is improved compared to the straight edge. Below each line

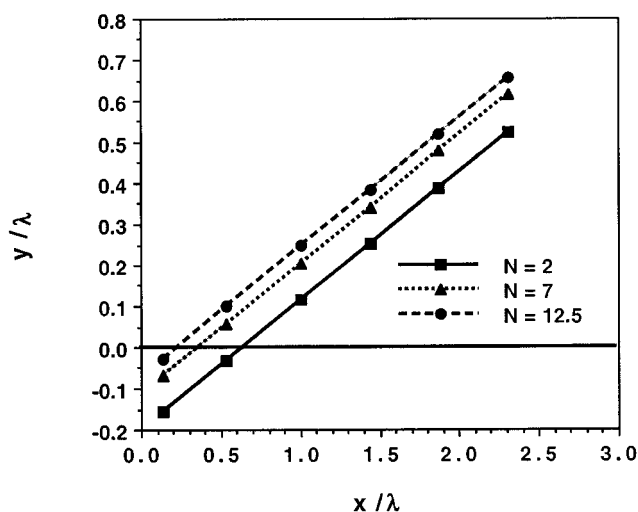


FIG. 10. Boundary between region in which straight-edge barrier has higher insertion loss and the region in which the jagged-edge barrier has a higher insertion loss. Here y/λ is the maximum height variation permitted, normalized by the wavelength of sound, and x/λ is the horizontal spacing between random height assignments in forming the jagged edge, normalized by the wavelength of sound.

the straight edge barrier performs better. Note that Fig. 10 shows that the improvement scales with Fresnel number, meaning it is frequency dependent. Higher-frequency signals require a more jagged edge in order to achieve improvement.

III. CONCLUSIONS AND RECOMMENDATIONS

Based on the results shown in this article, it is clear that the acoustical performance of noise barriers can be significantly improved by introducing a random fluctuation to the height of the barrier edge. This result is quite important commercially, as it suggests that one could make highway noise barriers more effective for the same cost, or reduce their average height (and cost) without loss of sound attenuation.

The results described here came from a large number of tests run on small scale barriers set in a laboratory. A phenomenological model has been developed, based on these tests, which indicates that for random-edge profile barriers, the more jagged the edge of a barrier, the higher the insertion loss. While this result makes sense intuitively, a great need exists for a theoretical model capable of producing a prediction for jagged-edge barriers. We have just begun to tackle this difficult problem. We believe that such a theoretical model may be of interest in other areas of acoustics, for example sound scattering at the rough surface of a body of water.

Also needed are full scale tests, under controlled conditions, of jagged-edge barriers. We have received generous offers of materials from a number of manufacturers of highway noise barriers, and intend to proceed with these tests shortly.

Finally, our tests show that it is possible to make the insertion loss of a barrier worse rather than better at low frequencies through introduction of a random-edge profile. Although we are not certain of the cause for this negative effect, we intend to study it further.

ACKNOWLEDGMENTS

This work was supported by the Texas Department of Transportation through Contract No. 1471. We are grateful to our program manager, John Chase, for permitting us to pursue an unusual research avenue. We are also grateful for the assistance of our professional colleagues working on a highway barrier design guide for the state of Texas: Dr. Richard Klingner, Dr. Michael McNerney, Mr. Ronald Peron, and Mr. Paul Calamia.

- ¹D. C. Hothersall, D. H. Crombie, and S. N. Chandler-Wilde, "The performance of T-profile and associated noise barriers," *Appl. Acoust.* **32**, 269 (1991).
- ²K. Fujiwara and N. Furuta, "Sound shielding efficiency of a barrier with a cylinder at the edge," *Noise Control Eng. J.* **37**, 5 (1991).
- ³J. B. Keller and D. G. Magiros, "Diffraction by a semi-infinite screen with a round end," *Commun. Pure Appl. Math.* **14**, 457 (1961).
- ⁴S. Ise, H. Yano, and H. Tachibana, "Basic study on active noise barrier," *J. Acoust. Soc. Jpn. (E)* **12**, 299 (1991).
- ⁵A. D. Pierce, *Acoustics: An Introduction to its Physical Principles and Applications* (McGraw-Hill, New York, 1981).
- ⁶A. D. Pierce, "Diffraction of sound around corners and over wide barriers," *J. Acoust. Soc. Am.* **55**, 941 (1974).
- ⁷Y. Kawai and T. Terai, "The application of integral equation methods to the calculation of sound attenuation by barriers," *Appl. Acoust.* **31**, 101 (1990).
- ⁸T. Isei, T. F. W. Embleton, and J. E. Piercy, "Noise reduction by barriers on finite impedance ground," *J. Acoust. Soc. Am.* **67**, 46 (1980).
- ⁹A. L'Esperance, J. Nicolas, and G. A. Daigle, "Insertion loss of absorbent barriers on ground," *J. Acoust. Soc. Am.* **86**, 1060 (1989).
- ¹⁰S. I. Hayek, "Mathematical modeling of absorbent highway noise barriers," *Appl. Acoust.* **31**, 77 (1990).
- ¹¹U. J. Kurze, "Noise reduction by barriers," *J. Acoust. Soc. Am.* **55**, 504 (1974).
- ¹²U. J. Kurze and G. A. Anderson, "Sound attenuation by barriers," *Appl. Acoust.* **4**, 35 (1971).
- ¹³Z. Maekawa, "Noise reduction by screens," *Appl. Acoust.* **1**, 157 (1968).
- ¹⁴M. R. Bailey, D. T. Blackstock, and E. L. Carstensen, "Isolation of a negative pressure pulse by means of diffraction," *J. Acoust. Soc. Am.* **92**, 2359(A) (1992).
- ¹⁵M. R. Bailey and D. T. Blackstock, "Isolation of a negative-pressure pulse by means of diffraction," in *Advances in Nonlinear Acoustics—13th ISNA*, edited by H. Hobaek (World Scientific, Singapore, 1993), pp. 158–163.
- ¹⁶M. R. Bailey, "Isolation of a negative-pressure pulse and studies of irregular-edge diffraction," M.S. thesis, Mechanical Engineering Department, The University of Texas at Austin, May 1994.
- ¹⁷L. S. Wirt, "The control of diffracted sound by means of thnadners (shaped noise barriers)," *Acustica* **42**, 73 (1979).
- ¹⁸L. S. Wirt, "The control of diffracted sound by means of thnadners," Technical Report, Lockheed-California Company, November 1976.
- ¹⁹A. I. Papadopoulos and D. G. Din, "A study of barrier attenuation by using acoustic impulses," *J. Acoust. Soc. Am.* **90**, 1011 (1991).
- ²⁰Y. W. Lam and S. C. Roberts, "A simple method for accurate prediction of finite barrier insertion loss," *J. Acoust. Soc. Am.* **93**, 1445 (1993).
- ²¹R. Pirinchieva, "Model study of the sound propagation behind barriers of finite length," *J. Acoust. Soc. Am.* **87**, 2109 (1990).
- ²²K. Takagi, R. Hotta, and K. Yamamoto, "A simple method for the calculation of noise attenuation by a finite length barrier," *Appl. Acoust.* **43**, 353 (1994).
- ²³P. C. Badavas, *Real-Time Statistical Process Control* (Prentice-Hall, Englewood Cliffs, NJ, 1993).
- ²⁴A. V. Feigenbaum, *Total Quality Control* (McGraw-Hill, New York, 1983), 3rd ed.
- ²⁵J. M. Juran and F. M. Gryna, *Quality Planning and Analysis* (McGraw-Hill, New York, 1993), 3rd ed.
- ²⁶R. J. Brook and G. C. Arnold, *Applied Regression Analysis and Experimental Design* (Marcel Dekker, New York, 1985).
- ²⁷D. C. Montgomery, *Design and Analysis of Experiments* (Wiley, New York, 1976), 2nd ed.

Annoyance caused by sounds of wheeled and tracked vehicles

Niek J. Versfeld and Joos Vos^{a)}

TNO Human Factors Research Institute, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands

(Received 22 January 1996; revised 15 July 1996; accepted 17 December 1996)

In a laboratory experiment, the relationship between the annoyance caused by sounds of tracked and wheeled vehicles was investigated. Sounds of various vehicles, ranging from a main battle tank (MBT) to a passenger car, were recorded at several distances. Subjects were presented with these sounds and were asked to judge the annoyance if they were exposed to them at home on a regular basis. Overall, sounds emanated from military tracked vehicles were judged to be less annoying than civil passenger cars, provided that the indoor A-weighted sound-exposure levels (ASELs) were the same. It should be noted that receiver distance and operating conditions were not specifically matched for military and civil vehicles but were chosen to represent typical ranges of values found in practice. For steady-speed conditions, the sounds were judged to be equally annoying if the (indoor) ASELs of the tracked armored personnel carrier (APC) and the MBT exceeded that of a passenger car by 5 and 10 dB, respectively. In specific driving conditions such as alternate accelerations and decelerations, however, these differences were smaller. For ASELs below 55 dB, passenger cars were found to be more annoying than heavier wheeled vehicles, such as a bus or a truck with a trailer. The data indicate that the difference between the high-frequency part and the low-frequency part of the spectrum might play a role in the annoyance. With respect to noise-zoning procedures it is highly interesting that the outdoor ASEL could predict the annoyance, as rated indoors, better than could any indoor noise measure. The 5- and 10-dB bonuses to the indoor level of the APC and the MBT, respectively, were reduced to 1 and 3 dB if the dose was measured outdoors. © 1997 Acoustical Society of America. [S0001-4966(97)00705-4]

PACS numbers: 43.50.Qp, 43.50.Ba, 43.50.Lj [GAD]

INTRODUCTION

A significant part of the sound exposure near military training fields may originate from heavy tracked vehicles such as the armored personnel carrier (APC) and the main battle tank (MBT). At present it is not clear how communities around military training fields are affected by these heavy vehicle sounds: To our knowledge no dosage-effect relationships for annoyance caused by military vehicle sounds have been reported. Such dose-response relationships are not necessarily equal to those for road-traffic sounds, as have been accumulated by, among others, Schultz (1978) and Fidell *et al.* (1991). This may occur due to differences in both the spectral and temporal properties of the various sounds.

Relevant differences between training military vehicles and civil road traffic are, e.g., the property of the surface on which they are typically driven (sandy soil versus asphalt covered roads), the distribution and range of vehicle weight, and engine power. These differences all have an effect on the spectral contents of the sounds. The differences in distance between the source and the receiver, as well as the driving speed, on the other hand, affect the temporal envelope. For example, the duration of an individual vehicle passby is short especially for traffic sounds produced in residential streets, whereas during training in a military zone, vehicles are at relatively large distances from residential areas and vehicle noise levels show less temporal variation.

To explore possible procedures for the assessment of annoyance due to military vehicle sounds, the present laboratory study was designed. Subjects were presented with sounds from wheeled and tracked military vehicles. As a reference, sounds of various civil wheeled vehicles, such as passenger cars, buses, and trucks, were included. The sounds were presented in a simulated indoor home environment. For each sound, subjects were asked to rate their annoyance assuming they were exposed to it at home in the living room on a regular basis.

I. METHOD

A. Stimuli

Sound recordings were made of military vehicles passing by on a sandy track of a military training field. The ground was essentially flat and the intervening ground cover was sand. Microphones were located at 50, 100, 200, and 400 m from this track, along a line perpendicular to the track, at approximately 1.5 m above ground level. Recordings were made of a tracked main battle tank (Leopard-2), a tracked armored personnel carrier (YPR-765), a wheeled 4-ton truck, and a four-wheel drive car (Jeep).

In a second session, sound recordings of civil road traffic were made, using four microphones located at 12.5, 25, 50, and 100 m from the center of a two-lane road, along a line perpendicular to the road, again at about 1.5 m above ground level. The intervening ground cover was grass. Recordings were made of two passenger cars, a delivery van, a bus, and a truck with a trailer, all using their top gears.

^{a)}Electronic mail: joos@tm.tno.nl

During both recording sessions, the wind speed was moderate. Recordings were performed in downwind conditions. There were no obstacles between, or in the vicinity of the sandy track and the road and the microphones that could have had a noise screening effect. Both surfaces were dry.

In both sessions, $\frac{1}{2}$ -in. condenser microphones were used to record the sounds on different tracks of DAT recorders at a sampling rate of 48 kHz and an amplitude resolution of 16 bits. The axes of the microphones were positioned vertically.

Information about the vehicle types and operating conditions is given in the first three columns of Table I. The so-called “washboard” denotes a condition in which a vehicle alternately accelerates and decelerates, in order to simulate the driving over a wavelike terrain. The peaks of such a washboard track were about 2 m in height, and about 25 m apart. The remaining columns of Table I will be discussed below. Portions of the recorded sounds were sampled down to 16 kHz, and stored as files on a hard disk of a personal computer. With a signal processing card sampled sounds could be filtered and converted into analog signals. In order to simulate the frequency-dependent outdoor-to-indoor reduction in sound level, the signals were attenuated with the help of a finite impulse response (FIR) filter on the DSP card. With the windows entirely closed, the attenuation relevant to Dutch dwellings would range from about 10 dB for the 16- and 31.5-Hz octave bands up to 35 dB for the 8 kHz octave band (Vos, 1996). Such a sound reduction was also found to be representative to the houses located in the greater Washington, DC area (Yaniv *et al.*, 1982). In some countries, such as in Switzerland, application of special window glazing might result in higher sound-level reductions for frequencies higher than 2 kHz (Meloni and Rosenheck, 1995).

In the present experiment, however, it was intended to simulate an indoor condition in which the windows were slightly opened, as a result of which for frequencies higher than 125 Hz, the attenuation was about 10 dB lower than had been the case if the windows had been completely closed. To obtain the intended reduction, the signals were attenuated by 10 dB, and low-pass filtered at 125 Hz with a -2 -dB/octave slope.

The filter on the DSP card was also used to compensate, as much as possible, for both the resonances due to room dimensions (since the listening booth was relatively small compared to a standard living room) and the nonflat frequency characteristic of the speaker and the audio chain. At the subject's listening position the response was flat within 2.5 dB over a frequency range from 50 to 6000 Hz. Low-frequency wind noise was removed with a high-pass filter at 17 Hz and a 48-dB/octave slope.

Each stimulus had a duration of 45 s, including a 2-s onset and offset, and contained only one (military or civil) vehicle passage. The maximum level occurred near the middle of the sample. For each vehicle type and operating condition, four recordings (one from each microphone position) were available. They were used to create, at the subject's listening position, stimuli with ASEL between 40 dB (just audible) and 75 dB (rather loud). Conditions in which $L_{AE} < 40$ dB were omitted (as was, e.g., the case with the truck and the jeep at a recording distance of 400 m). In order

to increase the dynamic range and to create a greater overlap between sound levels of the different vehicle types, additional stimuli were created by utilizing signals recorded at the largest distance, and attenuating them in steps of approximately 5 dB (as was done with conditions of the MBT and APC). The ASELs of most conditions were rounded to the nearest multiple of 5 dB, to enable adequate comparisons among the various conditions by means of analyses of variance. Rows 4 and 5 of Table I yield, for each stimulus, the recording distance d (m), from which the stimulus was derived, and the corresponding indoor ASEL, respectively. To illustrate similarities and differences in spectral contents among the various vehicle types, Fig. 1 displays, for one condition of each vehicle type (*viz.*, drive-by at a constant speed), and at a recording distance of 50 m, the 1/3-octave band indoor (linear) sound exposure level as a function of frequency. For the sounds produced by the military vehicles [Fig. 1(a)] and for those produced by the civil vehicles [Fig. 1(b)], most sound energy was situated near 50–100 Hz, and a secondary maximum was present in the region of 500–2000 Hz. This secondary maximum, however, was more prominent for the wheeled vehicles (*cf.* Cermak and Cornillon, 1976).

In order to make the acoustic environment more realistic, a background noise of remote road traffic was continuously present throughout the experiment at an A-weighted average level as low as 30 dB (dotted line in Fig. 1). Only for technical reasons was the background noise not filtered, such that the spectrum contained some resonances due to the room dimensions.

B. Apparatus

The experiment was entirely computer controlled. The sounds were reproduced in two acoustically identical rooms ($4.7 \times 2.6 \times 2.9$ m) by means of an amplifier and two loudspeakers (one for each room). The listener was seated near the middle of the room, behind a table on which a monitor and keyboard were placed. The loudspeaker was placed in front of the listener at a distance of about 2 m, and was hidden behind a curtain. The rooms were well sound-insulated, and the reverberation time was shorter than 500 ms for frequencies above 100 Hz.

C. Subjects

A group of 20 subjects was asked to participate. Two of them were discarded because analyses showed that they did not give consistent responses. They were replaced by two new subjects. The new group consisted of eight females between 20 and 30 years of age (median 21) and 12 males between 16 and 30 years of age (median 23). None of these subjects had a hearing loss larger than 15 dB (over 125–8000 Hz) in their best ear. Subjects were paid for their services.

D. Procedure

Two subjects participated simultaneously. Each session started by screening the subjects' hearing. Next, they were seated in separate listening rooms, and were asked to read

TABLE I. Description of the stimuli of the present experiment. For each stimulus (1) vehicle type, (2) operating condition, (3) estimated driving speed v (km/h), (4) recording distance d (m), (5) A-weighted sound-exposure level L_{AE} (dB), as measured indoors at the subject's listening position, (6) difference ΔL (dB) between the level in a 1/3-octave band near 100 Hz and that in a 1/3-octave band near 1000 Hz, (7) A-weighted sound-exposure level L_{AE} (dB), as measured outdoors, and (8) mean annoyance rating is given.

Vehicle type	Operating condition	ν (km/h)	d (m)	L_{AE} (indoors)	ΔL (dB)	L_{AE} (outdoors)	Annoyance
A. Military vehicles							
Main battle tank	idling	0	50	55	31	66	4.3
			100	50	41	61	2.0
Main battle tank	accelerating	0→50	50	76	44	86	9.1
			100	70	41	80	7.2
			200	65	42	75	5.8
			400	61	45	71	3.7
			400	56	45	66	2.7
			400	51	45	61	2.2
Main battle tank	“washboard”	...	50	76	33	88	9.6
			100	70	37	81	8.1
			200	66	39	76	6.7
			400	61	39	71	6.2
			400	56	39	67	5.2
			400	51	39	60	4.0
Main battle tank	steady speed	15	50	74	35	85	8.4
			100	68	38	79	6.9
			200	64	38	74	5.0
			400	60	37	71	4.7
			400	55	37	66	3.7
			400	49	37	60	2.3
Main battle tank	steady speed	50	50	76	35	88	9.2
			100	70	38	81	7.4
			200	64	42	74	4.9
			400	60	31	70	3.7
			400	54	31	64	2.6
			400	50	31	60	1.8
Armored personnel carrier	steady speed	15	50	67	34	80	7.5
			100	60	38	73	5.8
			200	56	39	69	4.5
			400	51	39	64	2.5
			400	46	39	59	1.5
			400	40	39	53	0.8
Armored personnel carrier	steady speed	50	50	70	21	85	8.9
			100	65	25	79	6.8
			200	59	25	73	5.2
			400	55	27	69	4.5
			400	50	27	64	2.7
			400	40	39	53	0.8
Truck (4 tons)	steady speed	30	50	61	31	75	6.2
			100	56	36	69	3.2
			200	52	39	64	2.1
Jeep	steady speed	15	50	61	34	73	5.3
			100	56	38	67	2.7
			200	51	38	62	2.0
B. Civil vehicles							
Passenger car 1	steady speed	80	12.5	54	21	72	5.4
			25	49	25	66	4.9
			50	45	29	61	2.7
			100	40	32	56	1.6
Passenger car 2	steady speed	80	12.5	55	17	73	5.4
			25	50	22	67	4.0
			50	45	27	61	2.5
			100	41	31	56	1.5
Delivery van	steady speed	80	12.5	56	19	73	5.4
			25	50	22	66	4.0
			50	46	26	61	2.7
			100	41	29	56	1.5
Bus	steady speed	80	12.5	67	35	79	6.5
			25	62	39	73	5.1
			25	57	39	68	3.5
			50	54	43	65	2.5
			50	49	43	60	1.6
			100	44	43	54	1.2
Truck with trailer	steady speed	80	12.5	66	24	81	7.8
			25	60	25	75	6.3
			25	56	25	70	5.3
			50	51	30	64	2.7
			100	46	30	58	1.3

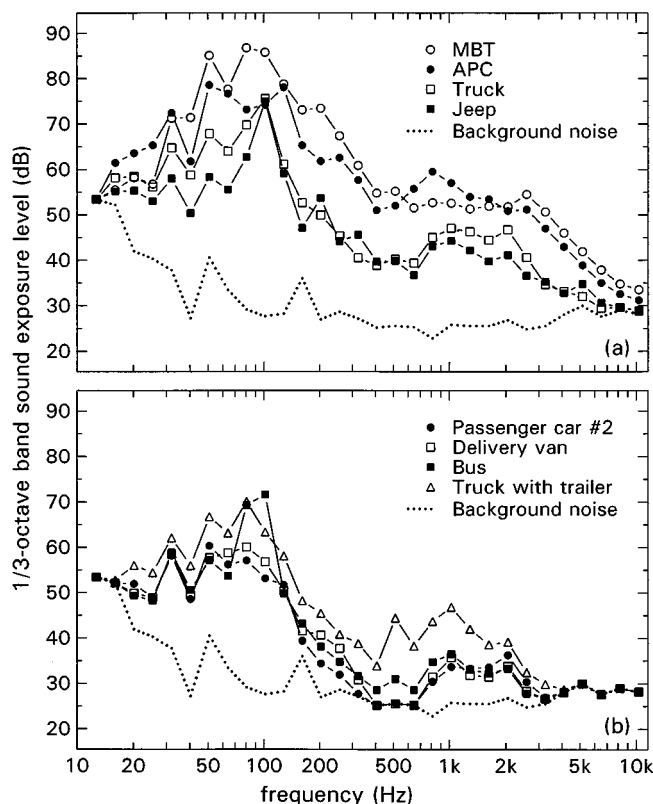


FIG. 1. The 1/3-octave band sound-exposure level (dB) as a function of frequency for a number of vehicles. (a) Military vehicles; (b) civil vehicles. Recording distance was 50 m. Vehicles were driven past at a steady speed.

the instructions. Their main task was to respond after each stimulus presentation to the question: “How annoying would you find the sound in the preceding period if you were exposed to it at home on a regular basis?” They responded by pressing a button corresponding to a digit from 1 (“not annoying at all”) to 10 (“extremely annoying”). Buttons with digits from 2–9 were not verbally labeled.

Subjects received five practice trials, followed by 65 experimental trials. To test the subject’s responses for consistency, 15 of the 65 stimuli were presented twice. The stimuli and replications were mixed, but never presented directly after each other. The stimuli were assigned to ten blocks; presentation order of the blocks was balanced, that of the eight stimuli within each block was quasirandomized. The total duration of the experimental session was about two hours, including a break in between.

II. RESULTS

A. Within- and between-subject correlation

Correlation coefficients, r , between the first and the second ratings of the 15 stimuli that were replicated were calculated for each subject individually. The 20 subjects produced high r values between 0.64 and 0.93. For the two discarded subjects, r was only 0.19 and 0.40. The first and second ratings were not significantly different ($p > 0.10$). The correlation matrix, based on all responses from all subjects showed that the average between-subject correlation ranged between 0.60 and 0.88; the average r value was 0.75.

With these analyses it was ascertained that the (final) group of 20 subjects produced consistent results, and reasonably agreed upon the degree of the annoyance caused by the different stimuli, as a result of which averaging across replication and subjects is allowed.

B. Annoyance of individual vehicle sounds

For each vehicle, operating condition, and recording distance, the mean annoyance ratings are given in the last column of Table I. For each vehicle type and operating condition, the function

$$S(L_{AE}) = 1 + 9 \cdot \Phi \left[\frac{L_{AE} - \mu}{\sigma} \right] \quad (1)$$

was fitted to the data. $S(L_{AE})$ denotes the score on the annoyance rating scale, and is a function of ASEL. $\Phi(z)$ denotes the cumulative normal (Gaussian) distribution, given by

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-(1/2)t^2} dt. \quad (2)$$

This function has been rescaled such that it covers the entire rating scale (varying from 1 to 10). The value for μ denotes the L_{AE} value at which an annoyance rating of 5.5 is obtained; σ is a measure for the steepness of the function: A small value for σ denotes a steep function. The values of μ and σ were estimated by a least-squares fit of Eq. (1) to the raw data (replications included). For each vehicle type and operating condition, μ and σ , as well as the proportion of explained variance, r^2 , is given in Table II.

Given that the functions were fitted to the raw data, the r^2 values were very high in almost all conditions ($0.48 \leq r^2 \leq 0.81$). Only for the idling condition a lower r^2 value of 0.29 was obtained. F tests indicated that, for all cases, the data were not described significantly worse if the value of σ was fixed at the average value of 10.5 dB. The values for μ and r^2 , given that σ is set to $L_{AE} = 10.5$ dB, are given in the last two columns of Table II. Indeed, r^2 dropped only marginally, and μ changed hardly, meaning that the functions may be considered as being parallel. Note that stimulus conditions are arranged according to descending values of μ , i.e., from most to least annoying. *Post hoc* Tukey’s paired-comparison tests were performed on subsets of the data. Vertical bars on the entire left-hand side in Table II indicate that the values for μ did not differ significantly ($p > 0.05$) between the conditions within the group.

1. Military vehicles

Figure 2 displays the mean annoyance ratings as a function of (indoor) ASEL for each condition with military vehicles. Bars denote one standard error. Figure 2(a) shows that for corresponding ASELs, the dashboard condition was found to be more annoying than the other MBT conditions, which on their turn yielded ratings that were rather similar. The results of the *post hoc* tests confirmed that, at a given ASEL, the annoyance ratings for the different operating conditions of the MBT shown in Fig. 2(a) (the “dashboard” and idling condition excluded) were not significantly different. Figure 2(b) shows that the annoyance caused by the

TABLE II. Results of a least-squares fit of Eq. (1) to the data. For the various vehicle types and driving conditions μ and σ , expressed as the A-weighted sound-exposure level in decibels, as well as the proportion of explained variance r^2 , are given. The last two columns show μ and r^2 for the analyses in which σ is fixed at 10.5 dB. Vertical bars indicate that the annoyance did not differ significantly among conditions ($p > 0.05$).

Vehicle type	Operating condition	ν (km/h)	μ	σ	r^2	μ	r^2
Passenger car 1	steady speed	80	53.5	10.1	0.55	53.6	0.55
Passenger car 2	steady speed	80	54.6	10.0	0.55	54.8	0.55
Delivery van	steady speed	80	55.5	10.4	0.48	55.5	0.48
Main battle tank	"washboard"	...	57.4	15.2	0.49	57.8	0.48
Truck with trailer	steady speed	80	58.2	9.3	0.71	58.3	0.71
Armored personnel carrier	steady speed	50	59.1	11.8	0.62	59.0	0.62
Main battle tank	idling	0	56.4	5.3	0.29	59.1	0.29
Armored personnel carrier	steady speed	15	59.3	10.0	0.69	59.4	0.69
Truck	steady speed	30	60.3	6.4	0.62	61.7	0.62
Jeep	steady speed	15	61.1	6.9	0.53	62.8	0.52
Bus	steady speed	80	63.5	10.5	0.72	63.5	0.72
Main battle tank	steady speed	15	63.5	13.0	0.66	63.7	0.64
Main battle tank	steady speed	50	64.5	10.3	0.81	64.6	0.81
Main battle tank	accelerating	0→50	64.5	9.8	0.79	64.6	0.79
<i>Averaged across vehicle type</i>							
Passenger cars	steady speed	80	54.6	10.2	0.52	54.7	0.52
Armored personnel carrier	steady speed	15 and 50	59.3	10.9	0.67	59.2	0.67
Main battle tank	steady speed and accelerating	15 and 50	64.2	13.6	0.76	64.2	0.76
<i>All conditions</i>	61.2	13.6	0.59	60.6	0.59

truck and the jeep sounds was similar, and somewhat less than that of the tracked APC. Neither for the MBT, nor for the APC did the annoyance at a given ASEL depend on the driving speed.

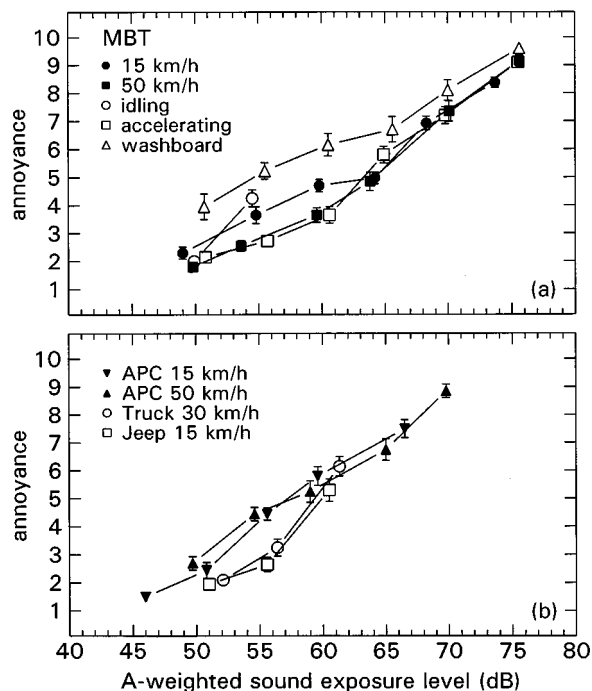


FIG. 2. Annoyance ratings as a function of the A-weighted sound-exposure level (dB) for (a) the main battle tank in five operating conditions, and (b) the armored personnel carrier, the truck, and the jeep.

2. Civil vehicles

Figure 3 shows the mean annoyance ratings for each civil condition separately. Bars again denote one standard error. The mean annoyance ratings for the two passenger cars and the delivery van coincided, whereas the mean annoyance ratings for the bus passby sounds were consistently lower than those for the passenger cars and the delivery van. For $L_{AE} < 55$ dB, the annoyance caused by sounds of the truck with a trailer tended to be similar to that of the bus. For $L_{AE} > 55$ dB, the annoyance caused by sounds of the truck with a trailer was comparable to the degree of annoyance that for the same level range would be expected for the passenger car sounds.

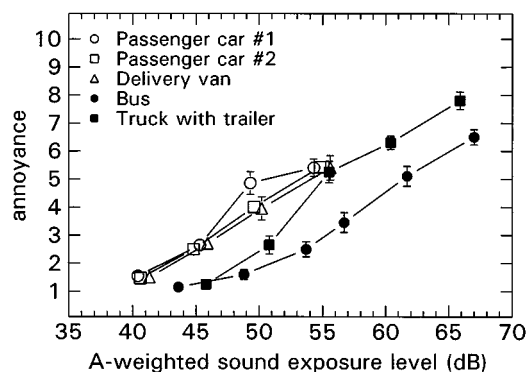


FIG. 3. Annoyance ratings as a function of the A-weighted sound-exposure level (dB) for the two passenger cars, a delivery van, a bus, and a truck with a trailer.

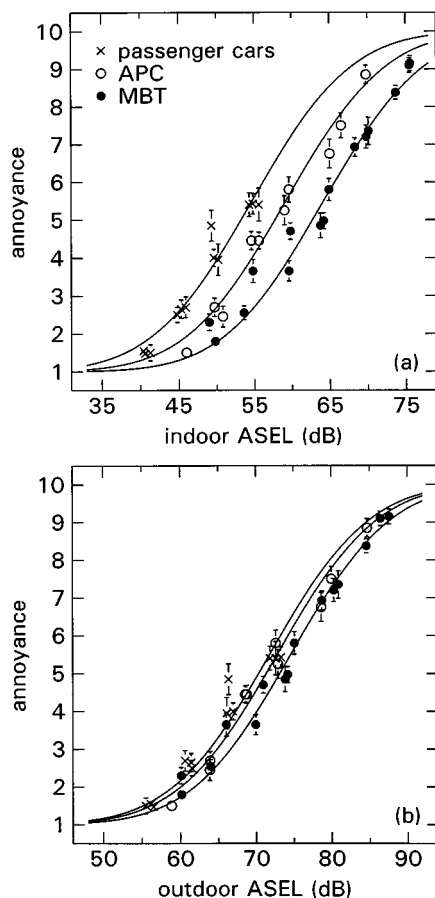


FIG. 4. Annoyance ratings as a function of the A-weighted sound-exposure level (dB) for three groups of vehicles: Passenger cars and the delivery van (crosses), tracked armored personnel carrier at two driving speeds (open symbols), and tracked main battle tank at two driving speeds or accelerating (filled symbols). (a) Indoor-measured A-weighted sound-exposure level, (b) outdoor-measured A-weighted sound-exposure level. The solid lines denote a fit of Eq. (1) to the data sets, with σ fixed at 10.5 dB.

3. Comparison among vehicle types

Figure 4(a) summarizes the present results for three groups of vehicles. The results for the first group (indicated by crosses) were obtained with sounds from the two passenger cars and the delivery van; those for the second group (open circles) were obtained for the tracked APC, at two driving speeds; those for the last group (filled circles) were obtained with the tracked MBT (two driving speeds; one accelerating condition). Figure 4(b) will be discussed later on. Tukey *post hoc* tests showed that averaging across these subgroups was allowed. To the data of each group a sigmoid function [Eq. (1)] was fitted, both with σ free or fixed at $L_{AE} = 10.5$ dB. Obtained values of μ and σ are given in the last rows of Table II. It is clear, both from Fig. 4(a) and the data in Table II that at a given ASEL the drive-by sounds of different vehicle types yielded very different annoyance ratings. To obtain the same annoyance rating, overall differences in the ASEL between military and civil vehicles could be as large as 10 dB.

III. DISCUSSION

The main conclusion drawn from the results of the present experiment is that for corresponding ASELs, sounds

of different vehicles and/or operating conditions can give rise to different annoyance scores. The data in Fig. 4(a) and in Table II suggest that the annoyance decreases with increasing engine power, or vehicle weight.

Cermak and Cornillon (1976) and Cermak (1979) asked their subjects to compare a wide range of road-traffic sound recordings. Despite the use of a variety of time- and/or spectrum-dependent measures, Cermak and Cornillon (1976) and Cermak (1979) did not find a significant contribution of measures that could differentiate between the passenger cars and heavy trucks. This indicates that there probably was little difference between the annoyance of passenger cars and heavier vehicles at the same noise exposure level.

Rasmussen (1979) measured the annoyance caused by simulated road-traffic noise, as produced by wheeled vehicles. He used synthetic sounds to minimize the influence from nonacoustical factors such as associations. The frequency spectra of the sounds were modeled according to the "average" car or truck spectrum. Contrary to the data in the present paper (cf. Fig. 1), the sound spectrum of the simulated passenger car passby contained stronger low-frequency components than did the spectrum of the simulated truck. Noise from passenger cars was found to be more annoying than that from trucks. When both sounds were equally annoying, the difference in the A-weighted average sound level was about 7 dB, which is, despite the different spectra, in line with the present results.

In the experiment reported by Yaniv *et al.* (1982) the samples contained either passenger cars, trucks, or both. They found that for comparable A-weighted average sound levels, the annoyance was not significantly affected by vehicle type.

Although the results of all papers cited above indicate that A-weighted noise-rating indices are good predictors for annoyance, there is, within the category of wheeled vehicles, no consensus on whether annoyance depends on vehicle type.

A recent field-laboratory study carried out by Schomer *et al.* (1994), included a few conditions in which the annoyance caused by the sounds from tracked vehicles (MBT and APC) was compared with the annoyance caused by wheeled vehicles, both vehicle types driven on graded dirt roads. In paired comparison tests the listeners were presented with wheeled and tracked vehicle passby sounds and were asked, for each pair, to indicate which sound was more annoying, the first or the second one. The outdoor ASELs of the MBT and APC sounds were always about 80 and 73 dB, respectively, whereas the level of the wheeled vehicle sounds was changed by the use of different vehicles such as a passenger van, a Jeep, a bus, a large tow truck, etc. The results of these preliminary tests showed that for indoor listening conditions (with windows closed or slightly opened), the sounds were judged equally annoying if the indoor ASEL of the tracked vehicle sounds was 2–3 dB higher than that of the wheeled vehicle sounds.

Both in the present study and in the study of Schomer *et al.* (1994), the annoyance depended on vehicle type; an alternative acoustic measure, possibly in combination with another frequency weighting network, might therefore be

more appropriate. On the other hand, it might be possible that ASEL is appropriate indeed, but that additional acoustic factors play a role, such as differences in temporal variation or differences between frequency spectra. Lastly, *outdoor* measures might be superior to indoor measures in predicting the indoor perceived annoyance. These alternatives will be discussed below.

A. Other acoustic measures

Figure 4(a) showed that the annoyance of heavy vehicles [containing a relatively large amount of low-frequency energy, cf. Fig. 1(a)] was lower than that of light vehicles [containing a relatively large amount of high-frequency energy, cf. Fig. 1(b)]. If the sound levels would be measured with other frequency weighting networks, such as B, C, or linear, the data of the different vehicle groups would spread even further, since these weighting networks emphasize the low-frequency part more than does the A-weighting network. This is consistent with the results reported by Watts and Nelson (1993). In fact, only a weighting network that emphasizes the high-frequency part of the spectrum can account for the differences in annoyance.

A second possibility is that ASEL is not the most appropriate measure for describing the annoyance of these sounds. To investigate this, all stimuli were analyzed with respect to six A-weighted noise-rating indices. Each signal, as measured indoors, was sliced into 450 100-ms portions, and the A-weighted level in each portion was used for the calculation of the noise-rating indices. The noise-rating indices were [see Schultz (1982) for a detailed description of these noise-rating indices] (1) the 45-s average sound level, L_{Aeq45s} , which is 16.5 dB lower than ASEL; (2) the noise pollution level (NPL) which is similar to L_{Aeq45s} , but contains a correction term to account for the variability in the sound level; (3) the primed 45-s average sound level, L'_{Aeq45s} , which is similar to L_{Aeq45s} , but contains a correction term to account for the rate of change in the sound level; (4) the maximum A-weighted sound level, L_{max} , reached in one or more 100-ms slices; (5) the A-weighted sound level that is exceeded in 10% of the time, L_{10} ; and (6) the traffic noise index (TNI) which is a linear combination of the A-weighted level that is exceeded 90% of the time and the level that is exceeded 10% of the time.

The correlations between the different acoustic measures were without exception high, ranging from 0.94 (between L_{max} and TNI) to 0.996 (between L_{Aeq45s} and L'_{Aeq45s}). This result is in agreement with that of Cermak and Cornillon (1976) and Yaniv *et al.* (1982).

Next, for each acoustic measure, a function similar to that given in Eq. (1) was fitted to the total data set. The higher proportions of explained variance were obtained for L_{max} (0.61), NPL (0.59), and L_{Aeq45s} (0.58). The differences in the goodness-of-fit among these three acoustic measures were not statistically significant. Since, in the present study, the difference between ASEL and L_{Aeq45s} was just a constant of 16.5 dB, it must be concluded that for the set of A-weighted measures investigated here, there is no real better alternative to ASEL. Apparently, the differences among the dose-response relationships for the different vehicle

groups, as found in the present study, cannot be explained by a relatively simple acoustic measure.

B. Additional factors

Since neither conventional frequency weighting networks nor noise-rating indices accounted for the systematic differences in annoyance caused by the various groups of vehicle sounds, it is worthwhile to examine whether there are other factors in the signal that contribute to the annoyance.

1. Spectral differences

Informal listening sessions led the authors to believe that the level difference between the low-frequency part and the high-frequency part of the spectrum might account for a portion of the variance. To that end, the difference ΔL between the highest sound exposure level in the 1/3-octave band near 100 Hz and that in the 1/3-octave band near 1000 Hz was calculated for all stimulus conditions. Values of ΔL are given in Table I, and were used to fit the equation

$$S(L_{AE}, \Delta L) = 1 + 9 \cdot \Phi \left[\frac{L_{AE} - \mu - \delta(\Delta L / \Delta L_0)}{\sigma} \right] \quad (3)$$

to the entire data set. Here, μ , σ , and δ are free parameters, and ΔL_0 is a constant, set to 1 dB. A least-squares fit to the data yielded L_{AE} values for μ , σ , and δ of 48.5, 12.2, and 0.37 dB, respectively ($r^2 = 0.65$). Fisher's r to Z transformation showed that the increase of about 6% in the variance was significant ($Z = 2.8$, $p < 0.01$).

2. Distance of recording

Subjects might have rated remote sound sources as being less annoying than nearby sources, because remote sound sources are perceived as being less threatening than nearby sources. Furthermore, a remote passby is characterized by a gradual increase and decrease of the level; a near passby, on the other hand, is rather peaked. Lastly, due to the frequency-dependent air absorption, distance will also be related to the spectral content. To the knowledge of the authors, no paper has reported on the ability to estimate recording distance, but it is believed that such a task can be performed with reasonable accuracy. To capture all these spectro-temporal properties in the signal (and others, such as the Doppler shift), the recording distance d was taken as a predictor.

Distance of recording d (in m) was used to perform a least-squares fit to the annoyance scores, according to the equation

$$S(L_{AE}, d) = 1 + 9 \cdot \Phi \left[\frac{L_{AE} - \mu - \gamma \log_2(d/d_0)}{\sigma} \right], \quad (4)$$

where μ , σ , and γ are free parameters, and $d_0 = 12.5$ m, the shortest recording distance (see Table I). The least-squares fit yielded L_{AE} values for μ , σ , and γ of 57.2, 13.7, and 1.38 dB, respectively ($r^2 = 0.63$). The addition of recording distance as a predictor increased the explained variance by 3%, but this was not a significant improvement ($Z = 1.52$, $p > 0.05$).

C. Annoyance as a function of the outdoor A-weighted sound exposure level

A common practice in noise zoning procedures is to measure (or calculate) the noise level outside, near the façade of a building (e.g., see Schomer *et al.*, 1994). In order to investigate the relationship between indoor noise annoyance ratings and outdoor noise measurements, the outdoor noise levels were determined for all conditions by correcting for the (frequency dependent) outdoor-to-indoor sound attenuation. For the entire set of conditions, a best fit of Eq. (1) to the annoyance ratings and corresponding outdoor ASELs was obtained with L_{AE} values for μ and σ of 73.3 and 11.8 dB, respectively. The proportion of explained variance was equal to 0.67, significantly higher than that obtained with the indoor ASEL ($r^2 = 0.59$), or the best A-weighted sound measure (L_{max} , $r^2 = 0.61$).

A model that could describe the mean annoyance data perfectly, still explains only about 69% of the total variance: the remaining 31% is due to differences among subjects. The fact that only 2% remains unexplained means that outdoor ASEL is a much better predictor than is indoor ASEL. Similar conclusions about the adequacy of outdoor levels in the prediction of indoor annoyance ratings were drawn by, among others, Watts and Nelson (1993).

The present results imply that knowledge of the outdoor ASELs is almost sufficient for prediction of the community response to sounds from wheeled and tracked vehicles. This is illustrated in Fig. 4, where for three groups of vehicles (passenger cars, APC, and MBT) annoyance scores have been plotted both as a function of the indoor ASEL [panel (a)] and as a function of the outdoor ASEL [panel (b)]. In all cases, the cumulative normal distributions were fitted with σ fixed at 10.5 dB.

The bonuses of 5 and 10 dB [Fig. 4(a)] that had to be subtracted from the indoor levels of the tracked armored personnel carrier and the tracked main battle tank, respectively, were reduced to 1 and 3 dB if the dose was measured outdoors. Future research should focus in more detail on the effects on bonus size of alternative outdoor-to-indoor reductions in noise level, as obtained in listening conditions in which the windows are completely opened or completely closed.

The preliminary results obtained by Schomer *et al.* (1994) suggest that the differences in annoyance between wheeled and tracked vehicle sounds, as obtained in the condition with slightly opened windows, are equal to those obtained in the condition in which the windows are completely closed. For outdoor listening conditions, which are spectrally comparable to listening indoors with the windows widely opened, the relative effects determined by Schomer *et al.* (1994) suggest that the differences in annoyance between the wheeled and tracked vehicle sounds will be smaller than those obtained in the condition with slightly opened windows.

IV. CONCLUSIONS

The primary aim of the present laboratory study was to relate the annoyance caused by the sounds of wheeled

and tracked military vehicles to that caused by (much lighter) wheeled civil vehicles, such as passenger cars. It should be noted that receiver distance and operating conditions were not specifically matched for military and civil vehicles but were chosen to represent typical ranges of values found in practice. From the results it was concluded that:

(1) Overall, the sounds produced by the military tracked vehicles were rated to be less annoying than those produced by the light wheeled civil vehicles at the same A-weighted indoor sound-exposure level (ASEL). For steady-speed conditions, bonuses of 5 and 10 dB had to be subtracted from the indoor ASELs of the tracked armored personnel carrier (APC) and the tracked main battle tank (MBT), respectively, in order to obtain the ASEL of equally annoying civil passenger cars.

(2) In specific driving conditions of the military tracked vehicles, such as alternate accelerations and decelerations, however, a bonus should not be applied.

(3) For relatively low ASELs ($L_{AE} < 55$ dB), heavier wheeled vehicles, such as a bus and a truck with a trailer, were rated as being less annoying than the passenger cars.

(4) The prediction of the annoyance caused by the various vehicle sounds is enhanced if the difference in level between the low- and the high-frequency part of the spectrum is added as a second factor: The smaller this level difference, the higher the annoyance.

(5) A finding relevant to noise-zoning procedures was that the ASEL measured outdoors could predict the annoyance experienced indoors very well. The bonuses of 5 and 10 dB that had to be subtracted from the indoor levels of APC and MBT, respectively, were reduced to 1 and 3 dB if the dose was measured outdoors. Future research should focus in more detail on the effects on bonus size of alternative outdoor-to-indoor reductions in noise level, as obtained in listening conditions in which the windows are completely opened or completely closed.

ACKNOWLEDGMENTS

This research was financed by the Dutch Ministry of Defence, in close cooperation with Pieter I. J. van der Weele. The authors are grateful to Frank W. M. Geurtsen for his assistance with stimulus preparation and data collection. The constructive comments made by the anonymous referees on a previous version of this manuscript are also gratefully acknowledged.

- Cermak, G. W. (1979). "Exploratory laboratory studies on the relative aversiveness of traffic sounds," *J. Acoust. Soc. Am.* **65**, 112–123.
- Cermak, G. W., and Cornillon, P. C. (1976). "Multidimensional analyses of judgements about traffic noise," *J. Acoust. Soc. Am.* **59**, 1412–1420.
- Fidell, S., Barber, D. S., and Schultz, T. J. (1991). "Updating a dosage-effect relationship for the prevalence of annoyance due to general transportation noise," *J. Acoust. Soc. Am.* **89**, 211–233.
- Meloni, T., and Rosenheck, A. (1995). "Choice of frequency weighting for the evaluation of weapon noise," *J. Acoust. Soc. Am.* **97**, 3636–3641.
- Rasmussen, K. B. (1979). "Annoyance from simulated road traffic noise," *J. Sound Vib.* **65**, 203–214.

- Schomer, P. D., Wagner, L. R., Benson, L. J., Buchta E., Hirsch, K.-W., and Krahé, D. (1994). "Human and community response to military sounds: Results from field-laboratory tests of small-arms, tracked-vehicle, and blast sounds," *Noise Control Eng. J.* **42**, 71–84.
- Schultz, T. J. (1978). "Synthesis of social surveys on noise annoyance," *J. Acoust. Soc. Am.* **64**, 377–405.
- Schultz, T. J. (1982). *Community noise rating* (Applied Science, London, UK).
- Vos, J. (1996). "Annoyance caused by impulse sounds produced by small, medium-large, and large firearms," *Proceedings Internoise 1996* (Institute of Acoustics, St. Albans, UK), Book 5, pp. 2231–2236.
- Watts, G. R., and Nelson, P. M. (1993). "The relationship between vehicle noise measures and perceived noisiness," *J. Sound Vib.* **164**, 425–444.
- Yaniv, S. L., Danner, W. F., and Bauer, J. W. (1982). "Measurement and prediction of annoyance caused by time-varying highway noise," *J. Acoust. Soc. Am.* **72**, 200–207.

A parametric error analysis of the backward integration method for reverberation time estimation

Dennis R. Morgan

Bell Laboratories, Lucent Technologies, 700 Mountain Avenue, Murray Hill, New Jersey 07974-0636

(Received 14 August 1996; revised 26 November 1996; accepted 31 December 1996)

Backward integration of a room impulse response has long been used in room acoustics for the estimation of reverberation time. However, the inevitable noise floor limits the minimum level of the measured impulse response, thereby leading to errors. This paper shows how the error is influenced by the selection of truncation time and evaluation range. A general guideline that emerges from this study is to truncate the measured impulse response at the knee where the main decay slope intersects the noise floor, then measure the slope of the backward integrated truncated impulse response down to a level about 5 dB above the noise floor. © 1997 Acoustical Society of America. [S0001-4966(97)04005-8]

PACS numbers: 43.55.Br, 43.55.Ka, 43.58.Gn [JDQ]

INTRODUCTION

Backward integration of a room impulse response has long been used in room acoustics for the estimation of reverberation time.¹ To illustrate and analyze this method, we postulate a simple model for the measured squared impulse response:

$$\tilde{f}(t) \equiv h^2(t) = [e^{-t/2}r(t) + \sqrt{a}n(t)]^2, \quad (1)$$

where the basic response is an exponential with unity $1/e$ time constant, $r^2(t)$ is a unit-mean multiplicative random process that models fluctuations due to multiple reflected acoustic paths, a is the noise floor power, and $n(t)$ is a unit-variance noise process. Since the time constant is normalized, time values obtained using this model are to be interpreted in units of the actual basic time constant for the particular application at hand. (Note that for mathematical convenience, we deal here with $1/e$ time constants, whereas in applications, one is usually concerned with the reverberation time, which is defined as the time for the response to decay 60 dB; however, these units are readily convertible using the scaling factor $6 \ln 10 = 13.8155$.)

As a simple illustrative example, suppose that $r(t)$ is modeled as a Rayleigh fading process, which exhibits many of the characteristics of actual room responses.² Thus we postulate a random variable R , such that R^2 has an exponential probability density

$$p_{R^2}(r^2) = e^{-r^2}, \quad r \geq 0. \quad (2)$$

[R can be easily generated as $R = \sqrt{(X^2 + Y^2)/2}$, where X and Y are two independent zero-mean, unit-variance, normal random variables.] Suppose further that the additive noise model is Gaussian with probability density

$$p_N(n) = \frac{1}{\sqrt{2\pi}} e^{-n^2/2}. \quad (3)$$

The solid line in Fig. 1 shows a plot of (1) using the above models that was simulated over 100 points using Matlab with $a = 0.0001$ (−40 dB), which is a typical noise floor level. For comparison, the mean exponential response e^{-t} is also

plotted as a dotted line. This simulated response is typical of an actual room impulse response which often exhibits large ragged fluctuations. Note that (2) and (3) are merely for illustration; the results in this paper apply more generally to model (1) with arbitrary processes $r(t)$ and $n(t)$.

Backward integration of (1) from some (normalized) reference time t_1 , initially corresponding to the record length, gives the so-called “decay curve”

$$\tilde{F}(t) \equiv \int_t^{t_1} \tilde{f}(u) du. \quad (4)$$

The rationale for this effective and efficient method can be understood from two viewpoints. Heuristically, backward integration smooths a noisy decaying exponential, obtaining another exponential with the same time constant; this is a basic feature of the exponential, whereby

$$\int_t^\infty e^{-u} du = e^{-t}. \quad (5)$$

Thus by this interpretation, backward integration acts to filter out the randomlike fluctuations of the impulse response while preserving the basic exponential shape. A more fundamental interpretation, given originally by Schroeder,¹ is that backward integration is mathematically equivalent to the ensemble average of the classical interrupted white noise method, whereby a room is excited with white noise which is then suddenly shut off and the decaying signal recorded.

The backward integration (4) using the model of (1) is plotted in Fig. 2 for the illustrative randomizations (2) and (3) with $a = 0.0001$ (−40 dB) and a record length $t_1 = 20$, corresponding to 20 time constants (−86.9-dB dynamic range without the noise floor). As can be seen, the random fluctuations have been smoothed. However, the noise floor has the effect of distorting the desired exponential shape (which is linear on a dB plot). Therefore, if one attempts to fit the experimentally obtained backward integrated curve, the estimated time constant will be biased. (For real room impulse responses, the first 5 dB or so of decay is usually

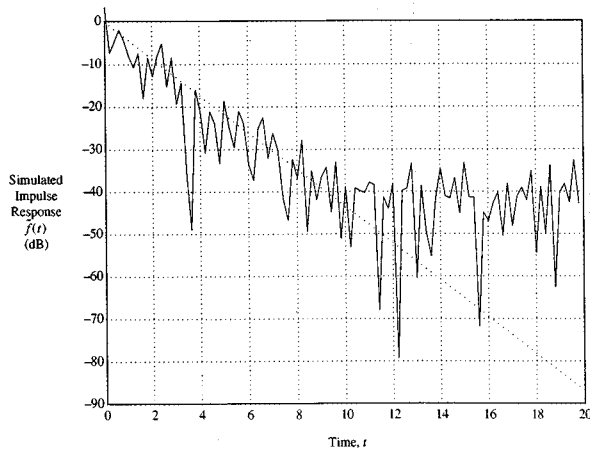


FIG. 1. Simulated impulse response with -40 -dB noise floor.

ignored in this fitting to reduce the anomalous effect of early returns; however, this is of no concern for the idealized model considered in this paper.)

The main results of this paper are relevant to measurements over the largest practical evaluation range (typically about 30 dB) where the noise floor has a significant effect. However, actual rooms may not exhibit strictly exponential decay, especially over the early part, and measurements are also desired over shorter evaluation ranges, e.g., 10 or 20 dB; in those cases, the noise floor effects may be negligible, requiring no special treatment.

A well-known remedy for the noise floor problem is to truncate the impulse response data record length beyond a certain point. Kürer and Kurze³ study one case of a -40 -dB noise floor and recommend truncation at one-half of the reverberation time. Most modern treatments^{4–6} advocate the intuitive and more general rule of truncation at a point where the main slope of the measured impulse response is estimated to intersect the noise floor, as for example could be ascertained from inspection of Fig. 1. In terms of model (1), this procedure would ideally set $t_1 = \ln(1/a)$ and we will refer to this as *truncating at the knee of the curve*. A related problem is how to select the range over which the slope of

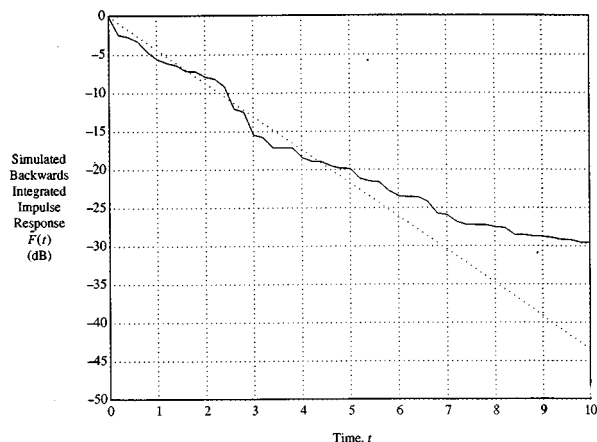


FIG. 2. Simulated backward integrated impulse response with -40 -dB noise floor and data record length $t_1 = 20$ (no truncation).

the backward integrated impulse response (4) is evaluated. Only qualitative or implicit discussion on this point appears in Kürer and Kurze³ and Chu.⁴ Vorländer and Bietz⁵ recommend an evaluation range down to a level at least 10 dB above the noise floor, but give no quantitative rationale.

Other more complicated techniques have also been proposed to ameliorate the noise floor problem. Following the two-step analog technique of Kürer and Kurze,³ Lundeby *et al.*⁶ propose a compensation technique whereby the estimated noise contribution in (4) is subtracted out. More generally, Xiang⁷ developed a nonlinear regression technique for simultaneous estimation of the initial level, slope, and noise floor. These advanced methods^{6,7} have a certain theoretical appeal; however, the inherent modeling and estimation uncertainties leave some doubt as to their accuracy for real applications. The added computational complexity of these methods is also an issue.

In this paper, we limit our investigation to the basic backward integration method because of its robustness, simplicity, and general availability of software packages (e.g., SYSid⁸). A comprehensive analysis will be presented for the effects of the truncation time and evaluation range on the estimated decay time constant. Also, we will examine how well the backward integration decay curve matches the ideal exponential shape. While it can be argued that all of this “fine tuning” may not greatly affect the accuracy associated with real measurements, it does eliminate a small systematic bias, costs nothing to implement, and establishes a methodical process.

For mathematical tractability, the analysis developed here will be formulated in terms of the mean-squared impulse response

$$f(t) \equiv E\{h^2(t)\} = e^{-t} + a, \quad (6)$$

where the expectation E is understood to be over the ensemble of random processes $r(t)$ and $n(t)$ in (1). Thus, we investigate the deterministic consequences of various parameter interactions on the basis of the mean-square response. At the end we will revisit the random aspects using the aforementioned illustrative simulation models.

For simplicity, the analysis will assume a simple two-point fit to the decay curve over the evaluation range from 0 dB down to some specified level, as is implemented, for example, by the SYSid software package.⁸ More complicated mean-square fits, with or without constraints on the starting point, were also investigated, but were found to produce similar results concerning the effect of parameter variations.

I. TIME CONSTANT BIAS

Backward integration of (6) gives

$$F(t) \equiv \int_t^{t_1} f(u) du = e^{-t} - e^{-t_1} + a(t_1 - t), \quad (7)$$

which is similar to previously derived forms.^{3,7} If $t_1 = \infty$ and $a = 0$, then $F(t) = e^{-t}$ and the exact exponential shape is preserved. However, in general, the finite record length and limiting minimum value a will distort the desired exponential shape. Figure 3 shows the response (7) for $a = 0.0001$

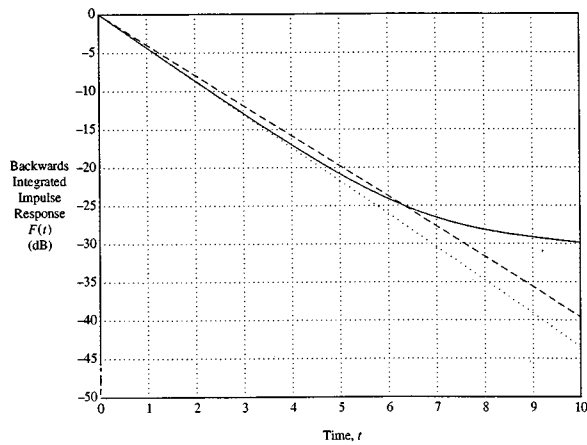


FIG. 3. Backward integrated impulse response with -40 -dB noise floor and data record length $t_1 = 20$ (no truncation).

(-40 dB), corresponding to the mean of Fig. 2. A straight line fit over a 25 -dB range of this curve is shown as a dashed line in this figure. As can be seen, the estimated slope is shifted away from the true dotted line slope and therefore results in a positive bias of the estimated time constant.

If the impulse response is truncated, then the upper limit t_1 of the backward integration is reduced to a value shorter than the actual data record length. This has the effect of reducing the influence of the noise floor a because its cumulative effect is lessened. For example, consider the effect on the example of Fig. 3 if t_1 is reduced to value 6 . As shown in Fig. 4, this causes the curve to bend down, going to zero ($-\infty$ dB) as $t \rightarrow t_1 = 6$. A straight line fit over a 25 -dB range of this curve shows that the positive bias of Fig. 3 has now been traded for a negative bias. So the trick is to choose t_1 appropriately so that a reasonably accurate, unbiased straight line fit to the actual slope can be obtained. Also, the range over which the fit is obtained must be selected. One would like to choose this range as large as possible in order to make the most use of the data and to minimize the sensitivity to random perturbations, as modeled in (1) and visualized in Fig. 2. However, due to the bending induced by noise

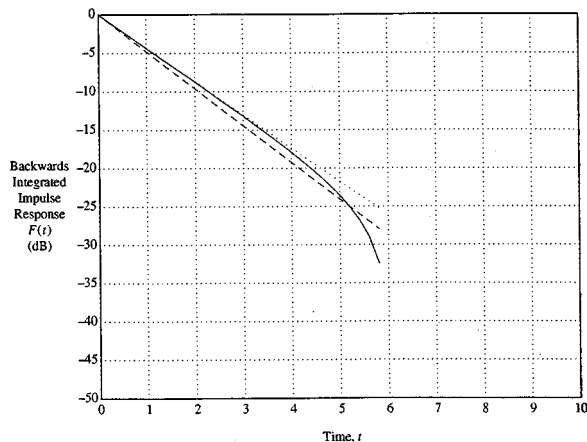


FIG. 4. Backward integrated impulse response with -40 -dB noise floor and truncation time $t_1 = 6$.

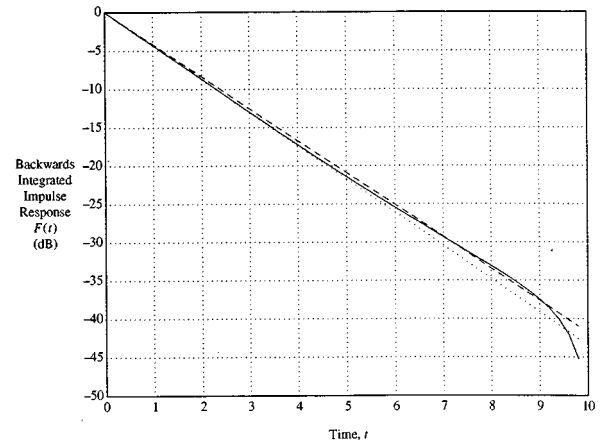


FIG. 5. Backward integrated impulse response with -40 -dB noise floor and truncation time $t_1 = 10$.

floor effects, this range should not be too large. It is the object of this discussion to determine a rule for selecting the truncation length and evaluation range to achieve good results.

Note that we are *not* advocating using the noise level to make the decay curve straight; rather, the method is to be taken as accommodating a given noise level that we can do nothing about. Therefore, as previously stated, the main intent of this paper applies to the determination of decay over the largest practical range, where the noise floor effects are significant. As a matter of practice, it is sometimes desirable to use the same evaluation range in order to enable meaningful comparisons of different halls or rooms. In that case the noise level should be controlled accordingly.

As a first attempt to choose the right value of the truncation interval, suppose that one selects t_1 so as to make the backward integration curve $F(t)$ as straight as possible on a logarithmic scale. Figure 5 shows the result of one such attempt which achieves this objective for $t_1 = 10$. As can be seen, the dashed line fit is quite good to $F(t)$ shown by the solid line. However, there is some bias between the dashed line slope and the actual slope shown by the dotted line. Therefore this strategy clearly fails.

Let us examine the time constant bias that truncation induces on the fitted slope over a given range. The straight line fit is written

$$\hat{F}(t) = (1 - e^{-t_1} + at_1)e^{-t/\hat{\tau}}, \quad (8)$$

where it is assumed that the gain constant is adjusted so that the two curves are equal at time $t = 0$, and $\hat{\tau}$ is the estimated ($1/e$) time constant. Suppose that the fit is obtained over the time interval $(0, t_0)$, i.e., over a range

$$\rho \equiv \frac{F(0)}{F(t_0)} = \frac{1 - e^{-t_1} + at_1}{e^{-t_0} - e^{-t_1} + a(t_1 - t_0)}. \quad (9)$$

Then equating (7) and (8) for $t = t_0$ obtains the time constant estimate

$$\hat{\tau} = \frac{t_0}{\ln \rho}, \quad (10)$$

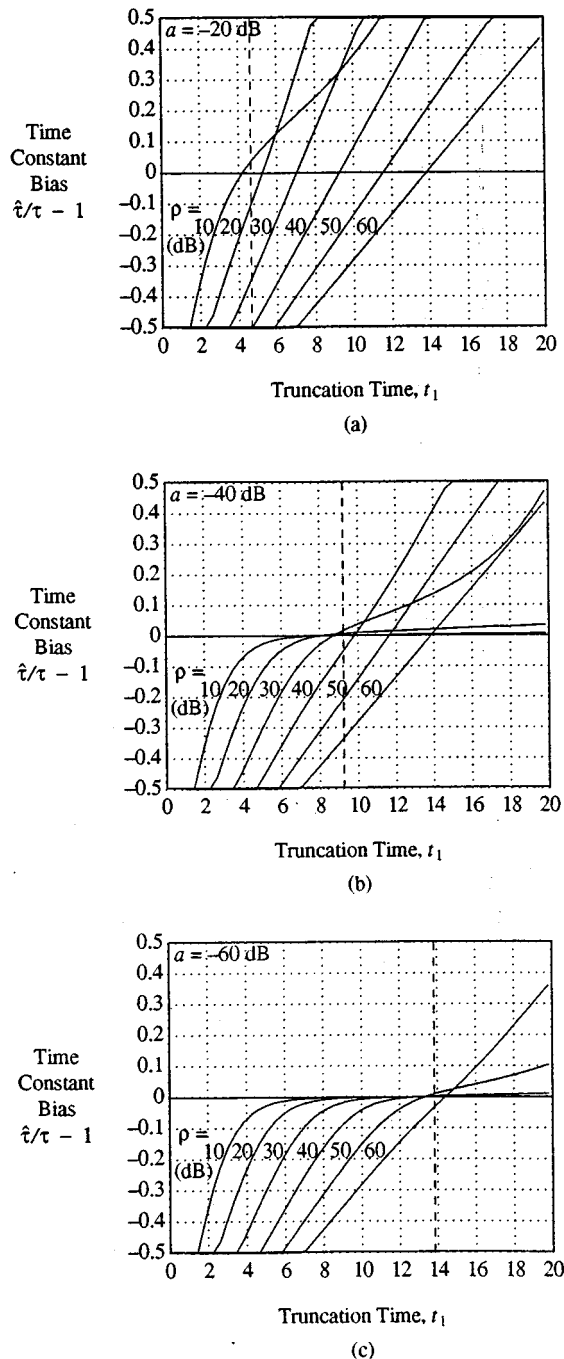


FIG. 6. Normalized time constant bias ($\hat{\tau}/\tau - 1$) as a function of truncation time t_1 for various evaluation ranges ρ and noise floor levels a . (a) $a = -20$ dB, (b) $a = -40$ dB, (c) $a = -60$ dB.

where t_0 is obtained for a given range ρ by solving (9). As an example, consider again the example of Fig. 5, for which $a = 0.0001$ (-40 dB), $t_1 = 10$, and $\rho = 1000$ (30 dB). Solution of (9) gives $t_0 = 7.1769$, which when substituted into (10) gives $\hat{\tau} = 1.0390$. Thus, the bias in this case is 0.0390 or 3.90% , which is in agreement with what is observed in Fig. 5. Figure 6 shows plots of the normalized bias ($\hat{\tau}/\tau - 1$) as a function of truncation time t_1 for several noise floor values a .

II. TIME CONSTANT SENSITIVITY

If one had perfect knowledge of the noise floor a and the actual time constant (normalized to 1 here), one could use the above results to choose the truncation interval t_1 to give zero bias over a given range. In reality, even if the noise floor is accurately measured, the time constant is that which we are trying to estimate, which is by definition unknown. Consequently, we can usually only hazard a guess at the actual time constant, and the “optimal” value of t_1 that we estimate will involve some uncertainty. Therefore, we would like the estimated time constant $\hat{\tau}$ not to be too sensitive to the value of t_1 selected. We will now calculate the sensitivity of the estimate (10) obtained for the idealized model (7). This sensitivity will be defined in the usual way as $(d\hat{\tau}/\hat{\tau})/(dt_1/t_1)$, i.e., the fractional change in $\hat{\tau}$ that results from a fractional change in t_1 . For the methodology employed in this paper, the sensitivity will be calculated so that the range (9) remains fixed, i.e., we choose a certain range over which to fit the decay curve. Since ρ is assumed to be fixed, (10) gives $d\hat{\tau}/\hat{\tau} = dt_0/t_0$, and dt_0/dt_1 can be calculated by setting the total differential of (9) equal to zero. Thus, after some calculus, we obtain

$$\left. \frac{d\hat{\tau}/\hat{\tau}}{dt_1/t_1} \right|_{\rho \text{ fixed}} = \frac{(a + e^{-t_1})t_1}{(a + e^{-t_0})t_0} \left(1 - \frac{1}{\rho} \right). \quad (11)$$

As an example, again consider Fig. 5 for which $a = 0.0001$ (-40 dB) and $t_1 = 10$ with a straight line fit over the range $\rho = 1000$ (30 dB). From before, $t_0 = 7.1769$ and substitution of these values into (11) gives a sensitivity of 0.2342 , or about 23% . Therefore, if t_1 is increased by 20% to the value $t_1 = 12$, then the estimated time constant would increase by about 4.7% ; this value is affirmed by inspection of Fig. 6(b). Figure 7 shows plots of the sensitivity $(d\hat{\tau}/\hat{\tau})/(dt_1/t_1)$ as a function of truncation time t_1 for several noise floor values a . Ideally, we would like the sensitivity to be as small as possible. However, a tradeoff must be made with the bias, which works in the opposite direction; further discussion of this point will follow later.

III. MAXIMUM CURVE FITTING ERROR

Finally, it is of interest to know how close the backward integration decay curve $F(t)$ is to the actual decay curve. We assume that the gain of the true decay curve is scaled to obtain alignment at the initial time $t = 0$ and form the error ratio

$$\epsilon \equiv \frac{F(t)}{(1 - e^{-t_1} + at_1)e^{-t}}. \quad (12)$$

Differentiation of (12) obtains the stationary point

$$t_{\max} = t_1 - 1 - e^{-t_1}/a. \quad (13)$$

Therefore, the maximum error ratio over the evaluation range from $t = 0$ to $t = t_0$ is determined as

$$\epsilon_{\max} = \begin{cases} 1, & t_{\max} \leq 0, \\ \epsilon(t_{\max}), & 0 < t_{\max} < t_0, \\ \epsilon(t_0), & t_{\max} \geq t_0. \end{cases} \quad (14)$$

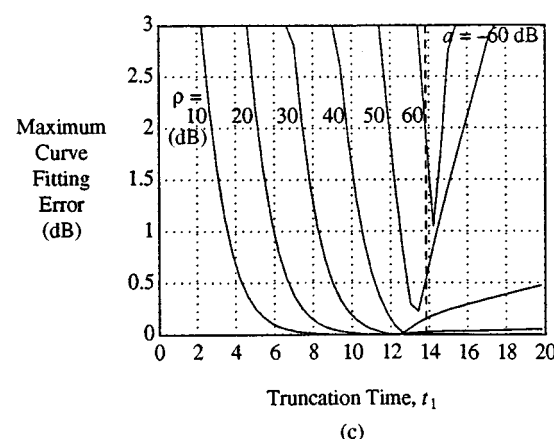
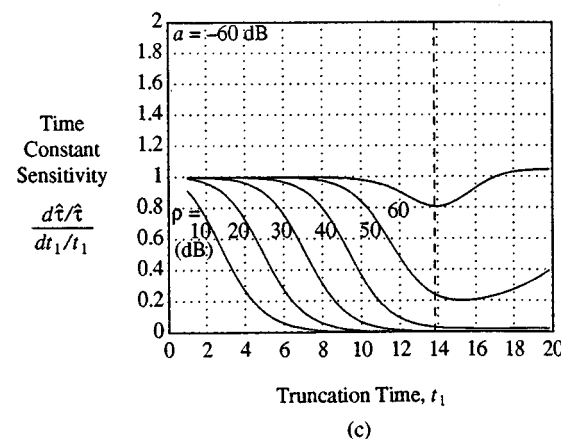
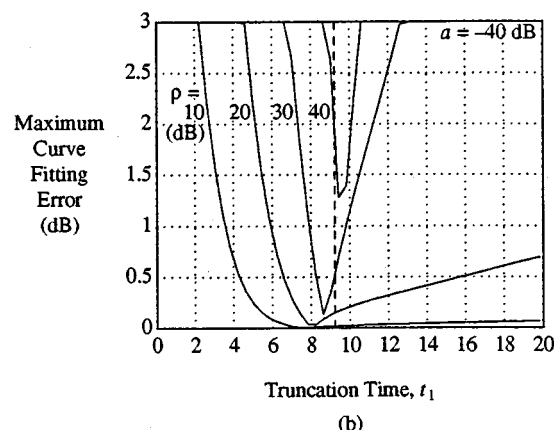
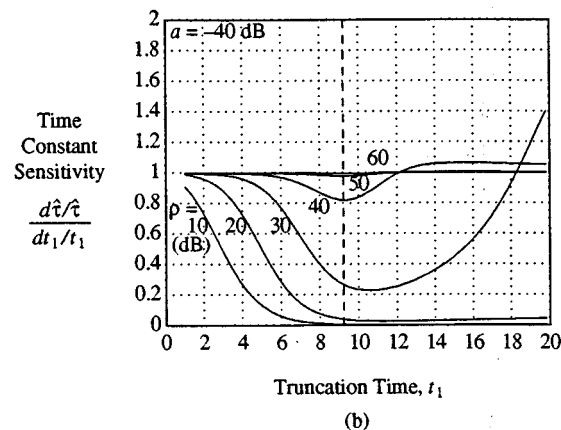
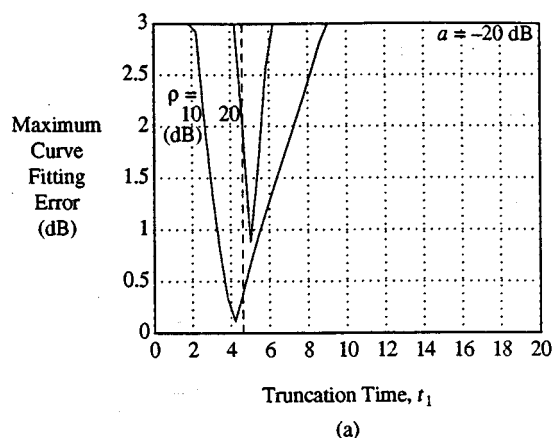
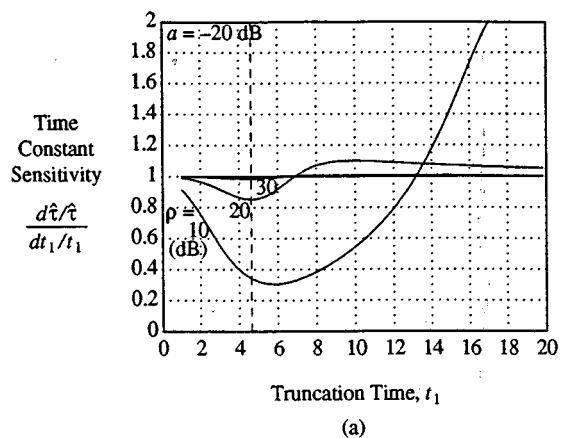


FIG. 7. Time constant sensitivity ($d\hat{\tau}/\hat{\tau})/(dt_1/t_1)$) as a function of truncation time t_1 for various evaluation ranges ρ and noise floor levels a . (a) $a = -20$ dB, (b) $a = -40$ dB, (c) $a = -60$ dB.

FIG. 8. Maximum curve fitting error as a function of truncation time t_1 for various evaluation ranges ρ and noise floor levels a . (a) $a = -20$ dB, (b) $a = -40$ dB, (c) $a = -60$ dB.

Also, the minimum error ratio is determined from the boundary values as

$$\epsilon_{\min} = \min[1, \epsilon(t_0)]. \quad (15)$$

The maximum absolute error in dB is therefore calculated as

$$10 \log_{10}[\max(\epsilon_{\max}, 1/\epsilon_{\min})], \quad (16)$$

and is plotted in Fig. 8 as a function of truncation time t_1 for several noise floor values a .

The significance of the maximum curve fitting error is judged in relationship to the randomlike fluctuations exhib-

ited by the backward integrated impulse response, as for example in Fig. 2. Thus a maximum curve fitting error of 1 or 2 dB will not severely degrade the visual quality of the decay curve. (For early decay time measurements, smaller curve fitting error may be desirable; however, in that case, the evaluation range is much smaller so the noise floor effects are negligible.)

IV. SELECTION GUIDELINES

In the foregoing, we have shown how the choice of truncation time t_1 and evaluation range ρ for a given noise floor

level a influences the bias and sensitivity of the time constant estimate. The choice of truncation interval also influences the maximum curve fitting error which is also of interest in visual interpretations. Recall, however, that t_1 is defined as the truncation time *relative* to the actual time constant (assumed equal to 1 in the analysis). Therefore, in an actual application, one would have to guess at the time constant, choose the truncation time based on that guess, obtain a refined estimate of the time constant, and continue in this fashion, iterating toward a final solution. Moreover, the choice of ρ depends on a and t_1 , making for a very complicated procedure. All of this is very clumsy and time consuming. Therefore, we seek general guidelines for choosing t_1 and ρ that do not require iteration and are simpler to apply.

As a first step, consider the simple rule of truncating at the knee of the curve, i.e., at the intersection of the main decay slope with the noise floor, which has been recommended by several studies, as previously mentioned. In terms of the model parameters, this procedure corresponds to taking $t_1 = \ln(1/a)$, which is indicated by a vertical dashed line in Figs. 6–8. Reexamination of these figures shows that this is a reasonable choice: For evaluation ranges that do not extend below the noise floor, the bias is small, the sensitivity is moderate, and the maximum curve fitting error is less than about 2 dB. Thus, we have objectively reaffirmed this heuristic guideline by analytical means.

With the above choice of truncating at the knee, there is now the question of how to choose the evaluation range. Again, re-examination of Figs. 6–8 at the truncation level indicated by the dashed vertical line shows that “optimal” ranges tend to track more or less at a fixed level above the noise floor. Here “optimal” means small time constant bias with moderate sensitivity and small maximum curve fitting error, all within the aforementioned practical constraint of not making the evaluation range “too small,” in consideration of the random perturbations that will affect the actual time constant estimate. Therefore, we are motivated to examine these performance measures over a range of noise floor levels when the evaluation range is set at a given margin above the noise floor. For example, when the noise floor is -40 dB, we see, following along the dashed line in Fig.

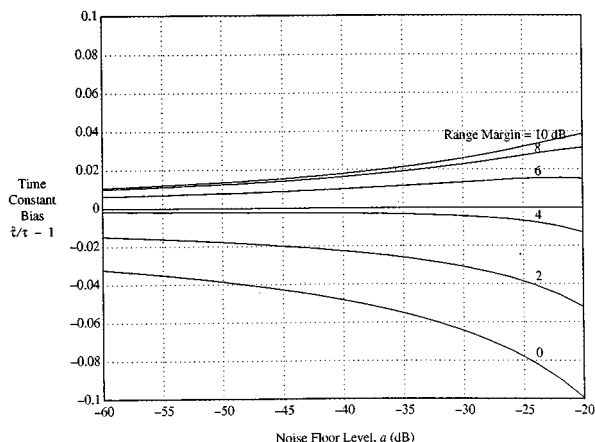


FIG. 9. Normalized time constant bias ($\hat{\tau}/\tau - 1$) as a function of noise floor level a for various evaluation range margins.

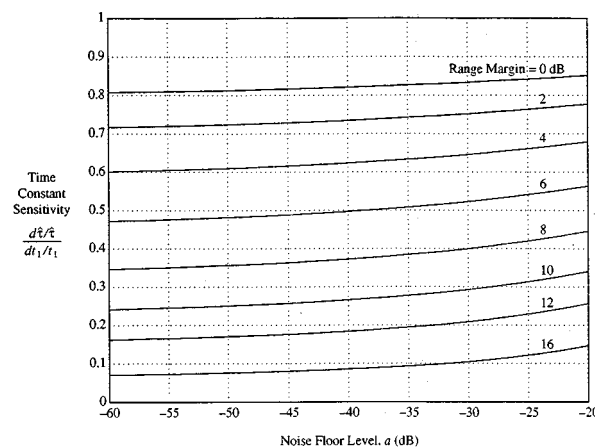


FIG. 10. Time constant sensitivity ($d\hat{\tau}/\hat{\tau}/(dt_1/t_1)$) as a function of noise floor level a for various evaluation range margins.

6(b), that the bias will be close to zero for the evaluation range between 30 and 40 dB, i.e., a margin of 0–10 dB. Figure 9 shows a plot of the bias versus noise floor level for several values of range margin. As can be seen, the bias is uniformly minimized over all noise floor levels for a margin between 4 and 6 dB, somewhat smaller than the 10 dB advocated by previous studies. Figure 10 shows a similar plot for the sensitivity, which is seen to be nearly uniformly dependent on the range margin. For margins greater than about 5 dB, the sensitivity is less than 0.5, which is small enough for many applications. Finally, the maximum curve fitting error appears in Fig. 11 and is seen to be uniformly equal to about 0.5 dB for range margins between about 4 dB and 12 dB.

V. APPLICATION

Figure 12 shows impulse responses simulated from (1) for the illustrative randomizations (2) and (3) with several representative noise floor levels, viz., -20 , -40 , and -60 dB. As a general rule of thumb, peak values of a Gaussian random variable over a moderate number of samples are about 2.5 times or 8 dB above the rms value. Therefore, the

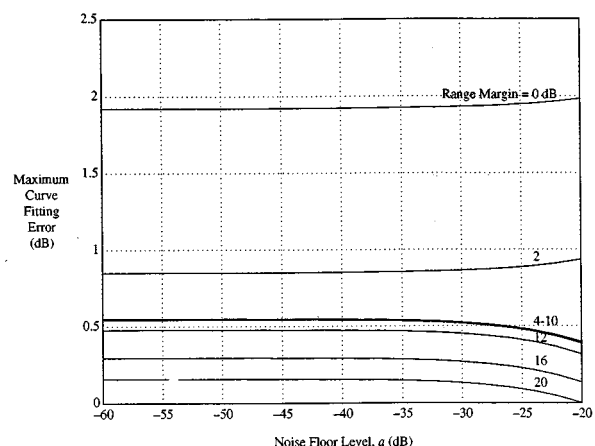


FIG. 11. Maximum curve fitting error as a function of noise floor level a for various evaluation range margins.

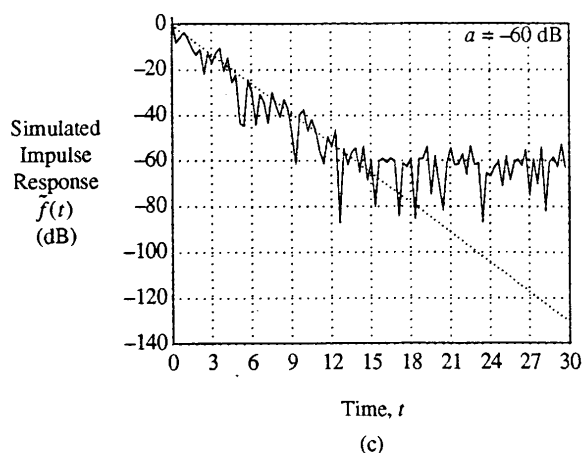
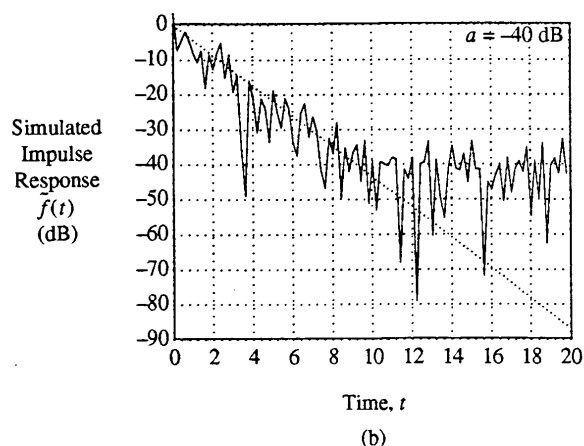
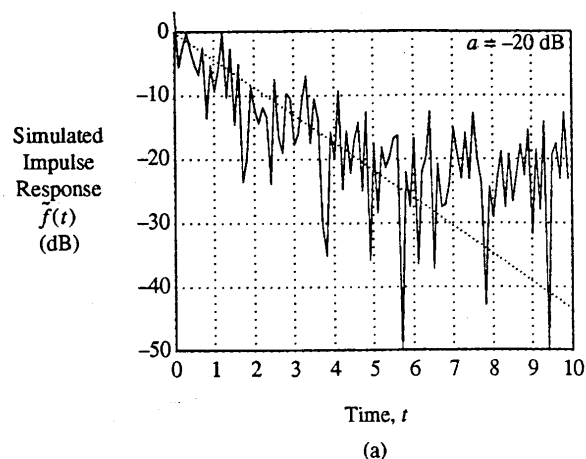


FIG. 12. Simulated impulse responses with noise floor (a) $a = -20$ dB, (b) $a = -40$ dB, (c) $a = -60$ dB.

noise floor level associated with the tail of experimental impulse responses limited by Gaussian noise should be estimated at about 8 dB below the average peak value in that region. Further, since the main decay of the measured impulse response is well characterized by Rayleigh fading, which is equivalent to the mean-squared sum of two independent Gaussian random variables, the mean slope in that region for such a model should be estimated at about $8 - 3 = 5$ dB below the trend of the peak values. With these guidelines, visual inspection of the plots in Fig. 12 determines an

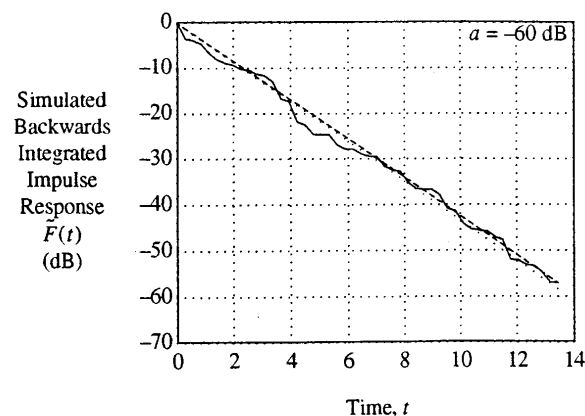
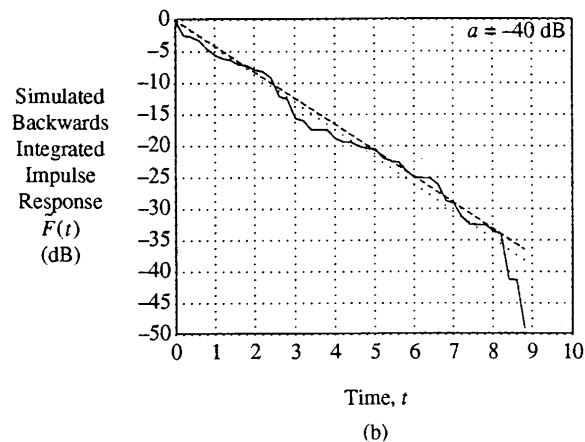
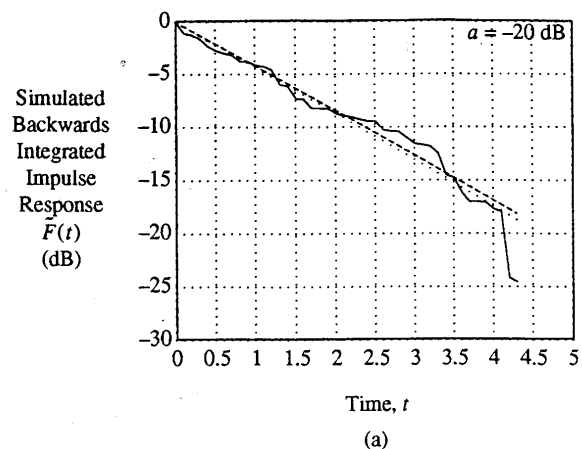


FIG. 13. Simulated backwards integrated impulse responses with various noise floor a and truncation time t_1 . (a) $a = -20$ dB, $t_1 = 4.5$. (b) $a = -40$ dB, $t_1 = 9$. (c) $a = -60$ dB, $t_1 = 14$.

estimate of the knee of the curve, which we guess here as occurring at $t_1 = 4.5, 9, 14$, respectively, at each corresponding noise floor level. After truncating the impulse responses at these endpoints, we calculate the backward integral (4), choose a range evaluation margin of 5 dB, and evaluate the slope over a range of 15, 35, and 55 dB, respectively, resulting in the plots of Fig. 13. As can be seen, for each case, we obtain a fairly good estimate of the slope with a maximum curve fitting error less than about 2 dB.

Further refinement is possible by short-term averaging the measured impulse response so as to presmooth the data.

For example, the procedure in Lundeby *et al.*,⁶ advocates an initial averaging over 10–50 ms intervals, followed by refined averaging over 3–10 intervals for each 10 dB of decay. This technique facilitates accurate determination of the knee of the curve and also improves the fitted slope estimate obtained from the backwards integrated curve.

VI. CONCLUSIONS

An analysis of the backward integration method for reverberation time estimation has quantified the effects of noise floor level, truncation time, and evaluation range on the accuracy. A general guideline that emerges from this study is to truncate the measured impulse response at the knee where the main decay slope intersects the noise floor, then measure the slope of the backward integrated truncated impulse response down to a level about 5 dB above the noise floor. Simulations using models with known parameter values have validated the efficacy of this procedure.

ACKNOWLEDGMENT

The author wishes to thank Gary Elko for many useful discussions and helpful guidance in the course of developing this work.

¹M. R. Schroeder, "New method of measuring reverberation time," *J. Acoust. Soc. Am.* **37**, 409–412 (1965).

²M. R. Schroeder, "Die statistischen parameter der frequenzkurven von grossen räumen," *Acustica* **4**, 594–600 (1954) [English transl.: "Statistical parameters of the frequency response curves of large rooms," *J. Audio Eng. Soc.* **35**, 299–306 (1987)].

³R. Kürer and U. Kurze, "Integrationsverfahren zur nachhallauswertung," *Acustica* **19**, 313–322 (1967/68).

⁴W. T. Chu, "Comparison of reverberation measurements using Schroeder's impulse method and decay-curve averaging method," *J. Acoust. Soc. Am.* **63**, 1444–1450 (1978).

⁵M. Vorländer and H. Bietz, "Comparison of methods for measuring reverberation time," *Acustica* **80**, 205–215 (1994).

⁶A. Lundeby, T. E. Vigran, M. H. Bietz, and M. Vorländer, "Uncertainties of measurements in room acoustics," *Acustica* **81**, 344–355 (1995).

⁷N. Xiang, "Evaluation of reverberation times using a nonlinear regression approach," *J. Acoust. Soc. Am.* **98**, 2112–2121 (1995).

⁸*User's Manual for the SYSid Audio-Band Measurement and Analysis System Version 4.0* (Ariel Corporation, Highland Park, NJ, 1992).

Comparison of methods for determining specific acoustic impedance

Michael G. Jones

Lockheed Martin Engineering and Sciences Company, NASA Langley Research Center, M.S. 463, Hampton, Virginia 23681-0001

Patricia E. Stiede

NASA Langley Research Center, M.S. 463, Hampton, Virginia 23681-0001

(Received 5 February 1996; accepted for publication 14 November 1996)

An investigation was conducted to explore potential improvements provided by alternate implementations of two impedance-measurement methods, known as the two-microphone and multipoint methods. This investigation is part of a continuing technology development to find more efficient and convenient methods for impedance measurements in harsh environments at high frequencies. The two alternate methods are compared with other methods to determine relative merits. As expected, the “best method” depends upon trade-offs between convenience and accuracy. The single-microphone method eliminates one of the two microphones typically used in the two-microphone method. The single-microphone method should be useful where mechanical constraints allow only one microphone to be placed into the test environment. It is found to be quite accurate with a single discrete frequency source, but the most inaccurate with a random noise source. An alternate implementation of the multipoint method, achieved by replacing the single-tone source with a pseudo-random noise source, requires significantly less time with minimal loss of accuracy. Typical agreement between this implementation and the accepted standard is within 0.1 ρc units. However, slightly larger deviations at frequencies above 2.5 kHz indicate challenges in the proposed extension to high-frequency (up to 25-kHz) applications. © 1997 Acoustical Society of America. [S0001-4966(97)06504-1]

PACS numbers: 43.58.Bh, 43.50.Gf [SLE]

LIST OF SYMBOLS

a_0	ratio of duct area to perimeter
c	sound speed in tube, m/s
c_p	specific heat at a constant pressure, J/(kg K)
j	counter of complex pressures
k	wave number, m^{-1}
L	range of distance in axial direction, m
M	mach number
m	counter of weighting functions
m_{\max}	maximum number of weighting functions
N	number of measured complex acoustic pressures
P_a	sound-pressure amplitude, Pa
P_i	complex incident acoustic pressure
P_j	complex acoustic pressure (amplitude and phase) at x_j
P_r	complex reflected acoustic pressure
R	complex reflection factor
W_m	weighting function
x	distance in the axial direction, m
x_j	axial location of acoustic pressure measurement, m
Greek	
β_v	viscothermal dissipation of acoustic energy at the tube wall, m^{-1}
Γ	plane-wave propagation constant, m^{-1}
γ	specific heat ratio
ζ	test specimen surface impedance, Pa s/m

θ	specific acoustic resistance at test specimen surface, Pa s/m
κ	heat conduction coefficient, kg s K/m ⁷
ν	kinematic viscosity of the fluid, m ² /s
ρ	density of the fluid, kg/m ³
ϕ	sound-pressure phase, deg
χ	acoustic reactance at test specimen surface, Pa s/m
ω	angular frequency, s ⁻¹

Abbreviations

AWG	arbitrary waveform generator
FFT	fast Fourier transform
NIT	normal incidence impedance tube
OASPL	overall sound-pressure level, dB (<i>re</i> : 20 μ Pa)
RNG	random noise generator
SPL	sound-pressure level, dB (<i>re</i> : 20 μ Pa)
TMM	two-microphone method
MPM	multipoint method
SMM	single-microphone method
SWM	standing wave method

Suffixes

DF	discrete frequency source
PR	pseudo-random noise source
R	random noise source

Note 1: All impedances assume $e^{-i\omega t}$ time convention unless otherwise noted.

Note 2: All parameters are dimensionless unless otherwise noted.

INTRODUCTION

A number of measurement techniques have been developed over the past 80 years to determine the normal incidence acoustic impedance of passive linear materials.¹ Table I lists strengths and weaknesses of NASA Langley implementations of some of the more commonly accepted methods. The standing wave method (SWM) has been a commonly used procedure for impedance measurements for at least 80 years.² This method is based on the unique relationship between the standing wave parameters and the test specimen impedance. It is implemented by means of a tube that guides plane waves such that normal incidence is achieved on the test specimen surface. In its original usage, a continuous trace of sound-pressure level (SPL) versus distance was measured, and the standing wave ratio and null location were determined from this analog trace. This procedure was inherently time consuming. Also, it was limited to frequencies below the cuton frequency of higher-order propagating modes in the tube.

About 20 years ago, the development of fast Fourier transform algorithms which could be implemented on laboratory computers provided the impetus for the development of impedance measurement methods that made use of two or more point-type measurements in a standing wave field. The two-microphone and multipoint methods (TMM and MPM, sometimes called the transfer function and least-squares methods) grew out of this development³⁻⁸ with a consequent tremendous productivity increase in impedance measurement capability. Since that time, further modifications have been

incorporated to accommodate increasingly specialized impedance measurement requirements.

The TMM was initially implemented with two fixed microphones flush-mounted on the side wall of a normal incidence acoustic impedance tube;⁶ thus, the same high-frequency limit holds as for the SWM. However, since no movement of a microphone probe is involved, the data can be acquired in a greatly reduced amount of time. In addition to discrete tones, this method can be implemented with random noise, which reduces data acquisition time by at least an order of magnitude. However, the TMM requires very precise transfer function measurements, which implies accurate amplitude and phase calibrations for each of the two microphones. To circumvent the potentially tedious and time consuming calibration process, the standard TMM implementation was later modified to incorporate a microphone switching technique.⁵

For the last eight years, the multi-point method (MPM) for determining normal incidence acoustic impedance of passive liner materials has been used in the Flow Impedance Test Laboratory at NASA Langley Research Center.⁹ This method differs from the TMM by including at least three measurement points in the standing wave field, thus requiring a least-squares fit to the data. The MPM requires acoustic pressure measurements at a minimum of three selected distances from the surface of the test specimen, whether by multiple fixed-location microphones or a movable traversing probe microphone. The latter version (traversing probe) is currently in use in the Flow Impedance Test Laboratory,

TABLE I. Strengths and weaknesses of commonly used methods.

Method	Strengths	Weaknesses
Standing wave method (SWM)	(a) good accuracy over 0.1 to 10 ρc impedance range (b) can automate via probe and SWR and null-searching code (c) standing wave parameters directly measured	(a) accuracy limited by SWR and null location resolution (b) not suited to grazing incidence/mean flow environments (c) tedious—time consuming
Two-microphone method without microphone switching (TMM) ^a	(a) significant time saving (b) no hardware movement (c) ideal for quick screening tests with random noise source	(a) assumes 1-D wave propagation model (b) accurate microphone amplitude and phase calibrations required (c) accuracy degrades at microphone separations near one half-wavelength (d) accuracy degrades for large wavelengths
Two-microphone method with microphone switching (TMM) ^b	(a) same as above (b) precise microphone calibrations are not needed	same as above except (b)
Multipoint method (MPM)	(a) accuracy improved via least-squares fitting (b) no restriction on microphone separation relative to wavelength (c) enhanced frequency range	(a) hardware movement required (b) more time consuming

^aOriginal implementation.

^bModified implementation, currently in use at NASA Langley Research Center.

which limits the usage to single or multiple discrete frequency noise sources.

For certain test requirements, it is inconvenient to implement either of these methods (TMM or MPM). This may be due to geometric constraints (e.g., high-frequency impedance tubes) or time and accuracy requirements (e.g., usage of TMM with discrete frequencies is slow, usage of TMM with random noise excitation limits source spectrum control and may also compromise accuracy).

One purpose of the present investigation was to explore the above methods to examine their usefulness at high frequencies (e.g., in small apparatus). The single-microphone method (SMM),¹⁰⁻¹² an alternative implementation of the TMM, was investigated to determine measurement capabilities in situations where it would be highly convenient to place only one microphone into a test environment. Essentially, the SMM references the measured signal at two different locations back to the signal source, instead of to a reference microphone placed at another location in the acoustic wave field. This implies mechanical movement of the microphone between the two chosen locations. A potential advantage for the SMM is high-frequency impedance measurement, where small diameter impedance tubes are required. In this case, space limitations may not allow judicious spacing of two microphones.

The multiple frequency version of the multipoint method¹ (MPM-PR) was included as an extension of the MPM. This extension incorporates pseudo-random noise,¹¹ generated by the combination of discrete frequencies (each at a selected amplitude and phase), to reduce the time necessary to complete an acoustic impedance measurement by a factor equal to the number of frequencies simultaneously measured. For harsh environment applications, the increased efficiency may be critical to the survival of the measurement probe.

Another purpose of this study was to compare each of the normal incidence acoustic impedance measurement methods in use at NASA Langley Research Center for accuracy, consistency, and efficiency. This information should be useful in determining which method should be used for high-frequency applications. It was expected that the "best method" would depend upon the amount of time available and the accuracy necessary for the final results.

I. THEORY

The TMM can be viewed as a deterministic means of modeling a 1-D acoustic standing wave field using the ratio of the complex acoustic pressures at two locations (i.e., the transfer function), caused by the interference between incident and reflected waves at the face of the test specimen. Because this wave field is uniquely related to the test specimen impedance, this information allows the acoustic impedance to be computed directly. However, as noted in Table I, the method is subject to large error when the separation between the two microphone locations is near one half-wavelength of the test frequency. The MPM reduces this inaccuracy by employing a weighted residual technique to fit a wave propagation model to three or more point measurements. As used in the Flow Impedance Test Laboratory, the MPM uses complex acoustic pressure data acquired at six

equally spaced measurement locations, with the spacing based on the wavelength of the selected test frequency. In essence, the TMM uses the same equations as the MPM, with the number of points reduced to two. The SMM, which is simply a variation of the TMM, uses the equations in the same form as the TMM. Thus the significant analytical difference between these methods is that the MPM is an over-determined method, while the TMM and SMM are deterministic methods.

The acoustic pressure distribution (assume 1-D propagation) in a normal incidence impedance tube is given by

$$P_j = [P_i e^{i\Gamma x_j} + P_r e^{-i\Gamma x_j}] e^{-i\omega t}. \quad (1)$$

The plane-wave propagation constant, Γ , is defined as

$$\Gamma = k + i\beta_v, \quad (2)$$

where the wave number, k , and the viscothermal dissipation of acoustic energy at the tube wall, β_v , are given by Ingard and Singhal¹³ as

$$k = \frac{\omega}{c} \quad (3)$$

and

$$\beta_v = \frac{\sqrt{2}k}{4a_0} \left[\sqrt{\frac{\nu}{\omega}} + (\gamma - 1) \sqrt{\frac{\kappa/c_p}{\rho\omega}} \right]. \quad (4)$$

If N measured complex acoustic pressures, P_j , are known over the range of $0 \leq x \leq L$, the set of complex residuals formed by $(P_j - P_i e^{i\Gamma x_j} - P_r e^{-i\Gamma x_j})$ can be minimized by applying the method of weighted residuals.¹⁴ This method requires that the residuals be orthogonal to an appropriate set of weighting functions, W_m . Thus

$$\sum_{j=1}^N (P_j - P_i e^{i\Gamma x_j} - P_r e^{-i\Gamma x_j}) W_m = 0, \quad (5)$$

where W_m ($m = 1, 2, \dots, m_{\max}$) must be a complete set and may also be complex. For the specific application here, a pair of weighting functions are chosen using Galerkin's approach as

$$W_1 = e^{i\Gamma x_j} \quad (6)$$

and

$$W_2 = e^{-i\Gamma x_j}. \quad (7)$$

The resulting two equations can then be written as

$$\sum_{j=1}^N (P_j - P_i e^{i\Gamma x_j} - P_r e^{-i\Gamma x_j}) e^{i\Gamma x_j} = 0 \quad (8)$$

and

$$\sum_{j=1}^N (P_j - P_i e^{i\Gamma x_j} - P_r e^{-i\Gamma x_j}) e^{-i\Gamma x_j} = 0. \quad (9)$$

Solving for the incident and reflected acoustic pressures yields

$$P_i = \frac{CE - AB}{DE - A^2} \quad (10)$$

and

$$P_r = \frac{C - P_i D}{A}, \quad (11)$$

where

$$A = \sum_{j=1}^N (1), \quad B = \sum_{j=1}^N P_j e^{-i\Gamma x_j}, \quad C = \sum_{j=1}^N P_j e^{i\Gamma x_j}, \quad (12)$$

$$D = \sum_{j=1}^N e^{2i\Gamma x_j}, \quad E = \sum_{j=1}^N e^{-2i\Gamma x_j}.$$

In actual usage, the amplitude, P_a , and phase, ϕ , are determined by Fourier transform calculations based on the measured voltage time history. The measured acoustic pressures, P_j , are then computed by using

$$P_j = P_a e^{i\phi}. \quad (13)$$

In addition, since the functional behavior of the standing wave is assumed to be characterized by the 1-D propagation model, only a minimal number of points are needed for the determination of P_i and P_r . The MPM, as currently used at NASA Langley Research Center, uses six measurement points, whereas the TMM uses only two.

The standing wave generated by the reflection process at the test specimen surface is conveniently described by the complex reflection factor, R , given by

$$R = \frac{P_r}{P_i}, \quad (14)$$

which is related to the test specimen impedance, ζ , by

$$\zeta = \frac{1+R}{1-R} = \theta + i\chi, \quad (15)$$

where θ and χ represent the acoustic resistance and reactance at the test specimen surface, respectively.

One procedure to attain phase information, ϕ , as input data to the MPM is to position a microphone very near the surface of the test specimen. As the probe is located at successive positions, x_j , in the normal incidence tube (NIT), transfer function measurements between the probe microphone and the reference microphone provide relative amplitudes and phases versus axial distance from the test specimen surface. The TMM requires only the relative amplitudes and phases between two fixed locations, therefore, a rotating plug (described in a later section) is used to precisely position the microphones and transfer function measurements are used to determine the relative amplitudes and phases.

The SMM uses only one microphone in the rotating plug. This method was included to determine if it would be possible to acquire meaningful data in cases where the standard TMM is inconvenient. Since this method was developed for possible usage in test environments, where only one microphone can be placed into the wave field, the signal at this microphone is referenced to the electronic source, either a function generator or a random noise generator. A transfer function between the measurement microphone and the source is measured, then the plug is rotated to place the microphone in another known location and another transfer function is measured. If TF_{1R} represents the transfer function between the measurement microphone at location one and

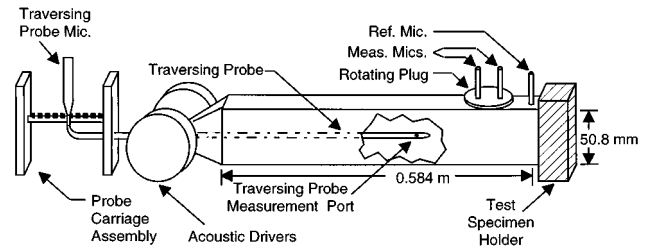


FIG. 1. Schematic of normal incidence impedance tube (NIT).

the electronic noise source, and TF_{R2} represents the transfer function between the electronic source and the measurement microphone at location two, the transfer function between the two measurement locations, TF_{12} , is then given by

$$TF_{12} = \frac{TF_{1R}}{TF_{R2}}. \quad (16)$$

Obviously, each of these methods requires the sound field to be statistically stationary throughout the test sequence.

II. EXPERIMENTAL INSTRUMENTATION AND PROCEDURE

A. Impedance tube description

The data used in the current study were acquired in a normal incidence impedance tube (NIT) located in the Flow Impedance Test Laboratory of NASA Langley Research Center. A schematic illustrating this apparatus is shown in Fig. 1. The cross section of the tube measures 50.8 by 50.8 mm, and the length is 0.584 m. The test specimen is carefully clamped to the exit plane of the NIT such that there is no leakage of sound around the edges of the test specimen.

Two 120-W acoustic drivers generate acoustic plane waves over a frequency range of 0.3 to 3.0 kHz, with sound-pressure levels (SPLs) up to 140 dB at the test specimen surface. Acoustic waves which impinge upon the surface of the test specimen are reflected to create a standing wave pattern that uniquely characterizes the normal incidence acoustic impedance of the test specimen. The various impedance measurement methods used in the current study involve the use of three types of acoustic sources (single discrete frequency, random noise, pseudo-random noise) and two different types of measurement apparatus (traversing probe and rotation plug) for obtaining acoustic pressures.

B. Acoustic source description

1. Single discrete frequency

The first type of acoustic source used in this investigation was a single discrete frequency (tone dominated) source. Because it is a very repeatable source, the discrete frequency source is often the source of choice. In the current investigation, a function generator was used to generate the discrete frequency. The source frequency was sequentially stepped through the frequency range of 0.5 to 3.0 kHz, in steps of 0.1 kHz. At each frequency, the amplitude was then adjusted until a level of 120 ± 0.5 dB was achieved at the reference microphone location.

2. Random noise

The second type of acoustic source was the random noise source. A white-noise random noise generator was used to generate this sound field. Obviously, this source allows all frequencies in the range of interest to be studied simultaneously. However, the disadvantage is that the pressure spectrum over the desired frequency range at the face of the test specimen (the reference microphone location) cannot be held to a specified target amplitude. Thus precise control of the spectrum shape is replaced by a requirement on the overall sound-pressure level (OASPL). Restated, the total energy in the frequency range of 0.5 to 3.0 kHz was set to 120 ± 1 dB. Because of test specimen impedance variability and source-tube coupling efficiency problems, there is always inherent potential for signal-to-noise ratio deficiency over unpredictable portions of the spectrum. Regardless, this source is often used for screening purposes since it provides for quick measurements.

3. Pseudo-random noise

The last acoustic source type was the pseudo-random noise source. This noise source was generated by combining a prescribed number of discrete frequencies. To accomplish this, a computer code generated a digitized pattern as input to an arbitrary waveform generator. The signal created from this digitized pattern was repetitively output from the waveform generator. By adjusting amplitude and phase parameters in the computer code (performed automatically by the code), the combination of discrete frequencies could be adjusted until the SPL of each of the input frequencies was within ± 1 dB of the target of 120 dB at the reference microphone. The advantage of this procedure is the precise control of a relatively broadband sound spectrum incident upon the surface of the test specimen.

In the current study, a combination of 13 frequencies (0.5 to 2.9 kHz or 0.6 to 3.0 kHz, in steps of 0.2 kHz) was used to generate a pseudo-random noise spectrum. The time taken to achieve stability (0.5 to 2.0 min for these test specimens) was dependent upon test specimen absorption characteristics. This somewhat lengthy setup time was offset by the fact that only two test runs (one for odd frequencies, second for even frequencies) were necessary. It should also be noted that the choice of frequency combinations was done in this manner only for ease of usage. Other combinations were also tested briefly, with essentially identical results.

C. Measurement point locations

1. Reference microphone

The reference microphone location is shown in the schematic of Fig. 1. As indicated in the sketch, this 6.35-mm-diam condenser-type microphone is flush mounted in the top wall of the NIT, 6.35 mm from the face of the test specimen. This microphone was used for two purposes in the current investigation; (1) as a set-point for absolute SPL near the specimen surface, and (2) as a reference for the transfer function measurements used to determine the acoustic pressures at selected locations.

In an attempt to ensure that each of the impedance measurement methods was tested under equivalent conditions, the sound-pressure level at the reference microphone location was set to 120 dB for each of the methods implemented with discrete frequency (single or multiple) sources. For those methods which incorporated a random noise source, the overall sound-pressure level over the frequency range of interest was set to 120 dB at this location. The SPL amplitude of 120 dB was chosen in an attempt to avoid nonlinear effects which occur at higher amplitudes, while also avoiding signal-to-noise limitations associated with lower amplitudes.

For the MPM implementation, the reference microphone was used as prescribed in Eq. (16), as a reference for each of the measurements with the traversing probe microphone. Therefore, the relative amplitude and phase between the fixed reference microphone location and the traversing probe microphone location could be used to determine the amplitude and phase at each measurement location.

2. Traversing probe description

The traversing probe consisted of an 0.84-m-long stainless steel tube with outer and inner diameters of 1.6 and 0.8 mm, respectively. This probe was coupled to a 12.7-mm-diam condenser-type microphone via a flexible tube, and could be positioned within 0.025 mm of a specified location. Four 0.71-mm-diam measurement ports, equally spaced around the probe in a single measurement plane located 14.2 mm from the tip of the probe, could be positioned over a range of 17.8 to 305.8 mm from the test specimen face. Previous tests⁹ showed that the flanking sound transmitted through the solid probe walls is at least 32 dB below that which enters through the measurement ports at the measurement plane. Thus the effective dynamic range of the probe is at least 22 dB.

3. Rotating plug description

An innovative feature of this apparatus is a rotating plug implementation of the microphone switching for the TMM. As indicated in Fig. 1, a rotating plug is installed in the top wall of the NIT. This plug, made of stainless steel, is designed to allow the positions of two 6.35-mm-diam condenser-type microphones to be switched in a very convenient and precise manner. The microphones are flush mounted in the rotating plug 31.7 mm apart. Thus, the difficulties associated with mounting conditions for typical TMM measurements (e.g., differences in channel calibrations, microphone sensitivities, and sensitivity changes due to varied mounting conditions), in which the microphones are physically removed from their original ports and installed in the alternate ports, are completely eliminated.

D. Electronic instrumentation

A block diagram of the signal conditioning instrumentation and data processing system for the acoustic signals is shown in Fig. 2. Three sources are available by means of two instruments. The first source instrument is the arbitrary waveform generator (AWG). This instrument is capable of an output of standard functions (sine wave, square wave,

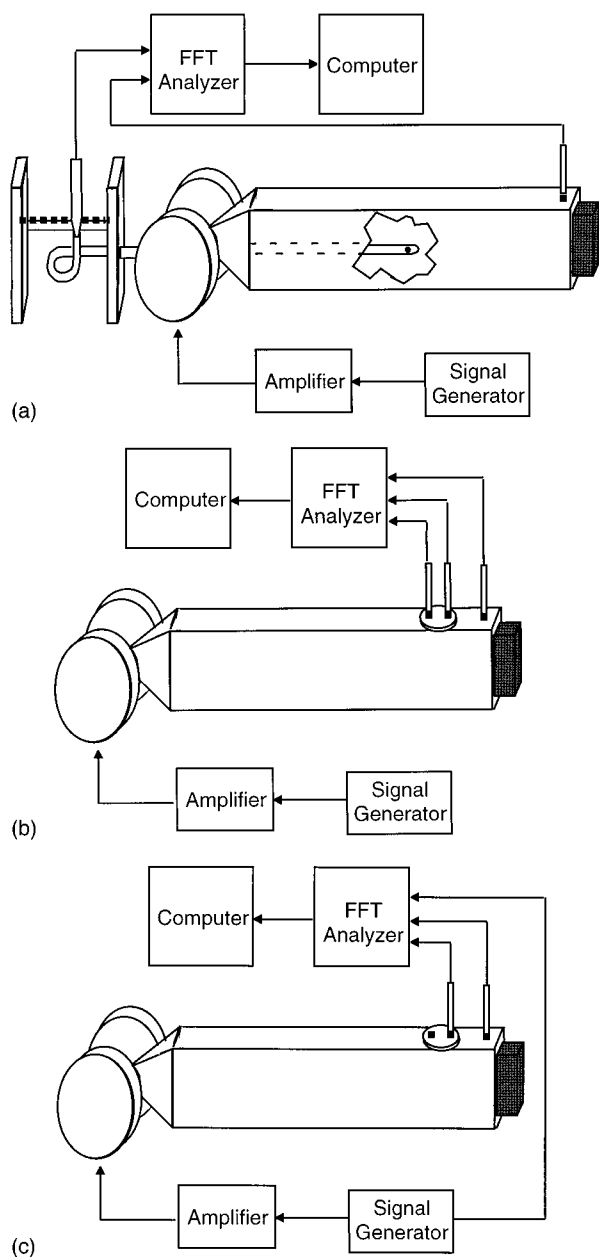


FIG. 2. Instrumentation diagram for impedance measurement methods. (a) Multipoint method (MPM). (b) Two-microphone method (TMM). (c) Single-microphone method (SMM).

etc.) or arbitrary waveforms (created using a computer generated sequence). Thus the AWG was used for the single discrete frequency source and the pseudo-random noise source (combination of multiple discrete frequencies). The other source instrument was the random noise generator (RNG), which was used to output broadband noise. For the current study, the selected source signal was amplified and input to the acoustic drivers connected to the normal incidence impedance tube (NIT). Figure 2 depicts the instrumentation connection sequences for each of the three impedance measurement methods. In each implementation, the reference microphone was used to set the amplitude of the acoustic signal near the face of the test specimen.

Figure 2(a) shows the instrumentation for the MPM. The reference microphone and the traversing probe microphone

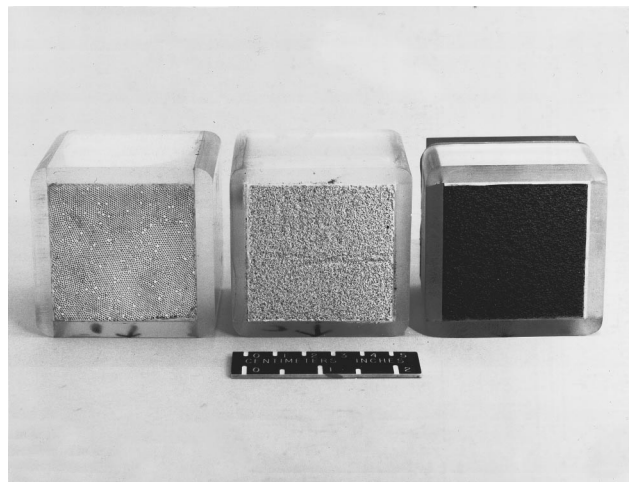


FIG. 3. Photograph of three test specimens: left to right; ceramic tubular, ceramic foam, silicon carbide.

are connected to the FFT analyzer, which is used to record the transfer function between the two inputs as the traversing probe is positioned to prescribed locations. The results are stored onto the computer. The sample rate of the FFT analyzer was set to 25.6 kHz for all of the data reported in this study. In addition, anti-aliasing filters were incorporated to prevent aliasing of the measured data.

Figure 2(b) shows the instrumentation for the TMM. Each of the two measurement microphones (A and B) is mounted in the rotating plug. The FFT analyzer is used to record the transfer function between the two inputs, and the results are stored by the computer. The plug is then rotated to switch the microphone locations and the data acquisition process is repeated.

Figure 2(c) shows the instrumentation for the SMM. For this implementation, the second microphone in the rotating plug is replaced by an AWG reference signal. The FFT analyzer is used to record the transfer function between the two inputs (measurement microphone and AWG source), and the results are stored by the computer. The measurement microphone is then moved to a second location (by rotating the plug) and the data acquisition process is repeated.

E. Description of test specimens

The current study was conducted with three test specimens, as shown in Fig. 3. These test specimens were chosen to cover a range of typical acoustic impedance values. Each of the test specimens was cut into 50.8 by 50.8-mm square samples and placed into Plexiglass holders such that they could be mounted to terminate the NIT.

The first specimen tested was constructed with a ceramic tubular material, which consisted of a distribution of equal-length cylindrical channels embedded in a solid ceramic matrix. The length and porosity of the ceramic tubular specimen were 49.8 mm and 57%, respectively. This sample has been used as a standard test specimen for a number of years in the Flow Impedance Test Laboratory. It has been shown to be relatively independent of mean flow velocity¹⁵ (up to approximately $M=0.5$), and is capable of withstanding harsh thermal (elevated temperature) environments. Experience

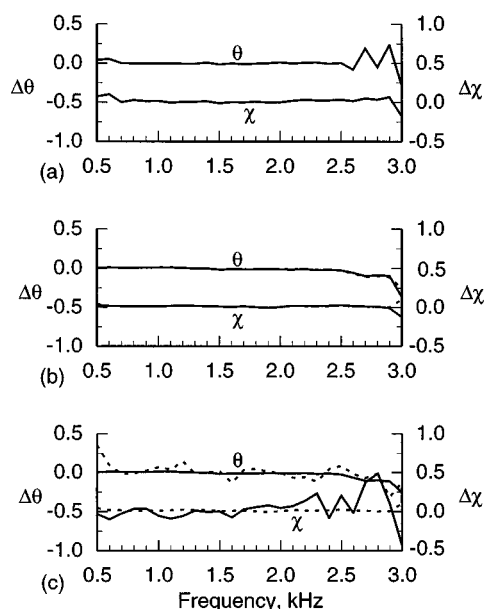


FIG. 4. Measured normalized impedance comparisons against standard method (MPM-DF) for ceramic tubular test specimen. (a) MPM-PR; (b) TMM-DF (solid) and TMM-R (dashed); and (c) SMM-DF (solid) and SMM-R (dashed).

with this material has shown it to provide highly repeatable results. Further, the inherent linearity of this material makes it suitable as a 'calibration' material.

The other two test specimens consisted of ceramic foam and silicon carbide materials. Results of impedance measurements of these two materials in the Flow Impedance Test Laboratory have not previously been reported. The ceramic foam sample had a length and porosity of 51.2 mm and 85%, respectively. The silicon carbide sample had a length and porosity of 53.7 mm and 45%, respectively. Therefore, the three test specimens were essentially the same length, and covered a porosity range of 45%–85%.

III. EXPERIMENTAL RESULTS

In this section, results from normal incidence impedance measurements for each of the three test specimens, using each of the three methods (MPM, TMM, and SMM) and (where applicable) each of the three source types (single discrete frequency, pseudo-random, random), are discussed. Figures 4, 5, and 6 contain differences between results obtained with each of the individual methods and those obtained with the MPM-DF (selected as the baseline) for each of the test specimens investigated. Figures 7, 8, and 9 combine results of all of the available methods.

The data for this study were acquired over a period of about one month. During this time, the temperature was observed to vary from 70 to 76 °F. The atmospheric pressure ranged from 753 to 771 mmHg (Torr), and the relative humidity remained below 50%. While the effects of temperature variability were included in the calculations, the effects of variability in the atmospheric pressure and the relative humidity were not. However, it should be noted (although not shown in this report) that multiple tests were conducted, which showed the results to be very precise (or repeatable).

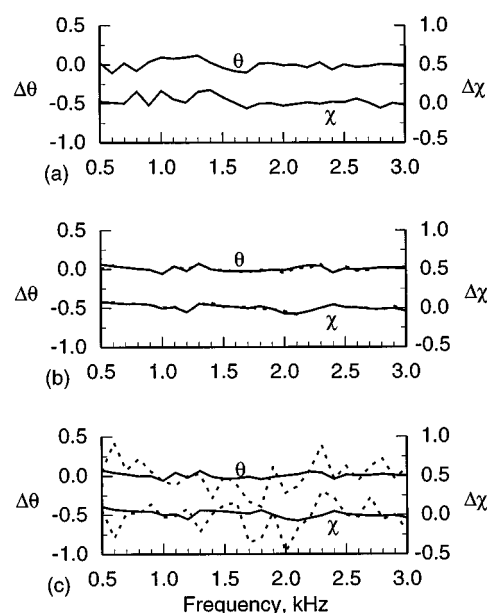


FIG. 5. Measured normalized impedance comparisons against standard method (MPM-DF) for ceramic foam test specimen. (a) MPM-PR; (b) TMM-DF (solid) and TMM-R (dashed); and (c) SMM-DF (solid) and SMM-R (dashed).

It is therefore believed that the variability of the atmospheric pressure and the relative humidity are not the cause of deviations between the results of the different methods.

As mentioned above, multiple repeat tests showed the results to be very precise. The following discussion will therefore address only the accuracy of the measurements. In Figs. 4–9, the result of applying the MPM-DF (one test frequency at a time) is taken to be the baseline, or standard, against which other results are compared. This choice is based upon documented comparisons with other methods.⁹

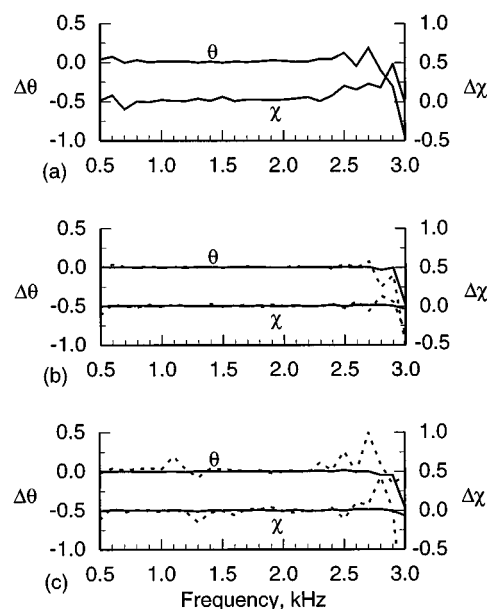


FIG. 6. Measured normalized impedance comparisons against standard method (MPM-DF) for silicon carbide test specimen. (a) MPM-PR; (b) TMM-DF (solid) and TMM-R (dashed); and (c) SMM-DF (solid) and SMM-R (dashed).

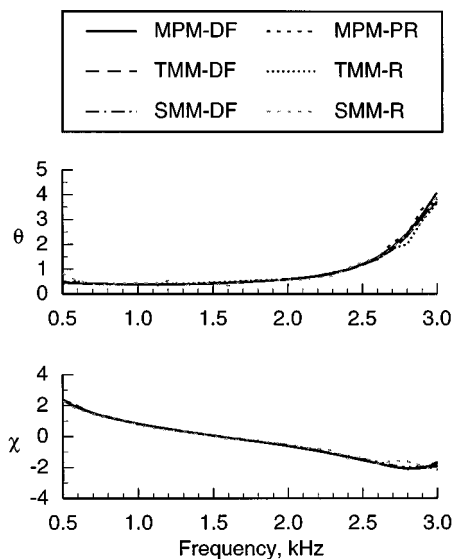


FIG. 7. Measured normalized impedance for ceramic tubular test specimen.

Thus results acquired with each of the other methods discussed below will be compared against the MPM-DF results.

A. Description of MPM results

The multipoint method (MPM) can be implemented with two of the source types, single discrete frequency (MPM-DF) or pseudo-random (MPM-PR, created by combination of multiple discrete frequencies).

1. Single discrete frequency source (MPM-DF)

The MPM-DF is generally considered to be the most accurate of all of the methods discussed here because it eliminates the half-wavelength microphone separation problem and provides a mechanism (weighted residual technique) for minimizing the effects of minor errors in the calibrations of the measurement microphones. Thus the MPM-DF will be used as the standard of comparison.

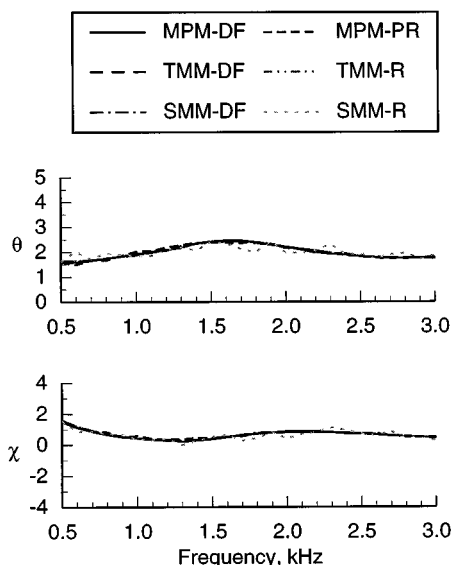


FIG. 8. Measured normalized impedance for ceramic foam test specimen.

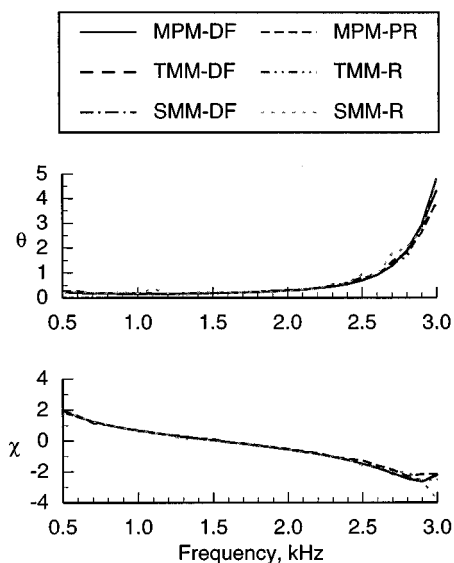


FIG. 9. Measured normalized impedance for silicon carbide test specimen.

2. Pseudo-random noise source (MPM-PR)

As stated earlier, the MPM-PR (using combination of multiple discrete frequencies) is more time-efficient than the MPM-DF. The MPM-PR results coincide almost exactly with the MPM-DF results. With the exception of a few frequencies (different frequencies for each test specimen), the normalized resistance results measured with the MPM-PR and MPM-DF agree closely. Typically, the agreement is within $0.1 \rho c$ units. The few outlying results, those where the MPM-PR differed from the MPM-DF results by more than $0.1 \rho c$ units, typically occurred at or above test frequencies of 2.5 kHz. However, it should be noted that these deviations from the standard results were never larger than about $0.2 \rho c$ units. Regardless, the difficulty at test frequencies above 2.5 kHz indicates significant challenges in the proposed extension to high-frequency (up to 25 kHz) applications.

There are a number of possibilities for these discrepancies. Potential candidates include: (1) nonlinearity of the source excitation; (2) changes in SPL (versus frequency) at test specimen surface introduce nonlinearities; and (3) changing mounting conditions between tests. Care was taken to conduct these tests at source amplitudes such that the nonlinearities indicated in the first two of these candidates were minimized. Also, each test specimen was left in place until each of the measurement methods were completed. It should be noted here that previous unpublished measurements of the ceramic foam sample also indicated inconsistent behavior. Regardless, the results shown in Figs. 4–6 indicate that additional studies of the causes for the differences between the different methods at higher frequencies should be conducted. Clearly, the consistency of the results are specimen and/or method dependent.

As can be seen in Figs. 4 and 5, the MPM-PR normalized reactance results are almost identical to the MPM-DF results for the ceramic tubular and ceramic foam test specimens. For the silicon carbide test specimen (Fig. 6), the comparison continues to be excellent from 0.5 kHz to approxi-

mately 2.5 kHz. At frequencies above 2.5 kHz, the comparison deteriorates somewhat. The MPM-PR results range as much as 0.4 ρc units above the MPM-DF results. It is not immediately clear as to why this test specimen would have noticeably different comparison results from those of the ceramic tubular and ceramic foam test specimens. However, it is important to note that the MPM-PR does provide a very good description of the overall trends of the normalized reactance. Also, it should be noted that nonlinearities of the source excitation are most prevalent at the higher frequencies.

B. Description of TMM results

The two-microphone method (TMM) can be implemented with two of the source types, single discrete frequency (TMM-DF) or random noise (TMM-R). As discussed earlier (Table I), this method is attractive because of its ease of usage, with only a short analysis time necessary.

1. Single discrete frequency source (TMM-DF)

The two-microphone method with a single discrete frequency source is conducted using the rotating plug discussed earlier. As the test frequency is sequentially cycled through the range of interest, transfer function measurements between the two microphones are acquired first with the microphone in one orientation, then with the positions interchanged. Since this method requires only two measurements, it is much faster than the MPM-DF.

As seen in Figs. 4–6, the TMM-DF results (both normalized resistance and normalized reactance) are almost identical to the standard (MPM-DF) results. The only frequency where a difference can be discerned is at 3.0 kHz for the silicon carbide test specimen. Thus, it appears that this method is quite capable of supplying quality results. This should come as no surprise, since this method has been used in the aeroacoustics community for well over 15 years. It should be remembered that this method has a difficulty (model equations are ill-conditioned) when the microphone spacing is nearly one half-wavelength of the test frequency. However, if the test frequency range of interest does not include frequencies near this value, this method continues to be an excellent choice.

2. Random noise source (TMM-R)

Of all the methods included in the current study, the two-microphone method with a random noise source (TMM-R) is the most time efficient. The ill-conditioning associated with the microphone separation of nearly one half-wavelength for any of the frequencies within the range of interest must always be considered. In addition, random noise type source spectra are generally nonuniform at the test specimen surface. This can cause dynamic range difficulties if the amplitude levels of the spectral components at each of the frequencies of interest are too mismatched.

In spite of these problem areas for the TMM-R method, Figs. 4–6 show that this method can approach the quality of our ‘*de facto*’ standard. With the exception of the high-frequency results (2.8–3.0 kHz) for the ceramic tubular and silicon carbide test specimens, the TMM-R results are very

nearly identical to the MPM-DF results. Even in these high-frequency regions, the results acquired with the two methods (TMM-R and MPM-DF) are in close agreement. It should be noted here, however, that high-frequency applications will require increased accuracy over that attained with the TMM-R method. It would thus appear that this method may not be suited to high-frequency applications without further enhancements.

C. Description of SMM results

Efforts are underway at NASA Langley Research Center to implement a high temperature normal incidence impedance tube facility. Because of the extremely high temperatures in this facility (perhaps eventually as high as 1600 K), it was initially thought that single microphone probe measurements might be the best approach, in which a microphone probe would be inserted into the high temperature environment at preselected locations for only a short period of time. Plans are also underway for the design of a much smaller (high frequency) normal incidence impedance tube. The single-microphone method (SMM) was investigated in this study to explore the possibility of acquiring quality impedance data in these harsh, or space limited (e.g., smaller diameter tube), environments.

For the current study, the SMM was implemented with either of two source types, the single discrete frequency (SMM-DF) or the random noise (SMM-R). This method is applied in essentially the same manner as the TMM, with the second microphone replaced by the electronic source reference signal (function generator or random noise generator) in the transfer function measurements. Although only a single microphone is placed into the test environment, the difficulties associated with the TMM remain with the SMM implementations. This is a result of the fact that measurements are made with the single microphone at only two locations; thus, the system is deterministic.

1. Single discrete frequency source (SMM-DF)

Results for the ceramic foam and silicon carbide test specimens with the SMM-DF are almost identical to the standard (MPM-DF) data, for both the normalized resistances and the normalized reactances. The same is true for the ceramic tubular test specimen results with the exception of the data near 3.0 kHz. The SMM-DF normalized reactance result is about 0.2 ρc units below that of the standard method at this test frequency. Overall, the results for this method are quite encouraging, but the deterioration at 3.0 kHz indicates serious potential problems for using this method up to 25 kHz.

2. Random noise source (SMM-R)

The last method included in the current comparison study was the single-microphone method with the random noise source. It was expected that this would be the least reliable of the methods of this study for the following two reasons: (1) only one microphone is placed in the test environment; therefore, difficulties due to reduced coherence (between noise source and measurement microphone) may be evident; and (2) the problems previously listed for a random

TABLE II. Time and accuracy rankings of impedance measurement methods.

Measurement method	Time ranking	Accuracy ranking
MPM-DF	6	1
MPM-PR	3	4
TMM-DF	4	2
TMM-R	1	5
SMM-DF	4	2
SMM-R	1	6

noise source are also present for this method. When two microphones are used (TMM), only the acoustic environment (over a short separation distance) and the individual microphone signal conditioners are contributors to the coherence measurement quality. When one of the microphones is replaced with the electronic noise source (SMM), the contributions of the removed microphone signal conditioner are replaced with contributions due to the electronic noise source, amplifiers, acoustic drivers and the acoustic environment (over a much larger distance). As a result, the coherence for the SMM implementation is significantly less than that of the TMM.

As expected, the results for this method were somewhat less impressive than the other methods included in this study. For the ceramic tubular and silicon carbide test specimens, the comparisons of normalized resistances and normalized reactances between the two methods (MPM-DF and SMM-R) are quite acceptable, with discrepancies at or below approximately $0.1 \rho c$ units for all but a few frequencies. The results for the ceramic foam test specimen, however, show the SMM-R data to oscillate about the standard data, in both components of the normalized impedance. Multiple test runs were conducted, which indicated that this behavior is repeatable. Since the SMM-R results for the other two test specimens were reasonably smooth and well-matched to the MPM-DF values, the results for the ceramic foam sample were somewhat puzzling. However, as previously mentioned, similar difficulties had been observed with this particular test specimen material in the past. It is therefore the belief of the authors that the poorer quality of the results for this particular sample may be more a function of the test material than of the test procedure.

D. Comparison of MPM, TMM, and SMM

As stated earlier, one purpose of this study was to provide a comparison of each of the normal incidence acoustic impedance measurement methods in use at NASA Langley Research Center. This was done for the purpose of determining relative merits of each of the methods in more demanding test conditions. Figures 7–9 provide an overall comparison of results acquired with all of the methods. As expected, the “best method” is dependent upon a tradeoff between the time availability and accuracy considerations. Table II provides a relative merit listing on a 1–6 scale (1 is best, 6 is worst) of each of the methods included in the current study. Each method is ranked according to the amount of time required for completion and the accuracy of the results. Together with the information provided in Table I, which lists

the strengths and weaknesses of each of the methods, this table summarizes the results of this investigation for use in future tests.

As is shown in Table II, the TMM-R and SMM-R methods require the least amount of data acquisition time. The pseudo-random implementation of the MPM (MPM-PR) takes some additional time, mainly because of the setup time needed for the pseudo-random noise source. The single discrete frequency implementations of the TMM and the SMM take considerably longer, while the single discrete frequency implementation of the MPM takes the longest.

The accuracy rankings are almost inversely proportional to the data acquisition time, as might be expected. The MPM-DF is the most accurate method.⁹ The other two methods (TMM-DF and SMM-DF) which implement the single discrete frequency source are essentially equal in their accuracy, followed by the pseudo-random noise implementation of the MPM (MPM-PR). The two least accurate methods use the random noise source. Possibly because the TMM-R has two microphones in the test environment, whereas the SMM-R compares a single microphone in the test environment against the electronic random noise source, the increased coherence between the two input signals causes the accuracy of the TMM-R to be slightly greater than that of the SMM-R.

IV. CONCLUDING REMARKS

Alternate implementations of the two-microphone and multipoint impedance measurement methods (TMM and MPM) were investigated to determine their potential usefulness in more difficult environments. The single-microphone method (SMM), a modification of the TMM, was included as a potential candidate to provide measurement capabilities in situations where only one microphone could be placed into the test environment. The pseudo-random version of the multi-point method was included as a more time efficient implementation. This method incorporates pseudo-random noise, generated by the combination of a number of discrete frequencies (each at a selected amplitude and phase), to significantly reduce the time necessary to complete an acoustic impedance measurement.

The MPM with a single discrete frequency source (MPM-DF) is the most accurate, but most time demanding, method currently in use. Because of the accuracy of the MPM-DF method, it was chosen as the standard against which the other methods would be compared. Comparisons of each of the other normal incidence acoustic impedance measurement methods with this “*de facto* standard” method in use at NASA Langley Research Center were obtained for three different test specimens to characterize the strengths and weaknesses of each method. As expected, the relative merits of each of the other methods were shown to depend upon a trade-off between the amount of time available and the accuracy necessary.

With minimal loss of accuracy (typically within $0.1 \rho c$ units, slightly greater at or above 2.5 kHz), the pseudo-random noise source implementation of the MPM requires significantly less time. Regardless of the source type, the MPM is an ‘overdetermined’ method which requires a mean

square fit with accompanying residual error. This limitation is, however, offset by the freedom from specific sensor placement restrictions.

The SMM was found to be quite accurate with a single discrete frequency source, but the most inaccurate with a random noise source. Both the SMM and the TMM are deterministic methods; i.e., the impedance of a test specimen can be determined from just two measurements. These two measurements must, however, exclude separations of near one half-wavelengths.

Further research is recommended for the ceramic foam material to determine the nature of the apparent specimen dependent accuracy degradation associated with this test specimen. This should be followed with a test series on non-uniform surface distributions of impedances. Finally, general degradation (albeit slight) at the higher frequencies (2.5 to 3.0 kHz) should be further investigated in a systematic manner in order to extend the appropriate method to 25 kHz, which will most likely be a challenging task.

¹W. T. Chu, "Impedance tube measurements—a comparative study of current practices," *Noise Control Eng. J.* **37**, 37–44 (1991).

²L. L. Beranek, *Acoustical Measurements* (revised edition) (American Institute of Physics, Woodbury, NY, 1988).

³D. A. Blaser and J. Y. Chung, "A transfer function technique for determining the acoustic characteristics of duct systems with flow," *Inter-Noise*, 901–908 (1978).

⁴A. F. Seybert and T. L. Parrott, "Impedance measurement using a two-microphone, random-excitation method," NASA TM-78785 (1978).

⁵J. Y. Chung and D. A. Blaser, "Transfer function method of measuring in-duct acoustic properties. I. Theory and II. Experiment," *J. Acoust. Soc. Am.* **68**, 907–921 (1980).

⁶A. F. Seybert and D. F. Ross, "Experimental determination of acoustic properties using a two-microphone, random-excitation technique," *J. Acoust. Soc. Am.* **61**, 1362–1370 (1977).

⁷"Standard test method for impedance and absorption of acoustical materials using a tube, two microphones, and a digital frequency analysis system," ASTM E 1050-90.

⁸"Standard test method for impedance and absorption of acoustical materials by the impedance tube method," ASTM C 384-90a.

⁹M. G. Jones and T. L. Parrott, "Evaluation of a multi-point method for determining acoustic impedance," *Mech. Syst. Signal Process. J.* **3**, 15–35 (1989).

¹⁰F. J. Fahy, "Rapid method for the measurement of sample acoustic impedance in a standing wave tube," *J. Sound Vib.* **97**, 168–170 (1984).

¹¹W. T. Chu, "Single-microphone method for certain applications of the sound intensity technique," *J. Sound Vib.* **101**, 443–445 (1985).

¹²W. T. Chu, "Transfer function technique for impedance and absorption measurements in an impedance tube using a single microphone," *J. Acoust. Soc. Am.* **80**, 555–560 (1986).

¹³U. Ingard and V. K. Singhal, "Sound attenuation in turbulent pipe flow," *J. Acoust. Soc. Am.* **55**, 535–538 (1974).

¹⁴F. B. Hildebrand, *Methods of Applied Mathematics* (Prentice-Hall, Englewood Cliffs, NJ, 1965), p. 286.

¹⁵T. L. Parrott, W. R. Watson, and M. G. Jones, "Experimental validation of a two-dimensional shear-flow model for determining acoustic impedance," NASA TP-2679 (1987).

Spectral cues for sound localization in cats: A model for discharge rate representations in the auditory nerve

Bradford J. May^{a)}

Department of Otolaryngology-HNS, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205-2195

Aileen Y. Huang^{b)}

Department of Biomedical Engineering, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205-2195

(Received 15 October 1996; revised 6 January 1997; accepted 11 January 1997)

Neural representations of pinna-based spectral cues for sound localization were modeled by simulating auditory nerve discharge rates to noise bursts that had been shaped by filtering properties of the cat's head-related transfer functions (HRTFs) at 179 locations in the frontal field. The auditory nerve model transformed spectral differences between HRTFs into simulated neural rate differences. Linear equations for this transformation were developed from actual auditory nerve responses to a limited subset of HRTF-filtered noise bursts [Rice *et al.*, J. Acoust. Soc. Am. **97**, 1764–1776 (1995)]. Signal detection methods were used to investigate simulated neural responses to pairwise changes between HRTFs. The quality of neural representation for these changes, in terms of d' values, declined when the reference HRTF was moved from a central location (0° AZ, 0° EL) to a large positive azimuth in the horizontal plane (75° AZ, 0° EL) or a high elevation in the median plane (0° AZ, 75° EL). Most simulated responses exhibited large d' values for comparisons of contralateral versus ipsilateral azimuths, or eccentric versus frontal elevations. This rate information resulted from directionally dependent changes in the overall gain of HRTFs. In addition, fibers with best frequency (BF: the frequency of greatest sensitivity for individual fibers) between 5 and 18 kHz showed large d' values for HRTF contrasts in the immediate frontal field because of the effects of spectral notches (i.e., sharp drops in gain over a narrow frequency range). Spectral notches also played a prominent role in simulations that required identification of HRTF location in the absence of a fixed reference stimulus. These modeling results correspond well with previously described patterns in the cat's localization behaviors. © 1997 Acoustical Society of America. [S0001-4966(97)01505-1]

PACS numbers: 43.64.Bt, 43.64.Pg, 43.66.Qp [RDF]

INTRODUCTION

Human psychophysical experiments have established the fundamental importance of pinna-based spectral cues in directional hearing. Spectral information contributes to the resolution of front/back confusions when different sound source locations create the same interaural cues, and is critical for accurate localization of elevation in the median plane where interaural differences are presumed to reduce to zero (Butler and Belenduk, 1977; Hebrank and Wright, 1974). Sounds presented through closed acoustic systems can be made to appear to arise from specific external locations if stimuli are shaped by appropriate pinna-filtering effects (Plenge, 1974; Wightman and Kistler, 1989a, b). Although the importance of pinna-based directional information is not disputed, the precise acoustic basis of spectral cues for sound localization is still the subject of some debate. Various researchers have offered psychophysical evidence supporting the importance of directional cues conveyed by spectral

peaks (Blauert, 1983; Middlebrooks, 1992) or, in contrast, spectral notches (Hebrank and Wright, 1974; Bloom, 1977; Watkins, 1978).

The filter function describing the transformation of a free-field sound to energy at the eardrum is known as the head-related transfer function (HRTF). The HRTF of the cat has been the subject of extensive acoustic analyses (Wiener *et al.*, 1966; Phillips *et al.*, 1982; Calford and Pettigrew, 1984; Musicant *et al.*, 1990). Transfer functions shown in Fig. 1 are taken from the database (Rice *et al.*, 1992) that serves as the source of stimuli for modeling work described here. Although localization cues exist throughout the HRTF spectrum, the Rice *et al.* study proposed a special significance for directional information in the 5–18 kHz region. One primary spectral notch is found at these frequencies and it may attain a depth of over 30 dB for sound locations in the frontal field. It has been hypothesized that mid-frequency notches play an important role in the cat's sound localization behavior because these features show systematic frequency changes in response to movement of a sound source. If the mid-frequency spectrum is removed from a broadband stimulus, cats show poorly directed sound orientation behaviors (Huang and May, 1996a).

In order to provide the CNS with effective sound local-

^{a)}Corresponding author: Bradford J. May, Ph.D., Dept. of Otolaryngology-HNS, 505 Traylor Research Bldg., The Johns Hopkins University, 720 Rutland Avenue, Baltimore, MD 21205-2195. Electronic mail: bmay@bme.jhu.edu

^{b)}Present address: Southwest Research Institute, San Antonio, TX.

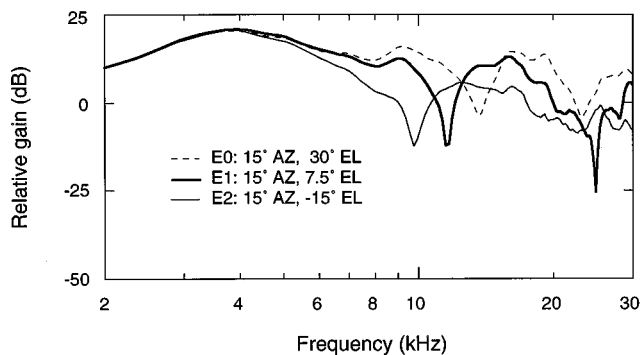


FIG. 1. Head-related transfer functions (HRTFs) at three elevations in the frontal sound field. Each HRTF exhibits a prominent spectral notch at mid-frequencies (5–18 kHz). Positive azimuths indicate locations ipsilateral to the microphone during HRTF recordings (redrawn from Rice *et al.*, 1992).

ization cues, directional properties of the HRTF must be adequately represented by activity in the auditory nerve. Poon and Brugge (1993) measured discharge rates of auditory nerve fibers (ANFs) to broadband noise with a spectral notch that was swept in frequency across the response area of individual fibers. Vigorous responses to the passbands of the noise fell to rates near spontaneous activity as the notch reached BF, showing that ANFs are sensitive to spectral notches.

Rice *et al.* (1995) provided a more complete description of the auditory nerve representation of HRTF-filtered noise using population measures that were originally applied to human speech sounds (Sachs and Young, 1979). Their analysis was based on the assumption that frequency components of complex spectra are encoded by a profile of discharge rates as a function of BF across a population of ANFs. In the context of sound localization, spectral differences between HRTFs are presumed to create distinct patterns of discharge rates in the auditory nerve as a sound source shifts from one location to another. Rice *et al.* (1995) demonstrated that differences in the discharge rate profiles produced by two HRTFs provide a good representation of differences in the HRTF spectra, and therefore of a change in sound source direction. Unfortunately, their direct electrophysiological measures are available only for a narrow range of elevations in the frontal field.

Our current modeling efforts are designed to extend the electrophysiological results of Rice *et al.* throughout the frontal sound field. We use an auditory nerve model to predict the representation of pinna-based spectral cues in terms of discharge rates, which are computed from the linear relationship between single fiber discharge rates and HRTF-filter shapes. A similar relationship has been previously observed for auditory nerve responses to vowel spectra (Conley and Keilson, 1995; May *et al.*, 1996). Interaural time differences (ITDs), another source of important sound localization cues, are ignored by this analysis. When the auditory nerve model was used to estimate the discriminability of the neural representation of HRTFs at 179 locations in the frontal field, greatest directional acuity was observed for locations directly in front of the cat (0° AZ, 0° EL), where mid-frequency notches are prominent features of HRTF spectra. These mod-

eling results correspond well with behavioral patterns observed in previous studies of sound localization in cats (May and Huang, 1996; Huang and May, 1996a,b).

I. AUDITORY NERVE MODEL

Rice *et al.* (1995) obtained electrophysiological recordings from barbiturate anesthetized cats using a closed-field sound delivery system. Closed-field simulations of free-field directional cues were made by shaping the amplitude spectrum of broadband noise with digital filters based on the cat's head-related transfer functions (HRTFs) at three locations in the frontal sound field: E0 (15° AZ, 30° EL), E1 (15° AZ, 7.5° EL), and E2 (15° AZ, -15° EL). The acoustic data that served as the source of their digital filters are shown in Fig. 1. Notice that each HRTF displays a sharp drop in gain at frequencies between 5 and 18 kHz. The frequency of this spectral notch changes systematically with sound location (Rice *et al.*, 1992) and is thought to provide useful cues for sound localization in cats (May and Huang, 1996; Huang and May, 1996a).

The analyses to be described estimate the perceptual discriminability of pairwise contrasts of sound source locations by evaluating the magnitude of neural rate differences evoked by stimuli at the two locations. Signal detection methods for obtaining these estimates are illustrated by data in Fig. 2. The left panel of Fig. 2(a) shows the ratio of the magnitude spectra for stimuli E0 and E1 (i.e., $|E0/E1|$), which is computed as the difference between the log (dB) plots in Fig. 1. Large positive and negative amplitude differences between 9 and 13 kHz reflect the mid-frequency spectral notches of the HRTFs underlying these stimuli. Amplitude difference contours for the HRTF pairs $|E0/E2|$ and $|E1/E2|$ exhibit similar features, as shown in the remaining panels of Fig. 2(a).

The auditory nerve encoding of the $|E0/E1|$ contrast is shown in the left panel of Fig. 2(b). Each symbol indicates the response of a single auditory nerve fiber (ANF), which is plotted at the fiber's BF on the abscissa. Different symbols distinguish ANFs with the following spontaneous rates (SRs): high ($SR \geq 18$ spikes/s), medium ($1 < SR < 18$ spikes/s), and low ($SR \leq 1$ spike/s). Responses are measured in terms of rate differences. For the $|E0/E1|$ contrast, these rate differences are calculated by subtracting each fiber's discharge rate during presentations of the E1 stimulus from responses during E0 presentations. When this analysis is performed for a large number of fibers with differing BFs, the resulting rate-difference population profile has a shape similar to the dB difference between the stimuli, with peaks and valleys at BFs corresponding to spectral notches in the HRTFs. Similar results are seen for the $|E0/E2|$ and $|E1/E2|$ contrasts in the remaining panels of Fig. 2(b). Note that neurons showing the largest rate differences are situated at different BF regions for each comparison because of the unique frequency location of the mid-frequency notch for each sound source location. These results suggest a rate-place representation of sound direction in the auditory nerve.

Signal detection theory (Green and Swets, 1966) offers a method for quantitative evaluation of the rate differences in

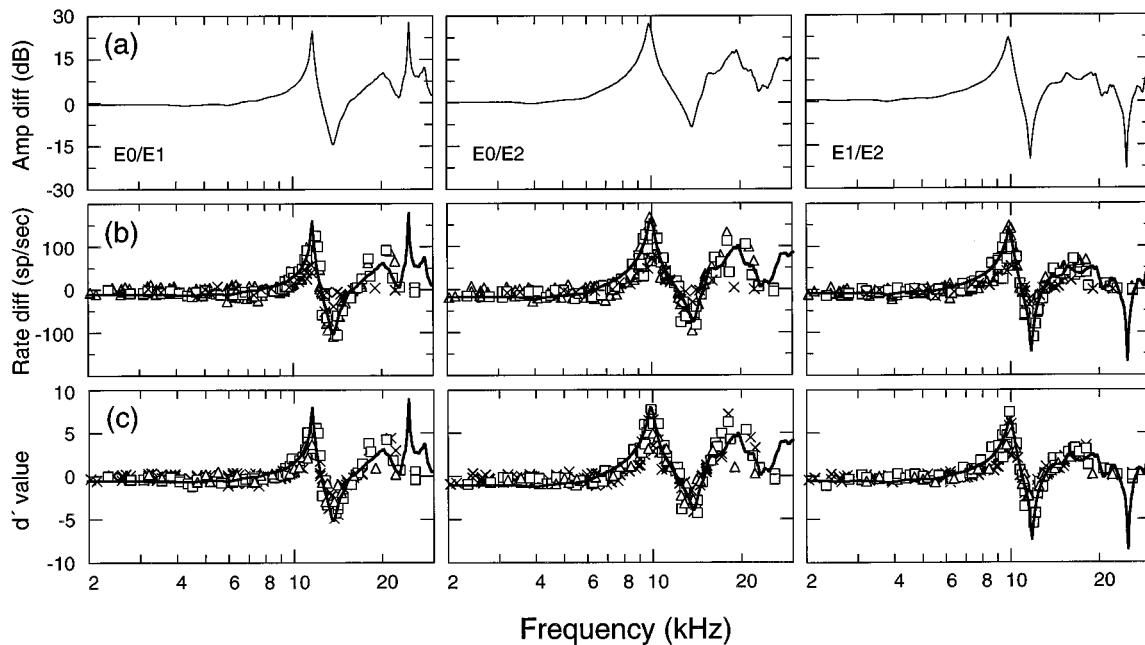


FIG. 2. Representation of differences between HRTFs by changes in auditory nerve discharge rates. (a) Amplitude difference contours for pairwise comparisons of the three HRTFs in Fig. 1, shown in terms of the dB difference between the HRTFs. (b) Rate difference profiles for single fiber responses to each pair of HRTFs. Each symbol indicates one fiber's difference in discharge rate to the two stimuli, plotted at that fiber's best frequency (BF). Fibers with medium spontaneous rates ($1 < \text{SR} < 18$ spikes/s) are marked by squares. Fibers with higher SRs are shown by \times 's, and fibers with lower SRs by triangles. (c) d' values for rate differences in (b) (see text). For comparison, amplitude difference contours in upper panels are superimposed on neural data in lower panels (bold lines). (Replotted from Rice *et al.*, 1995.)

Fig. 2(b). The d' index of neural discrimination is computed as follows:

$$d' = \frac{R_{E0} - R_{E1}}{\sqrt{\sigma_{E0}^2 + \sigma_{E1}^2}}, \quad (1)$$

where R_{E0} and R_{E1} are the average rates of a single fiber in response to multiple presentations of E0 and E1, and σ_{E0} and σ_{E1} are the standard deviations of those rates. The d' statistic represents the z transform of the difference in the fiber's spike counts for responses to the stimuli, and as such is monotonically related to the probability of correct discrimination between stimuli based on the observation of discharge rates for a single paired presentation of E0 and E1 (Rice *et al.*, 1995). d' values for auditory nerve responses to the three stimulus contrasts are shown in Fig. 2(c). Note that the largest positive and negative d' values are evoked by mid-frequency spectral notches.

The data in Fig. 2 suggest that a linear relationship exists between rate differences and spectral differences for HRTF stimuli. Such a relationship has been demonstrated previously for speech stimuli (Conley and Keilson, 1995), and provides a basis for estimating auditory nerve responses to stimuli that were not tested in the physiological study of Rice *et al.* (1995). The linear relationship between rate and HRTF-based spectral differences is illustrated for high and medium SR fibers in Fig. 3(a), and for low SR fibers in Fig. 3(b). The data shown here are taken from Fig. 2(a) and (b); in this figure each fiber's rate difference is plotted against the spectral amplitude difference at the fiber's BF. This analysis is based only on fibers with BFs between 5 and 18 kHz. At frequencies below 5 kHz, large amplitude differences are not

observed between the HRTFs; at frequencies above 18 kHz, a sparse sampling of fibers makes it difficult to quantify the precise nature of high-frequency response patterns.

Linear regressions fit to each set of data in Fig. 3 achieved a statistically significant correlation coefficient ($R^2 > 0.70$). Parameters of the regression lines are given in Table I. No significant differences in either slope or y intercept were observed for regressions of high versus medium SR fibers (upper panels of Fig. 3), nor were differences found between HRTF contrasts. Therefore, our modeling analysis uses two linear regressions: one for high/medium SR fibers and one for low SR fibers. Both regressions are based on pooled data across the three HRTF contrasts. Note that the y intercepts of pooled regressions are not statistically significant from 0; therefore, y intercepts of 0 were used in calculations of rate differences on the assumption that level differences of 0 dB should produce no rate differences.

The estimates in Table I are based on data at 0-dB approximate level (see Rice *et al.*, 1995, for details of their method for specifying stimulus level). As is shown by May *et al.* (1996) the slopes obtained for rate difference versus spectral difference functions vary with overall sound level of the stimuli. At 0-dB approximate level, low SR fibers with high thresholds show a narrow range of discharge rates (i.e., a shallow regression function) because they are not sensitive to low level spectral features. Although auditory nerve responses to high level stimuli are not modeled in this report, low SR fibers typically provide the best representation of complex spectra at high stimulus levels because high/medium fibers show rate saturation effects (Sachs and Young, 1979; Le Prell *et al.*, 1996).

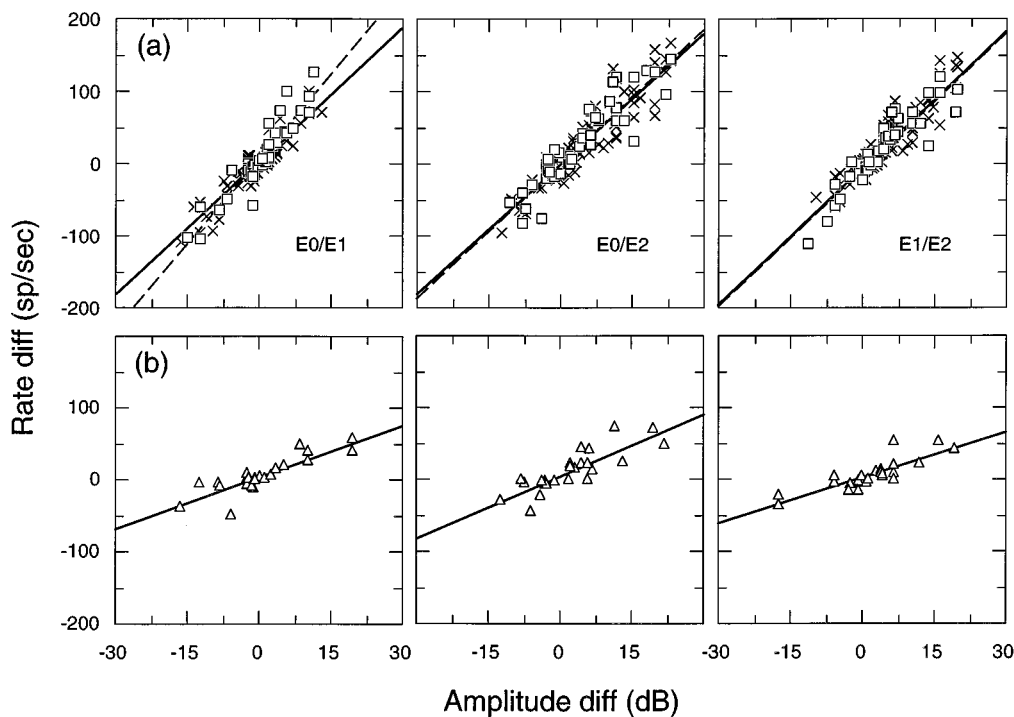


FIG. 3. Rate difference versus spectral difference functions for auditory nerve responses to paired combinations of HRTFs. These sensitivity functions were created by plotting the single fiber responses in Fig. 2(b) according to the amplitude of the spectral difference at each fiber's BF in Fig. 2(a). Responses and linear fits to data of fibers with high or medium SRs are shown in (a); low SR responses are presented in (b). Linear fits for high and medium SR fibers are indicated by solid and dashed lines, respectively. Parameters of the linear fits are given in Table I; symbol conventions are described in Fig. 2.

The linear regressions shown in Table I can be used to estimate auditory nerve rate differences for the difference spectrum of any pairwise comparison of HRTFs. d' can then be estimated from the rate differences, given a measure of σ . Figure 4(a) shows σ as a function of BF for all fibers in the Rice *et al.* study. Across most BFs, high/medium SR fibers show σ values that are evenly distributed around a median value of 16.7 spikes/s (horizontal line). A trend toward lower σ values is observed among fibers with high BFs (>20 kHz). The sample of low SR fibers produced a median σ of 9.5 spikes/s (dashed line), and showed an even sharper decline in variability at high BFs. This pattern of responses can be interpreted by inspection of Fig. 4(b), which plots σ as a function of the total rates elicited by HRTF-filtered noise. As noted in our previous studies (May *et al.*, 1996), σ is related to total rate by a power function, which for these data is

$$\sigma = 2.6 \times (\text{total rate})^{0.38}. \quad (2)$$

Although σ values of both high/medium and low SR fibers are fit by Eq. (2), low SR fibers operate at the lower extreme of the power function because of their relatively low driven rates. The threshold effects that yield these low rates are particularly evident among fibers with BFs at high frequencies where most HRTFs show low gain (Fig. 1).

For convenience, d' calculations for our simulated auditory nerve responses use constant values of σ that are based on the median scores of fibers in Fig. 4. This computational method ignores the effects of variation in σ , which in the case of high SR fibers shows an interquartile range that extends from 13.9 to 18.9 spikes/s. In the context of Eq. (1), changes in σ of this magnitude can alter predicted d' scores by approximately $\pm 15\%$ relative to values obtained with the 16.7 spikes/s median σ score. That the assumption of constant σ does not grossly affect model calculations is shown in Fig. 5, where simulated d' profiles (solid lines) are compared with actual physiological responses (symbols) for the $|E0/E1|$ contrast. The physiological data used in this comparison are taken from Fig. 2(c). Our model uses a linear transformation to convert HRTF-based spectral differences into rate differences; consequently, simulated d' profiles resemble scaled versions of the difference spectrum shown in Fig. 2(a). Although high/medium SR fibers are on average more variable than low SR fibers (i.e., have a higher σ value), these simulations yield larger d' values because of the steeper slopes of the high/medium SR linear regressions.

TABLE I. Linear regressions for rate differences of three HRTF contrasts.

Contrast	Regression	High SR	Medium SR	Low SR
E0/E1	y intercept	0.66	4.53	2.49
	Slope	6.22	7.87	2.46
	R^2	0.85	0.86	0.75
E0/E2	y intercept	0.18	-1.90	0.99
	Slope	6.04	6.26	2.88
	R^2	0.87	0.84	0.70
E1/E2	y intercept	-3.81	-7.14	0.16
	Slope	6.35	6.37	2.13
	R^2	0.84	0.84	0.70
Pooled	y intercept	(-2.28)		(1.33)
	Slope	6.19		2.46
	R^2	0.84		0.70

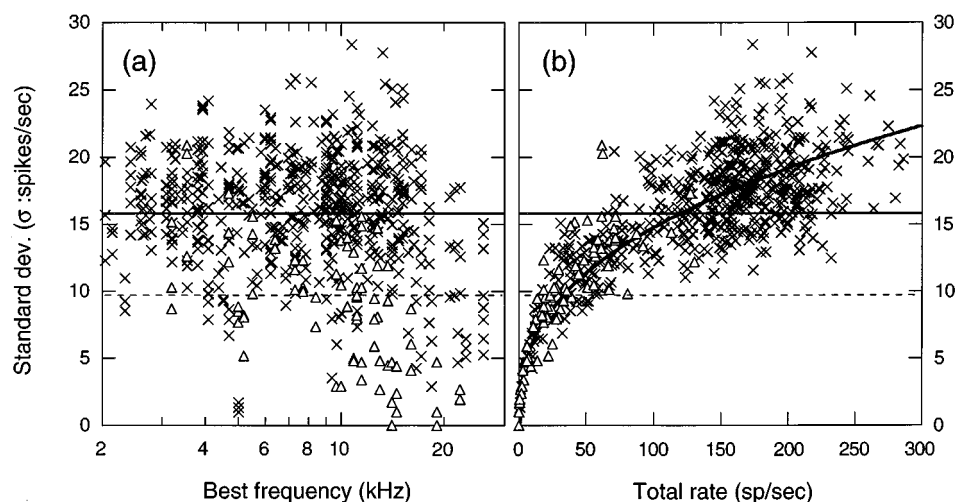


FIG. 4. Standard deviation (σ) of auditory nerve responses as a function of best frequency (a) and total discharge rate (b). Fibers with high or medium SRs are indicated with \times 's; fibers with low SR are shown by triangles. Median σ is indicated by the horizontal solid line for high/medium SR fibers and the horizontal dashed line for low SR fibers. The bold line in (b) is the power function described by Eq. (2).

High-frequency ANFs integrate sound energy over broader bandwidths than mid-frequency fibers and therefore may be less sensitive to amplitude differences that result from sharp spectral features. This is a potential source of error for our model because regressions applied to all simulated fibers were derived solely from mid-frequency re-

sponses. One method to correct for BF-dependent changes in frequency tuning is to filter HRTF spectra prior to applying the linear model. Dashed lines in Fig. 5 show modeling results obtained with $|E_0/E_1|$ after frequency components of each HRTF were filtered by a bank of 512 Gaussian filters with bandwidths equivalent to the 10-dB bandwidth of ANFs at that frequency (Evans, 1975). Filtered simulations exhibit a good correspondence to actual responses, except at mid-frequency notches which are a prime source of spectral information. Subsequently modeling results are based on unfiltered stimuli.

Horizontal lines in Fig. 5 point out regions where d' values exceed ± 1 , which is a common signal detection criterion for the threshold of detectable neural rate differences. Our description of modeling results uses this convention to identify those frequency regions that elicit good ($d' \leq -1$ or $d' \geq 1$) versus poor ($-1 < d' < 1$) rate representations of directional changes. Notice that filtering the HRTF changes the magnitude of d' values but not the overall pattern of frequencies showing largest rate differences.

Rate information conveyed by a small number of optimally selected auditory-nerve fibers can account for the sensitivity of behavioral performances in psychophysical tasks such as absolute detection (Young and Barta, 1986), intensity discrimination (Winslow *et al.*, 1987), and formant frequency discrimination (May *et al.*, 1996). To allow comparison of these optimal responses across a variety of stimulus conditions, Fig. 6 shows the statistical distribution of d' values in terms of "box plots" (Tukey, 1977). Details of this plotting convention are described in the legend of Fig. 6. Maximum d' values for a change between HRTFs can be judged by the vertical placement of the upper error bar on the box plots, which marks the largest d' value that falls within the 1.5 interquartile range above the top of the box. In some cases, a few neural responses yield d' values that exceed the 1.5 interquartile range. These outliers are plotted as individual symbols. Comparison of the box plots in Fig. 6 indicate that maximum d' values obtained with the model are

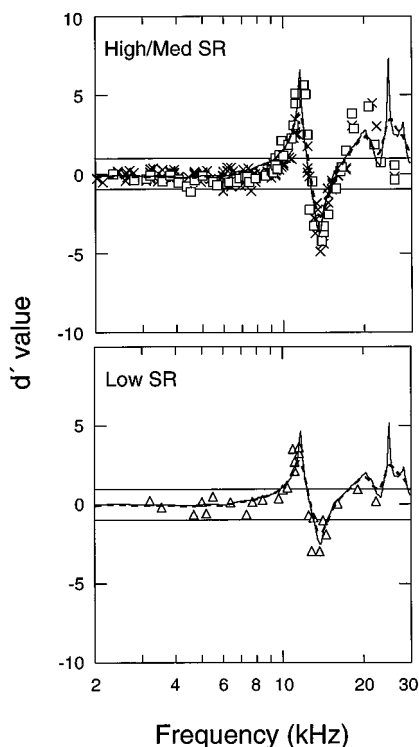


FIG. 5. Simulated d' values for $|E_0/E_1|$. Modeling results for high/medium SR simulations are shown in the upper panel; low SR data are presented in the lower panel. Results obtained with unfiltered HRTFs are indicated by solid lines; simulations for filtered HRTFs are shown by dashed lines. For comparison, d' values of actual auditory nerve fibers are superimposed as individual symbols. Physiological data are taken from Fig. 2(c); symbol conventions are described there. Horizontal lines mark d' values of ± 1 .

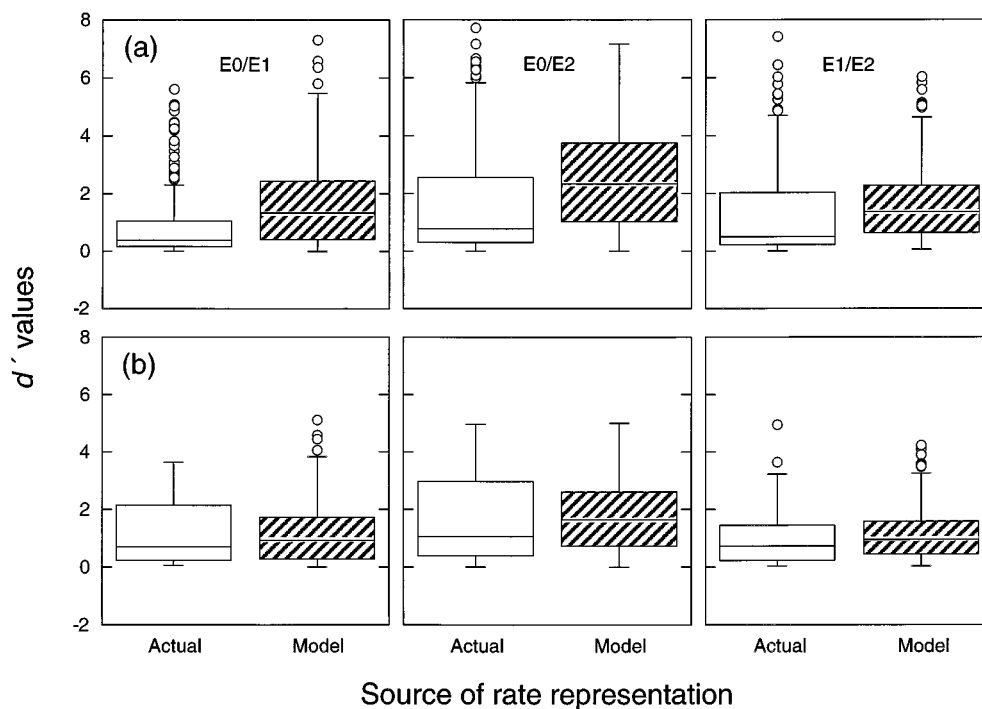


FIG. 6. Box plots showing the distribution of actual and simulated d' values for high/medium SR (a) and low SR auditory nerve fibers (b). Each panel shows responses for a different HRTF contrast. The lower and upper limit of each box indicates the first and third quartile of the distribution of neural d' values; this is known as the interquartile range of the distribution. The box is bisected at the second quartile, which is the median of the distribution. Upper error bars are drawn to the highest data point that falls within the 1.5 interquartile range above the box; lower error bars show lowest data point within 1.5 interquartile range below the box. Points lying outside error bars are plotted as individual symbols.

equivalent to actual neural data under most stimulus conditions, and that high/medium SR simulations yield higher maximum d' values than low SR simulations.

Because the d' distribution of the actual electrophysiological database reflects the random isolation of fibers in the auditory nerve, some BF regions are represented by many fibers and others not at all. Such sampling limitations are avoided by our modeling results because all comparisons are based on a uniformly-spaced (100-Hz steps), nonrepetitive sampling of the complete HRTF spectrum. However, it is important that the model simulates response patterns that are similar in magnitude to actual auditory nerve discharge rates. In Fig. 6, maximum d' values for actual and simulated rate representations (as indicated by the upper error bars of box plots) do show a good correspondence for $|E0/E2|$ and $|E1/E2|$, but the model yields a relatively large maximum d' for $|E0/E1|$ in comparison to actual values. Examination of auditory nerve data in Fig. 2(b) suggests the discrepancy lies in an incomplete sampling of fibers with BFs near the peaks of the $|E0/E1|$ difference spectrum. As a result, the limited sample of actual responses may underestimate the full range of spectral cues for the HRTF contrast.

II. RESULTS

Auditory nerve encoding of pinna-based spectral cues was simulated at 179 locations in the frontal field. Azimuthal locations ranged from -75° to 75° , in increments of 15° . Elevations ranged from -30° to 90° , in increments of 7.5° . HRTFs for these simulations were collected from cat 1206 in an earlier acoustic study of pinna-based spectral cues (Rice

et al., 1992). Transfer functions from five cats in that study showed general similarities in spectral features such as overall gain and the frequency location of prominent notches.

A. Localization cues for a central reference HRTF (0° AZ, 0° EL)

Figure 7(a) shows a reference HRTF at 0° AZ, 0° EL (bold line) and comparison HRTFs at $\pm 15^\circ$ AZ in the interaural horizontal plane (left panel) and $\pm 15^\circ$ EL in the median plane (right panel). Note the change in notch frequency between the HRTFs and the overall gain increase as the sound source moves from a contralateral (-15° AZ, solid line) to an ipsilateral location (15° AZ, dashed line). HRTFs in the median plane show essentially identical gain at low frequencies (<5 kHz) and a complex pattern of spectral notches at high frequencies (>18 kHz). Like HRTFs in the horizontal plane, transfer functions in the median plane also exhibit directionally dependent spectral notches at mid-frequencies (5–18 kHz).

Figure 7(b) shows d' profiles that were obtained when the auditory nerve model was applied to the HRTFs in Fig. 7(a). Difference spectra were calculated by taking the ratio (dB difference) of each comparison HRTF and the reference HRTF at 0° AZ, 0° EL. These spectra were multiplied by the pooled linear regression slopes for high/medium SR fibers (6.19 spikes/s/dB) to create a rate-difference profile and then divided by a constant σ (16.7 spikes/s) to compute neural d' values [Eq. (1)]. The high/medium SR model was chosen for this analysis because it produced larger d' values than the low SR model at the sound level used (Fig. 5). Both SR

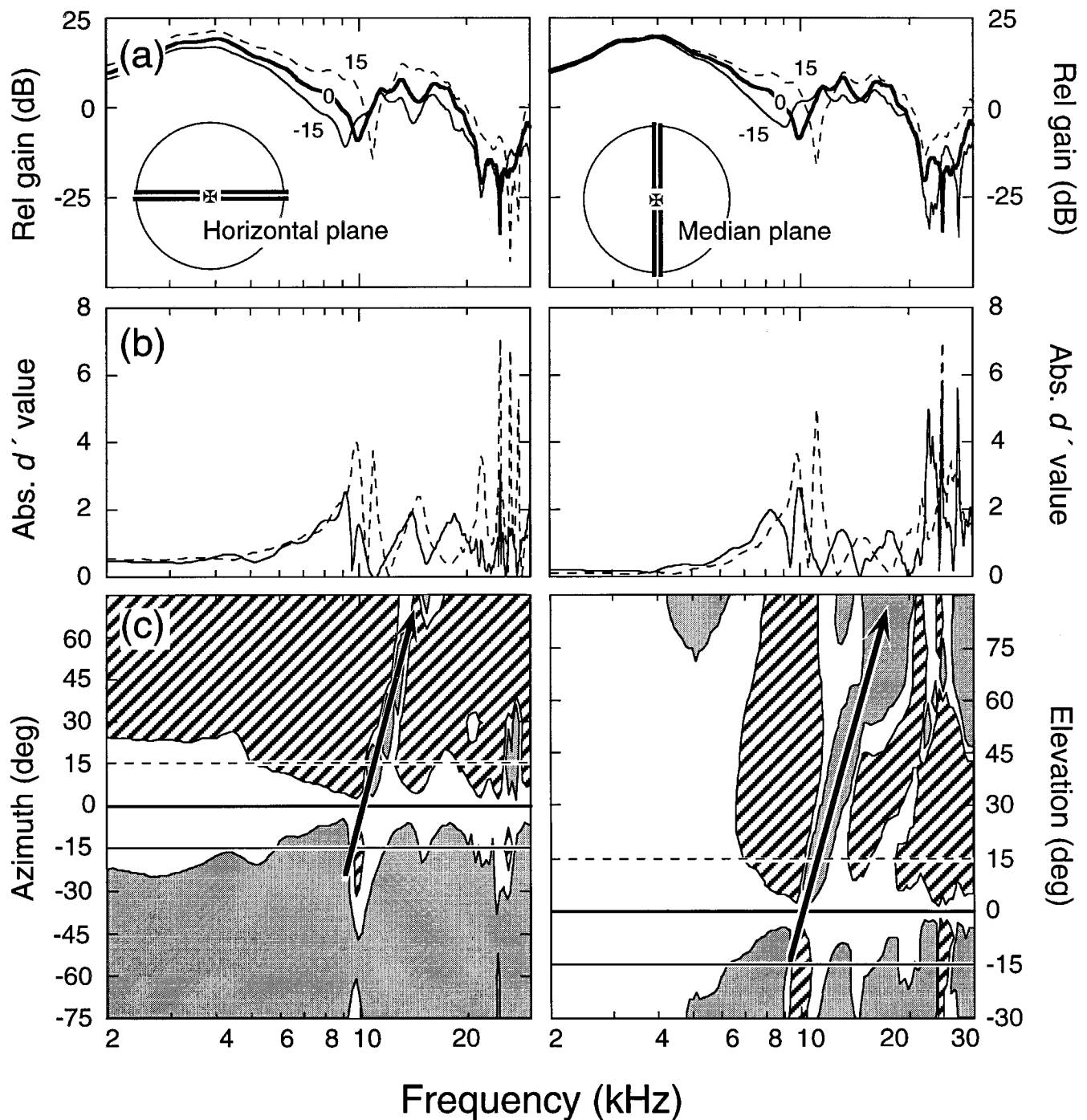


FIG. 7. Neural representation of spectral cues at 0° AZ, 0° EL. (a) HRTFs for sound locations in the horizontal plane are shown on the left; HRTFs in the median plane on the right. Numbers indicate HRTF location in degrees AZ or EL. (b) Simulated d' values produced by applying the model to spectral differences between HRTFs in upper panels. The function at 0° AZ, 0° EL served as the reference for each HRTF contrast; the comparison HRTF used in the simulation is identified by the same line pattern as HRTFs in upper panels. (c) Contour plots of simulated d' values for comparisons of the 0° AZ, 0° EL reference HRTF with other transfer functions at 10 azimuths in the horizontal plane (left panel) and 16 elevations in the median plane (right panel). Solid filled regions illustrate combinations of frequency and location where spectral differences between a comparison HRTF (location shown on ordinate) and the reference HRTF created rate decreases with d' values of less than -1 . Striped regions indicate rate increases with d' greater than 1 . Horizontal lines mark the location of transfer functions shown in (a). Arrows follow the frequency of the mid-frequency spectral notch across HRTFs.

classes yield the same general pattern of responses.

Negative and positive rate differences (i.e., negative and positive values of d') are presumed to encode spectral cues with equal resolution; therefore, unsigned d' values are plotted in Fig. 7(b) to emphasize the clustering of large d' values at particular BF regions. Changes in both azimuth and eleva-

tion elicited substantial rate differences from fibers with BFs near mid-frequency or high-frequency notches, but in all cases, the maximum d' was conveyed by fibers with BFs above 20 kHz. High-frequency rate information was substantially reduced for the comparison HRTF in the contralateral sound field (-15° AZ, solid line in left panel). Remaining

HRTFs produced maximum d' values of similar magnitude.

Contour plots in Fig. 7(c) provide a graphical representation of the combined effects of stimulus frequency and location on neural representations of HRTF contrasts. These results were obtained by applying the high/medium SR model to spectral differences between a fixed reference HRTF at 0° AZ and transfer functions at each of 10 comparison azimuths in the horizontal plane (left panel), or 16 comparison elevations in the median plane (right panel). Locations of HRTFs in Fig. 7(a) are marked with horizontal lines. Solid filled areas show combinations of frequency and location where d' values fell below -1 ; striped areas delineate d' values above 1. In terms of signal detection theory, fibers in these marked regions convey discriminable rate differences as the source of sound moves from the reference location to the location of the comparison stimulus indicated on the ordinate.

Two patterns of localization cues are obvious for sound sources in the horizontal plane. First, model fibers show decreases in discharge rates for contralateral sound sources ($AZ < 0$) and increases in rate for ipsilateral sources ($AZ > 0$). These rate differences arise from changes in the overall gain of the HRTFs, which can be seen in the left panel of Fig. 7(a). This acoustic variation is the basis for the traditional interaural level difference (ILD) cue. At low frequencies (< 5 kHz), larger changes in azimuth are required to produce rate differences of equal magnitude because HRTF-filtering effects are less pronounced. The second localization cue apparent in Fig. 7(b) is a distinct reversal in the d' profile across a narrow range of frequencies between 9 and 13 kHz (arrow). Fibers tuned to these frequencies show rate decreases even when sounds arise from large ipsilateral azimuths because of the presence of mid-frequency notches in the HRTF spectra. Mid-frequency notches are also present at contralateral azimuths but these features are obscured in this plotting format by negative gains at other frequencies of the comparison HRTF. The region of negative d' values that is consistently seen near 10 kHz for contralateral azimuths is produced when the mid-frequency notch of the reference HRTF falls below the gain of the comparison HRTFs. These filtering effects can lead to situations in which narrow-band sounds arrive at the ipsilateral ear with less intensity than the contralateral ear. Perceptual consequences of this ILD ambiguity will be discussed later.

Over a limited range of frequencies, d' values in the median plane display many of the same properties as those in the horizontal plane. High BF fibers exhibit rate increases ($d' \geq 1$) as the sound source moves to more positive elevations and rate decreases ($d' \leq -1$) for movement toward negative elevations. These changes are not reflected by responses of low BF fibers. Once again, mid-frequency fibers display the effects of spectral notches as a directionally dependent decrease in discharge rates (arrow).

Effects of frequency domain on simulated auditory nerve representations were examined by separating d' values in Fig. 7(b) according to fiber BF. Box plots in Fig. 8(a) compare the resulting d' distributions for a 15° change in azimuth; responses to a 15° change in elevation are shown in Fig. 8(b). Under both stimulus conditions, maximum d' val-

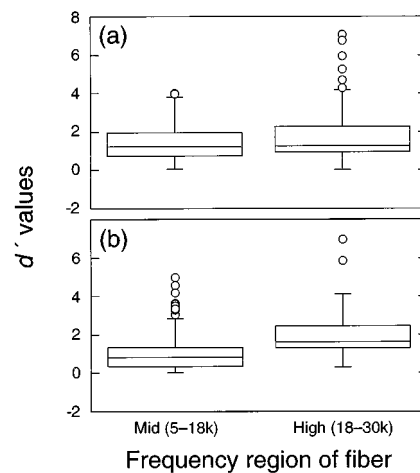


FIG. 8. Effects of frequency domain on the distribution of d' values for a 15° change in azimuth (a) or elevation (b). Boundaries between mid- and high-frequency spectral cues were proposed by Rice *et al.* (1992). Box plot conventions are described in Fig. 6.

ues of high BF fibers attained larger values than responses of mid-frequency fibers (compare upper error bars), which suggests that the most effective directional cues for changes in sound location near 0° AZ, 0° EL are conveyed by high-frequency spectral features [Fig. 7(a)].

B. Localization cues for a reference HRTF at a large ipsilateral azimuth (75° AZ, 0° EL)

Effects of source location on neural representations of pinna-based sound localization cues were investigated by shifting the reference HRTF to a large positive azimuth in the interaural horizontal plane (75° AZ, 0° EL). Representative HRTFs for these contrasts are shown in Fig. 9(a). Mid-frequency notches are still present at these more eccentric locations, but the spectral features do not display the same directional dependency. Notice that there is a mid-frequency notch which moves systematically in azimuth (left panel), but not in elevation (right panel). The overall gain of the HRTFs are nearly identical at lower frequencies, and complex spectral differences are observed at higher frequencies. d' profiles for these HRTF contrasts are shown in Fig. 9(b). Relative to representations of changes in sound source location at central sound locations [Fig. 7(b)], the ipsilateral reference HRTF produced smaller d' values among a more restricted range of BFs.

Figure 9(c) shows d' values for high/medium SR simulations based on the contrast between the ipsilateral reference HRTF and comparison HRTFs at 10 azimuthal locations ranging from -75° to 60° in the horizontal plane (left panel) and 16 elevations along the vertical plane passing through 75° AZ (right panel). Locations of the five HRTFs in Fig. 9(a) are marked by horizontal lines.

The d' plots in Fig. 9(c) suggest changes in azimuth and elevation at large positive azimuths are for the most part represented by rate decreases in the auditory nerve. Comparison HRTFs near the reference location elicited fewer rate differences that exceeded $d' = 1$, and these differences were restricted to BFs above 10 kHz. As the comparison HRTF

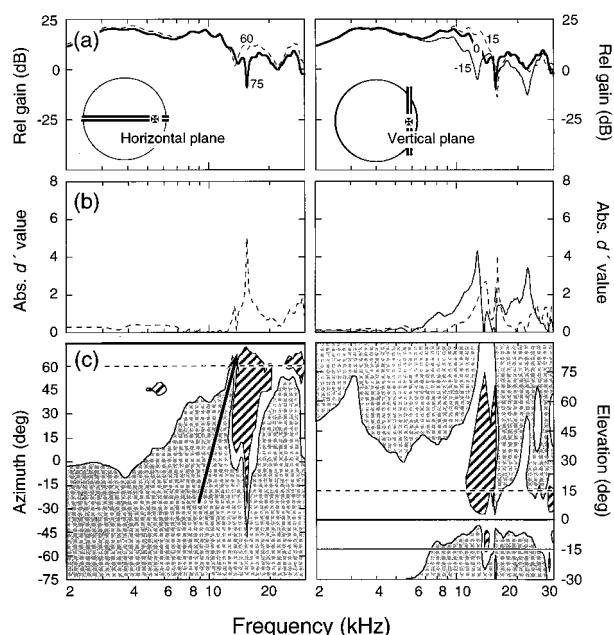


FIG. 9. Neural representation of changes in azimuth (left panels) and elevation (right panels) when the reference HRTF is moved to an ipsilateral azimuth in the horizontal plane (75° AZ, 0° EL). (a) Magnitude spectrum of the reference HRTF with select comparison HRTFs. (b) Simulated d' values for HRTF contrasts in (a). (c) Patterns of rate differences produced by additional combinations of stimulus frequency and location. Plotting conventions are described in Fig. 7.

moved toward smaller positive azimuths (left panel), acoustic shadowing by the head and pinna lowered and the overall gain of the HRTF and discharge rates decreased. The drop in activity reached the criterion for detection ($d' \leq -1$) first among high BF fibers, then at progressively lower BFs as the location of the comparison HRTF passed into the contralateral sound field (AZ < 0). Additional localization cues resulted from the deep mid-frequency notch of the reference HRTF. Since this feature was of lower amplitude than most comparison HRTFs, the change to the comparison location created an island of rate increases among fibers with BFs near 14 kHz (striped area, $d' \geq 1$). Similar patterns of responses were noted as the comparison HRTF moved in elevation away from the interaural horizontal plane (right panel).

C. Localization cues for a reference HRTF at a large positive elevation (0° AZ, 75° EL)

Neural encoding of HRTF contrasts was also measured with the reference location placed at a high elevation in the median plane (0° AZ, 75° EL). As shown in Fig. 10(a), HRTFs do not display mid-frequency spectral notches at these high elevations. In fact, the transfer function of the reference HRTF (bold lines) is essentially identical to comparison HRTFs within $\pm 15^\circ$ AZ (right panel) and EL (left panel). Although spectral differences are seen at high frequencies, the range of frequencies conveying this information is small relative to filtering properties of reference HRTFs in the horizontal plane (Figs. 7 and 9). The minimal spectral differences between HRTFs at high elevations were reflected by relatively low d' values [Fig. 10(b)].

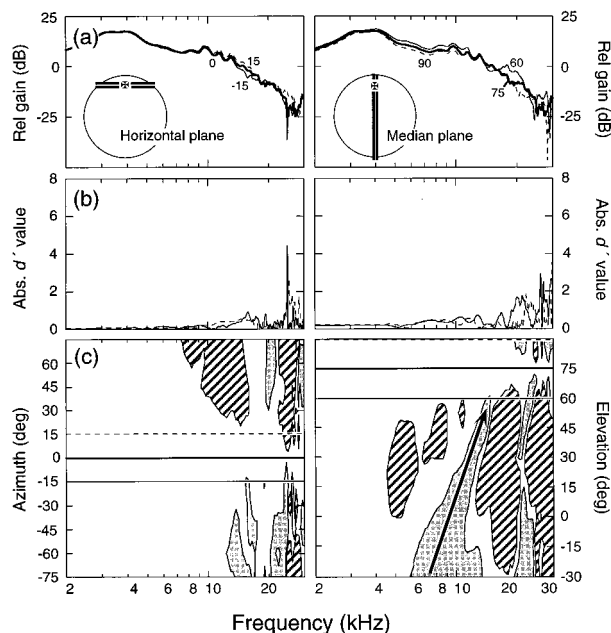


FIG. 10. Neural representation of changes in azimuth (left panels) and elevation (right panels) when the reference HRTF is placed at a high elevation in the median plane (0° AZ, 75° EL). (a) Reference HRTF and nearby comparison HRTFs. (b) d' values obtained by contrasting the reference HRTF with transfer functions in (a). (c) Patterns of rate differences for additional HRTF contrasts with the 0° AZ, 75° EL. See Fig. 7 for plotting details.

The pattern of rate differences for spectral contrasts with both reference and comparison HRTFs in the high horizontal plane is shown in the left panel of Fig. 10(c). Under these stimulus conditions, poor sound localization accuracy is predicted by the relatively limited combinations of frequency and location that elicit rate changes with d' values exceeding ± 1 . Results of contrasts with comparison stimuli in the median plane are presented in the right panel of the figure. Unlike changes in azimuth, the range of BFs showing discriminable rate differences increased considerably as the location of the comparison HRTF dropped below 60° EL. This improvement in the representation of centrally located sound sources was largely due to an increase in overall gain at frequencies above 5 kHz, as well as the presence of mid-frequency notches (arrow). Of course, results obtained for the contrast between the reference HRTF at 75° EL and a comparison stimulus at 0° EL were identical in magnitude but opposite in sign as results obtained when the HRTF at 0° EL served as the reference stimulus (Fig. 7).

The distribution of d' for a 15° change in azimuth is compared at three reference locations in Fig. 11(a); d' scores for a 15° change in elevation are described in Fig. 11(b). Maximum d' values for changes in both azimuth and elevation declined progressively as the reference location moved from a central location, to an ipsilateral azimuth, and then to a high elevation. In part, the single-pole coordinate system that we use to describe HRTF location contributed to the lowered neural discriminability of eccentric sound locations by compressing the actual distance between sound sources at high elevations and more lateralized azimuths (Middlebrooks *et al.*, 1989). This coordinate system has been used in most

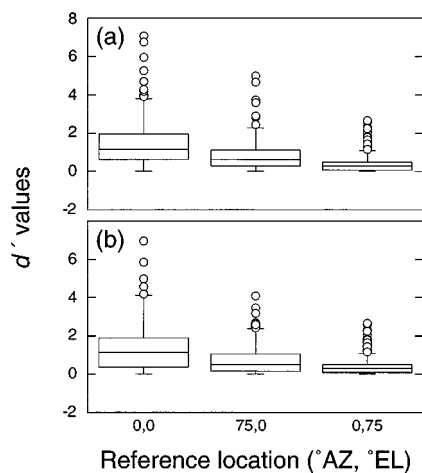


FIG. 11. Effects of reference location on the distribution of d' values for a 15° change in azimuth (a) or elevation (b). Box plot conventions are described in Fig. 6.

psychophysical studies of sound localization, and we retain its plotting conventions to allow comparison of our modeling results with behavioral performances described in other published reports.

D. Spectral cues for identification of sound source location

Thus far, our auditory nerve simulations have been based on spectral contrasts for paired combinations of spe-

cific HRTFs. This analysis offers useful insights into the acoustic cues that contribute to performance in spatial discrimination tasks, like the minimum-audible angle procedure (Mills, 1958), where the stimulus shifts from one location to another. However, directional hearing often involves identification of source position without the advantage of cues that follow a change in sound location. To achieve accurate localization under these conditions, auditory nerve responses must be able to convey patterns of discharge rates that are unique to the location of individual sound sources. In our final simulation, sound source identification was modeled by contrasting the amplitude spectrum of each HRTF to filtering properties that are common to all HRTFs. This analysis is based on the assumption that localizable spectral cues are those that show large acoustic variation relative to general features of the HRTF. Bronkhorst (1995) used a similar approach to investigate patterns of spectral variability in human HRTFs.

The reference HRTF for simulations of sound source identification is indicated by bold lines in the upper panels of Fig. 12. This general transfer function was computed by averaging the dB magnitude spectra of the 179 transfer functions used as comparison stimuli. Dashed lines indicate ± 1 standard deviation (s.d.) from this average. Notice that variability between HRTFs increases progressively from low to high frequencies because of the complex high-frequency spectrum of many HRTFs. The HRTF at 0° AZ, 0° EL is shown in the upper left panel of the figure. Although this

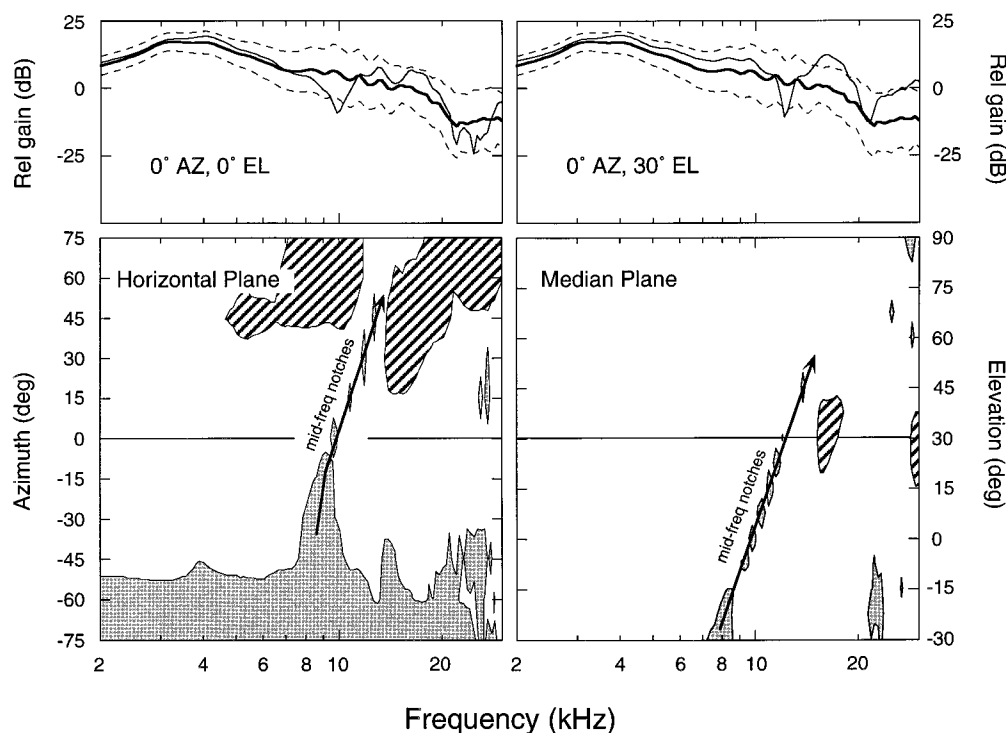


FIG. 12. Neural encoding of spectral cues for the identification of sound sources. Simulated auditory nerve responses to HRTFs in the interaural horizontal plane are shown in the left panels; responses to transfer functions in the median plane are presented in the right panels. Representative HRTFs (solid lines) are compared with the average amplitude of all HRTFs (bold lines) in the upper panels. Dashed lines indicate the standard deviation (s.d.) of amplitude values across HRTFs at each frequency. Contour lines in the lower panels show combinations of frequency and sound source location where spectral contrasts produced rate increases that reached d' values greater than 1 (striped regions) or rate decreases with d' values less than -1 (solid filled regions). Mid-frequency notches were the major source of these rate differences at central locations in the frontal field (arrows).

transfer function follows the sample average at most frequencies, a prominent notch falls below the average gain at 10 kHz by more than 1 s.d. From a statistical perspective, the negative gain of this mid-frequency notch is uncommon among the majority of HRTFs and therefore is a potentially distinguishing localization cue for sound sources at 0° AZ, 0° EL. The cat's HRTF at a slight positive elevation in the median plane (0° AZ, 30° EL) is shown in the upper right panel. This transfer function displays a prominent spectral notch at 13 kHz. At higher frequencies, the HRTF shows two broad amplitude peaks that exceed average gains by more than 1 s.d.

Unique spectral features in the upper panels of Fig. 12 can only be considered useful localization cues if they are capable of eliciting discriminable patterns of activity in the auditory nerve. Neural encoding of these potential localization cues was examined by applying the auditory nerve model to spectral differences between the average transfer function and comparison HRTFs. To account for frequency-dependent changes in variability between HRTFs, difference spectra used in these simulations were computed in terms of the dB gains that extended beyond ± 1 s.d. of average values at each frequency. For example, the HRTF at 0° AZ, 0° EL (upper left panel) produced a difference spectrum of 0 dB at all frequencies but 10 kHz, where the mid-frequency notch produced a negative gain of 5 dB beyond the s.d. boundary. As in previous simulations, the linear model for high/medium SR fibers was used to transform these spectral differences to rate difference profiles, which were then evaluated in terms of d' statistics.

d' values for spectral differences based on the average transfer function are shown in the lower panels of Fig. 12. The main effect of this analysis was to reduce the contribution of level difference cues that distinguish ipsilateral and contralateral azimuths. In contrast with simulations shown in Fig. 7, identification of ipsilateral sound locations in the horizontal plane failed to elicit rate decreases with d' values less than -1 at negative azimuths of less than 45° (left panel of Fig. 12). Positive locations show a similar loss of rate information. Although a narrow range of high-frequency BFs show discriminable rate decreases for azimuths as small as 15°, discharge rates of low BF fibers are relatively unchanged at all positive azimuths. For the most part, rate representations of azimuths in the immediate frontal field (i.e., $\pm 30^\circ$) are limited to negative rate differences that result from mid-frequency spectral notches (arrows). Localization cues derived from mid-frequency notches appear to play an even more important role in the identification of sound source elevation. For sounds in the median vertical plane (right panel of Fig. 12), the only consistent localization cue is the neural representation of a systematic change in the frequency of the mid-frequency notch. No clear representation of spatial differences is seen at higher elevations where HRTFs lack these spectral features (Fig. 10).

III. DISCUSSION

Our auditory nerve simulations predict a decrease in the quality of neural representations for changes in sound location as the reference location moves to more eccentric azi-

muths and elevations (Fig. 11). Spectral cues for directional changes in the frontal field are conveyed by rate information at all frequencies of the auditory nerve fiber array, but fibers with high-frequency BFs often provide the most sensitive representation of small changes in azimuth or elevation (Fig. 8). High-frequency rate information becomes even more significant for encoding small directional changes at large positive azimuths (Fig. 9) or elevations (Fig. 10), where transfer functions exhibit similar low-frequency filter shapes across large spatial regions. In tasks that require absolute identification of sound source location, auditory nerve encoding of mid-frequency spectral notches provides essential directional information (Fig. 12).

The following discussion compares our modeling results with the cat's psychophysical threshold in a variety of sound localization tasks. These comparisons identify stimulus conditions where both neural and behavioral performances demonstrate general trends toward better or worse localization accuracy. More quantitative comparisons are made difficult because the independent behavioral studies discussed in this review may fail to account for other localization strategies such as head or pinna movements, and also may employ different threshold criteria than the signal detection methods used in our modeling efforts.

A. Comparison of neural representations with behavioral performance

Casseday and Neff (1973) measured localization of pure tones using a sound orientation procedure in which cats approached a goal box at the perceived location of the sound source to obtain food rewards. Across stimulus frequencies of 0.25–8 kHz, average thresholds were relatively constant until they climbed abruptly to values approaching 25° for 4-kHz tones. A similar trend toward poor performance has been demonstrated at frequencies between 1 and 2 kHz in humans (Stevens and Newman, 1936) and Old World monkeys (Brown *et al.*, 1978), and is interpreted as a transitional state from low-frequency ITD cues to high-frequency ILD cues. ILD cues are established at higher frequencies in cats than in humans and monkeys because cats have smaller heads. At a test frequency of 8 kHz, thresholds in cats improved relative to performance at 4 kHz, which suggests an increase in the availability of ILD information. This interpretation is supported by modeling results shown in Fig. 7. Notice in this figure how the change in azimuth required to produce a rate difference with d' exceeding ± 1 steadily decreases in the 4–8 kHz frequency region (direct evidence for these frequency effects are discussed later). Although Casseday and Neff do not give details of the cat's localization performance for broadband sounds, they do report average threshold values of less than 5° for noise bursts that served as the initial training stimuli in their behavioral paradigm. Presumably, cats required less training and produced better thresholds when spectral cues for sound localization were available.

A subsequent behavioral study by Martin and Webster (1987) provided direct comparison of sound localization thresholds for broadband noise and pure tones (0.5–32 kHz). Noise thresholds in the horizontal plane were less than 4°,

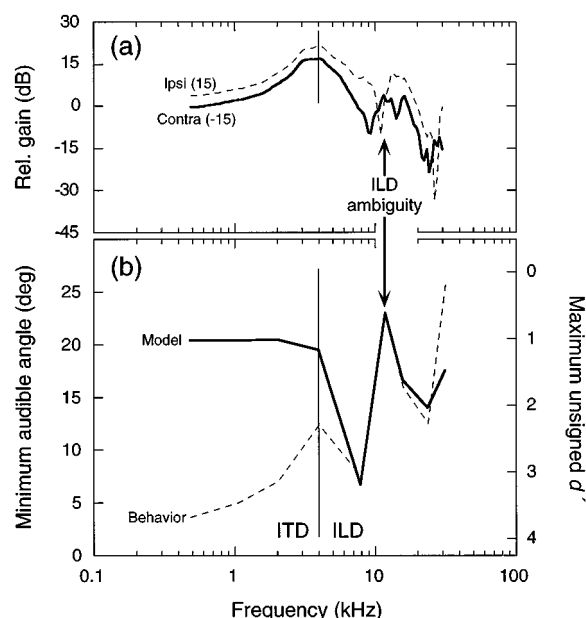


FIG. 13. Localization of pure tones in the horizontal plane. (a) HRTFs for sound sources at $\pm 15^\circ$ AZ. Differences in gain between the HRTFs are the basis of ILD cues for a single ipsilateral sound source at 15° AZ. Although the gain of the transfer function for the ipsilateral ear is generally larger than the contralateral transfer function at frequencies above 4 kHz (vertical line), spectral notches can drop the gain of the ipsilateral function below that of the contralateral ear (arrow). (b) Behavioral and neural discrimination of changes in azimuth for pure tone stimuli. At frequencies above 4 kHz, trends in the behavioral MAA (dashed line) are predicted by changes in neural d' values (bold line). ILD ambiguities that arise from spectral notches can lead to elevated pure-tone thresholds (arrow). (Behavioral data are replotted from Fig. 2 in Martin and Webster, 1987.)

the smallest angular separation possible in their testing apparatus. Average pure-tone MAAs obtained by Martin and Webster followed the same trends as previous results of Casseday and Neff (1973). That is, behavioral performance showed a general decline as a function of increasing stimulus frequency, with discontinuities observed for pure tones at 4 and 12 kHz.

HRTF-filter shapes in Fig. 13(a) offer an acoustic basis for the behavioral observations of Martin and Webster (1987). The two transfer functions represent spectral cues in the ipsilateral and contralateral ear for a sound source at 15° AZ in the horizontal plane. This location has been selected because it falls near the average of all MAAs in the Martin and Webster study. Notice how ILD cues are established by the difference in gain between transfer functions in the two ears. The magnitude of this directional information increases at frequencies above 4 kHz [vertical line in Fig. 13(a)], which may explain why Martin and Webster observed a decrease in behavioral thresholds as the frequency of pure-tone stimuli increased from 4 to 8 kHz [vertical line in Fig. 13(b)]. In spite of the head's sound shadowing effects, spectral notches can create ILD ambiguities in which the level of a narrow range of frequencies is actually louder in the contralateral ear (Irvine, 1987). These ambiguities may lead to poor localization of pure tones. For example, mid-frequency notches around 9–13 kHz [arrow in Fig. 13(a)] correspond to a sharp loss of directional acuity in the Martin and Webster study [arrow in Fig. 13(b)].

Auditory nerve representations of pure-tone ILD cues were estimated by applying the high/medium SR model to the HRTF contrast shown in Fig. 13(a). For this analysis, unsigned d' values were calculated at discrete frequency locations along the difference spectrum, using the same frequencies as those tested by Martin and Webster (1987). The results of these simulations are shown by the bold line in Fig. 13(b). Because maximum d' values are inversely related to the magnitude of the minimum audible angle, neural and behavioral responses are plotted on reversed ordinates in the figure. A sharp increase in d' values at frequencies above 4 kHz supports the assumption that ILD cues begin to exert strong influences on the cat's spatial acuity in this frequency region. Across higher frequencies, the quality of the neural representation is a good predictor of trends in behavioral performance, including high thresholds at 12 kHz.

Threshold differences between broadband noise and pure tones are even more pronounced when ILD cues are minimized by testing in the median plane. Martin and Webster (1987) examined the cat's ability to discriminate changes in elevation using broadband noise and pure-tone frequencies of 2, 8, and 16 kHz. Three of four cats in the vertical MAA task failed to achieve criteria of stable performance when tested with tones. In contrast, these same subjects produced thresholds of less than 4° when tested with noise bursts. The similarity of behavioral thresholds in the horizontal and median planes for broadband stimuli suggests equivalent neural representations for the spectral cues that signal small changes in azimuth and elevation. This interpretation is supported by the distribution of d' values in Fig. 8, which shows that 15° changes in azimuth or elevation both yield a maximum unsigned d' of approximately 7.

Huang and May (1996b) investigated the effects of frequency domain on the cat's sound localization thresholds. Lowest thresholds in the horizontal plane (MAA = 3.1°) were obtained for high-frequency noise bursts (18–50 kHz). MAAs increased to an average of 5.9° when the passbands of noise stimuli were lowered to mid-frequencies (5–18 kHz). High-frequency noise bursts also conveyed more effective localization cues than mid-frequency noise in the median plane. The MAA at positive elevations was 7.3° for high-frequency noise bursts and 11.6° for mid-frequency noise. Modeling results shown in Fig. 8 indicate maximum d' values of mid-frequency fibers are consistently lower than those of high-frequency fibers.

Heffner and Heffner (1988) measured the effects of azimuth on sound localization acuity in the cat. They observed average MAAs of approximately 5.9° for tests with 100-ms noise bursts when sound sources were centered around 0° AZ. The threshold of directional acuity declined to 10.7° when tests were conducted at a positive azimuth of 90° . Human subjects in the Heffner and Heffner study showed even greater sensitivity to changes in azimuth. Average thresholds increased monotonically from a threshold of 1.3° for centrally located sound sources to 9.7° at a reference location of 90° AZ. A loss of discrimination accuracy at eccentric sound locations has also been observed in humans (Mills, 1958) and nonhuman primates (Brown *et al.*, 1982). In much the same manner as behavioral performance, maximum d' val-

ues decline when the reference stimulus moves in the horizontal plane from a central to a more ipsilateral azimuth [Fig. 11(a)].

Huang and May (1996a) measured the cat's ability to orient toward a single brief burst of noise that was presented from randomized locations in the frontal sound field. Unlike the spatial discrimination procedures described above, this task required the identification of actual sound source locations. Cats showed well directed orientation responses toward the source of broadband noise bursts when tested with this task (May and Huang, 1996). This pattern of accurate localization behavior was relatively unaffected by limiting the bandwidth of noise to mid-frequencies where spectral notches provide sound localization cues. In contrast, mid-frequency tones or high-frequency noise bursts produced poorly directed and highly variable responses. In Fig. 12, our analysis of spectral cues for the identification of sound sources suggests a powerful influence of mid-frequency spectral notches at azimuths and elevations less than $\pm 45^\circ$. This "frontal focus" of spectral notches may explain a consistent under-estimation of more eccentrically located sound sources in orientation tasks (May and Huang, 1996; Huang and May, 1996a).

B. Interpreting maximum d' statistics

The predictive power of our sound localization model has been tested by comparing model representations and behavioral performances under a variety of stimulus conditions. For the most part, auditory nerve simulations have been characterized in terms of maximum d' values. This analysis implies behavioral thresholds are closely linked to those small number of fibers that provide the best rate representation of an acoustic event. As sophisticated optimal processing models have been developed for combining rate information across populations of neurons (Siebert, 1968; see Winslow *et al.*, 1987, for a review), it has become clear that performance accuracy predicted by large arrays of auditory nerve fibers far exceeds behavioral capabilities for pure-tone detection (Young and Barta, 1986), intensity discrimination (Delgutte, 1986; Winslow and Sachs, 1988), and frequency discrimination (Siebert, 1968; Barta, 1985). In the context of speech processing, both Conley and Keilson (1995) and May *et al.* (1996) have noted that vowel formant discrimination based on the combined auditory nerve response also overestimates actual behavioral thresholds.

Neural discrimination of a change between two HRTFs can be computed in terms of Q , a statistical method for summing neural d' values across the entire population of auditory nerve fibers (Barta, 1985). In its current form (May *et al.*, 1996), Q is weighted according to the density of fibers along the basilar membrane (Keithley and Schreiber, 1987) and the fraction of total fibers in each SR class (Liberman, 1978). Decision rules developed by Barta define the spatial just-noticeable-difference (JND) as the change in location that yields $Q=0.68$. The smallest change in azimuth (15°) or elevation (7.5°) in our sampling of HRTFs produced Q values approximately 200-fold larger than this criterion. Although our HRTF contrasts involved spatial changes that

were much larger than the MAA, interpolation to the criterion proposed by Barta predicts JNDs in the range of 0.04° . No species has demonstrated directional acuity near this range of accuracy.

C. The dorsal cochlear nucleus: A specialized system for processing spectral notches?

Spectral notches may play a key role in directional hearing if specialized mechanisms exist for processing sharp spectral features. Electrophysiological evidence suggests that DCN neurons may serve in this capacity (Young *et al.*, 1992). Type IV units in the DCN show excitatory responses to broadband sounds but are largely inhibited by pure tones or narrow bands of noise (Young and Brownell, 1976; Spirou and Young, 1991); inhibitory responses are also observed when excitatory input is weakened by placing a narrow spectral notch at frequencies near the type IV unit's BF (Nelken and Young, 1994). Thus, the frequency location of narrow peaks or notches in broadband sounds can be encoded by changes in the discharge rates of type IV neurons relative to responses to stimuli with flat spectra.

Thompson and Masterton (1978) measured the effects of auditory lesions on a cat's reflexive head orientations toward a loud unexpected sound. They found that deep lesions of the lateral lemniscus, which also compromised fibers projecting from the DCN via the dorsal acoustic stria (DAS), eliminated orientation responses to contralateral sounds in the horizontal plane. Sutherland (1991) studied localization of vertical sound sources with similar behavioral paradigms and also found promising results in preliminary DAS-lesioning experiments.

Sutherland (1994) pursued his initial discovery of DAS-lesioning effects using a conditioned suppression task. Cats were trained to avoid a shock by suppressing a licking response when the source of a noise burst changed position along a vertical speaker array. DAS lesions had no effect on the discrimination of large speaker changes ($\geq 15^\circ$) in these experiments. Sutherland's negative findings are predicted by our modeling results which suggest that spectral notches contribute more effective directional cues in the context of absolute sound localization. These conclusions are supported by our own behavioral studies which have used spectral manipulations to investigate the acoustic basis of the cat's sound orientation and spatial discrimination behaviors. Although high-frequency spectral information dominated discrimination between sound sources in a minimum-audible angle task (Huang and May, 1996b), mid-frequency spectral notches were a critical localization cue in testing procedures where cats were required to point out the actual location of sound sources with head orientation responses (May and Huang, 1996; Huang and May 1996a).

Our simulated auditory nerve responses indicate that filtering effects of the head-related transfer function (HRTF) provide straightforward and predictable localization cues for broadband stimuli with flat spectra, but most natural acoustic signals have complex spectral shapes. Is the auditory system capable of using spectral localization cues when the subject has no knowledge of a complex signal's spectral characteristics? Using a computational theory of spectral cue localiza-

tion that was based on human HRTFs (Martens, 1991), Zakarauskas and Cynader (1993) demonstrated that robust monaural localization can be achieved with complex unknown sounds if the listener assumes a "locally constant slope" for the source spectrum. The amplitude spectra of most natural sounds meet this criterion. Under these conditions, effective use of spectral cues in the Zakarauskas and Cynader model depends on the presence of rapidly changing features, such as the mid-frequency notches produced by HRTF-filtering effects.

ACKNOWLEDGMENTS

This research was supported by research Grant No. 5 R29 DC00954 (B. J. May) and training Grant No. 2 P60 DC00979 (E. D. Young) from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health. HRTF spectra and auditory nerve data were generously provided by E. D. Young, as were computational details of the Q analysis. The authors thank E. D. Young, D. B. Moody, and an anonymous reviewer for comments on this manuscript.

- Barta, P. (1985). "Testing stimulus encoding in the auditory nerve," Dissertation Thesis, Johns Hopkins University.
- Blauert, J. (1983). *Spatial Hearing: The Psychophysics of Human Sound Localization* (TUMIT Press, Cambridge, MA).
- Bloom, P. J. (1977). "Creating source elevation illusions by spectral manipulation," *J. Audiol. Eng. Soc.* **25**, 560–565.
- Bronkhorst, A. W. (1995). "Localization of real and virtual sound sources," *J. Acoust. Soc. Am.* **98**, 2542–2553.
- Brown, C. H., Beecher, M. D., Moody, D. B., and Stebbins, W. C. (1978). "Localization of pure tones by Old World monkeys," *J. Acoust. Soc. Am.* **63**, 1484–1492.
- Brown, C. H., Schessler, T., Moody, D. B., and Stebbins, W. C. (1982). "Vertical and horizontal sound localization in primates," *J. Acoust. Soc. Am.* **72**, 1804–1811.
- Butler, R. A., and Belendui, K. (1977). "Spectral cues utilized in localization of sound in the median plane," *J. Acoust. Soc. Am.* **61**, 1804–1811.
- Calford, M. B., and Pettigrew, J. D. (1984). "Frequency dependence of directional amplification at the cat's pinna," *Hear. Res.* **14**, 13–19.
- Casseday, J. H., and Neff, W. D. (1973). "Localization of pure tones," *J. Acoust. Soc. Am.* **54**, 365–372.
- Conley, R. A., and Keilson, S. E. (1995). "Rate representation and discriminability of second formant frequencies for /e/-like steady-state vowels in cat auditory nerve," *J. Acoust. Soc. Am.* **98**, 3223–3234.
- Delgutte, B. (1986). "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 (Nijhoff, Dordrecht), pp. 333–353.
- Evans, E. F. (1975). "Cochlear nerve and cochlear nucleus," in *Handbook of Sensory Physiology: Vol. 5*, edited by W. D. Keidel and W. D. Neff (Springer, Berlin), pp. 1–108.
- Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics* (Krieger, Huntington, NY).
- Hebrank, J., and Wright, D. (1974). "Spectral cues used in localization of sound sources on the median plane," *J. Acoust. Soc. Am.* **56**, 1829–1834.
- Heffner, R. S., and Heffner, H. E. (1988). "Sound localization acuity in the cat: Effect of azimuth, signal duration, and test procedure," *Hear. Res.* **36**, 221–232.
- Huang, A. Y., and May, B. J. (1996a). "Sound orientation behavior in cats. II. Mid-frequency spectral cues for sound localization," *J. Acoust. Soc. Am.* **100**, 1070–1080.
- Huang, A. Y., and May, B. J. (1996b). "Spectral cues for sound localization in cats: Effects of frequency domain on minimum audible angles in the median and horizontal planes," *J. Acoust. Soc. Am.* **100**, 2341–2348.
- Irvine, D. R. F. (1987). "Interaural intensity differences in the cat: Changes in sound pressure level at the two ears associated with azimuthal displacements in the frontal horizontal plane," *Hear. Res.* **26**, 267–286.
- Keithley, E. M., and Schreiber, R. C. (1987). "Frequency map of the spiral ganglion in the cat," *J. Acoust. Soc. Am.* **81**, 1036–1042.
- Le Prell, G. S., Sachs, M. B., and May, B. J. (1996). "Representation of vowel-like spectra by discharge rate responses of individual auditory-nerve fibers," *Aud. Neurosci.* **2**, 275–288.
- Lieberman, M. C. (1978). "Auditory-nerve response from cats raised in a low-noise chamber," *J. Acoust. Soc. Am.* **63**, 442–454.
- Martens, D. L. (1991). "Directional hearing on the frontal plane: Necessary and sufficient spectral cues," Dissertation Thesis, Northwestern University.
- Martin, R. L., and Webster, W. R. (1987). "The auditory spatial acuity of the domestic cat in the interaural horizontal and median vertical planes," *Hear. Res.* **30**, 239–252.
- May, B. J., and Huang, A. Y. (1996). "Sound orientation behavior in cats. I. Localization of broadband noise," *J. Acoust. Soc. Am.* **100**, 1059–1069.
- May, B. J., Huang, A. Y., Le Prell, G. S., and Hienz, R. D. (1996). "Vowel formant frequency discrimination in cats: Comparison of auditory nerve representations and psychophysical thresholds," *Aud. Neurosci.* **3**, 135–162.
- 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.
- Mills, A. W. (1958). "On the minimum audible angle," *J. Acoust. Soc. Am.* **30**, 237–246.
- Musicant, A. D., Chan, J. C. K., and Hind, J. E. (1990). "Direction-dependent spectral properties of cat external ear: New data and cross-species comparisons," *J. Acoust. Soc. Am.* **87**, 757–781.
- Nelken, I., and Young, E. D. (1994). "Two separate inhibitory mechanisms shape the responses of dorsal cochlear nucleus type IV units to narrow-band and wideband stimuli," *J. Neurophysiol.* **71**, 2446–2462.
- Phillips, D. P., Calford, M. B., Pettigrew, J. D., Aitkin, L. M., and Semple, M. N. (1982). "Directionality of sound pressure transformation at the cat's pinna," *Hear. Res.* **8**, 13–28.
- Plenge, G. (1974). "On the differences between localization and lateralization," *J. Acoust. Soc. Am.* **56**, 944–951.
- Poon, P. W. F., and Brugge, J. F. (1993). "Sensitivity of auditory nerve fibers to spectral notches," *J. Neurophysiol.* **70**, 655–666.
- Rice, J. J., May, B. J., Spirou, G. A., and Young, E. D. (1992). "Pinna-based spectral cues for sound localization in cat," *Hear. Res.* **58**, 132–152.
- Rice, J. J., Young, E. D., and Spirou, G. A. (1995). "Auditory nerve encoding of pinna-based spectral cues: Rate representation of high frequency stimuli," *J. Acoust. Soc. Am.* **97**, 1764–1776.
- Sachs, M. B., and Young, E. D. (1979). "Encoding of steady-state vowels in the auditory nerve: Representation in terms of discharge rate," *J. Acoust. Soc. Am.* **66**, 470–479.
- Siebert, W. M. (1968). "Stimulus transformations in the peripheral auditory system," in *Recognizing Patterns*, edited by P. A. Kolars and M. Eden (MIT, Cambridge), pp. 104–133.
- Spirou, G. A., and Young, E. D. (1991). "Organization of the dorsal cochlear nucleus type IV unit response maps and their relationship to activation by bandlimited noise," *J. Neurophysiol.* **66**, 1750–1768.
- Stevens, S. S., and Newman, E. B. (1936). "Localization of actual sources of sound," *Am. J. Psychol.* **48**, 297–306.
- Sutherland, D. P. (1991). "A role of the dorsal cochlear nucleus in the localization of elevation of sound sources," *Abst. Assn. Res. Otolaryngol.* **14**, 33.
- Sutherland, D. P. (1994). "Elevation discrimination in cats: Role of central auditory structures," Dissertation Thesis, Florida State University.
- Thompson, G. C., and Masterton, R. B. (1978). "Brainstem auditory pathways involved in reflexive head orientation to sound," *J. Neurophysiol.* **41**, 1183–1202.
- Tukey, J. W. (1977). *Exploratory Data Analysis* (Addison-Wesley, Reading, MA), p. 688.
- Watkins, A. J. (1978). "Psychoacoustical aspects of synthesized vertical locale cues," *J. Acoust. Soc. Am.* **63**, 1152–1165.
- Wiener, F. M., Pfeiffer, R. R., and Backus, A. S. N. (1966). "On the sound pressure transformation by the head and auditory meatus of the cat," *Acta Oto-Laryngol.* **61**, 255–269.

- Wightman, F. L., and Kistler, D. J. (1989a). "Headphone simulation of free-field listening. I. Stimulus synthesis," J. Acoust. Soc. Am. **85**, 858–867.
- Wightman, F. L., and Kistler, D. J. (1989b). "Headphone simulation of free-field listening. II. Psychophysical validation," J. Acoust. Soc. Am. **85**, 868–878.
- Winslow, R. L., and Sachs, M. B. (1988). "Single-tone intensity discrimination based on auditory-nerve rate responses in backgrounds of quiet, noise, and with stimulation of the crossed olivocochlear bundle," Hear. Res. **35**, 165–190.
- Winslow, R. L., Barta, P. E., and Sachs, M. B. (1987). "Rate coding in the auditory nerve," in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 212–224.
- Young, E. D., and Barta, P. E. (1986). "Rate responses of auditory-nerve fibers to tones in noise near masked threshold," J. Acoust. Soc. Am. **79**, 426–442.
- Young, E. D., and Brownell, W. E. (1976). "Responses to tones and noise of single cells in dorsal cochlear nucleus of unanesthetized cats," J. Neurophysiol. **39**, 282–300.
- Young, E. D., Spirou, G. A., Rice, J. J., and Voigt, H. F. (1992). "Neural organization and responses to complex stimuli in the dorsal cochlear nucleus," Philos. Trans. R. Soc. London, Ser. B **336**, 407–413.
- Zakarauskas, P., and Cynader, M. S. (1993). "A computational theory of spectral cue localization," J. Acoust. Soc. Am. **94**, 1323–1331.

Neural contributions to the perstimulus compound action potential: Implications for measuring the growth of the auditory nerve spike count as a function of stimulus intensity

John R. Doucet^{a)} and Evan M. Relkin

Institute for Sensory Research and Department of Bioengineering and Neuroscience, Syracuse University, Merrill Lane, Syracuse, New York 13244-5290

(Received 22 July 1996; revised 6 January 1997; accepted 10 January 1997)

The perstimulus compound action potential (PCAP), unlike the more familiar compound action potential (CAP), can be recorded in response to *asynchronous* as well as synchronous auditory nerve activity. When all neurons contribute equally to the PCAP, the area under the PCAP (the PCAP area) is proportional to the number of action potentials fired by auditory nerve neurons (the auditory nerve spike count). The auditory nerve spike count is one proposed code for stimulus intensity, and our goal is to use the PCAP to test this hypothesis. In this study, two independent tests were developed to measure the contributions of neurons to the PCAP as a function of their characteristic frequency (CF). The test results were verified using a model of the auditory periphery designed to calculate the auditory nerve spike count as a function of pure tone intensity and frequency. In nearly all experiments, neurons having CFs that span contiguous three or four octave bands contribute equally to the PCAP. For pure tones that stimulate only those neurons contributing equally to the PCAP, the PCAP area grows over intensity ranges frequently exceeding 80 dB, and in one case equaling 108 dB. These results demonstrate that the auditory nerve spike count, at least for pure tones, is capable of encoding changes in stimulus intensity over the entire dynamic range of the auditory system. © 1997 Acoustical Society of America. [S0001-4966(97)05305-8]

PACS numbers: 43.64.Pg, 43.66.Cb, 43.64.Bt [RDF]

INTRODUCTION

Recently, we described a new compound potential, called the perstimulus compound action potential (PCAP), that can be recorded from the chinchilla auditory nerve (Doucet and Relkin, 1995). The fundamental difference between the PCAP and the more familiar compound action potential (CAP) lies in the shape of the voltage waveform, or unit response (UR), contributed by a single action potential to the two compound potentials. The CAP UR is biphasic (e.g., Teas *et al.*, 1962; Kiang *et al.*, 1976), and unsynchronized biphasic URs tend to cancel with one another; therefore, recording a CAP is limited to stimuli (e.g., clicks) that produce synchronized firing by many neurons. The PCAP UR is monophasic (Doucet and Relkin, 1995) and unsynchronized monophasic URs simply summate; as a result, the PCAP can be recorded in response to *asynchronous* as well as synchronous neural activity. Although the potential uses for the PCAP are many, for nearly every application it is important to know whether or not all auditory nerve neurons contribute equally to the PCAP. Consequently, the focus of this study was to measure the contributions of auditory nerve neurons to the PCAP as a function of their characteristic frequencies (CFs).

The importance of equivalent contributions from all neurons is immediately apparent when trying to interpret the growth of the PCAP as a function of stimulus intensity. In a

previous publication (Doucet and Relkin, 1995), we described a simple relationship between the area under the PCAP (the **PCAP area**) and the total number of spikes fired by all neurons (the **auditory nerve spike count**). The auditory nerve spike count is a candidate code for stimulus intensity, and we are interested in using the PCAP to test this hypothesis. If all neurons contribute equally to the PCAP, the proportional relationship between the PCAP area and the auditory nerve spike count can be expressed as follows:

$$\int_{\Delta t} P(t)dt = \int_{\Delta t} F(t)dt \int_{\Delta t} U_p(t)dt. \quad (1)$$

The integral involving $P(t)$ is equal to the PCAP area evaluated over the time interval Δt . This time interval is usually chosen to coincide with the duration of a particular stimulus. $F(t)$ is equal to the instantaneous firing rate of the entire auditory nerve, and its integral over Δt is equal to the auditory nerve spike count. $U_p(t)$ is the PCAP UR, and the area under $U_p(t)$ is equal to a constant that we refer to as the unit area (UA). If all neurons *do not* contribute equally to the PCAP, the growth of the PCAP area would still depend on the auditory nerve spike count, but also on the sizes of the UAs contributed by neurons recruited as stimulus intensity increases. In practice, correcting for this variation in UR size is impossible, and therefore this and most applications for the PCAP require that all neurons responding to a particular sound contribute equally to the PCAP.

Two factors typically result in UR size variation: (1) individual fibers lie at varying distances from the recording electrode and/or (2) axon diameters vary across neurons

^{a)}Current address: Johns Hopkins School of Medicine, Center for Hearing Sciences, 510 Traylor Research Building, 720 Rutland Avenue, Baltimore, MD 21205. Electronic mail: jdoucet@bme.jhu.edu

within the bundle. It is well known that UR size is positively correlated with axon diameter (e.g., Hakansson, 1956). In the auditory nerve, the distribution of axon diameters for the cat (Arnesen and Osen, 1978) and guinea pig (Gleich and Wilson, 1993) is unimodal and extremely narrow. Studies in cat (Liberman and Oliver, 1984) and guinea pig (Gleich and Wilson, 1993) have reported that low-spontaneous-rate (SR) neurons have smaller axon diameters than high-SR neurons. Near the cochlear nucleus (the PCAP recording site), cat low- and high-SR neurons have average axon diameters that differ by less than $0.5\ \mu\text{m}$ (Liberman and Oliver, 1984). In the guinea pig, no statistically significant difference between the diameters of the two SR classes was found (Gleich and Wilson, 1993). Assuming these observations also apply to the chinchilla (our experimental animal), large differences in PCAP UR size are unlikely to be caused by differences in axon diameter.

Auditory nerve neurons that lie at varying distances from the PCAP electrode is a much more likely reason for variation in the PCAP UR. The PCAP is recorded with a glass electrode placed on the surface of the auditory nerve. Since the amplitude of field potentials decreases with distance from the source, axons located deep within the nerve may contribute smaller PCAP URs than those axons near the PCAP electrode. From our own experience in recording from the chinchilla auditory nerve, microelectrode tracks encounter a regular arrangement of characteristic frequencies (CFs), implying that, like the cat (Sando, 1965; Arnesen and Osen, 1978), the chinchilla auditory nerve is tonotopically organized. The orderly spatial arrangement of neurons according to CF, combined with the placement of the PCAP electrode on the surface of the auditory nerve, led us to suspect that the size of the PCAP UR might depend on CF.

The UR contributed by single neurons to a compound potential can be measured directly using spike-triggered averaging (Kiang *et al.*, 1976). However, this technique requires that the PCAP and single neuron activity be recorded simultaneously, an unfeasible requirement at this time given that the PCAP electrode entirely covers the accessible portion of the chinchilla auditory nerve. Instead, two tests were developed to measure the CF dependence of the PCAP UR. For reasons that will become obvious once they are described in Sec. I we named the two measures the *area ratio test* and the *peak ratio test*.

Because of the novelty of the two tests, most of the paper describes experiments designed to validate the ratio tests. Two approaches were used. One was to design the two ratio tests so that they were *independent* measures of the dependence of the PCAP UR on CF. Then, the results of the two tests could be compared for consistency. The second approach involved, in several experiments, performing the ratio tests and recording the growth of the PCAP area as a function of pure tone intensity and frequency. Any systematic variation in the size of PCAP URs suggested by the ratio tests should be reflected in the rate-of-growth of the PCAP area as a function of tone intensity and frequency. The reason is that the contribution of each neuron to the PCAP area is equal to the number of spikes fired by that neuron multiplied by its UA [Eq. (1)]. So for example, a pure tone that at

high intensities excites neurons contributing relatively large URs to the PCAP should result in a more rapid growth of the PCAP area as compared to the growth when all stimulated neurons are contributing equally. These qualitative predictions were made quantitative by constructing a model of the auditory periphery to compute the auditory nerve spike count in response to tones, allowing us to compare the measured and computed growth of the PCAP area as a function of pure tone intensity and frequency. Based on the results presented here, we conclude the following: (1) in most experiments, contiguous CF ranges spanning three or four octaves contribute equivalent URs to the PCAP and (2) the PCAP is an accurate measure of the summed output of the auditory nerve when stimuli are chosen to excite only those neurons contributing equally. Portions of this work have been presented previously in abstract form (Doucet and Relkin, 1990).

I. METHODS

A. General methods

A total of 24 chinchillas, all less than 2 years old, were the experimental animals used in this study. The care and use of the animals were approved by the Syracuse University Institutional Animal Care and Use Committee, which adheres to the guidelines of the Declaration of Helsinki as reflected by compliance with an institutional Letter of Assurance (# A368701) approved by the Office of Protection from Research Risks at NIH (U.S.P.H.S.).

The methods used to record the PCAP have been described in detail previously (Doucet and Relkin, 1995; Doucet, 1995). Briefly, chinchillas were anesthetized initially with an intraperitoneal injection of Dial in urethane ($0.55\ \text{ml/kg}$) and supplemental doses were given to maintain areflexia of the hind paw. After performing initial surgical procedures a sound system was inserted and calibrated, and CAP thresholds were measured for pure tones. Subsequently, the left vestibulocochlear (VIII) nerve was exposed by removing the occipital bone, cutting and removing the dura covering the cerebellum, and aspirating the portion of the cerebellum that overlies the VIII nerve. The brain stem was retracted medially from the temporal bone until we observed the cochlear-vestibular (C–V) cleavage plane, a landmark that approximately divides the VIII nerve into its cochlear and vestibular components.

Recording the PCAP requires the production of a conduction block of auditory nerve activity. To do this, a custom made glass electrode (hereafter referred to as the PCAP electrode) was lowered with a micromanipulator onto the surface of the auditory nerve (using the C–V cleavage plane as a guide). The tip of the PCAP electrode was employed both to apply pressure on the nerve and to record the PCAP. To produce a conduction block of auditory nerve activity, we manually advanced the electrode with the micromanipulator and gradually applied more and more pressure on the nerve. Pressure was increased until a complete conduction block of auditory nerve activity was achieved (our definition of a complete conduction block is described below).

For 14 experiments, the PCAP was recorded differentially between the glass electrode and a silver wire glued to

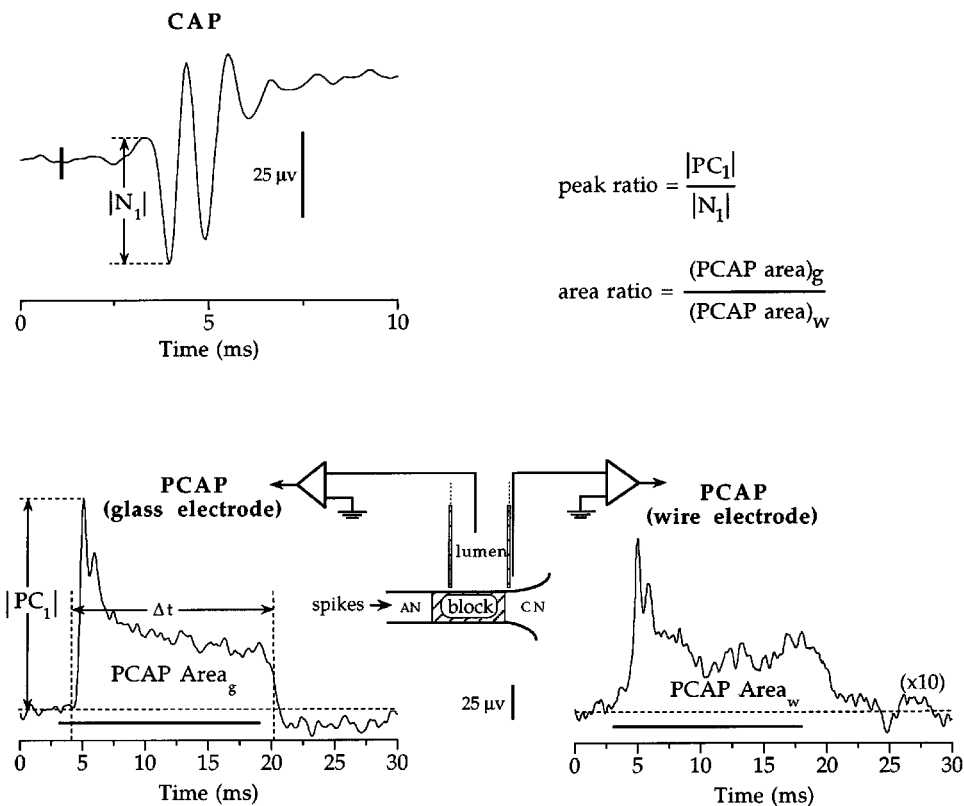


FIG. 1. Definitions for the CAP and PCAP measures used to calculate area and peak ratios. The three waveforms displayed are (1) the CAP recorded at the round window, (2) the PCAP recorded with the glass electrode, and (3) the PCAP recorded with a wire electrode that was glued to the outside of the glass electrode. All potentials were recorded in response to a 4 kHz pure tone. Tone onset is denoted by the thick vertical line for the CAP waveform, and tone duration is represented by the thick horizontal line for the two PCAP waveforms. The schematic displays the position of the glass and wire electrodes with respect to the cochlear nucleus (CN), the conduction block, and the direction of propagated spike activity in the auditory nerve (AN). The time interval (Δt) used to calculate the PCAP area, and its relationship to the PCAP waveform and tone duration (solid black line) is shown for the PCAP recorded with the glass electrode. PCAP areas and the amplitude of PC_1 ($|PC_1|$) were measured with respect to the pre-stimulus baseline (dashed lines). With the filter settings we use, the CAP usually rides on the summing potential (positive potential observed at 10 msec. When the summing potential could be observed prior to N_1 , the amplitude of N_1 ($|N_1|$) was always measured as shown in the figure. Otherwise, $|N_1|$ was measured with respect to the pre-stimulus baseline.

the outside of the PCAP electrode, relative to a vertex ground electrode. The bipolar recording was passed through a differential amplifier (DC–10 kHz; gain, 10), a custom made high-pass filter (cutoff, 0.0016 Hz; gain, 100), an anti-aliasing filter (two 8-pole Bessel filters connected in series; cutoff, 4 kHz; gain, 1), and finally sampled at a rate of 20 kHz. For ten earlier experiments the PCAP was recorded with the glass electrode referenced to the vertex ground electrode, with the bandwidth and gain in the signal pathway exactly the same as described above. We observed little difference when comparing the average amplitude of the PCAP recorded using both electrode configurations; therefore, we pooled data from all 24 experiments. Recordings from the round window (RW) were measured with respect to the vertex electrode and passed through a P511K Grass amplifier (0.01 Hz–10 kHz; gain, 1000), a second anti-aliasing filter (type and settings the same as for the PCAP signal pathway), and sampled at a rate of 20 kHz.

In order to measure the growth of the PCAP as a function of pure tone intensity, it is necessary that there be no noise or distortion product that varies in level with tone intensity. For the frequencies used here, energy at all higher harmonics was at least 70 dB down from the fundamental.

Furthermore, for the range of intensities tested, energy at frequencies lower than the fundamental was not measurable.

B. Criteria for presenting data

All PCAP data presented here met a number of criteria that were described in detail previously (Doucet and Relkin, 1995). Briefly, prior to exposing the auditory nerve, an automated computer-tracking procedure was used to measure CAP thresholds (10 μ V criterion) for tones ranging in frequency from 500 Hz to 22 kHz. All animals had initial CAP thresholds within two standard deviations of averaged normal CAP thresholds from our laboratory. For most experiments, we used the same automated procedure to measure CAP thresholds at the end of data collection and thresholds at all test frequencies had increased by less than 10 dB. The animal occasionally died before we could remeasure CAP thresholds, but for every experiment the amplitudes of N_1 of the CAP and PC_1 of the PCAP (for definitions of both peaks, see Fig. 1) in response to a 50-dB sensation level (SL) click were monitored throughout the experiment. Zero dB SL was equal to the click intensity necessary to evoke a visually detectable N_1 prior to exposing the auditory nerve. The am-

plitudes of N_1 and PC_1 varied by less than 20% with respect to baseline values throughout each experiment.

For the 24 experiments described in this study, the conduction block was defined as complete and stable throughout every experiment. In 16 animals, the block was monitored using the click-evoked brainstem auditory-evoked potential (BAEP). The amplitude of the signal-averaged BAEP was measured in response to a high-intensity click and compared to the amplitude of the residual noise measured when no click was presented. When pressure on the nerve resulted in the amplitudes of the BAEP and residual noise being approximately equal, the conduction block was classified as complete (for further details, see Doucet and Relkin, 1995). For the remaining experiments, we used an alternative method for monitoring the conduction block since setting up and performing the BAEP recordings required a significant amount of time. When the BAEP was abolished, we noticed that the tone-evoked PCAPs for low-intensity tones at all frequencies resembled single neuron PST histograms. The PCAPs rose to a maximum within 10 ms after tone onset (for tones with 2 msec rise/fall times) and remained positive throughout the duration of the tone (see Fig. 4 of Doucet and Relkin, 1995). In 14 animals, whose data are not presented in this report, pressure reduced but did not abolish the click-evoked BAEP. In these 14 experiments, we *always* recorded a PCAP in response to at least one tone frequency that would either dip below the prestimulus baseline during the tone or would reach a maximum well after (10–15 msec) tone onset. Our interpretation of this correspondence between the BAEP and the tone-evoked PCAPs is that the shape of the PCAP reflects whether or not all neurons are blocked and contributing monophasic URs to the PCAP. Since tone-evoked PCAPs having a wide range of frequencies can be examined in less than 5 min, their shape was used to monitor the completeness and stability of the conduction block during experiments where the BAEP was not measured.

C. The ratio tests

Two tests are now described, called the *peak ratio test* and *area ratio test*, that are designed to measure the relative sizes of PCAP URs contributed by neurons as a function of CF. We designed two ratio tests because the interpretation of both depends on assumptions that were not directly verified. However, a key point is that the assumptions associated with the peak ratios are *independent* of those associated with the area ratios. Therefore, one strategy used to validate the area and peak ratios was to compare them for consistency. If they agree, then the most parsimonious explanation would be that the assumptions underlying each test are reasonable.

1. The peak ratio

PC_1 and N_1 (for definition, see Fig. 1) are produced by the synchronous firing of many neurons in response to an acoustic transient. In order to ease the following explanation of the peak ratio, assume that neurons responding to a low-intensity tone fire in *perfect* synchrony at tone onset. Then, the amplitudes of N_1 and PC_1 are simply related to the amplitudes of the CAP and PCAP URs as shown below:

$$|PC_1(f)| = N(f)|PC_0(f)|, \quad (2)$$

$$|N_1(f)| = N(f)|N_0|. \quad (3)$$

Here, $|PC_1(f)|$ and $|N_1(f)|$ represent the amplitudes of PC_1 and N_1 , respectively, while f is the tone frequency. In Fig. 1, we demonstrate how the amplitudes of the two peaks were measured. $N(f)$ represents the number of neurons stimulated by the tone. A crucial concept is that for a low-intensity tone, the neurons contributing to N_1 and PC_1 have CFs equal or similar to the tone frequency (Dallos and Cheatham, 1976; Ozdamar and Dallos, 1978). Therefore, as tone frequency is changed from 16 kHz, to 8 kHz, etc.; this has the desired effect of sampling the synchronized URs of neurons having CFs equal to or near 16 kHz, 8 kHz, etc.

Following the convention of Kiang *et al.* (1976), $|N_0|$ represents the amplitude of the first negative peak of the CAP UR. $|N_0|$ is independent of f (i.e., CF) because spike-triggered averaging of the CAP in cats (Kiang *et al.*, 1976; Wang, 1979) and guinea pigs (Versnel *et al.*, 1992) has demonstrated that all neurons make approximately equal contributions to the CAP. Since the cochlear anatomy of the chinchilla is similar to the cat and guinea pig, we assume that $|N_0|$ is independent of CF for the chinchilla as well. The PCAP UR is monophasic and positive (Doucet and Relkin, 1995). $|PC_0(f)|$ represents the dependence of the positive peak amplitude of the PCAP UR on f (i.e., CF), and is the function we would like to measure.

$|PC_0(f)|$ can be measured by computing the *peak ratio* that is defined in Eq. (4),

$$\text{peak ratio} = \frac{|PC_1(f)|}{|N_1(f)|}, \quad (4)$$

since substituting Eqs. (2) and (3) into Eq. (4) gives

$$\text{peak ratio} = \frac{N(f)|PC_0(f)|}{N(f)|N_0|}, \quad (5)$$

and simplifying and rearranging terms allows us to write

$$|PC_0(f)| = (\text{peak ratio})|N_0|. \quad (6)$$

In Eq. (4), the peak ratio is defined as the ratio of the amplitudes of PC_1 and N_1 measured in response to a pure tone. From Eq. (6), plotting the peak ratio as a function of tone frequency (f) reveals the dependence of the PCAP UR amplitude on f (i.e., CF). Notice that only the *relative* contributions of different neurons to the PCAP are obtained because of the unknown constant $|N_0|$.

Peak ratios as a function of tone frequency were collected in 16 experiments. The PCAPs and CAPs were recorded in response to 20 dB SL tones having frequencies equal to 16, 8, 4, 2, 1, and 0.5 kHz. The SLs are referenced to pure-tone CAP thresholds measured at the beginning of the experiment. In terms of dB SPL, tone intensities averaged 35 dB for 1–8-kHz tones, 40 dB for the 0.5-kHz tones, and 50 dB for the 16-kHz tones. Tones were either 20 or 34 msec in duration (onset to offset), shaped by \cos^2 functions with rise/fall times of 2 msec, and presented once every 503 msec. The PCAPs and CAPs were sampled at a rate of 20 kHz for 100 msec including a 17 msec prestimulus baseline. Responses to at least 32 tone bursts were averaged.

2. The area ratio

The area ratio was computed using tone-evoked PCAPs measured for two monopolar electrode configurations: (1) the glass electrode referenced to ground and (2) the silver-wire electrode that is glued to the glass electrode, referenced to ground. Recall that the glass and wire electrodes are used to record the PCAP differentially. The relationship of the two electrodes with respect to the auditory nerve and the conduction block is shown in Fig. 1. The drawing is highly schematic, but its only purpose is to display that the wire electrode was attached to the side of the PCAP electrode furthest from the active portion of the auditory nerve. The PCAP can be recorded using the wire electrode (Fig. 1); however, in all experiments it is at least a factor of 10 smaller than the PCAP measured with the glass electrode. This size difference is at least partially due to the greater distance of the wire electrode from the active portion of auditory nerve neurons.

Assuming that the wire electrode is far enough away that the active zones of auditory nerve neurons can all be treated as the same point in space; then the PCAP UR contributed at the *wire* electrode will be independent of CF. Combining this assumption with Eq. (1) from the Introduction, the PCAP area measured with the glass and wire electrode in response to a low-intensity tone can be expressed as follows:

$$(\text{PCAP area})_g = C(f)UA_g(f), \quad (7)$$

$$(\text{PCAP area})_w = C(f)UA_w. \quad (8)$$

The subscripts *w* and *g* denote the wire and glass electrodes, respectively. $C(f)$ represents the driven auditory nerve spike count as a function of tone frequency (*f*). Here UA denotes the unit area, the area under the PCAP UR, and $UA_g(f)$ represents the dependence of the UAs contributed at the glass electrode on *f* (i.e., CF), and is the function we would like to measure.

$UA_g(f)$ can be measured by computing the *area ratio* that is defined below;

$$\text{area ratio} = \frac{(\text{PCAP area})_g}{(\text{PCAP area})_w}, \quad (9)$$

since substituting Eqs. (7) and (8) into Eq. (9) gives

$$\text{area ratio} = \frac{C(f)UA_g(f)}{C(f)UA_w}, \quad (10)$$

and simplifying and rearranging terms allows us to write

$$UA_g(f) = (\text{area ratio})UA_w. \quad (11)$$

In Eq. (9), the area ratio is the ratio of the PCAP areas measured with the glass and wire electrodes in response to a pure tone. The PCAP area is defined in Fig. 1 as the area beneath the PCAP waveform with respect to the prestimulus baseline. Details of how the PCAP area was calculated are given shortly. From Eq. (11), plotting the area ratio as a function of tone frequency reveals the dependence of the PCAP UA on *f* (i.e., CF). Similar to the peak ratios, notice that only the *relative* contributions of different neurons to the PCAP are measured because of the unknown constant UA_w .

D. Recording the PCAP as a function of tone intensity

In 15 of the 24 experiments, PCAPs were recorded as a function of intensity for at least one tone frequency. Tone intensity was increased in steps of 5 dB and other stimulus and recording parameters have already been described. The growth of the PCAP was quantified by calculating the PCAP area. The lower time limit for the area calculation was equal to the onset of the electrical gating signal +1 msec (to account for neural latency). The upper time limit was equal to the lower time limit + the tone duration. The relationship of this time window with respect to the duration of the tone and the PCAP is shown in Fig. 1. The PCAP area was calculated in the following manner: (1) The average value of the pre-stimulus baseline (typically a 17-ms window) was subtracted from all the sampled voltages; (2) a straight line was fit between pairs of sampled voltages lying in the time window defined above; (3) the area under each line segment was calculated; and (4) all the areas were summed to equal the PCAP area.

II. RESULTS

A. Peak ratios and area ratios

In order to facilitate comparisons across experiments, the ratios measured in each experiment were normalized to the average value of the ratios at all frequencies (area and peak ratios were normalized separately). As a result, normalized ratios that equal one for all frequencies would imply that all auditory nerve neurons contribute equally to the PCAP. This would be the ideal recording situation for most envisioned applications of the PCAP. To compare the ratios with this ideal result, a broken horizontal line is positioned at a value of one when ratios are presented in Figs. 2–4.

In Fig. 2, normalized peak and area ratios obtained from all 24 experiments are plotted versus frequency. Ratios from individual experiments are connected by dotted lines. Clearly, for both ratio tests, there is a tendency for low-frequency ratios to be smaller than high-frequency ratios. Ratios from two experiments that exhibit this trend are highlighted (closed circles, solid lines). The largest and smallest normalized ratio in a particular experiment typically differed by a factor between 2 and 5, but factors as large as 11 were observed. In a few experiments, ratios were measured for half-octave steps in tone frequency and we observed no variation in the ratios that was not captured with the octave spacing of tone frequency. Occasionally, the normalized ratios were approximately independent of frequency (e.g., open circles, solid lines). If the description “ratios are independent of frequency” is defined as all normalized ratios being between 0.75 and 1.25, 4 of the 24 experiments meet this criterion.

In Fig. 3, we compare peak and area ratios when both were measured in the same experiment over at least a five-octave range. The normalized peak and area ratios have approximately the same dependence on frequency in each of the three experiments, as would be expected if they were *both* reflecting the relative size of the PCAP UR as a function of CF.

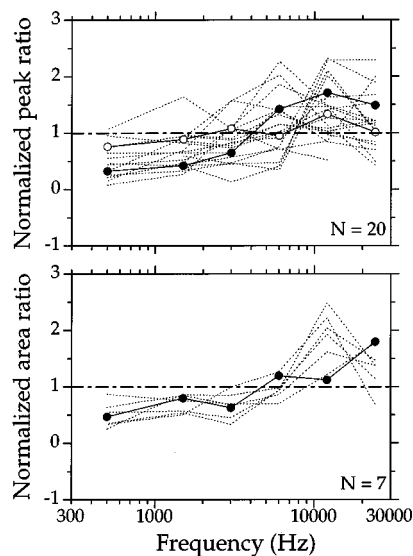


FIG. 2. Plot of the dependence of the area and peak ratios on frequency for all 24 experiments. Normalized ratios from individual experiments are connected by dashed lines except for three examples that are discussed in the text. In any particular experiment, normalized ratios that equaled one for all frequencies (denoted by broken horizontal line) would indicate that all neurons contribute equally to the PCAP. Notice that for both area and peak ratios, low frequency ratios are almost always smaller than high frequency ratios, suggesting that low CF auditory nerve neurons typically contribute smaller URs than high CF neurons.

1. Consequences of the normalized ratios dependence on frequency

Our observation that low-frequency ratios are consistently smaller than high-frequency ratios implies that low-CF neurons usually contribute smaller URs to the PCAP relative to high-CF neurons. There are alternative hypotheses besides variation in UR size that could explain or contribute to the observed dependence of either the area or peak ratios on tone frequency. However, since the two tests are *independent*, the similar dependence of the ratios on frequency for the two tests (Fig. 3) is evidence that variation in UR size is the dominant factor influencing the frequency dependence of the ratios. If all neurons do not contribute equally to the PCAP, does this preclude using the PCAP as a measure of the auditory nerve spike count? We argue that the answer to this question is no, for the following reason. As long as a contiguous *subset* of auditory nerve neurons contributes equally to the PCAP, and stimuli are chosen to stimulate only that subset of neurons, then the PCAP area is still linearly related to the auditory nerve spike count.

In most experiments, the ratios were approximately constant over contiguous three or four octave bands, suggesting that neurons having CFs within these bands contribute equally to the PCAP. We tested this interpretation of the ratios by measuring the growth of the PCAP area as a function of tone intensity and frequency. For those tone frequencies that stimulate neurons contributing equally according to the ratio tests, the PCAP area should grow in proportion to the auditory nerve spike count with increasing tone intensity [see Eq. (1)]. On the other hand, for tones that excite neurons contributing unequally to the PCAP, the rate-of-growth of the PCAP area with increasing intensity will depend on

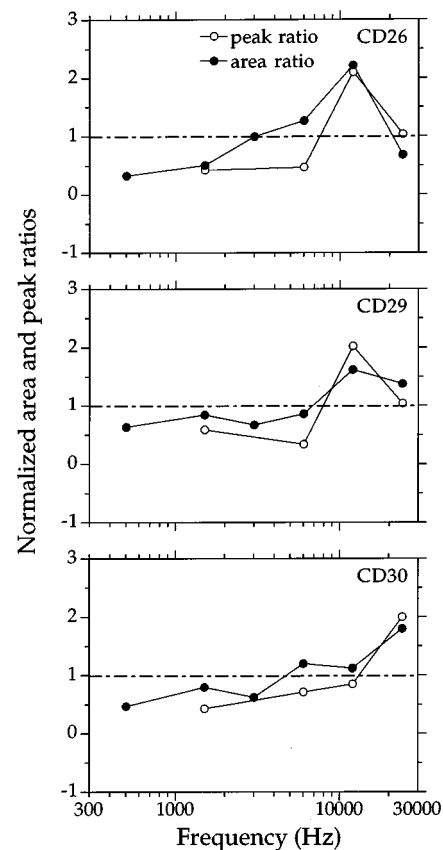


FIG. 3. Comparison between normalized area and peak ratios measured in the same experiment. Data from three animals are shown. In any particular experiment the normalized area and peak ratios usually are not equal, but they both show the same general dependence on frequency over a five octave range. Since independent measures are used to compute the area and peak ratios (see Fig. 1), their similar frequency dependence supports the idea that both ratio tests reflect the dependence of PCAP UR size on CF.

whether or not off-frequency neurons (i.e., neurons not tuned to the tone frequency) contribute larger or smaller UAs to the PCAP relative to on-frequency neurons (i.e., neurons tuned to the tone frequency). In other words, we now further examine the validity of the ratio test results by testing the idea that the variability in the growth of the PCAP area with tone intensity and frequency should be accounted for by the dependence of the ratios on tone frequency.

B. The growth of the PCAP area as a function of pure tone intensity and frequency: Dependence on ratios

In general, the ratios were an excellent predictor of differences in the shape of PCAP area-intensity functions observed across experiments. For example, in Fig. 4 the ratios and PCAP area-intensity functions measured for three experiments are shown. Common symbols across panels signify data obtained from the same experiment. To facilitate comparisons across experiments all PCAP areas were normalized to their values at 10 dB SL, and 0 dB SL refers to the CAP threshold measured at each tone frequency. For a pure tone PCAP area-intensity function, if off-frequency neurons have larger UAs with respect to on-frequency neurons, we expect a more rapid growth of the PCAP area at

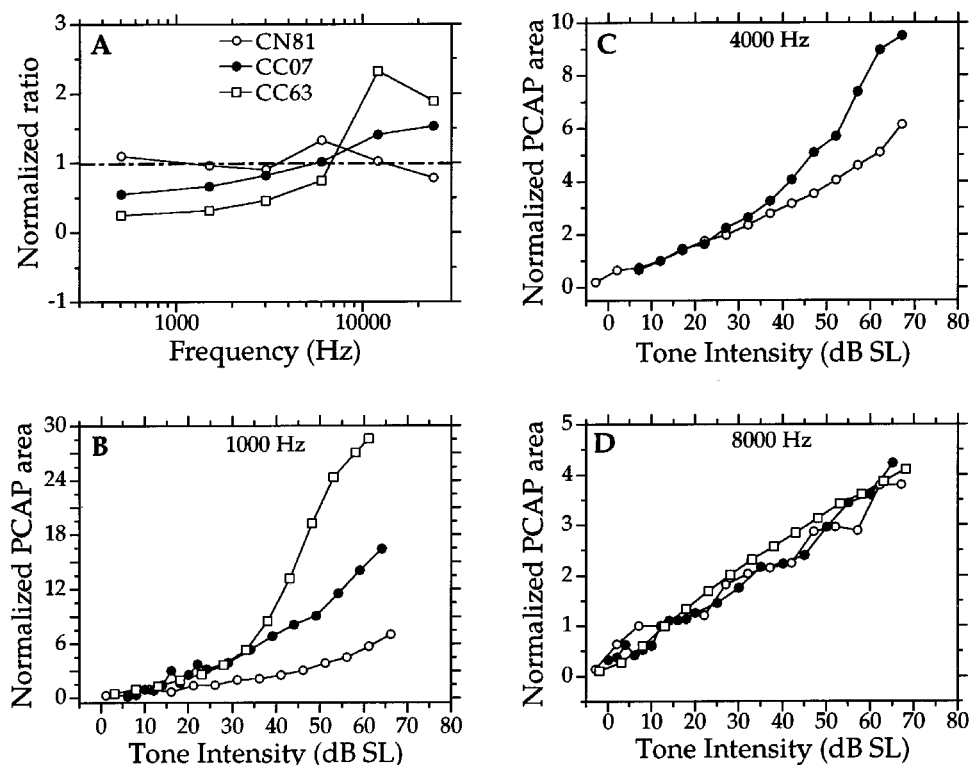


FIG. 4. Demonstration of the good correspondence between the frequency dependence of the normalized ratios and the rate-of-growth of the PCAP area at high intensities. Common symbols across all four panels denote data obtained from the same animal. (A) normalized ratios plotted versus frequency for three experiments and (B) PCAP area-intensity functions for 1-kHz tones. For each function, PCAP areas were normalized to their value at 10 dB SL, and 0 dB SL refers to the 1-kHz CAP threshold measured in that particular experiment. (C) Same as (B), only for 4-kHz tones; (D) same as (B), only for 8-kHz tones. Unlike CN81, the ratios for animals CC63 and CC07 indicate that the contributions of off-frequency neurons for 1 and 4 kHz tones are overemphasized when compared to on-frequency neurons. Notice that, for CC63 and CC07, the growth of the normalized PCAP area for 1-kHz tones is steeper than that observed for CN81. The same holds true for 4-kHz tones when CC07 is compared with CN81 [panel (C)]. In contrast, the normalized ratios at high frequencies are comparable within each experiment, and the rate-of-growth of the PCAP for 8-kHz tones is almost identical for the three cases.

high intensities than if all neurons contribute equally to the PCAP. This expectation is based on spread of excitation to off-frequency neurons at high intensities, and that the contribution of a single neuron to the PCAP is equal to its UA multiplied by its spike count. This prediction is borne out by the data in Fig. 4, where the rate-of-growth of the PCAP area for 1- and 4-kHz tones is steepest in those experiments where the contributions of high-CF neurons appear to be overemphasized (CC63 and CC07). In contrast, within each of the three experiments, the ratios computed for 8- and 16-kHz tones indicate that most neurons excited by an 8-kHz tone contribute nearly equally to the PCAP. Notice that the 8-kHz PCAP area-intensity functions from all three experiments grow at a similar rate.

We wanted to quantify comparisons between the ratios and the shape of pure tone PCAP area-intensity functions. Therefore, the ratios measured in any given experiment were reduced to a single-valued parameter that we call **Weight**. The equation used to calculate Weight was dependent on tone frequency. The purpose of calculating Weight was to capture with a single number the weighting (UR sizes) of off-frequency neurons relative to on-frequency neurons. Explaining how Weight was calculated is best accomplished using examples. The Weight equation for a 1-kHz tone is given below:

$$\text{Weight}_{1k} = \frac{\sum_{i=0}^5 (R_i / R_1) N_i}{\sum_{i=0}^5 N_i}, \quad (12)$$

R_i is the ratio calculated for the i th band, where the six ratios were designated as representing the contributions of neurons having CFs within six frequency bands indexed from 0 to 5 (i.e., 500-Hz ratio represents 0–1 kHz band = R_0 ; 1-kHz ratio represents 1–2-kHz band = R_1 , ..., 16-kHz ratio represents 16–32-kHz band = R_5). N_i is the number of neurons in the i th band and was computed using a chinchilla frequency-place map (Eldredge *et al.*, 1981) and chinchilla neural-innervation densities (Bohne *et al.*, 1982). In words, Weight_{1k} is calculated by dividing all ratios (contributions from on- and off-frequency neurons) by the 1-kHz ratio (contributions from on-frequency neurons). The result for each frequency band is then multiplied by the number of neurons in the band, the results are summed, and then divided by the total number of neurons. An equation analogous to Eq. (12), in that all ratios were included in the Weight equation, was used to calculate Weight for tone frequencies <3 kHz. For tone frequencies ≥ 3 kHz, ratios from lower frequency bands were dropped from the Weight equation, since the steep high-frequency slopes of single neuron tuning curves prevent low-CF neurons from making a significant

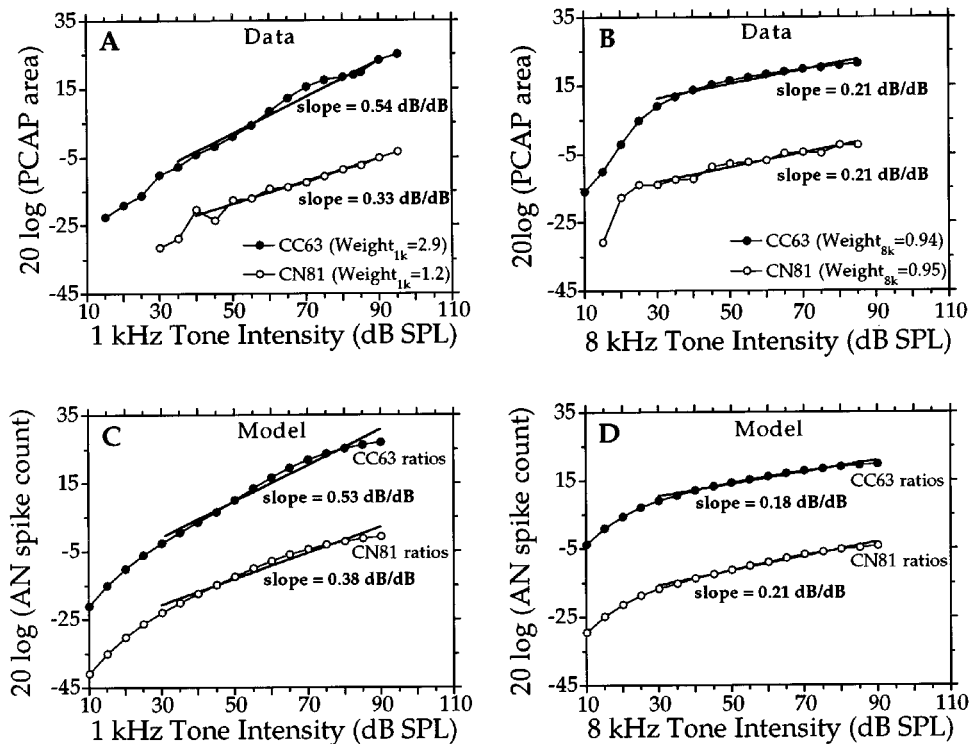


FIG. 5. Comparisons between Weight and the growth of the PCAP area at high intensities are shown for two experiments in panels (A) and (B). In panels (C) and (D), the computed growth of the auditory nerve spike count using a model of the auditory periphery is shown. In all four panels, both data and model results are plotted on log-log axes and the functions are fit by a straight line above 10 dB SL for the data, and between 30 and 90 dB SPL for the model. We refer to the slope of this line here and in the text as the high-intensity slope. (A) Weight_{1k} and PCAP area-intensity functions for 1-kHz tones are shown. Weight_{1k} for CC63 is nearly triple that for CN81, indicating that the contributions of off-frequency neurons for 1-kHz tones are overemphasized in CC63. Notice that the 1-kHz high-intensity slope for CC63 is nearly double that measured for CN81 (B) Weight_{8k} and PCAP area-intensity functions for 8-kHz tones are shown. For this tone frequency, Weight_{8k} is near one for both animals and their 8-kHz high-intensity slopes are identical. (C) Model output for a 1-kHz tone, after the ratios for CC63 and CN81 were incorporated into the model to simulate the effect of unequal contributions by different neurons to the PCAP (see text for details). (D) Same as in (C), only for an 8-kHz tone. Notice the close quantitative correspondence between the model and data slopes, implying that the ratios accurately reflect differences in the contributions of neurons to the PCAP as a function of their CF.

contribution to the PCAP at any tone intensity. For example, the equation for Weight_{8k} is given below:

$$\text{Weight}_{8k} = \frac{\sum_{i=4}^5 (R_i/R_4) N_i}{\sum_{i=4}^5 N_i}. \quad (13)$$

Notice that only the 8- and 16-kHz ratios are used to calculate Weight_{8k} .

If Weight equals 1 for any tone frequency, then the ratios indicate that on- and off-frequency neurons contribute equally to the PCAP, and therefore the PCAP area should grow in proportion to the auditory nerve spike count. If Weight is greater than 1 for a particular tone frequency, then the ratios imply that off-frequency neurons tend to contribute larger URs to the PCAP relative to on-frequency neurons, and the PCAP area-intensity function should grow more rapidly than the auditory nerve spike count. A similar argument leads to the prediction that Weights less than 1 should be associated with a slower growth of the PCAP area. Although these predictions are strictly valid only when the ratios are a monotonic function of tone frequency, in practice, this is true for the majority of our experiments (see Fig. 3).

In Fig. 5, we demonstrate how changes in Weight across different experiments correlated with the rate-of-growth of the PCAP area at high intensities. In the figure, 1- and 8-kHz PCAP area-intensity functions measured during two experi-

ments are shown in panels (a) and (b). The ratios for these two animals (CC63 and CN81) were presented in Fig. 4. The PCAP area-intensity functions are plotted on log-log axes, and they have been arbitrarily shifted along the y axis for illustration purposes. When plotted this way, PCAP area-intensity functions are well fit (R values greater than 0.94 for all fits in all experiments) by a straight line above 10 dB SL. We will hereafter refer to the slope of this line as the **high-intensity slope**. Weight_{8k} is very close to one in both experiments (0.94 for CC63 and 0.95 for CN81), and the 8-kHz high-intensity slopes for the two experiments are identical. For CC63, Weight_{1k} equals 2.9, whereas for CN81, Weight_{1k} equals 1.2. In accord with our predictions that Weights greater than 1 should correlate with steeper high-intensity slopes, the 1-kHz high-intensity slope for CC63 is nearly double that for CN81.

In order to assess if the dependence of the high-intensity slopes on Weight was reasonable, we constructed a model of the auditory periphery to calculate the auditory nerve spike count as a function of tone frequency and intensity. The model is described in detail in the Appendix, but the important characteristic of the model for this discussion is that the spike count was computed separately for 24 $\frac{1}{3}$ -oct CF bands. By partitioning the calculation in this manner, the effect of unequal weighting of different CF bands can be simulated by

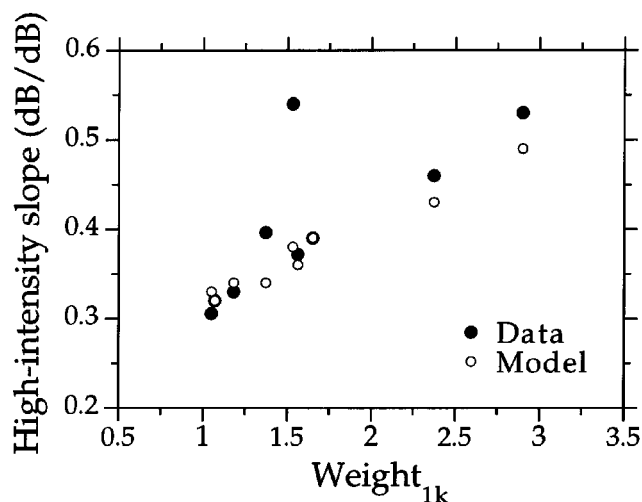


FIG. 6. Scatter plot illustrating the dependence of 1-kHz high-intensity slopes on Weight_{1k} . The closed circles show data obtained from nine experiments. Notice that except for one experiment there is a strong positive correlation between the high-intensity slope and Weight_{1k} ($R = 0.98$ if outlier is ignored), as is expected if the ratios indeed reflect the dependence of the PCAP UR on CF. The ratios from each of the nine PCAP experiments were, one at a time, incorporated into the model calculations to simulate unequal contributions to the PCAP (for more details of how this was done, see text). The open circles represent the high-intensity slopes computed for each set of ratios. The agreement between the model and data slopes is quite good, indicating that the ratios are quantitatively accurate in depicting differences in PCAP UR size for neurons differing in CF.

multiplying the spike counts for each band by different constants. Since the proposal being tested here is that the ratios measure the dependence of the UA on CF, the ratios for CC63 and CN81 were used to multiply the spike counts calculated for the appropriate model frequency bands prior to computing the auditory nerve spike count. Model calculations using the CN81 and CC63 ratios will be referred to, respectively, as the CN81 model and the CC63 model.

The CN81 and CC63 model results are displayed in panels (c) and (d) of Fig. 5. For a 1-kHz tone, not only do the model slopes in panel (c) show the same dependence on Weight_{1k} as the data slopes in panel (a), but each model slope is in close quantitative agreement with the data slopes as well. This near match between data and model is found when comparing the slopes for 8-kHz tones in panels (b) and (d). We find this correspondence between the model and data high-intensity slopes evidence in support of our interpretation of the ratios, since even though individual animals may differ with respect to numerous parameters that could affect the growth of the auditory nerve spike count, all model parameters were held constant. The only difference between the two animals incorporated into the model was the ratios and therefore, presumably, the relative size of the PCAP UAs contributed by neurons differing in CF.

In Fig. 6, the same comparisons described above were performed using data pooled from several experiments. In the figure, high-intensity slopes for 1-kHz PCAP area-intensity functions are plotted versus Weight_{1k} for nine experiments (closed circles). Also shown are the high-intensity slopes for the model (open circles), calculated using the nine sets of ratios obtained from the individual experiments. Only data for 1-kHz tones are presented because this is the sole

frequency where we measured PCAP area-intensity functions in several experiments and Weight_{1k} varied enough to make comparisons instructive. For eight out of the nine experiments the model and data high-intensity slopes were in close quantitative agreement, as long as the ratios for each experiment were incorporated into the model to weight the spike counts contributed by each CF band.

The data in Fig. 6 were employed to define acceptable variations in Weight (i.e., UR size) that would still allow the PCAP to be used as an accurate measure of the summed output of the auditory nerve. Boundary values for Weight were set at 0.75 and 1.25. These limits were chosen based on the three animals that had Weight_{1k} within these bounds (see Fig. 6) and also had 1-kHz high-intensity slopes that differed by less than 10%. Also, model calculations for 1-, 4-, and 8-kHz tones predict that this variation in Weight , tested using several different combinations of UR size variation across the different frequency bands, would result in less than a 15% change in the high-intensity slope for these tone frequencies.

C. PCAP area-intensity functions that pass the Weight criterion

At the highest intensities tested, which many times were near or above 100 dB SPL at all tone frequencies, we always observed that the PCAP area continued to grow with increasing intensity. In two experiments, nearly all neurons were contributing equally to the PCAP according to the Weight criterion defined above. For these two experiments, we plot the absolute value of the PCAP area versus dB SPL for several tone frequencies (CN81 and CN75 in Fig. 7). For almost all frequencies, notice that the PCAP area grows over at least a 70-dB range of intensity. In several experiments, we observed that the PCAP area grew over intensity ranges greater than 90 dB, and in one experiment for a 3-kHz tone this range equaled 108 dB.

Clearly, the absolute value and the rate-of-growth of the PCAP area depend on tone frequency. For example, low-frequency tones produce much larger PCAP areas at high intensities than do high-frequency tones. In addition, for frequencies ≤ 4 kHz, the growth of the PCAP area steepens at high intensities whereas for 8 and 16-kHz tones it grows at a nearly constant rate over the entire range of intensities tested. In the bottom panel of Fig. 7, we show that these differences in the PCAP area-intensity functions for high- and low-frequency tones are true for the model output as well, and in Sec. III we explain that the dependence of the rate-of-growth and the absolute value of the PCAP area on tone frequency and intensity are consistent with the skewed patterns of spread of excitation produced by pure tones in the auditory nerve (Kim and Molnar, 1979; Evans, 1981).

Experiments like CN81 and CN75 where *all* neurons appeared to be contributing equally to the PCAP were extremely rare. However, in almost all experiments, a contiguous *subset* of auditory nerve neurons spanning three or four octaves was contributing equally to the PCAP according to the Weight criterion. Earlier, we had suggested that even in those experiments where all neurons do not contribute equally to the PCAP, stimuli that excite only those neurons

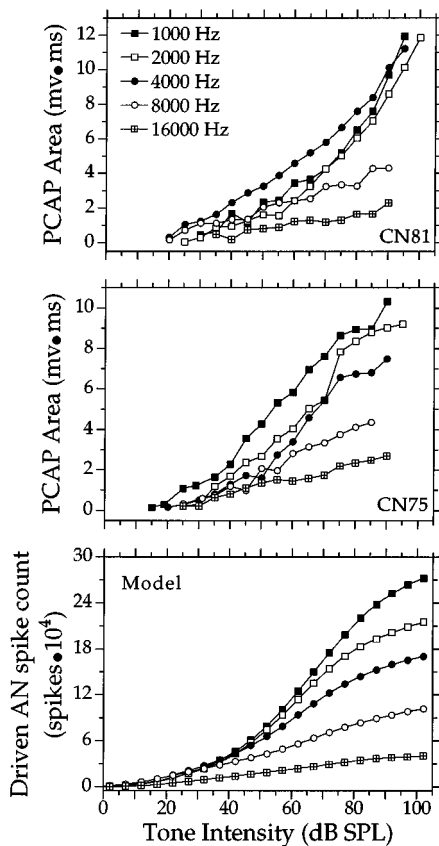


FIG. 7. The PCAP area-intensity functions collected in two animals (top two panels) where the weight criterion was met for a wide range of tone frequency. Unlike previous figures, here the absolute value of the PCAP areas are plotted versus dB SPL. In the bottom panel, the computed growth of the auditory nerve spike count is displayed when the spike counts for all model neurons were weighted equally. Common symbols across all three panels denote data obtained for the same tone frequency. For both CN81 and CN75 notice that at high intensities, low-frequency tones produce larger PCAP areas than do high frequency tones. This difference is found for the model output as well and is caused by the shrinking pool of auditory nerve neurons that can be stimulated as tone frequency is increased.

contributing equally should produce a PCAP area that is proportional to the auditory nerve spike count. To see if this was true, we screened the PCAP area-intensity functions measured in all our experiments using the Weight criterion, and the high-intensity slopes of all those that passed (signifying that all neurons contribute equally to the PCAP at all tone intensities) are plotted versus tone frequency in Fig. 8 (closed circles). The open circles are the high-intensity slopes for the model when the spike counts of all neurons were weighted equally. Considering that the data were pooled from 15 animals, the quantitative match between the model and data slopes is quite good at each tone frequency. For example, at 8 kHz, the average high-intensity slope for the data equals 0.22 dB/dB, whereas the model high-intensity slope equals 0.21 dB/dB. For a 16-kHz tone, the same comparison yields 0.19 dB/dB for the data versus 0.17 dB/dB for the model. We emphasize that all parameter values for the model were assigned using data for the chinchilla obtained from the literature or from experiments performed in our laboratory (see the Appendix).

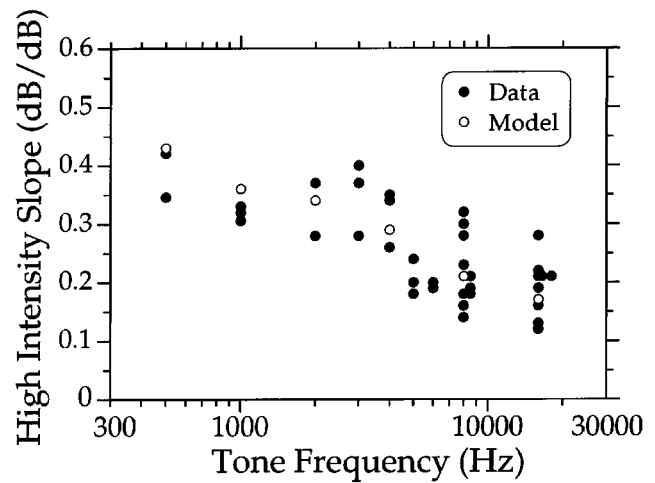


FIG. 8. Scatter plot of high-intensity slopes versus tone frequency. Closed circles are data pooled from 15 different experiments, and these high-intensity slopes were measured for pure tones thought to stimulate neurons contributing equally to the PCAP (according to the Weight criterion defined in the text). The open circles display high-intensity slopes for the model when the spike counts of all neurons are weighted equally. Notice that both the model and data slopes agree quite well at each tone frequency, supporting the idea that all neurons stimulated by the respective tones were contributing equally to the PCAP in these instances.

III. DISCUSSION

A. The ratio tests

The PCAP can only be related to auditory nerve activity if the relative contributions to the PCAP of different auditory nerve neurons are known. Ideally, the PCAP UR for all neurons would be identical, but the ratios tell us that is rarely the case. The area and peak ratios depend on tone frequency in nearly every experiment. Also, both ratio tests display the same trend across experiments (Fig. 2) and the two ratios have a similar dependence on frequency when they are measured in the same experiment (Fig. 3). This consistency between the two ratio tests is found even though they are calculated using different measures that are obtained from different compound potentials (see Fig. 1). The most parsimonious explanation for the agreement between the two tests is that the assumptions underlying their interpretation were reasonable, and that the ratios reflect the typical dependence of PCAP UR size on CF.

For almost all experiments, the ratios indicated that low-CF neurons contributed smaller PCAP URs than did high-CF neurons. In the cat, the auditory nerve is tonotopically organized with high-CF neurons located near the periphery of the nerve bundle and low-CF neurons within the core of the bundle (Sando, 1965; Arnesen and Osen, 1978). If this spatial organization is also true for the chinchilla auditory nerve, and given the location of the PCAP electrode on the surface of the nerve (near high-CF neurons), then distance from the recording electrode might be one reason for the usually smaller PCAP URs contributed by low-CF neurons. However, in order to address the practical usefulness of the PCAP, we focused on answering the following two questions: (1) Can we interpret differences in ratios as *quantitative* measures of variation in UR size across CF? and if we can, then (2) What variation in UR size is acceptable?

Acceptable, that is, while still allowing the PCAP to be used as an accurate measure of the auditory nerve spike count.

With respect to the first question, the most compelling evidence for the quantitative precision of the ratios is the comparisons between the model and data in Fig. 6. The data high-intensity slopes in Fig. 6 are pooled from nine experiments. Consider the parameters that can vary across animals and therefore might cause variability in the growth of the PCAP: density of neurons versus CF, the shapes of tuning curves, the growth of basilar membrane amplitude with tone intensity, etc. In spite of all these potential sources of variability in the high-intensity slopes, the model slopes in Fig. 6 were computed with all parameter values held constant. The lone exception was that the ratios for each animal were used to weight the spike counts computed for each CF band. That the model and data slopes agree quite well for eight out of the nine experiments implies that the dominant source of variability in the slopes across experiments was unequal contributions by different neurons to the PCAP. But even more importantly, the agreement demonstrates that the ratio tests can be used to quantify the relative contributions of different CF regions to the PCAP in every experiment.

Since unequal ratios accurately portray unequal UR sizes between groups of neurons, we wanted to use the ratios measured in any particular experiment to objectively define the phrase, “all neurons having CFs between ?–? kHz are contributing equally to the PCAP.” We thought that such a definition was possible, because for many experiments the ratios were nearly constant over wide frequency ranges. But what is an acceptable variation in UR size? We answered this question by using Weight to summarize the amount of UR variation between on- and off-frequency neurons, and then examining the dependence of 1-kHz high-intensity slopes on Weight (Fig. 6). Using this data and the model to set bounds on Weight (i.e., acceptable variation in UR size), we tested these boundaries by collecting the high-intensity slopes from all PCAP area-intensity functions that met the Weight criterion, and compared the data to model slopes when all model neurons were weighted equally (Fig. 8). The good agreement between the model and data slopes signals that our bounds on variation in PCAP UR size are sufficiently strict, and that when Weight for any tone frequency is between 0.75 and 1.25, all neurons stimulated by the tone are contributing effectively equally to the PCAP.

The bounds for acceptable variation in UR size were verified using the growth of the PCAP in response to pure tones, but it is almost certain that this UR variation would be acceptable for any acoustic stimulus. We found that in 18 of 24 experiments Weight_{4k} and Weight_{8k} were both within the 0.75 and 1.25 boundary values. Therefore, for these 18 experiments, the summed output of the nerve could have been measured for any stimulus whose excitation pattern was confined to neurons having CFs greater than 4 kHz. This CF range encompasses nearly half of all chinchilla auditory nerve neurons, and thus for most PCAP experiments a large portion of the auditory nerve contributes effectively equivalent URs to the PCAP. Since the ratios and Weights can be measured in approximately 15 min, the range of CFs contrib-

uting equally to the PCAP can be defined and monitored throughout every PCAP experiment.

B. The growth of the auditory nerve spike count as a function of pure tone intensity

We used the PCAP to measure the growth of the auditory nerve spike count for a wide range of tone frequency. Prior to the PCAP recording technique, these measurements were impossible. Given the importance of the auditory nerve spike count in several intensity coding schemes, numerous past investigators have used models to characterize its growth as a function of stimulus intensity (Zwislocki, 1965; Howes, 1974; Goldstein, 1974; Lachs *et al.*, 1984). A common goal of these studies was to test the hypothesis that loudness is linearly related to the auditory nerve spike count. We now compare our data and model results with these prior studies.

At high intensities, previous modelers have predicted that the growth of the auditory nerve spike count likely grows at a rate between 0.5–0.6 dB/dB when plotted versus SPL (Zwislocki, 1965; Howes, 1974; Goldstein, 1974; Lachs *et al.*, 1984). While these slopes agree with the reported growth of loudness (e.g., Hellman and Zwislocki, 1968), they are much greater than both the model and data slopes reported here, which even for low-frequency tones lie mostly between 0.3 and 0.4 dB/dB and are never larger than 0.45 dB/dB (Fig. 8). Since all the models are quite complicated, we were not able to uncover explanations for these discrepancies. However, at the time the prior models were developed, factors affecting the growth of the auditory nerve spike count such as the shapes of tuning curves and their dependence on CF (Kiang and Moxon, 1974), the distribution of thresholds for neurons having similar CFs (Liberman, 1978), accurate equations for the shape of rate-intensity functions (Sachs and Abbas, 1974; Sachs *et al.*, 1989; Yates, 1990), and the influence of basilar membrane mechanics on the shape of rate-intensity functions (Sachs and Abbas, 1974; Yates *et al.*, 1990; Yates, 1990) were either unknown or just beginning to be reported. Faced with this incomplete knowledge of peripheral physiology, prior modelers made reasonable assumptions in terms of model structure and parameter values that were consistent with the data available at the time. However, since then, many of these assumptions have turned out to be false, and any or all of these false assumptions could be responsible for their steeper high-intensity slopes.

In contrast, our model was designed so that all parameter values could be assigned using chinchilla auditory nerve data available in the literature or taken from experiments performed in our laboratory (see the Appendix). This design constraint was important, since our motivation for developing the model was to test our interpretation of the ratios, check the bounds we set on acceptable variation in UR size, and determine if the growth of the PCAP was consistent with the well-studied responses of auditory nerve neurons. Given these motives, if we freely adjusted the parameter values after comparing the model and PCAP data, then any agreement between the two would be meaningless. Instead, by using experimental data to set the parameter values, the

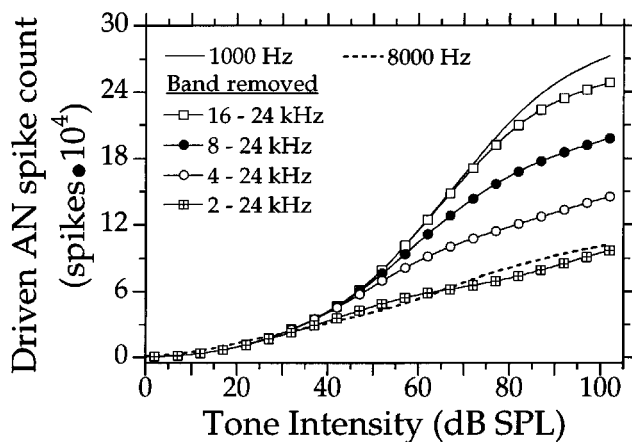


FIG. 9. The auditory nerve model is employed to examine the effects of spread of excitation on the rate-of-growth of the auditory nerve spike count. The solid line shows the computed growth of the auditory nerve spike count for a 1-kHz tone. Like 1-kHz PCAP area-intensity functions, the rate-of-growth of the auditory nerve spike count is steeper at high intensities than at low intensities. As spread of excitation is reduced by removing the spike counts contributed by high CF bands, the difference between the high- and low-intensity slopes is eliminated. In fact, when spread to high CF neurons is almost completely removed (cross-hatched squares), the auditory nerve spike count for a 1-kHz tone grows at a nearly constant rate with increasing tone intensity. This latter result is nearly identical to that observed for an 8-kHz tone (dashed line).

agreement between our model and PCAP high-intensity slopes indicates that the auditory nerve spike count for low frequency tones grows at a rate between 0.3–0.4 dB/dB, not 0.5–0.6 dB/dB as reported previously.

For 8- and 16-kHz tones the high-intensity slopes for both the PCAP data and our model are even shallower, the majority being between 0.1 and 0.2 dB/dB. The number of neurons stimulated by a pure tone is a decreasing function of tone frequency, as a result of the skewed pattern of spread of excitation towards high CF neurons with increasing tone intensity combined with the upper limit of the chinchilla's audible frequency range (approximately 24 kHz, Heffner and Heffner, 1991). A smaller pool of stimulated neurons explains why high-frequency tones produce smaller PCAP areas than do low-frequency tones (Fig. 7). In Fig. 9, we demonstrate that differences in spread also account for the shallower growth of the PCAP area observed for high-frequency tones. In the figure, the output of the model for a 1-kHz tone is shown as the spike counts contributed by high-CF neurons are systematically removed. Consistent with the explanation offered above, as spread is removed the rate-of-growth of the auditory nerve spike count for a 1-kHz tone approaches that observed for an 8-kHz tone.

The growth of the auditory nerve spike count in the absence of spread of excitation has also been modeled previously (Zeng and Turner, 1991). For a pure tone stimulus, Zeng and Turner computed the unweighted sum of the spike counts fired by the low- and high-SR neurons innervating a single hair cell. Given how little spread of excitation would be expected for 8- and 16-kHz tones, we expect that our model and data slopes at these frequencies would be either comparable or slightly larger than those computed for the output of a single hair cell. When the output of Zeng and Turner's model is plotted on log–log axes and fit with a

straight line over the range 30–90 dB SPL (the range we used to compute our high-intensity slopes), the computed spike count grows at a rate of 0.1 dB/dB. This slope agrees quite well with the slopes reported here for high frequency tones, and this shallow rate-of-growth should be expected for the auditory nerve spike count under conditions of restricted spread of excitation to off-frequency neurons. The reason is that, when spread is limited, an upper bound on the rate-of-growth of the auditory nerve spike count should be the rate-of-growth of basilar membrane (BM) amplitude for CF tones. For high-frequency tones and at moderate intensities (30–80 dB SPL), the rate-of-growth of basilar membrane amplitude has been found to lie between 0.1 and 0.3 dB/dB (Sellick *et al.*, 1982; Robles *et al.*, 1986; Ruggero and Rich, 1991; Ruggero, 1992).

C. Intensity coding and the auditory nerve spike count

In contrast to the over 100-dB dynamic range of the auditory system, nearly all auditory nerve neurons have dynamic ranges of approximately 30–50 dB (Kiang *et al.*, 1965; Evans and Palmer, 1980). For many years this mismatch puzzled auditory scientists, but recent studies have shown that combining the spike counts from high-, medium-, and low-SR neurons provides more than enough information to account for the dynamic range of hearing (Viemeister, 1983; Delgutte, 1987; Winslow and Sachs, 1988; Winter and Palmer, 1991). But an important remaining question is *how* are the spike counts from neurons having different SRs, or even CFs, combined to encode intensity? One hypothesis is that the outputs of neurons are combined by computing the unweighted sum of their spike counts (hereafter referred to as the **summation hypothesis**). This idea has been proposed many times in the past, but has been criticized because its success was thought to depend critically on spread of excitation to off-frequency neurons (e.g., Goldstein, 1974; Lachs *et al.*, 1984; Winslow and Sachs, 1988), whereas numerous psychophysical experiments have shown that spread of this type is unnecessary to explain the wide operating range of the auditory system (Viemeister, 1974; Moore and Raab, 1974; Hellman, 1978; Viemeister, 1983; Carlyon and Moore, 1984). For the complex stimulus conditions used in these psychophysical studies, the behavior of the auditory nerve spike count with changes in stimulus intensity is unknown. Now, using the PCAP, the auditory nerve spike count can be *measured* for a variety of stimulus conditions, thus allowing the summation hypothesis to be tested quantitatively and empirically.

The auditory nerve spike count grows over an enormous range of tone intensity, many times exceeding 80 dB and in one case equaling 108 dB. This result is found even for high-frequency tones, where the pool of neurons that can be stimulated at high intensities is greatly reduced when compared to the pool available for low frequency tones. For 8- and 16-kHz tones, the auditory nerve spike count grew at a nearly constant rate over the entire range of intensities tested, implying that this code provides a robust signal of changes in tone intensity. To draw firm conclusions with respect to the detectability of changes in the auditory nerve spike count

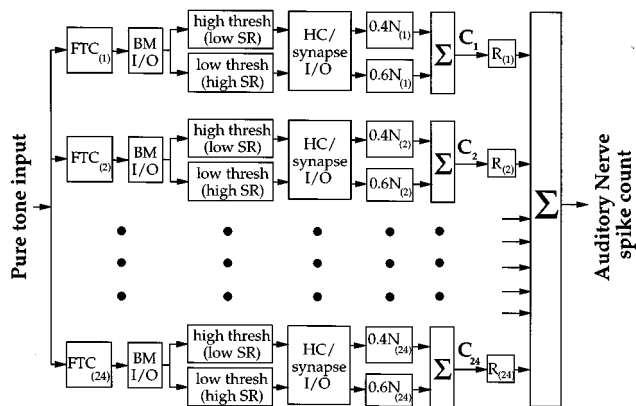


FIG. A1. Block diagram of the model used to compute the auditory nerve spike count as a function of pure tone intensity and frequency. The model was comprised of 24 sections that computed the spike count for neurons having CFs within $\frac{1}{3}$ -oct frequency bands. Twenty-four frequency-threshold curves (FTC) were used to define thresholds as a function of tone frequency for each section. A detailed description of the remaining blocks is given in the text. Abbreviations: BM I/O, basilar membrane input/output function; SR, spontaneous rate; HC/synapse I/O, hair cell/synapse input/output function; N_i , number of neurons in the i th band; C_i , spike count calculated for i th section; R_i , arbitrary constants used to multiply C_i s and thereby simulate the effect of unequal contributions of different sections (bands of neurons) to the PCAP.

requires that its variability also be measured, a topic we will address in the future. But for now these wide dynamic ranges for the auditory nerve spike count demonstrate that, at least for pure tones, this code is capable of signaling changes in stimulus intensity over the entire dynamic range of the auditory system.

ACKNOWLEDGMENTS

This research was supported by grants NSF IBN-9120262 and in its initial stages by NIH Grant DC00380. When writing the manuscript, the first author was supported by NIH grant DC00979. We thank Dr. Jozef Zwislowski, Dr. Robert Smith and Dr. Denis Pelli for helpful discussions and ideas. We also thank Richard Mitchell for constructing many of the custom tools used to record the PCAP and also Art Wixon and Dean Arpajian for building the custom electronics in our lab. We especially thank Anita Sterns for performing initial surgical preparations.

APPENDIX: COMPUTING THE AUDITORY NERVE SPIKE COUNT

We constructed a model of the auditory periphery with the sole purpose of computing the growth of the auditory nerve spike count as a function of pure tone frequency and intensity. Therefore, phenomena such as adaptation or suppression were not included in the model. Our objective was to keep the structure of the model as simple as possible so that (1) all parameter values could be set using data obtained from the literature or from experiments performed in our laboratory and (2) all parameter values would be specific to the chinchilla. After setting the values of the parameters, they remained fixed and the output of the model was compared to PCAP data (e.g., Figs. 6–8).

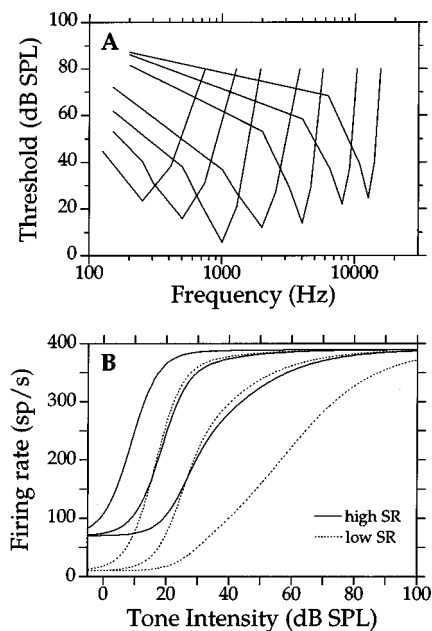


FIG. A2. (A) Model tuning curves for seven of the twenty-four CF bands comprising the auditory nerve model. (B) Rate-intensity functions for model high and low SR auditory nerve neurons stimulated by pure tones at their CF. The six functions represent model neurons having the lowest, average, and highest thresholds for each respective SR class. Notice that “saturating,” “sloping saturation,” and “straight” rate-intensity functions are all found among the model neurons. These shapes are due to the modeled nonlinear growth of basilar membrane amplitude in response to CF tones.

The model was comprised of 24, $\frac{1}{3}$ -oct bands that collectively ranged from 0.125–24 kHz. These boundaries set the minimum and maximum CFs for model neurons, and were based on (1) the range of CFs observed for hundreds of neurons recorded in our laboratory and (2) chinchilla behavioral audiograms obtained from other studies (Ades *et al.*, 1974; Miller, 1970; Heffner and Heffner, 1991). Pure tone thresholds for the most sensitive neurons (i.e., high-SR neurons) in each band were set using a frequency-threshold curve (FTC in Fig. A1). The FTC shape for each band was defined using averaged FTCs for chinchilla high-SR neurons. In Fig. A2, 7 of the 24 model FTCs are shown that cover a wide range of CFs. The CF thresholds for each of the 24 FTCs were calculated using a database of 238 high-SR neurons from our laboratory.

The equations used to calculate rate-intensity functions for single neurons were obtained from Yates (1990, Eqs. 8 and 9) which in turn are very similar to those of Sachs *et al.* (1989). These studies demonstrate that the shape of measured rate-intensity functions are well described by two equations thought to represent (1) the transformation of tone intensity into BM displacement and (2) the transformation of BM displacement into neural firing rate. For the first equation, BM displacement in a given frequency band (denoted by BM I/O in Fig. A1) was a *linear* function of intensity for tone frequencies *less* than the lower frequency limit of the band. For tone frequencies *within* the band, BM displacement was a *nonlinear* function of tone intensity (Eq. 9, Yates, 1990). Yates’ Eq. 9 requires values for two parameters that can be obtained by measuring pure tone rate-

intensity functions for single neurons: (1) the tone intensity where the BM response changes from linear to nonlinear growth and (2) the log-log slope of the nonlinear region of the BM amplitude-intensity function. Values for these two parameters were set using averaged data from fits to chinchilla pure tone rate-intensity functions (number of neurons = 57) measured in our laboratory (Jackson and Relkin, 1995). The average value for the first parameter equaled FTC threshold + 17.1 dB SPL and the log-log slope equaled 0.3 dB/dB. Both values are consistent with direct measurements of chinchilla BM motion (Robles *et al.*, 1986; Ruggero and Rich, 1991; Ruggero, 1992).

The equation used to transform BM displacement into neural firing rates was a sigmoidal function (Eq. 8; Yates, 1990) that represented a combined hair cell and synapse input/output function (denoted as HC/synapse I/O function in Fig. A1). Again, all parameter values were set using averaged data for rate/intensity functions measured for the chinchilla in our laboratory (Jackson and Relkin, 1995). Two groups of neurons were modeled, high-SR neurons and low-SR neurons. The SR for high-SR neurons equaled 70 sp/s and for low-SR neurons equaled 10 sp/sec. For a 34-ms tone burst, the maximum driven firing rate was 320 spikes/s for the high SR neurons and 380 spikes/sec for the low-SR neurons. The remaining parameter in Yates' Eq. 8 determines neural threshold, but it is not equivalent to the threshold measured by a FTC. Since Jackson and Relkin measured both the FTC threshold and the neural threshold as defined by Yates, the average difference between the two was known (7.7 dB for high-SR neurons and 6.5 dB for low-SR neurons) and incorporated into the model. For each band, the difference between the average threshold for low and high-SR neurons equaled 9.1 dB and thresholds for both groups were varied uniformly over a 20-dB intensity range. Therefore, the maximum difference between low- and high-SR thresholds in any band was 30 dB, a difference consistent with published chinchilla auditory nerve data (Salvi *et al.*, 1982; Relkin and Pelli, 1987). Sample CF rate-intensity functions for model low- and high-SR neurons are shown in Fig. A2.

Firing rates for all neurons were converted to spike counts for a 34 ms tone. Spike counts were multiplied by the number of neurons in each band (denoted as N in Fig. A1) that were computed using a chinchilla frequency-place map (Eldredge *et al.*, 1981) and chinchilla neural-innervation densities (Bohne *et al.*, 1982). High-SR and low-SR neurons comprised 60% and 40% of the neurons, respectively. Spike counts for each SR group were summed to equal the total spike count for each band (denoted as C in Fig. A1). Subsequently, the spike counts from all sections were summed to equal the auditory nerve spike count.

1. Simulating the effect of unequal weighting

One important reason for developing the model was to test the effects of unequal weighting (i.e., unequal PCAP UR size) on the growth of the PCAP area with tone intensity. The block diagram in Fig. A1 shows a constant gain of R_i at the output of each band. To simulate unequal weighting measured in a particular experiment using the ratio tests, each R_i was set equal to the normalized ratio calculated for the

i th band in a particular animal. The weighted spike counts for each band were then summed to equal the auditory nerve spike count.

- Ades, H. W., Trahiotis, C., Kokko-Cunningham, A., and Averbuch, A. (1974). "Comparison of hearing thresholds and morphological changes in the chinchilla after exposure to 4 kHz tones," *Acta Otolaryngol.* **78**, 192–206.
- Arnesen, A. R., and Osen, K. K. (1978). "The cochlear nerve in the cat: topography, cochleotopy, and fiber spectrum," *J. Comp. Neurol.* **178**, 661–678.
- Bohne, B. A., Kenworthy, A., and Carr, C. D. (1982). "Density of myelinated nerve fibers in the chinchilla cochlea," *J. Acoust. Soc. Am.* **72**, 102–107.
- 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.
- Dallos, P., and Cheatham, M. A. (1976). "Compound action potential (AP) tuning curves," *J. Acoust. Soc. Am.* **59**, 591–597.
- 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 (M. Nijhof, Dordrecht, The Netherlands), pp. 333–353.
- Doucet, J. R. (1995). "The Perstimulus Compound Action Potential and intensity coding in the chinchilla auditory nerve," Ph.D. thesis, Syracuse University, Syracuse, NY.
- Doucet, J. R., and Relkin, E. M. (1990). "Contribution of individual neurons to the peristimulus compound action potential estimated by narrow band responses to clicks," in *Abstracts of the Thirteenth Midwinter Research Meeting of the Association for Research in Otolaryngology*, p. 145.
- Doucet, J. R., and Relkin, E. M. (1995). "The Perstimulus Compound Action Potential: A new method for recording a compound potential from the chinchilla auditory nerve," *Audit. Neurosci.* **1**, 151–168.
- Eldredge, D. H., Miller, J., and Bohne, B. A. (1981). "A frequency-position map for the chinchilla cochlea," *J. Acoust. Soc. Am.* **69**, 1091–1095.
- Evans, E. F. (1981). "The dynamic range problem: Place and time coding at the level of cochlear nerve and nucleus," in *Neuronal Mechanisms in Hearing*, edited by J. Syka and L. Aitkin (Plenum, New York), pp. 69–85.
- Evans, E. F., and Palmer, A. R. (1980). "Relationship between the dynamic range of cochlear nerve fibres and their spontaneous activity," *Exp. Brain Res.* **40**, 115–118.
- Gleich, O., and Wilson, S. (1993). "The diameters of guinea pig auditory nerve fibers: distribution and correlation with spontaneous rate," *Hearing Res.* **71**, 69–79.
- Goldstein, J. L. (1974). "Is the power law simply related to the driven spike response rate from the whole auditory nerve?," in *Sensation and Measurement*, edited by H. R. Moskowitz, B. Scharf, and J. C. Stevens (Reidel, Dordrecht), pp. 223–229.
- Hakansson, C. H. (1956). "Conduction velocity and amplitude of the action potential as related to circumference in the isolated fiber of frog muscle," *Acta. Physiol. Scand.* **37**, 14–34.
- Heffner, R. S., and Heffner, H. E. (1991). "Behavioural hearing range of the chinchilla," *Hearing Res.* **52**, 13–16.
- Hellman, R. P. (1978). "Dependence of loudness growth on skirts of excitation patterns," *J. Acoust. Soc. Am.* **63**, 1114–1119.
- Hellman, R. P., and Zwislocki, J. J. (1968). "Loudness determination at low sound frequencies," *J. Acoust. Soc. Am.* **43**, 60–64.
- Howes, W. L. (1974). "Loudness function derived from data on electrical discharge rates in auditory nerve fibers," *Acustica* **30**, 247–259.
- Jackson, B. S., and Relkin, E. M. (1995). "Factors influencing the shape of rate-versus-intensity functions of auditory nerve fibers in the chinchilla," in *Abstracts of the Eighteenth Midwinter Research Meeting of the Association for Research in Otolaryngology*, p. 175.
- Kiang, N. Y. S., and Moxon, E. C. (1974). "Tails of tuning curves of auditory-nerve fibers," *J. Acoust. Soc. Am.* **55**, 620–630.
- Kiang, N. Y. S., Moxon, E. C., and Kahn, A. R. (1976). "The relationship of gross potentials recorded from the cochlea to single unit activity in the auditory nerve," in *Electrocochleography*, edited by R. J. Ruben, C. Elberling, and G. Salomon (University Park, Baltimore), pp. 95–115.
- Kiang, N. Y. S., Watanabe, T., Thomas, E. C., and Clark, L. F. (1965). *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve*, MIT Res. Monogr. No. 35 (Technology, Boston).

- Kim, D. O., and Molnar, C. E. (1979). "A population study of cochlear nerve fibers: Comparison of spatial distributions of average-rate and phase-locking measures of responses to single tones," *J. Neurophysiol.* **42**, 16–30.
- Lachs, G., Al-Shaikh, R., Bi, Q., Saia, R. A., and Teich, M. (1984). "A neural counting model based on physiological characteristics of the peripheral auditory system. V. Applications to loudness estimation and intensity discrimination," *IEEE Trans. Syst. Man Cybern.* **SMC-14**, 819–836.
- Liberman, M. C. (1978). "Auditory-nerve response from cats raised in a low-noise chamber," *J. Acoust. Soc. Am.* **63**, 442–455.
- Liberman, M. C., and Oliver, M. E. (1984). "Morphometry of intracellularly labeled neurons of the auditory nerve: correlations with functional properties," *J. Comp. Neurol.* **223**, 163–176.
- Miller, J. D. (1970). "Audibility curve of the chinchilla," *J. Acoust. Soc. Am.* **48**, 513–523.
- 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.
- Özdamar, Ö., and Dallos, P. (1978). "Synchronous responses of the primary auditory fibers to the onset of tone burst and their relation to compound action potentials," *Brain Res.* **155**, 169–175.
- 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.
- 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.* **1**, 215–220.
- Ruggero, M. A., and Rich, N. C. (1991). "Application of a commercially-manufactured Doppler-shift laser velocimeter to the measurement of basilar-membrane vibration," *Hearing Res.* **51**, 215–230.
- Sachs, M. B., and Abbas, P. J. (1974). "Rate versus level functions for auditory-nerve fibers in the cat: tone-burst stimuli," *J. Acoust. Soc. Am.* **56**, 1835–1847.
- Sachs, M. B., Winslow, R. L. and Sokolowski, B. H. A. (1989). "A computational model for rate-level functions from cat auditory nerve fibers," *Hearing Res.* **41**, 61–70.
- Salvi, R., Perry, J., Hamernik, R. P., and Henderson, D. (1982). "Relationships between cochlear pathologies and auditory nerve and behavioral responses following acoustic trauma," in *New Perspectives in Noise-Induced Hearing Loss*, edited by R. P. Hamernik, D. Henderson, and R. Salvi (Raven, New York), pp. 165–188.
- Sando, I. (1965). "The anatomical interrelationships of the cochlear nerve fibers," *Acta Otolaryngol.* **59**, 417–436.
- Sellick, P. M., Patuzzi, R., and Johnstone, B. M. (1982). "Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique," *J. Acoust. Soc. Am.* **72**, 131–141.
- Teas, D. C., Eldredge, D. H., and Davis, H. (1962). "Cochlear responses to acoustic transients: An interpretation of whole-nerve action potentials," *J. Acoust. Soc. Am.* **34**, 1438–1459.
- Versnel, H., Prijs, V. F., and Schoonhoven, R. (1992). "Round-window recorded potential of single-fiber discharge (unit response) in normal and noise-damaged cochleas," *Hearing Res.* **59**, 157–170.
- Viemeister, N. F. (1974). "Intensity discrimination of noise in the presence of band-reject noise," *J. Acoust. Soc. Am.* **56**, 1594–1600.
- Viemeister, N. F. (1983). "Auditory intensity discrimination at high frequencies in the presence of noise," *Science* **221**, 1206–1208.
- Wang, B. (1979). "The relation between the compound action potential and unit discharges of the auditory nerve," Ph.D. thesis, MIT, Boston, MA.
- Winslow, R. L., and Sachs, M. B. (1988). "Single-tone intensity discrimination based on auditory-nerve rate responses in quiet, noise, and with stimulation of the crossed olivocochlear bundle," *Hearing Res.* **35**, 165–190.
- Winter, I. M., and Palmer, A. R. (1991). "Intensity coding in low-frequency auditory-nerve fibers of the guinea pig," *J. Acoust. Soc. Am.* **90**, 1958–1967.
- Yates, G. K. (1990). "Basilar membrane nonlinearity and its influence on auditory nerve rate-intensity functions," *Hearing Res.* **50**, 145–162.
- Yates, G. K., Winter, I. M., and Robertson, D. (1990). "Basilar membrane nonlinearity determines auditory nerve rate-intensity functions and cochlear dynamic range," *Hearing Res.* **45**, 203–220.
- Zeng, F.-G., and Turner, C. W. (1991). "Binaural loudness matches in unilaterally impaired listeners," *Q. J. Exp. Psychol.* **43A** (3), 565–583.
- Zwislocki, J. J. (1965). "Analysis of some auditory characteristics," in *Handbook of Mathematical Psychology*, edited by R. D. Luce, R. R. Bush, and E. Galanter (Wiley, New York), pp. 1–97.

Is loudness simply proportional to the auditory nerve spike count?

Evan M. Relkin and John R. Doucet^{a)}

Institute for Sensory Research and Department of Bioengineering and Neuroscience Syracuse University, Syracuse, New York 13244-5290

(Received 13 August 1996; revised 9 January 1997; accepted 10 January 1997)

It is often asserted that the physiological correlate of loudness is the simple sum of the spike activity produced by all neurons in the auditory nerve (the auditory nerve spike count). We will refer to this hypothesis as the *spike count hypothesis*. The spike count hypothesis has been tested in the past using models of the auditory periphery and in almost all cases, the hypothesis has been supported. Our new technique for recording a compound potential from the chinchilla auditory nerve, the perstimulus compound action potential (PCAP), makes possible the measurement of the growth of the auditory nerve spike count, thus providing data that can be used to test the spike count hypothesis empirically. It was observed that the growth of the auditory nerve spike count in response to a 1-kHz pure tone (in dB/dB) is 33% shallower than the growth of loudness for a 1-kHz tone, and this discrepancy increases to 66% for an 8-kHz tone. In addition, "equal-count" contours were constructed in a manner analogous to equal-loudness contours. It was found that as reference intensity increases, equal-count contours become sharply curved upward at high frequencies whereas equal-loudness contours become increasingly flat. These differences are unlikely to be the result of the cross-species comparison, since the discrepancies are mostly attributable to the skewed pattern of spread of excitation along the basilar membrane, a property shared by humans and chinchillas. Therefore, we conclude that the simple sum of the spike activity in the auditory nerve cannot be the physiological correlate of loudness. © 1997 Acoustical Society of America. [S0001-4966(97)05405-2]

PACS numbers: 43.64.Bt, 43.64.Pg, 43.66.Ba, 43.66.Cb [RDF]

INTRODUCTION

Loudness is one of the most basic perceptual attributes of sound, but the neural correlate of loudness at any level of the auditory system is still unknown. One commonly held idea is that loudness is simply proportional to the total number of action potentials fired by all auditory nerve neurons (the **auditory nerve spike count**) (e.g., Fletcher and Munson, 1933). For convenience, we will refer to this hypothesis as the *spike count hypothesis*. In a companion research article (Doucet and Relkin, 1997), we described how a new compound potential, the perstimulus compound action potential (PCAP), can be used to *measure* the auditory nerve spike count. We establish in that article that when certain criteria are met, the area under the PCAP waveform is proportional to the auditory nerve spike count (see also Doucet and Relkin, 1995). Our goal here is to use the PCAP to test the spike count hypothesis empirically.

In the past, since measuring the auditory nerve spike count was not possible, the spike count hypothesis was tested using models of the auditory periphery (Zwislocki, 1965; Howes, 1974; Goldstein, 1974; Lachs *et al.*, 1984). These prior studies concluded that the spike count hypothesis was probably correct, since for pure tones the rate-of-growth of both loudness and the auditory nerve spike count agreed over a wide range of tone intensity. However, these prior models

were based on incomplete knowledge of auditory-nerve physiology and the models predicted rates-of-growth for the auditory nerve spike count (in dB/dB) that are almost a factor of 2 greater than those measured using the PCAP or calculated with a model using present day knowledge of auditory physiology (Doucet and Relkin, 1997). Furthermore, these previous tests have tended to concentrate on the growth of the auditory nerve spike count at a single tone frequency, but did not evaluate whether the spike count hypothesis holds over a wide range of tone frequency.

In addition, not all prior studies have supported the spike count hypothesis. For example, Pickles (1983) predicted the response of the total array of auditory nerve fibers based on physiological measurements from a few neurons. While many of the assumptions underlying this approach may be questioned, Pickles did find that the dependence of the predicted response of the array to noises of varying bandwidth was not the same as the dependence of loudness on bandwidth. In a more recent study (Jeng, 1992), a model of the auditory periphery was used to compute the auditory nerve spike count for a wide range of tone intensity and frequency. Although the model was similar in form to those used by the authors cited in the above paragraph, Jeng found that the growth of the auditory nerve spike count with tone intensity was more shallow at high intensities when compared to the growth of loudness. Also, Jeng reported that the shape of equal-count contours was markedly different from that of equal-loudness contours.

Differences in the assumptions and parameters used to

^{a)}Current address: Center for Hearing Sciences, 510 Traylor Research Building, 720 Rutland Avenue, Baltimore, MD 21205.

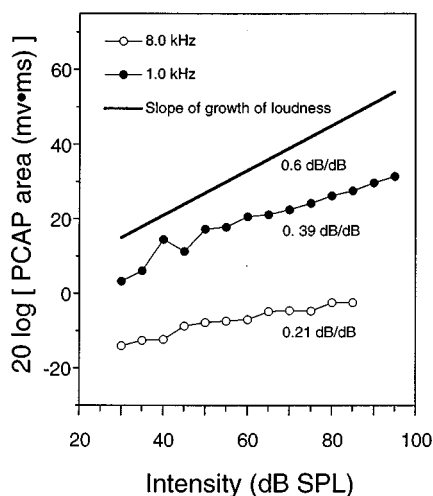


FIG. 1. The rate-of-growth of the PCAP area and loudness as a function of pure tone intensity are compared. Since the slopes of the three functions on log-log axes is the issue being addressed, all three functions have been arbitrarily shifted along the y axis for display purposes. The PCAP area-intensity functions demonstrate the rate-of-growth of the auditory nerve spike count, since the proportional relationship between the PCAP area and the auditory nerve spike count does not change the slopes measured on these axes. The rate-of-growth of the auditory nerve spike count depends on tone frequency, and this dependence is demonstrated with typical PCAP area-intensity functions measured for 1- and 8-kHz tones. In contrast, the growth of loudness is reported to be independent of tone frequency for frequencies greater than 1 kHz (Scharf, 1978). Also, notice that loudness grows much faster with increasing intensity than the auditory nerve spike count measured at either tone frequency.

construct models of the auditory periphery are the most likely reason that these prior studies have not all reached the same conclusions with respect to the spike count hypothesis. In contrast to these prior attempts, we are now able to use the PCAP to measure the growth of the auditory nerve spike count as a function of tone intensity and frequency. Thus, our approach to testing the spike count hypothesis was to compare the rate of growth of the PCAP area and loudness as a function of pure tone intensity for a wide range of tone frequency. In addition, we used PCAP area-intensity functions and our own model count-intensity functions to construct “equal-area” and “equal-count contours” in a manner directly analogous to how “equal-loudness contours” are measured for human listeners. Both comparisons uncover large problems with the spike count hypothesis, leading us to conclude that loudness cannot be proportional to the auditory nerve spike count.

I. METHODS

The PCAP and model data that are presented here were originally shown in a companion paper (Doucet and Relkin, 1997). Detailed methods for recording the PCAP, our protocol for assuring that all neurons stimulated contribute equally to the PCAP, and a full description of the auditory nerve model used to compute the auditory nerve spike count can be found in that paper. The PCAP data shown in Fig. 1, and used to construct the contours in Fig. 2, are from one chin-

chilla, CN81, for which it was established that all neurons were contributing equally to the PCAP within our previously defined limits.

II. RESULTS

The auditory nerve spike count in response to pure tones grows much more slowly than does the loudness of pure tones. This difference is illustrated in Fig. 1, where the growth of the PCAP area for 1- and 8-kHz tones is shown on log-log axes and compared to the typically reported growth of pure-tone loudness over this intensity range (Hellman and Zwisllocki, 1968; Hellman and Meiselman, 1988). Only one loudness slope is shown, since the growth of loudness is nearly independent of tone frequency between 1 and 8 kHz. The PCAP area is proportional to the auditory nerve spike count, but the proportionality factor does not affect the slope on these axes and, therefore, the slope of the PCAP function is equal to the rate of growth of the auditory nerve spike count. The slopes of the PCAP area-intensity functions are typical for these two tone frequencies, and notice that they are both much shallower than the loudness slope. This discrepancy is particularly apparent for an 8-kHz tone. Using the slopes in Fig. 1, consider that a human listener judges an 8-kHz tone at 90 dB SPL to be 64 times as loud as one at 30 dB SPL. If loudness were simply proportional to the auditory nerve spike count, then the listener would judge the 90 dB SPL tone to be only four times as loud as the 30 dB SPL tone.

An even more troublesome problem for the spike count hypothesis is the shape of equal-area and equal-count contours when compared to the shape of equal-loudness contours. In panels (A) and (D) of Fig. 2, we show, for humans, growth of loudness functions (redrawn from Scharf, 1978) and equal-loudness contours (redrawn from Fletcher, 1953). In panels (B) and (C), we replot PCAP area-intensity functions obtained in one chinchilla and model count-intensity functions that were both originally presented in our companion paper (Doucet and Relkin, 1997; Fig. 9). The data in panels (B) and (C) were then used to construct equal-area [panel (E)] and equal-count [panel (F)] contours. Because we were rarely able to achieve PCAP recording conditions that allowed us to define the growth of the auditory nerve spike count for a wide range of tone frequency and intensity in a single animal, data for the model are presented to provide a more complete picture of the pattern of the results. The modeled growth of the auditory nerve spike count was validated over a wide range of tone frequency and intensity, therefore, the model output is used to supplement the sparse PCAP data and demonstrate that our findings in a small number of animals are almost certainly valid in general. Equal-area and equal-count contours were constructed as follows: First, a reference intensity was chosen and the area (count) for the 1.0-kHz function at the reference intensity determined the reference area (count). Second, for each tone frequency different from 1 kHz, the intensity required for that tone to produce the reference area (count) was determined and plotted versus the frequency of the tone. This second step is illustrated graphically in Fig. 2 by the horizontal lines in panels (B) and (C). The points on the respective contour are

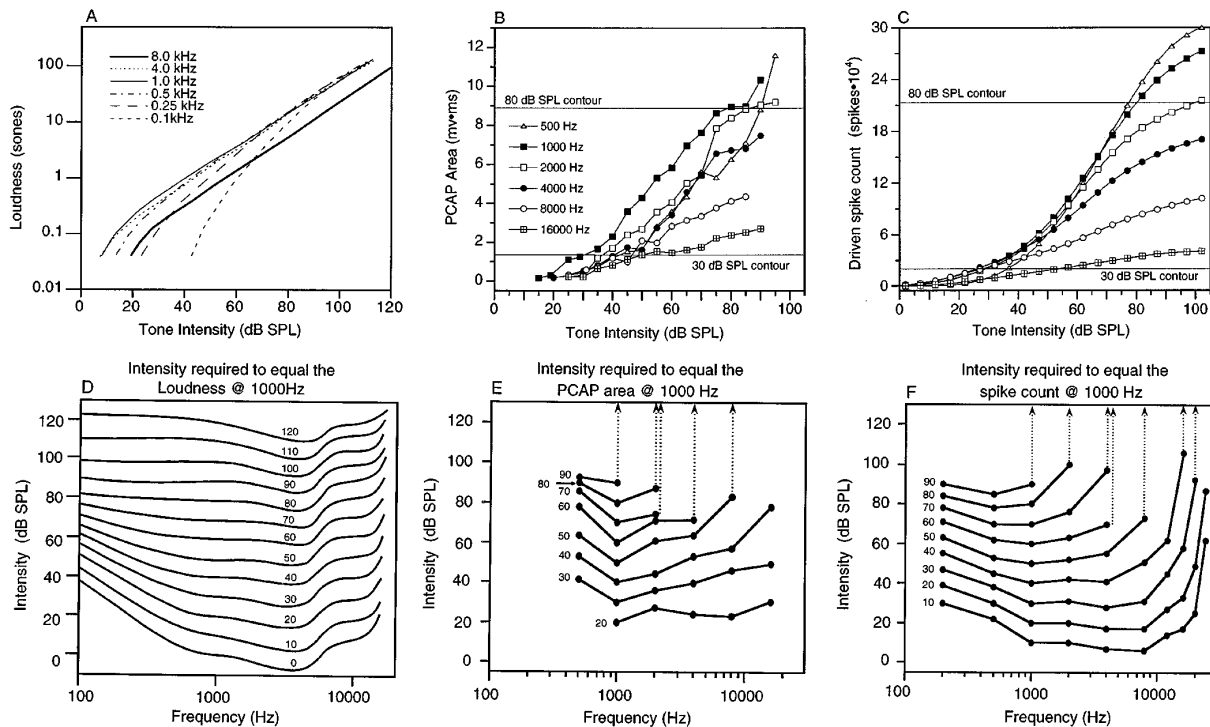


FIG. 2. Panels (A), (B), and (C) compare growth of loudness in humans [panel (A), redrawn from Scharf, 1978] to two measures of the auditory nerve spike count: (1) growth of PCAP area [panel (B)] measured in one chinchilla and (2) computed driven spike count for a model of the chinchilla periphery [panel (C)]. Notice that the high-intensity slope of the functions does not depend on frequency for the loudness data, but decreases drastically with frequency for the two measures of the auditory nerve spike count. Also note the good qualitative agreement between the empirical and model functions. The symbols in panel (C) correspond to the same frequencies as indicated in panel (B). The horizontal lines in panels (B) and (C) are included to demonstrate how the contours shown in panels (E) and (F) were constructed. See the text for details. Panels (D), (E), and (F) compare equal loudness contours for humans [panel (D), redrawn from Fletcher, 1953] to equal area contours for the PCAP [panel (E)] and equal driven count contours for our model [panel (F)] derived from the data shown in panels (B) and (C), respectively. The numbers next to each contour indicate the intensity of the reference 1.0-kHz tone in dB SPL. For further explanation of how the contours were constructed, see the text. The upward going arrows indicate that for all higher frequencies, it was not possible to match the value for the reference tone within the range of intensities for which data was available. Notice that while the equal loudness contours tend to flatten at all frequencies as reference intensity is increased, the equal auditory nerve count contours flatten at low frequencies while becoming increasingly steep at high frequencies.

determined by the intersection of the horizontal line with each of the intensity functions. This procedure was repeated for reference intensities from 10–90 dB SPL in 10-dB steps. Notice that if loudness were substituted for area (count) in the protocol described above, then we would be essentially describing the method used to construct the equal-loudness contours in panel (A).

The shape of equal-loudness contours at low intensities (<40 dB SPL) generally conform to the shape of the pure tone audiogram, but above 40 dB SPL equal-loudness contours gradually begin to flatten at all frequencies. Not so for the equal-area and equal-count contours, which flatten at low frequencies but become increasingly steep at high frequencies as the reference intensity increases. The significance of the dotted arrows attached to the area count contours is that no intensity less than the maximum for which data were available (100 dB SPL for the PCAP data, 130 dB for the model calculations) produces an area (count) equivalent to the reference area (count) for the respective contours for tones of higher frequency. Note the significance of these arrows for the spike count hypothesis. Using the equal-loudness contours in Fig. 2, human listeners consider a 70 dB SPL, 1-kHz tone and an 80 dB SPL, 8-kHz tone to be equally loud. If loudness were proportional to the auditory

nerve spike count and a listener were asked to match the intensity of an 8-kHz tone to a 70 dB SPL, 1-kHz tone, *the listener would report that they could not perform the task*. No matter what intensity they selected for the 8-kHz tone, the tone would always be perceived as less loud when compared to the 1-kHz tone.

III. DISCUSSION AND CONCLUSIONS

The drastic discrepancies between the predictions of the spike count hypothesis and loudness perception demonstrate that the loudness of a pure tone is not simply proportional to the auditory nerve spike count. While the growth of loudness at high intensities is often given as 0.6 dB/dB (e.g., Scharf, 1978), values as low as 0.1 dB/dB have been reported (Viemeister and Bacon, 1988), raising the possibility that the exact value of the slope may vary with methodology. Thus, the finding that the growth of the spike count is less than 0.6 dB/dB at high intensities is not, by itself, compelling evidence for the rejection of the spike-count hypothesis. However, equal-loudness contours are not subject to comparable methodological problems and the lack of agreement between equal-loudness contours and equal-count contours is much stronger evidence for rejecting the hypothesis. The latter re-

sult implies that any mathematical operation on the spike count that does not depend on frequency, i.e., which is the same for all frequency bands or, equivalently, which takes place after the sum is formed, will be inadequate to overcome the disagreement between the two sets of contours.

The reason the spike count hypothesis fails is that the pool of auditory nerve neurons that can potentially be stimulated by a pure tone is a decreasing function of tone frequency, as a result of the skewed pattern of spread of excitation along the basilar membrane towards neurons having high characteristic frequencies (CFs) (e.g., Kim and Molnar, 1979; Evans, 1981). Thus, in contrast to the dependence of loudness on tone frequency, low-frequency tones at moderate intensities produce rates-of-growth and values for the auditory nerve spike count that high-frequency tones can never match.

Our test of the spike count hypothesis involved using the chinchilla auditory periphery as a model for the human peripheral auditory system, but in this case the comparison across species does not require us to temper our conclusions. The skewed pattern of spread of excitation in the chinchilla auditory nerve that is responsible for the failure of the spike count hypothesis is caused by the asymmetric tuning of the chinchilla basilar membrane (e.g., Ruggero, 1992). Since spread of excitation along the basilar membrane towards the cochlear base was originally observed using human cochleas (Békésy, 1960), and masking patterns produced by pure tones for human listeners are always skewed towards high frequencies (e.g., see Moore, 1986), a skewed pattern of spread of excitation towards high-CF neurons almost certainly characterizes the population response of human auditory nerve neurons to pure tones. Therefore, we conclude that our negation of the spike count hypothesis would not change even if we could record the auditory nerve spike count in humans directly.

If the spike count hypothesis is wrong, what is the neural correlate of loudness in the auditory nerve? One possibility is that loudness is encoded in the spatio-temporal patterns of firing among populations of auditory nerve fibers (Carney, 1994). Groups of auditory nerve neurons having similar CFs and phase locking to the stimulus waveform fire more synchronously as tone intensity increases, and central auditory neurons sensitive to this coherence could be responsible for loudness perception. However, since phase locking in auditory nerve fibers is virtually absent for tone frequencies greater than 5 kHz (Johnson, 1980), changes in across-neuron synchrony cannot be the neural correlate of loudness in the auditory nerve for high frequency tones. Rather, the majority of psychophysical and physiological evidence agree with the idea that our perception of loudness and ability to discriminate small changes in stimulus intensity is linked to the number of spikes fired by peripheral neurons (for reviews of this argument, see Viemeister, 1988; Delgutte, 1996).

Another alternative to the spike count hypothesis proposes that spatial spread of excitation along the basilar membrane plays only a minor role in the perceived loudness of a pure tone, in other words, loudness is related to the summed spike counts produced only by those neurons tuned to the tone frequency. This idea is attractive because by restricting

the count to neurons innervating narrow cochlear regions, the large differences between the spike counts produced by low- and high-frequency tones [Fig. 2(B) and (C)] would be reduced. As a result, equal spike-count contours formed after summing over narrow cochlear regions should be nearly flat (like human equal-loudness contours) unlike the contours [Fig. 2(E) and (F)] obtained when summing over the entire cochlea. Although downplaying the role of spatial spread of excitation is appealing, and is consistent with some results in hearing impaired listeners (Hellman and Meiselman, 1986; Zeng and Turner, 1991) where spread was supposedly limited by a high-frequency hearing loss, numerous psychophysical studies of normal hearing listeners have demonstrated that the loudness of a pure tone is reduced when spatial spread of excitation is restricted using masking noises (Hellman, 1974, 1978; Moore *et al.*, 1985; Schneider and Parker, 1987; Schlauch, 1994). In a related series of experiments, intensity discrimination thresholds at high intensities are increased when spatial spread of excitation is reduced using masking noises (Viemeister, 1974; Moore and Raab, 1974; Viemeister, 1983; Carlyon and Moore, 1984; Schlauch, 1994). Thus, when tone intensity is increased and auditory nerve neurons not tuned to the tone frequency are stimulated, their activity does appear to play a role in the perceived loudness of a pure tone.

Collecting the evidence from past studies presented in the last two paragraphs, the perceived loudness of a pure tone appears to be linked both to the number of spikes fired by single neurons and to spatial spread of excitation in the auditory nerve. How can this be, considering our strong empirical evidence that the spike count hypothesis is false? The spike count hypothesis proposes that loudness is *proportional* to the auditory nerve spike count, that is, after peripheral processing, the remainder of the central auditory system plays only a minor role in the perceived loudness of an acoustic stimulus, i.e., it computes a simple, unweighted sum. Clearly, our findings do not preclude the possibility that the loudness of a pure tone is functionally *dependent* on the spike count for each auditory nerve neuron, as long as the spatial pattern of auditory nerve activity is taken into account when combining counts across neurons. Proposing a specific functional relationship between the individual spike counts and loudness is difficult at this time, since so little is known about how the central auditory nervous system processes the spatial firing patterns in the auditory nerve. However, the fact that equal-loudness contours are relatively flat at high intensities suggests that the central auditory system might behave as a kind of “equalizer,” boosting the effect of spike counts contributed by high-CF neurons when they are stimulated by high-frequency tones, and thus compensating for the skewed patterns of spread of excitation in the auditory nerve. If loudness perception is dependent on the auditory nerve spike count, it is probably important that the central auditory system perform this “equalization” procedure, otherwise potentially important high-frequency stimuli might be perceived as too “quiet” and thus go unnoticed when paired with low frequency stimuli. Whatever the functional relationship between the auditory nerve spike count and loudness, the results presented here clearly show that this rela-

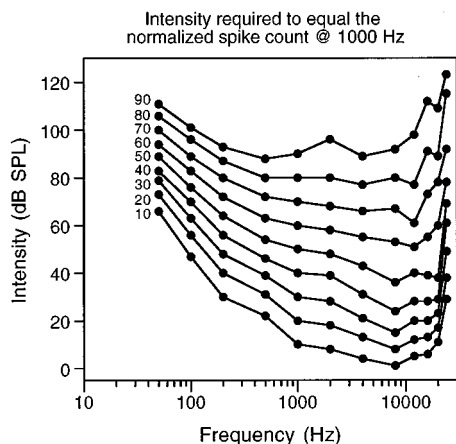


FIG. 3. Equal normalized-count contours for our model. Count/intensity functions for each test frequency were normalized by dividing by the count at 130 dB SPL, which was taken as the maximum count for the respective frequency. The contours plot the intensity required at each frequency to equal the normalized count at 1.0 kHz for the reference intensity indicated to the left of each contour. Note the closer similarity to equal-loudness contours [Fig. 2(D)] relative to the contours for the unequalized count [Fig. 2(F)].

tionship is more complicated than was typically assumed in the past.

We have tested, using our model, one equalization procedure that is conceptually similar to those proposed by Riesz (1933) and Houtsma (1980) for relating just-noticeable differences (jnd's) in intensity to growth of loudness. Specifically, the auditory nerve spike count evoked by a given tone frequency is normalized by dividing by the maximum number of spikes that could be fired in response to that tone frequency. We approximated the maximum spike count by the count computed at an intensity of 130 dB SPL, since above this intensity the spike count/intensity functions have slopes that approach zero. (A similar normalization procedure was not possible for the PCAP area data because we had not recorded responses at a sufficiently high intensity.) After normalization, equal-normalized-count contours were constructed as described previously for the equal-count contours and the results are shown in Fig. 3. Equal-normalized-count contours become increasingly flat as intensity increases and are compressed at low frequencies, comparing favorably to the equal-loudness contours in Fig. 2(D). However, for reference intensities above 70–80 dB SPL and at low frequencies, the normalized-count contours do not show the decompression seen in the equal-loudness contours. Also, the suggested normalization leaves the rate-of-growth (in dB/dB) of the auditory nerve spike count unchanged, and thus the disagreement between loudness slopes and the spike count slopes persists. However, given the relatively large range of slopes for growth of loudness that have been reported, as discussed above, perhaps this disagreement is not as important as the improved agreement between the contours.

Another important consideration is that the “equalization” function, whatever it may be, may depend on auditory experience, that is, it may be learned and also demonstrate plasticity. For example, the normalization factor used above

may change if neurons in the cochlea die or become inactive due to aging or acoustic trauma. Thus it is possible that the equalization function for a hearing-impaired listener may be different from that of normal listeners. When a hearing aid is prescribed for a hearing-impaired listener, the amplification of sound in different frequency regions may conflict with the person's learned “equalization” function, which could be a source of some of the discomfort associated with the new use of a hearing aid. It is further possible that, over time, listeners may be able to adapt their equalization functions to the new experience provided by use of a hearing aid.

ACKNOWLEDGMENTS

This research was supported by grant NSF IBN-9129262. When writing the manuscript the second author was supported by NIH Grant DC00979. We thank Anita Sterns for help preparing the figures. We also thank an anonymous reviewer for several helpful comments.

- Békésy, G. von (1960). *Experiments in Hearing* (McGraw-Hill, New York).
- 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.
- Carney, L. H. (1994). “Spatiotemporal encoding of sound level: Models for normal encoding and recruitment of loudness,” *Hearing Res.* **76**, 31–44.
- Delgutte, B. (1996). “Physiological models for basic auditory percepts,” in *Auditory Computation*, edited by H. L. Hawkins, T. A. McMullen, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 157–220.
- Doucet, J. R., and Relkin, E. M. (1995). “The peristimulus compound action potential: A new method for recording a compound potential from the chinchilla auditory nerve,” *Auditory Neurosci.* **1**, 151–168.
- Doucet, J. R., and Relkin, E. M. (1997). “Neural contributions to the per-stimulus compound action potential: Implications for measuring the growth of the auditory nerve spike count as a function of stimulus intensity,” *J. Acoust. Soc. Am.* **101**, 2720–2734.
- Evans, E. F. (1981). “The dynamic range problem: Place and time coding at the level of the cochlear nerve and nucleus,” in *Neuronal Mechanisms in Hearing*, edited by J. Syka and L. Aikin (Plenum, New York), pp. 69–85.
- Fletcher, H. (1953). *Speech and Hearing in Communication* (Van Nostrand-Reinhold, Princeton).
- Fletcher, H., and Munson, W. (1933). “Loudness, its definition, measurement, and calculation,” *J. Acoust. Soc. Am.* **5**, 82–108.
- Goldstein, J. L. (1974). “Is the power law simply related to the driven spike response rate from the whole auditory nerve?,” in *Sensation and Measurement—Papers in Honor of S. S. Stevens*, edited by H. R. Moskowitz, B. Scharf, and J. C. Stevens (Reidel, Dordrecht), pp. 223–229.
- Hellman, R. P. (1974). “Effect of spread of excitation on the loudness function at 250 Hz,” in *Sensation and Measurement—Papers in Honor of S. S. Stevens*, edited by H. R. Moskowitz, B. Scharf, and J. C. Stevens (Dordrecht, Reidel).
- Hellman, R. P. (1978). “Dependence of loudness growth on skirts of excitation patterns,” *J. Acoust. Soc. Am.* **63**, 1114–1119.
- Hellman, R. P., and Meiselman, C. H. (1986). “Is high frequency hearing necessary for normal loudness growth at low frequencies?” in *Proceedings of the Twelfth International Congress on Acoustics*, Vol. 1, pp. B11–B15 (Toronto, Canada).
- Hellman, R. P., and Meiselman, C. H. (1988). “Prediction of individual loudness exponents from cross-modality matching,” *J. Speech Hear. Res.* **31**, 605–615.
- Hellman, R. P., and Meiselman, C. H. (1990). “Loudness relations for individuals and groups in normal and impaired hearing,” *J. Acoust. Soc. Am.* **88**, 2596–2606.
- Hellman, R. P., and Zwillocki, J. J. (1968). “Loudness determination at low sound frequencies,” *J. Acoust. Soc. Am.* **43**, 1618–1627.
- Houtsma, A. J. M., Durlach, N. I., and Braida, L. D. (1980). “Intensity perception XI. Experimental results on the relation of intensity resolution to loudness matching,” *J. Acoust. Soc. Am.* **68**, 807–813.
- Howes, W. L. (1974). “Loudness function derived from data on electrical discharge rates in auditory nerve fibers,” *Acustica* **30**, 247–259.

- Jeng, P. S-F. (1992). "Loudness predictions using a physiologically based auditory model," Doctoral dissertation, City University of New York.
- Johnson, D. H. (1980). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," *J. Acoust. Soc. Am.* **68**, 1115–1122.
- Kim, D. O., and Molnar, C. E. (1979). "A population study of cochlear nerve fibers: Comparison of spatial distributions of average-rate and phase-locking measures of responses to single tones," *J. Neurophysiol.* **42**, 16–30.
- Lachs, G., Al-Shaikh, R., Bi, Q., Saia, R. A., and Teich, M. C. (1984). "A neural-counting model based on physiological characteristics of the peripheral auditory system. V. Application to loudness estimation and intensity discrimination," *IEEE Trans. Systems, Man Cybernet.* **SMC-14**, 819–836.
- Moore, B. C. J. (1986). *Frequency Selectivity in Hearing* (Academic, London).
- 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., Hess, R. F., and Birchall, J. P. (1985). "Effects of flanking noise bands on the rate of growth of loudness of tones in normal and recruiting ears," *J. Acoust. Soc. Am.* **77**, 1505–1513.
- Pickles, J. O. (1983). "Auditory-nerve correlates of loudness summation with stimulus bandwidth, in normal and pathological cochleae," *Hearing Res.* **12**, 239–250.
- Riesz, R. R. (1933). "The relationship between loudness and the minimum perceptible increment of intensity," *J. Acoust. Soc. Am.* **4**, 211–216.
- Ruggero, M. A. (1992). "Responses to sound of the basilar membrane of the mammalian cochlea," *Curr. Opin. Neurobiol.* **1**, 215–220.
- Scharf, B. (1978). "Loudness," in *Handbook of Perception, Volume IV*, edited by E. C. Carterette and M. P. Friedman (Academic, New York), pp. 187–242.
- Schlauch, R. S. (1994). "Intensity resolution and loudness in high-pass noise," *J. Acoust. Soc. Am.* **95**, 2171–2179.
- Schneider, B. A., and Parker, S. (1987). "Intensity discrimination and loudness for tones in notched noise," *Percept. Psychophys.* **41**, 253–261.
- Viemeister, N. F. (1974). "Intensity discrimination of noise in the presence of band-reject noise," *J. Acoust. Soc. Am.* **56**, 1594–1600.
- Viemeister, N. F. (1983). "Auditory intensity discrimination at high frequencies in the presence of noise," *Science* **221**, 1206–1208.
- Viemeister, N. F. (1988). "Psychophysical aspects of auditory intensity coding," in *Functions of the Auditory System*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 213–242.
- Viemeister, N. F., and Bacon, S. P. (1988). "Intensity discrimination, increment detection, and magnitude estimation for 1-kHz tone," *J. Acoust. Soc. Am.* **84**, 172–178.
- Zeng, F.-G., and Turner, C. W. (1991). "Binaural loudness matches in unilaterally impaired listeners," *Q. Rev. Exp. Psychol.* **43**, 565–583.
- Zwislocki, J. J. (1965). "Analysis of some auditory characteristics," *Handb. Math. Psychol.* **3**, 1–97.

Efferent projections of a physiologically characterized region of the inferior colliculus of the young adult CBA mouse

Robert D. Frisina,^{a)} Joseph P. Walton, Martha A. Lynch-Armour, and Darrel A. Klotz
Otolaryngology Division, Department of Surgery, University of Rochester School of Medicine and Dentistry,
601 Elmwood Avenue, Rochester, New York 14642-8629

(Received 20 December 1995; revised 4 November 1996; accepted 17 November 1996)

The present investigation is part of an ongoing series of studies aimed at discerning the neural bases of presbycusis. Presbycusis is a sensory perceptual disorder involving loss of high-pitch hearing and reduced ability to process biologically relevant acoustic signals in noisy environments. The purpose of the present experiment was to delineate the efferent projections of a functionally characterized region of the dorsomedial inferior colliculus (IC, auditory midbrain) in young, adult CBA mice. The CBA strain's progressive loss of hearing over its lifespan approximates many aspects of the mild-to-moderate hearing loss experienced by a significant number of humans suffering from presbycusis. Focal, iontophoretic injections of HRP were made in the 18–24 kHz region of dorsomedial IC of the CBA strain following physiological mapping experiments. Serial sections were reacted with a chromagen, counterstained and examined for anterogradely labeled fibers and boutons. Efferent projections were observed *ipsilaterally* in: medial and ventral divisions of the medial geniculate body (MGB); middle layers of the superior colliculus; central gray; and external nucleus (E), dorsal cortex (DC) and central nucleus of IC. *Contralaterally*, labeled fibers and boutons were seen in the IC at a location homologous to the injection site, as well as in E and DC. A small projection was noted in contralateral MGB. These findings in young, adult mice with normal hearing can now serve as a baseline for similar experiments being conducted in mice and animals of other species of older ages and with varying degrees of hearing loss. © 1997 Acoustical Society of America. [S0001-4966(97)01005-9]

PACS numbers: 43.64.Fy, 43.64.Qh [RLS]

ABBREVIATIONS

AR—Acoustic radiation
BF—Best frequency of a single unit
BIC—Brachium of IC
Cb—Cerebellum
CG—Central gray, periaqueductal gray
CIC—Intercollicular commissure
CP—Cerebral peduncle
D—Dorsal division of MGB
DAB—Diaminobenzidine
DC—Dorsal cortex of IC
E—External nucleus of IC
HRP—Horseradish peroxidase

IC—Inferior colliculus, auditory midbrain
ICC—Central nucleus of IC
LL—Lateral lemniscus
M—Medial division of MGB
MGB—Medial geniculate body, auditory thalamus
N.A.—Numerical aperture
NLL—Nuclei of LL
PP—Peripeduncular nucleus
SC—Superior colliculus
SG—Supragenicular nucleus of MGB
SPL—Sound pressure level
TMB—Tetramethylbenzidine
V—Ventral division of MGB

INTRODUCTION

Presbycusis—age-related hearing loss—is a common, progressive disorder involving functional deficits of the ear and brain. In terms of auditory perception, presbycusis is characterized by at least two perceptive difficulties: A high-pitch (high-frequency) loss in sensitivity, and an inability to discriminate speech and other biologically relevant sounds in noisy acoustic environments. The latter problem may involve auditory temporal processing deficits (e.g., McCroskey and Kasten, 1982; Willott, 1991).

Relative to the speech discrimination problem, much is

known about the progressive, high-frequency hearing loss that characterizes presbycusis. To a first approximation, it occurs in most humans and animals as they age (e.g., Human: Royster and Thomas, 1979; Royster *et al.*, 1980; Davis *et al.*, 1991; Pearson *et al.*, 1995; Animal: Henry, 1983). It always involves the loss of receptor cells (hair cells) of the inner ear, generally starting in the high-frequency (basal turn) part of the cochlea (Hinchcliffe, 1962; Bredberg, 1968; Maurer and Rupp, 1979). This loss of high-frequency output channels can then induce changes in the central auditory system where high-frequency sound information is normally processed (Willott and Bross, 1990; Willott *et al.*, 1992, 1993). Additional evidence suggests that other cochlear changes may occur that can contribute to this age-related

^{a)}Electronic mail: rdf@mother.ent.rochester.edu

decline in auditory sensitivity (Johnsson and Hawkins, 1972; Schuknecht *et al.*, 1974; Willott, 1991). For instance, Schmiedt and colleagues in animal physiological experiments have demonstrated that declines in the metabolic efficacy of the stria vascularis, the specialized organ of the cochlear outer wall that produces endolymph, are observed with age (Schulte and Schmiedt, 1992; Schmiedt, 1993). This disruption of endolymph production can reduce the endocochlear potential, the positive voltage of scala media that is necessary for proper transduction of sound waves into the code of the nervous system by cochlear hair cells.

In contrast to the high-frequency sensitivity loss characteristic of presbycusis, the neural etiology of the inability of presbycusis listeners to discriminate speech from noise is more poorly understood (Jerger, 1973). Broadly speaking, there are probably age-related changes in cochlear function and central neural processing that contribute to this speech-in-noise perceptual deficit (e.g., Jerger, 1973; Willott, 1991). The present study is part of a series of investigations aimed at determining the contributions of both the cochlea (peripheral) and brain (central) to this complex-sound perceptual problem.

The two most effectively used animal models for studying the neural bases of presbycusis have been the CBA and C57 strains of mouse (Henry and Chole, 1980; Willott, 1986). The CBA strain is characterized by a slow, progressive, cochlear-induced loss of sensitivity that has a time course similar to human presbycusis (when one corrects for the overall difference in lifespan of mouse and human). Physiological investigations of the central auditory system of CBA mice with *simple* sounds show minimal changes (Willott *et al.*, 1988). However, recent and ongoing investigations of *complex* sound processing in wideband noise have revealed age-related temporal processing deficits in single neurons of the CBA central auditory system (Walton *et al.*, 1995). Specifically, there are fewer neurons of the dorsomedial auditory midbrain (inferior colliculus) capable of encoding temporal gaps in noise in aged CBA mice relative to young adults.

The purpose of the present investigation was to begin to explore age-related changes in the neuroanatomy of the inferior colliculus (IC) where Walton and colleagues demonstrated the age-related functional deficit in complex sound processing. This particular article reports on some of the outputs of that region as delineated by anterograde transport of horseradish peroxidase (HRP) in young adult mice. Unlike most previous HRP studies where injections are made stereotactically, HRP injections of the present study were all made in the same neuroanatomical and physiologically characterized division of the dorsomedial IC based upon single-unit recordings. These findings are not only of use from a comparative anatomical perspective comprising one of the first reports on the outputs of the mouse IC, but they also serve as normative data for comparison with older CBA mice who are undergoing the same experimentation (Lynch-Armour *et al.*, 1995; Frisina *et al.*, 1995). In this way, results of the present study can be compared with output labeling in middle age and older mice to discover age-related changes in the connectivity of the auditory midbrain that may underlie

complex-sound processing deficits in noise, characteristic of presbycusis. These connectivity experiments complement our other ongoing aging studies of the structure and function of the IC in mice (Wilson *et al.*, 1995; Zettel *et al.*, 1995; Kazee *et al.*, 1995).

I. MATERIALS AND METHODS

Animals. Young (2.0–3.5 mon, $N=6$) CBA mice were obtained from Jackson Labs mouse colonies and allowed to age in *ad lib* conditions of food (Purina 5001 Lab Chow) and water on a 12-h light/dark cycle. Prior to each experiment, animals were briefly anesthetized with Metofane (methoxy-flurane) and the external auditory meatus was examined down to the tympanic membrane for blockage. Only animals found to have clear external canals were used.

A. Neurophysiological mapping

1. Surgical preparation

Prior to surgery, animals were anesthetized with methoxyflurane (Metofane) inhalation and kept in a deep state of anesthesia by using maintenance doses. Animals were surgically prepared under aseptic conditions. The skull was shaved and the cranium exposed by reflecting the scalp musculature, after which 2% lidocaine was applied topically. A small, threaded metal tube was attached to the skull surface using cyanoacrylic glue (Super Glue) and dental acrylic. A sharpened tungsten-wire indifferent electrode was implanted to just contact the dura mater. The IC was exposed by making a small craniotomy (0.5×0.5 mm) over the dorsal surface of the IC and covered with gelfoam. The animal was then allowed to recover at least one day prior to the experiment.

Before recording, mice were mildly tranquilized (Taractan 5–12 $\mu\text{l/gm}$) and placed in a form-fitting plastic restraint tube attached to a custom-built stereotactic frame (Schuller *et al.*, 1984). The frame was located in the middle of a heated ($27\text{--}30^\circ\text{C}$), double-walled sound-proofed room (IAC) lined with sound absorbing foam (Sonex). These conditions maintain normal body temperature in unanesthetized small mammals. The threaded tube was then screwed to a rigid bar fixed to the stereotaxic frame.

2. Stimulus presentation

Stimulus generation was computer controlled via a Micro-11/23+ computer system interfaced to programmable stimulus generation and spike acquisition equipment (Lesser *et al.*, 1990). The search stimulus consisted of 100 ms long noise bursts, gated on/off using a 1 ms cosine envelope (Wilsonics Programmable Gate) and produced by a Grasen-Stadler 1-MHz random noise generator. After a neuron was isolated the following protocol was followed: (1) Audio-visual determination of best frequency (BF), with frequency continuously variable from 1 to 80 kHz; (2) rate-intensity functions using noise and tone burst stimuli having 100 ms durations, with 1 ms linear rise-fall time presented at a rate of 4/s; (3) spontaneous rate measurement for 13 s; (4) gap-in-noise series presented at 30 dB *re*: threshold and 65 dB SPL (Walton *et al.*, 1995).

3. Recording, data acquisition and analysis

The responses of single units in the IC were recorded using glass micropipettes (A-M Systems) pulled on a programmable Sachs & Flaming pipette puller and filled with either 3 M KCl having tip impedances of 12–14 Mohm or an HRP solution with tip impedances of 12–14 Mohm. Electrodes were positioned over the IC using X and Y axis manipulators controlled by 1 μm resolution stepping motors. The Z axis was manually controlled for coarse positioning of the electrode over the surface of the IC, while a remote-controlled piezoelectric micropositioner (Burleigh Inchworm PZ-575) was used to advance the electrode through the IC. This stereotaxic coordinate system allowed for session-to-session electrode position accuracy of 200–300 μm . In each animal X, Y, and Z coordinates were referenced to a point on the stereotaxic frame. This allowed for control of electrode position from session to session, and for HRP injection.

Neural activity was amplified (Dagan 2400), filtered from 0.3 to 5 kHz (Krohn-Hite 3202) and input to a window discriminator (BAK MD1) for spike isolation. The TTL output of the discriminator was displayed in a dot raster format on a storage oscilloscope and also served as input to two real-time clocks used for time stamping of spike times. Custom software (Lesser *et al.*, 1990), allowed for spike times to be processed with a precision of 10 μs .

After multiple dorsal/ventral penetrations were made through the IC in which BFs of single neurons or multiunit activity were determined, 10% HRP (Sigma XII) in a 0.5 M solution of KCl, in 0.05 M Tris buffer, pH=7.3, was injected (1.5 μA for 15–20 min) into the center of the region recorded from: dorsomedial IC, tonotopic location of 18–24 kHz. The 18–24 kHz region is part of the tonotopic region defined in mouse by Stiebler and Ehret (1985) and Willott (1986), that extends from the dorsal cortex, through the dorsomedial area to the IC central nucleus. Animals were returned to their cage for 24 h, then given an overdose of sodium pentobarbital, and transcardially perfused with heparinized saline followed by 1.25% glutaraldehyde/1% paraformaldehyde fixative.

Since HRP is transported both anterogradely and retrogradely, interpretation of output fibers is problematical in regions where neural plexuses contain label due to staining of dendrites of retrogradely labeled cell bodies. Thus, HRP is no longer considered to be as useful as other, purely anterograde CNS tracers. Recognizing this limitation, the current study primarily reports on output labeling in regions of the brain where there was no significant retrograde labeling of cell bodies (less than one retrogradely-stained neuron/section). In these regions that lacked stained cells, retrogradely-labeled dendrites could not be confused with anterogradely stained fibers (e.g., MGB, SC). In one case, contralateral ICC, there were a few labeled cell bodies but not enough to account for the extensive fiber staining. Here, only statements about the laminar pattern of the label were made, not whether particular staining was retrograde or anterograde. Regions like the nuclei of the lateral lemniscus, where the anterograde and retrograde label was significantly intermixed, were not included in the present report due to this limitation of HRP. Although not optimal because of this

limitation, the anterograde label of the present report is quite meaningful when integrated with other aspects of this multidisciplinary study such as the input information (Frisina *et al.*, in press), neurophysiological functional characterization of the injection-site region (Walton *et al.*, in press), and behavioral studies done in the same strain (Ison *et al.*, 1996).

B. HRP histochemistry

These procedures were similar to previous studies (Frisina *et al.*, 1989; O'Neill *et al.*, 1989; Shore *et al.*, 1991). Following a wash in 3% (10 min) glycerol in phosphate buffer, the brain was dissected from the skull and stored at 4 $^{\circ}\text{C}$ for approximately 12 h periods in successive solutions of 10%, 15%, and 20% glycerol/buffer. Next, three serial sets of alternate sections were cut on a cryostat at 60 μm . Two were processed with TMB and one counterstained with safranin-O, the third was reacted with DAB (Mesulam, 1982) and counterstained with cresyl violet. Injection sites of 500 to 850 μm in diameter verified that recordings were all from the dorsomedial IC.

To determine the location of anterogradely labeled fibers and possible endings, HRP labeling was examined and drawn with a Nikon microscope equipped with drawing tube and 4x-N.A. 0.1, 10x-N.A. 0.25, 20x-N.A. 0.4, 40x-N.A. 0.65, and 60x-N.A. 0.85, plan achromat objectives. Sections were viewed under light-, and dark-field, and with polarized light. Cytoarchitectonic boundaries of mouse brainstem auditory nuclei were identified utilizing criteria in Willard and Ryugo (1983), Meininger *et al.* (1986) and Hofstetter and Ehret (1992).

Reconstructions shown here were taken from experiments where HRP was injected at the location of one particular unit. This unit's location, the presumed center of the injection site, was used to line up the computer-reconstructed unit BF-location map with the camera lucida drawings of the anatomical section containing the center of the injection site. In this way, the physiological map of unit BFs could be literally overlaid with the anatomical center of the injection site of the appropriate histological section [Fig. 1(b)], as determined by the dense core of HRP reaction product, after correcting for shrinkage (10% DAB, 20% TMB; O'Neill *et al.*, 1989).

Negative findings, such as the lack of labeled fibers or presumed endings in a particular location, will sometimes be reported. These negative results are meaningful in the sense that in the present study, because of the relatively small size of the mouse brain, *all* serial sections from the spinal cord to the frontal lobes were saved, processed and examined for presence of HRP reaction product, which was fairly consistent from section to section within an animal.

All animal procedures were approved by and conducted according to the guidelines for recovery surgery of the University of Rochester's Vivarium and Committee on Animal Resources and the National Institutes of Health (NIH Grant No. P01 AG9524).

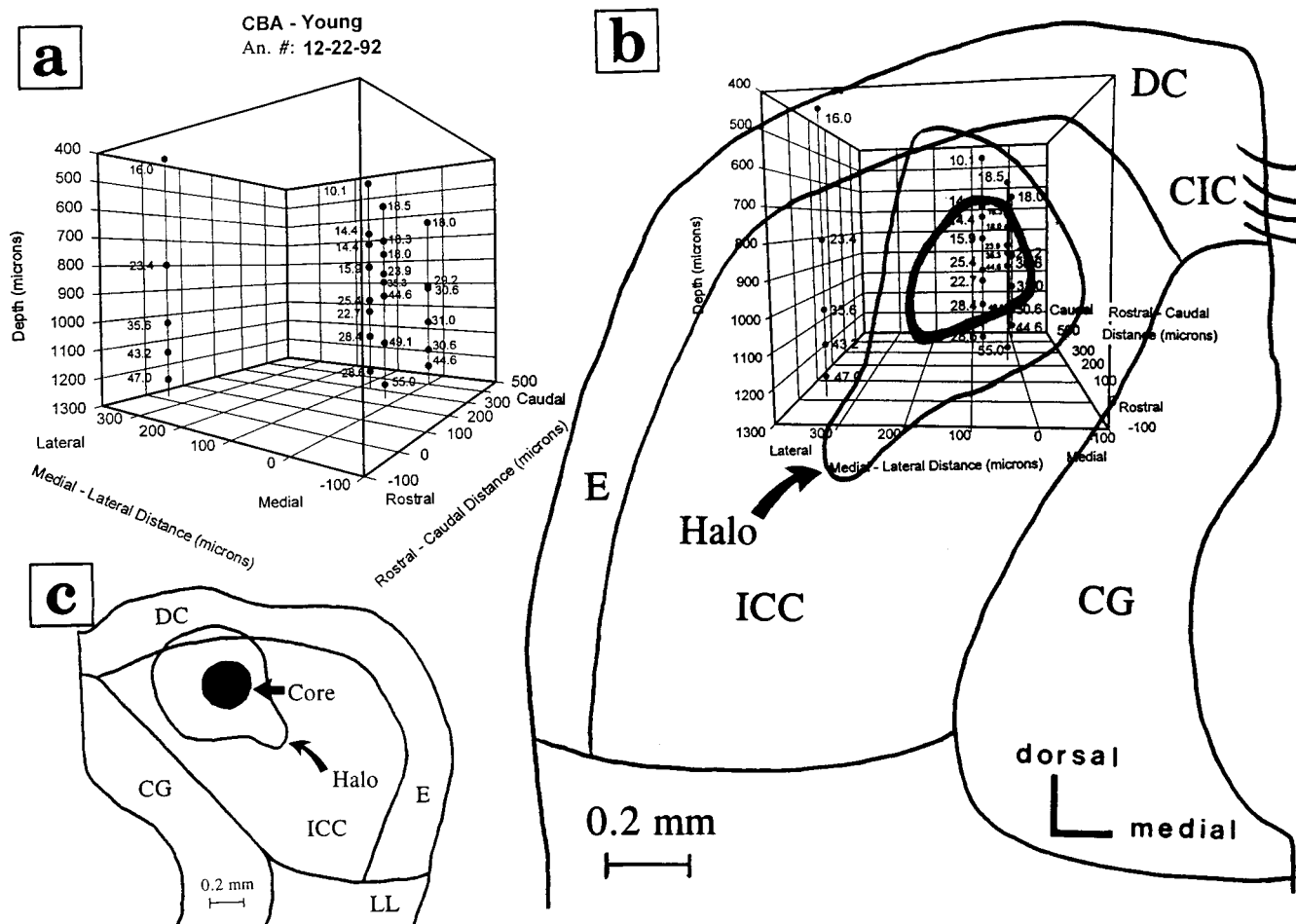


FIG. 1. Physiological recordings and iontophoretic injections of HRP were always centered in the 18–24 kHz region of the dorsomedial IC. Portions of the injection site halo and some of the recordings extended outside of this region as shown here. Boundaries between divisions within the IC are based upon Willard and Ryugo (1983). (a) Example of a physiological map from one of the animals represented in a 3-D graphical format. The BF of each single unit or multi-unit cluster is given in kHz. The graph is rotated so as to provide a medio-rostral vantage point. All axes are in arbitrary micron units, i.e., the zero points of the axes are arbitrary and do not correspond to the center of the injection site. As expected, the BFs increase as one moves deeper in the IC. An.: animal. (b) Correspondence of the physiological BF map (rotated to a paracoronal plane) and the neuroanatomical cytoarchitectonics. The dense core of the injection site has a **bold** border, and the halo a thin border (arrow). Injection sites were placed in dorsomedial IC and halos sometimes extended into DC. The 3-D physiological data are calibrated to the coronal, anatomical section at the 300 micron rostral-caudal plane, i.e., at the injection site center, both anatomically and physiologically. (c) Camera lucida drawing of the dense core (straight arrow) and halo (curved arrow) of the center of the injection site in dorsomedial IC from another animal (An. #: 12). Chromagen: DAB; Counterstain: cresyl violet; Section thickness: 60 μ m. Boundaries between ICC, DC and E should be considered approximate since they are based upon Nissl-stained material (not Golgi). CG: central gray (periaqueductal grey); CIC: commissure of the IC; DC: dorsal cortex of IC; E: external nucleus of IC; IC: inferior colliculus; ICC: central nucleus of IC; LL: lateral lemniscus.

II. RESULTS

A. Dorsomedial IC injection sites

All unit recordings of the present study were centered in the dorsomedial region of the IC, and the centers of the HRP injection sites were placed in the 18–24 kHz region. Figure 1(a) displays a 3-D plot of the BFs of the units from a representative animal. The vantage point of the 3-D plot is a medial position. To correlate the locations of the recording sites with the underlying IC cytoarchitectonics, the physiological map of Fig. 1(a) is rotated to a paracoronal plane (several degrees offset from the standard coronal plane) in Fig. 1(b). Since the anatomical sections were also cut in the coronal plane, this allows direct comparison of the recording sites with IC divisions and the extent of the injection site. The dense core of the injection site is centered in the 20–25 kHz region, and is completely contained in the dorsomedial

division of IC, as delineated by DAB reacted sections counterstained with cresyl violet [Fig. 1(b) and (c)]. The dorsal portion of the halo extends partly into the dorsal cortex (DC) (Morest and Oliver, 1984). As expected from previous physiological studies of mouse IC, neurons in dorsomedial IC, including DC and the central nucleus, form one tonotopic map with low frequencies found dorsally and high frequencies located ventrally (Willott and Urban, 1978; Stiebler and Ehret, 1985). The results of the present study, as exemplified by the data of Fig. 1(a) and (b), are consistent with these previous investigations.

B. Dorsomedial IC outputs to ipsilateral MGB

As determined by extent of HRP label, the primary output region of the murine dorsomedial IC is to the ipsilateral medial geniculate body of the thalamus (MGB). Figure 2(a)–

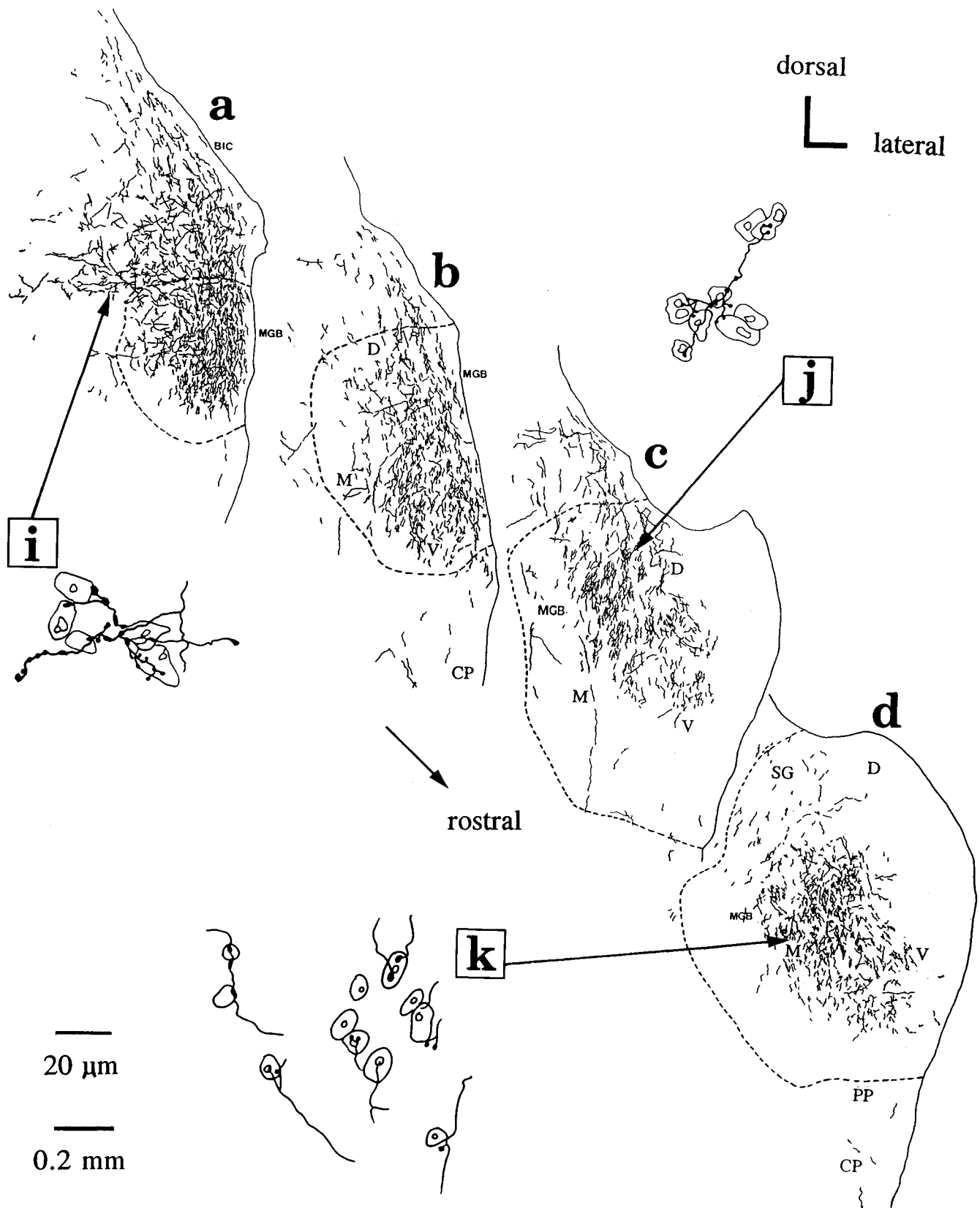


FIG. 2. Most processes and boutons in the ipsilateral MGB stained with anterograde label of HRP were found in the medial (M) and ventral (V) divisions. (a)–(h) Camera lucida drawings of 60 μm serial coronal sections extending from the caudal border of MGB to its most rostral extent. Sections are spaced 180 μm apart. 20x obj., 0.2 mm calibration bar. (i)–(n) High magnification drawings of HRP-stained fibers and boutons in the vicinity of MGB Nissl-stained neurons. 60x air obj.; 20 μm calibration bar. Coronal sections; DAB Chromagen; Cresyl violet counterstain; An.#: 19; ar: acoustic radiation; BIC: brachium of inferior colliculus; CP: cerebral peduncle; D: dorsal division of MGB; MGB: medial geniculate body; PP: peripeduncular nucleus; SG: supragenicular nucleus of MGB.

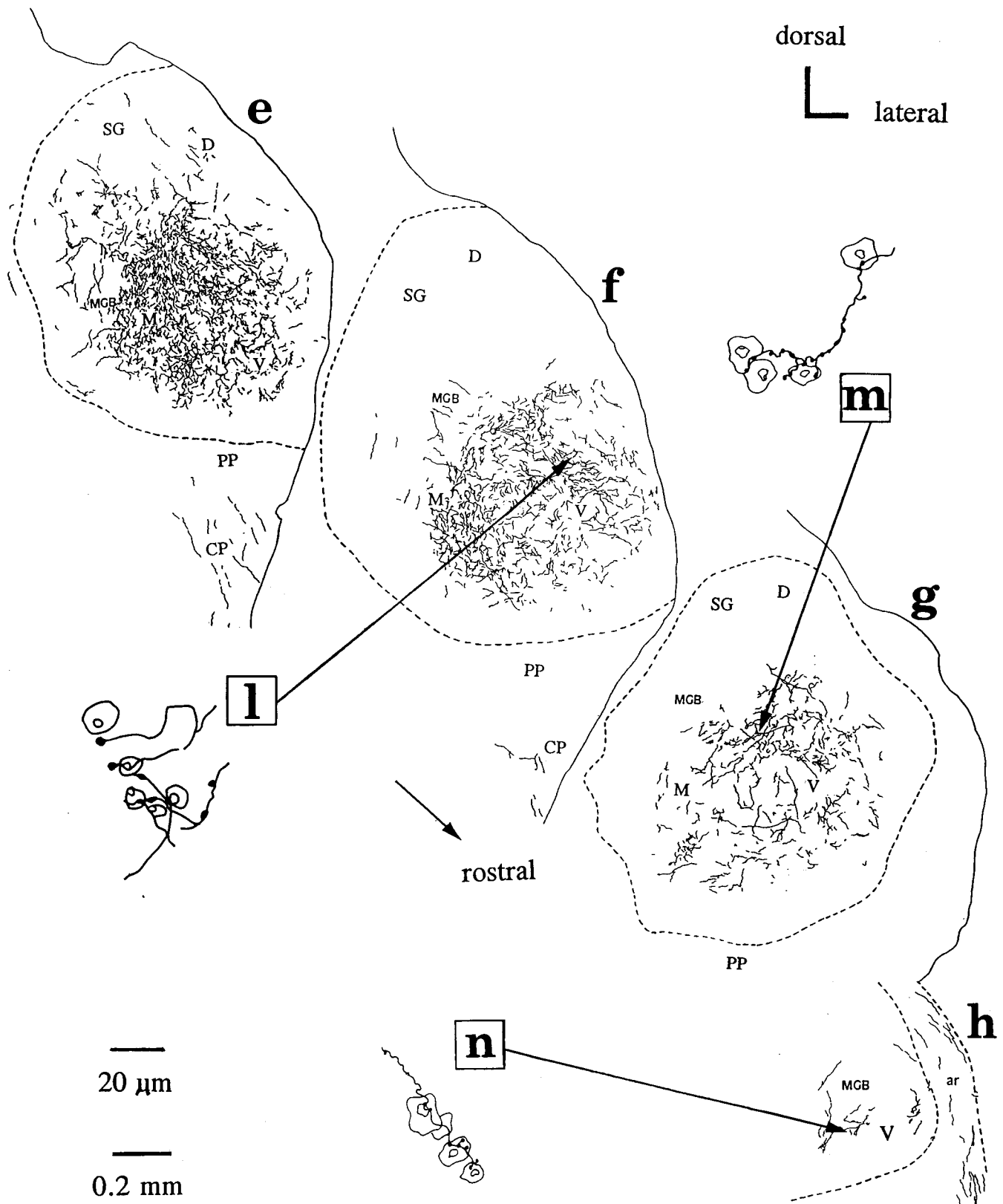


FIG. 2. (Continued.)

(h) shows a serial set of camera lucida drawings of DAB reacted sections counterstained with cresyl violet. Figure 2(a) is the most caudal of the coronal sections, and Fig. 2(h) the most rostral. Figure 2(a) includes the most caudal portion

of MGB, containing a large complement of HRP-stained processes, as does the brachium of the IC (BIC) overlying the MGB at this level. *Boutons en passant* and *boutons terminaux* can be found in the vicinity of, and directly overlay-

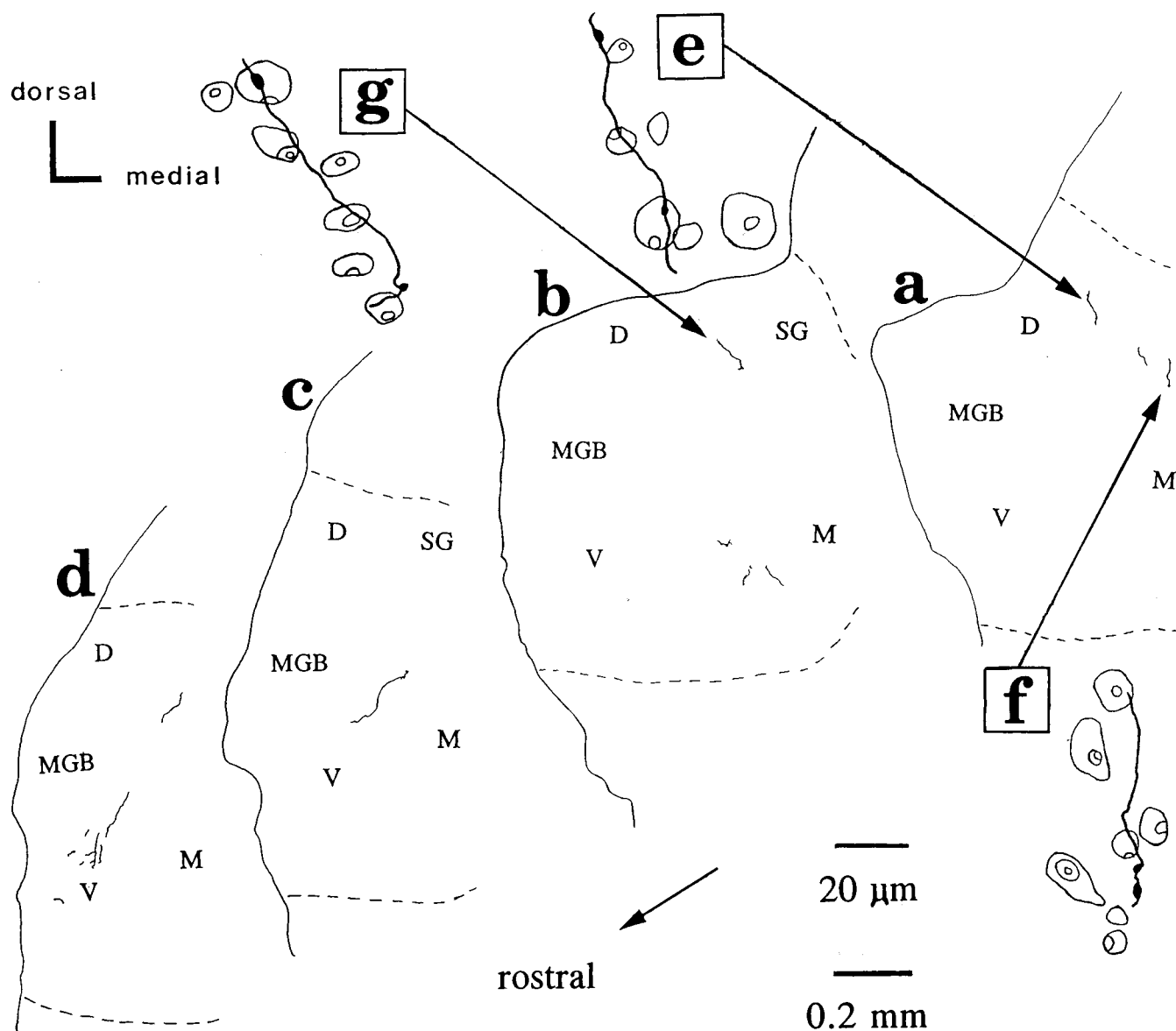


FIG. 3. A few processes and boutons in the contralateral MGB stained with anterograde label of HRP were found in the medial (M), dorsal (D) and ventral (V) divisions. (a)–(d) Camera lucida drawings of 60 μm serial coronal sections extending from a caudal level to a more rostral extent. Sections are spaced 180 μm apart. 20x obj., 0.2 mm calibration bar. (e)–(g) High magnification drawings of HRP-stained fibers and boutons in the vicinity of MGB Nissl-stained neurons. 60x air obj.; 20 μm calibration bar. Coronal sections; DAB Chromagen; Cresyl violet counterstain; An.#: 19; MGB: medial geniculate body; SG: supragenicular nucleus of MGB.

ing, counterstained cells of the ipsilateral BIC [Fig. 2(i)]. As one moves rostrally, as drawn in Fig. 2(b) and (c), labeled processes can still be found in BIC lying superior to the dorsal division of MGB (D), as well as in BIC fibers as they descend through D and into the medial (M) and ventral (V) divisions of MGB. Occasionally, branches of these descending BIC fibers give off stained collaterals exhibiting *boutons en passant* and *boutons terminaux* on or near the perikarya of counterstained MGB neurons of D [Fig. 2(j)].

Most HRP anterograde labeled processes and boutons were found in the middle to rostral regions of the ipsilateral MGB in V and M [Fig. 2(d)–(g)]. Examples of labeled processes characterized with *boutons en passant* and *boutons terminaux* were found frequently in V and M [Fig. 2(k)–(m)]. In these same sections [Fig. 2(d)–(g)], very little label was observed in D or in the supragenicular nucleus of MGB

(SG). A few labeled fibers were also observed in the cerebral peduncles (CP) ventral to V [Fig. 2(b), 2(d)–(f)].

Some labeling can also be seen in the most rostral portions of MGB where only V exists [Fig. 2(h)]. Sometimes a few labeled fibers extended into the acoustic radiation (ar) at this level. *Boutons en passant* and *boutons terminaux* can be found here on and around counterstained cell bodies [Fig. 2(n)].

C. Dorsomedial IC outputs to contralateral MGB

A very small projection to contralateral MGB was observed. The amount of label was not enough to definitively determine whether there was a differential projection to one or more of the MGB divisions, as was observed ipsilaterally [Fig. 3(a)–(d)]. Nevertheless, this projection may have some

functional significance for binaural processing in the sense that *boutons en passant* were observed in D [Fig. 3(e) and 3(g)] and M [Fig. 3(f)] overlying or nearby counterstained perikarya.

D. Dorsomedial IC outputs to contralateral IC

As displayed in Fig. 4(a)–(c), many labeled fibers were present in the contralateral IC. The distribution of label in the ICC dorsomedial division in a location homologous to the injection site was highly laminar in appearance, with the long axis of the labeled lamina extending in a dorsomedial-to-ventrolateral direction. Labeled fibers also projected into DC, and to a lesser extent into E. Labeled *boutons en passant* and *boutons terminaux* were quite plentiful in the dorsomedial region, of both ICC and DC [Fig. 4(e)–(f)]. Boutons were less frequent in E. In many instances boutons overlaid, or were adjacent to, the counterstained somata of IC principal cells [e.g., Fig. 4(e)–(f)].

E. Dorsomedial IC outputs to ipsilateral IC

Rich plexuses of labeled processes were also observed in the ipsilateral IC [Fig. 4(a)–(c)]. The regions containing HRP-labeled plexuses included ICC in a somewhat laminar way [Fig. 4(a)], DC [Fig. 4(a)–(c)] and E [Fig. 4(a)]. Theoretically, labeled processes in ipsilateral IC could come from several sources: (1) Anterogradely-labeled axons and endings from neurons with cell bodies located in the uptake region of the injection site; (2) Collaterals of axons of retrogradely-labeled perikarya of neurons from lower auditory centers that send axonal terminals into the effective uptake region; and (3) Dendritic processes of darkly-filled neurons whose perikarya are in the ipsilateral IC outside of the effective uptake region but that send dendritic or axonal processes into the effective uptake region. There were not enough of these darkly filled cells in the ipsilateral IC to account for the heavily stained neuropil as shown in Fig. 4(a). Thus it is likely that many of the labeled fibers of ipsilateral ICC, DC, and E [Fig. 4(a)–(c)] were axons of neurons projecting out of the injection site, or from lower auditory centers. Labeled fibers with *boutons en passant* and *boutons terminaux* were numerous in the ipsilateral ICC, DC, and E, with an appearance that resembled their counterparts in the contralateral IC [e.g., Fig. 4(e)–(f)].

F. Dorsomedial IC outputs to CG

Labeled fibers were seen bilaterally in CG [Fig. 4(a)–(c)]. They were consistent across animals, but relatively few in number, and concentrated in the dorsal portion of CG, coming in from the injection site or its homologous region in contralateral IC. This projection probably has functional significance since *boutons en passant* and *boutons terminaux* were observed, an example of which is presented in Fig. 4(d). As in other output regions, anterogradely-labeled boutons were found overlying or in the immediate vicinity of counterstained somata.

G. Dorsomedial IC outputs to ipsilateral SC

Auditory and visual information are integrated at the midbrain level. This was manifested in the present study by the presence of HRP labeled processes in the superior colliculus (SC). Caudally, as displayed in Fig. 5(a), stained fibers and labeled *boutons en passant* and *boutons terminaux* [Fig. 5(f)–(h)], were observed in the medial and lateral portions of the middle and deep layers of SC. Moving rostrally, label was concentrated in a laminar way in the central portions of SC [Fig. 5(b)–(e)]. Again, these stained fibers exhibited labeled boutons directly overlying or adjacent to counterstained SC neurons [Fig. 5(j) and (i)].

H. Dorsomedial IC outputs to LL

Many labeled processes were observed bilaterally in the lateral lemniscus (LL). Some of these processes may have been branches of ascending axons of cell bodies in the superior olivary complex and cochlear nucleus that were retrogradely labeled by HRP. Other labeled processes were portions of dendritic arborizations of neurons of the nuclei of LL (NLL) that were retrogradely stained with HRP. For determining if the NLL receive inputs from the injection site, the HRP procedure of the present investigation has a limitation: Due to the high density of labeled processes in the NLL it is not reliably possible to distinguish dendrites and axonal branches that were retrogradely labeled (inputs) from those fibers and endings that are anterogradely labeled (outputs). Thus, the nature of the outputs of the dorsomedial IC, 18–24 kHz region, to the NLL must await further investigation utilizing a tracer that is strictly anterograde.

III. DISCUSSION

A. Comparative Analysis

1. Murine

There have been no previous studies of the connectivity of the auditory system of the CBA strain. For mice in general there have been no previously published articles on the outputs of the IC, except for Gonzalez-Hernandez *et al.* (1986) who described some of the commissural connections of the albino mouse. Willard and Ryugo (1983) made a preliminary report that IC connections of mice are similar to rat. Thus, the present study serves as a basis of comparison for future mouse studies of IC connections in all strains (Fig. 6).

2. Mammalian MGB

Most previous studies of the outputs of the IC have been performed in cat or bat. The results of the present investigation in mouse are similar to previous feline findings (Morest, 1965a,b; Andersen *et al.*, 1980; Oliver, 1984; Morest and Winer, 1986) in the following ways. First, mouse and cat both show bilateral outputs to MGB, but with a very strong ipsilateral bias. Second, most of the ipsilateral MGB outputs project to the medial and ventral divisions, sparing the dorsal and supragenulate regions for the most part (Rasmussen, 1964). Since the present study examined just one region of

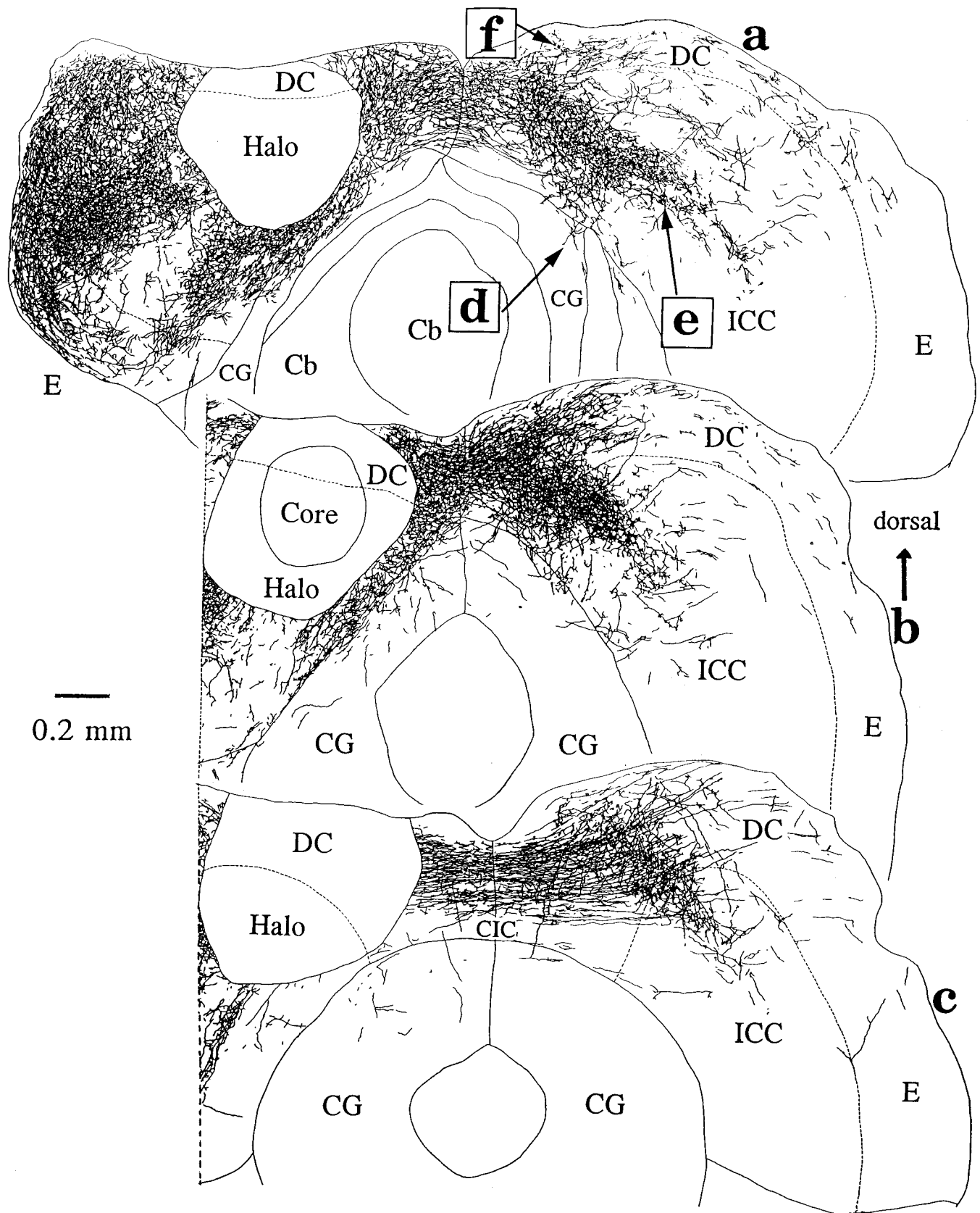


FIG. 4. The dorsomedial IC sent labeled processes to all major divisions of the ipsilateral and contralateral IC, as well as to CG bilaterally. (a)–(c) Camera lucida drawings of 60 μm serial coronal sections extending from a caudal level (a) to a more rostral extent (c). Sections are spaced 180 μm apart. Dashed lines that show the boundaries between ICC, DC and E should be considered approximate, since they are based upon Nissl-stained material (not Golgi). 20x obj., 0.2 mm calibration bar. (d)–(f) High magnification drawings of HRP-stained fibers and boutons in the vicinity of Nissl-stained neurons. 60x air obj.; 20 μm calibration bar. Coronal sections; DAB Chromagen; Cresyl violet counterstain; An.#: 14; Cb: cerebellum; CG: central gray; DC: dorsal cortex of IC; E: external nucleus of IC; IC: inferior colliculus; ICC: central nucleus of IC.

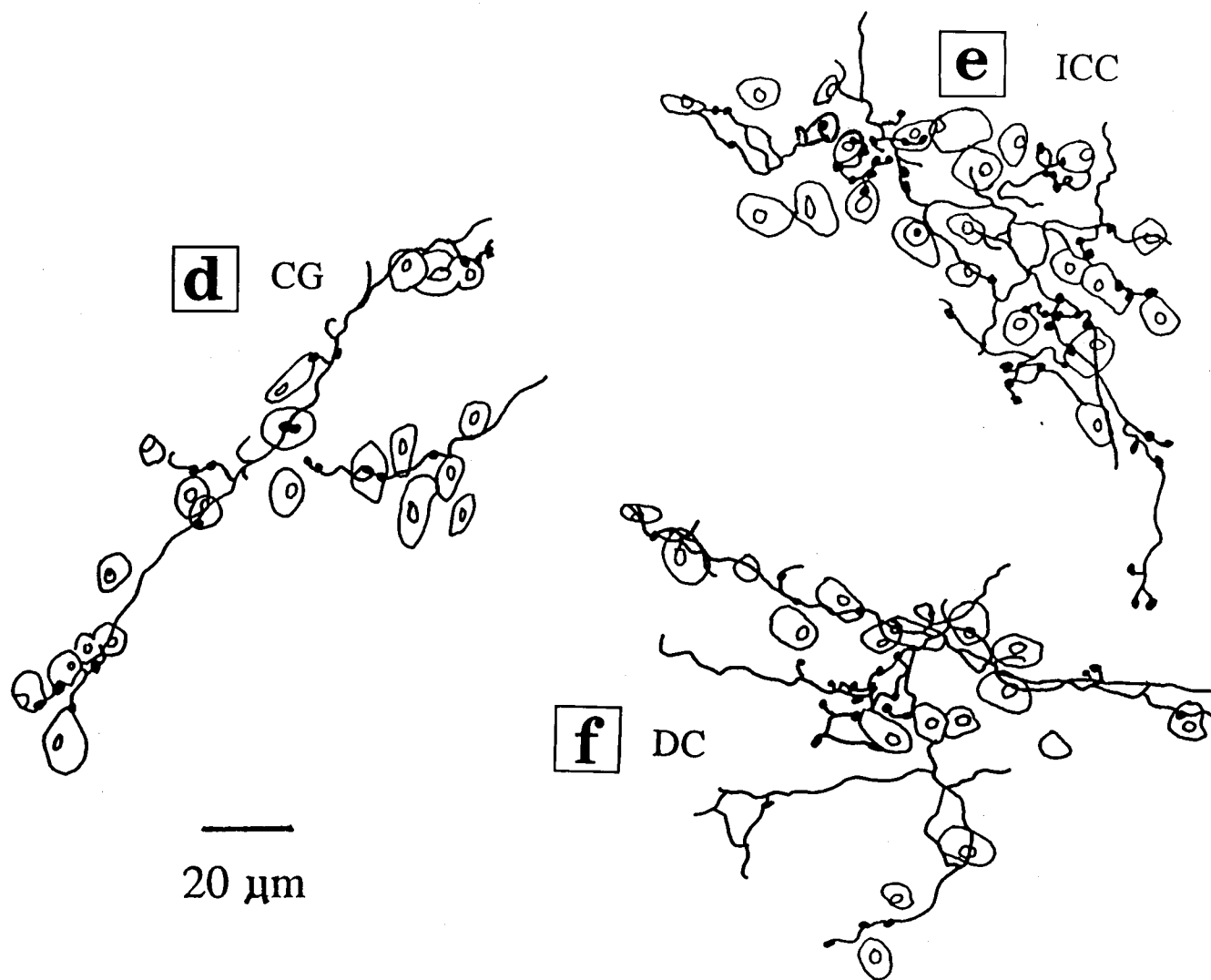


FIG. 4. (Continued.)

IC, it is possible that other regions of mouse IC could project more strongly to dorsal regions of MGB, as future investigations may reveal.

Studies in bats are similar to mouse and cat in the sense that the major MGB outputs from IC are much more extensive ipsilaterally (e.g., Schweizer, 1981). However, the bat studies reveal significant projections of certain regions of IC to the dorsal MGB, unlike the cat and mouse which highly favor the ventral and medial divisions. For example, Frisina *et al.* (1989) found that outputs of the FM₂ area of the anterolateral division of the mustached bat IC preferentially project to the dorsal and medial MGB. Wenstrup and colleagues (Wenstrup *et al.*, 1994; Wenstrup and Grose, 1995), in a series of anterograde and retrograde labeling experiments in mustached bat have demonstrated colliculothalamic pathways involving significant target zones in all three divisions of the MGB including the dorsal division. These MGB input differences between bats and the other mammals relate to the presence of combination sensitive neurons of the bat MGB which are intimately involved in the pulse-echo neural computations performed by the bat central auditory system for target ranging and other capabilities involved in biosonar

navigation. Lastly, as seen here in mouse, labeling in the ipsilateral cerebral peduncle just ventral to the MGB was noted in bat previously by Schweizer (1981).

3. Other mammalian output regions

Anterogradely labeled fibers and boutons of moderate number were also observed here in the mouse in the ipsilateral superior colliculus and bilaterally in central gray. These connections allow the integration of auditory information with other brainstem computational systems involving visual processing and vocalization production (e.g., Covey *et al.*, 1987; Suga *et al.*, 1974). These output regions of IC have been observed in previous studies of monkey (Moore and Goldberg, 1966; Moore *et al.*, 1977), cat (Powell and Hatten, 1969; Edwards *et al.*, 1979), bat (Schweizer, 1981; Frisina *et al.*, 1989) and kangaroo rat (Carey and Webster, 1971).

In the present experiment, a strong, laminar contralateral projection was observed to the homologous region of the dorsomedial IC, although in bats this reciprocal connection is not as strong as observed here for mouse (Schweizer, 1981; Frisina *et al.*, 1989).

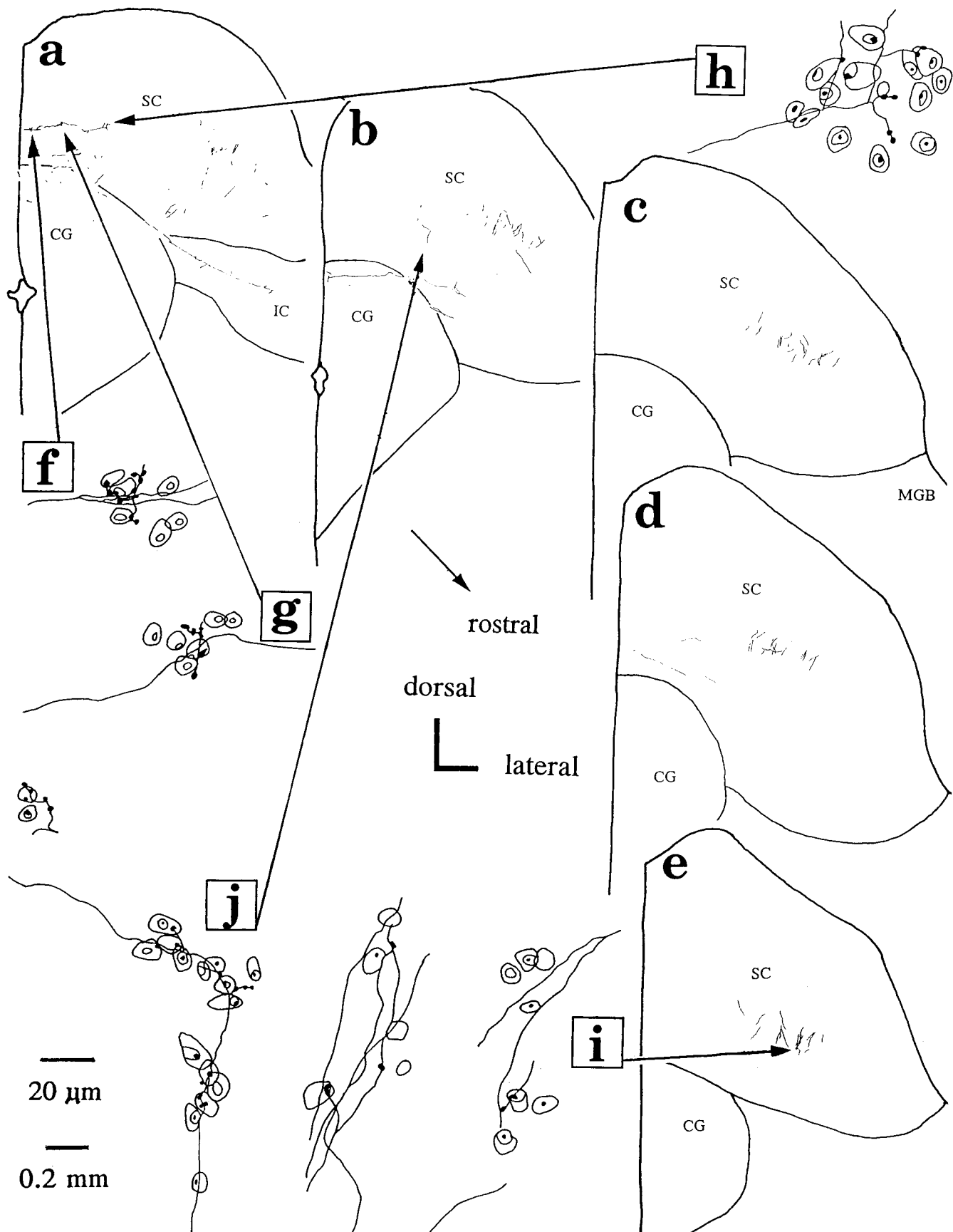


FIG. 5. The dorsomedial IC was found to provide inputs to the visual system via labeled fibers and boutons to middle layers of the ipsilateral SC. (a)–(e) Camera lucida drawings of 60 μm serial coronal sections extending from a caudal level to a more rostral extent. Sections are spaced 180 μm apart. 20x obj., 0.2 mm calibration bar. (f)–(j) High magnification drawings of HRP-stained fibers and boutons in the vicinity of SC Nissl-stained neurons. 60x air obj.; 20 μm calibration bar. Coronal sections; DAB Chromagen; Cresyl violet counterstain; An.#: 19; CG: central gray; IC: inferior colliculus; MGB: medial geniculate body; SC: superior colliculus.

CBA Mouse Model Outputs of Young Adult IC Dorsomedial Div., 18–24 kHz Region

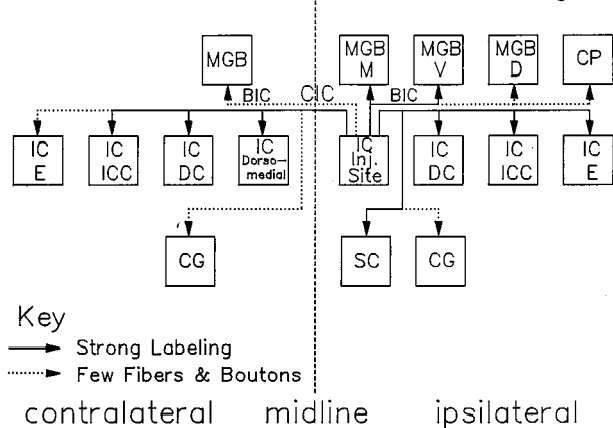


FIG. 6. Schematic summary diagram of the outputs of the CBA mouse dorsomedial IC, 18–24 kHz region as determined by focal, iontophoretic injections of HRP into physiologically-defined region. The outputs are primarily ipsilateral and at the mesencephalic (midbrain) and diencephalic (thalamic) levels of the brainstem. All abbreviations are given in the abbreviation list given above.

Lastly, some prior examinations of the outputs of the auditory midbrain have revealed a link between the auditory system and the cerebellum via the tectopontine tract through the lateral pontine nuclei (monkey: Moore and Goldberg, 1966; cat: Powell and Hatten, 1969; Kawamura, 1975; bat: Schweizer, 1981; Frisina *et al.*, 1989; kangaroo rat: Carey and Webster, 1971). This pathway was not observed in the present murine investigation. There are at least two reasons why this sensorimotor pathway was not labeled here. First, it may not exist in the CBA strain of mouse. Secondly and more probably, the pathway may not originate from the particular region studied here (dorsomedial IC, 18–24 kHz), but may originate from other cytoarchitectonic divisions of the CBA IC, or perhaps from another functional region having lower- or higher-frequency sound processing capability.

B. Implications for neural bases of presbycusis

By itself, the present investigation is limited in terms of its contribution to our understanding of age-related hearing loss. However, as a part of a systematic series of investigations into the functional anatomy of the central auditory system of CBA and C57 strains of mice, involving parallel human and animal experimentation, the present study is invaluable (Fig. 6). Knowledge of the outputs of the IC in the normal, young adult mouse now allows for determination of age-related and/or peripherally-induced deficits in the central auditory system that may contribute to the inability of older organisms to effectively process biologically-relevant acoustic features in noisy environments (Lynch-Armour *et al.*, 1995; Frisina *et al.*, 1995; Wilson *et al.*, 1995; Zettel *et al.*, 1995; Kazee *et al.*, 1995). Specifically, the extent, location or pattern of staining of the efferent projections of IC could change with age, thus altering the normal processing of acoustic information by the central auditory system.

ACKNOWLEDGMENTS

This study was supported by a grant from the National Institute of Aging of NIH: P01 AG09524, and the International Center for Hearing and Speech Research, Rochester, NY. We gratefully thank Jonathan Byrd and Patty Bardeen for their assistance in carrying out this study.

- Andersen, R. A., Roth, G. L., Aitkin, L. M., and Merzenich, M. M. (1980). "The efferent projections of the central nucleus and pericentral nucleus of the inferior colliculus," *J. Comp. Neurol.* **194**, 649–662.
- Bredberg, G. (1968). "Cellular pattern and nerve supply of the human organ of Corti," *Acta Otol. Suppl.* **236**, 1–127.
- Carey, Ch. L., and Webster, D. B. (1971). "Ascending and descending projections of the inferior colliculus in the Kangaroo rat (*Dipodomys merriami*)," *Brain Behav. Evol.* **4**, 400–412.
- Covey, E., Hall, W. C., and Kobler, J. B. (1987). "Subcortical connections of the superior colliculus in the Mustache bat *Pteronotus parnellii*," *J. Comp. Neurol.* **263**, 179–197.
- Davis, A. C., Ostri, B., and Parving, A. (1991). "Longitudinal study of hearing," *Acta Otol. (Stockholm)* **476**, 16–22.
- Edwards, St. B., Ginsburgh, Ch. L., Henkel, C. K., and Stein, B. E. (1979). "Sources of subcortical projections to the superior colliculus of the cat," *J. Comp. Neurol.* **184**, 309–330.
- Frisina, R. D., O'Neill, W. E., and Zettel, M. L. (1989). "Functional organization of mustached bat inferior colliculus: II. Connections of the FM₂ region," *J. Comp. Neurol.* **284**, 85–107.
- Frisina, R. D., Walton, J. P., Lynch-Armour, M. A., and Byrd, J. D. (1995). "Aging-related changes in connections to a functionally-characterized region of the inferior colliculus of the CBA mouse model of presbycusis," *Soc. Neurosci. Abst.* **21**, 908.
- Frisina, R. D., Walton, J. P., Lynch-Armour, M. A., and Byrd, J. D. (1997). "Inputs to a physiologically-characterized region of the inferior colliculus of the young adult CBA mouse," *Hearing Res.* (in press).
- Gonzalez-Hernandez, T. H., Meyer, G., and Ferres-Torres, R. (1986). "The commissural interconnections of the inferior colliculus in the albino mouse," *Brain Res.* **368**, 268–276.
- Henry, K. R. (1983). "Ageing and audition," in *The Auditory Psychobiology of the Mouse*, edited by J. F. Willott (Thomas, Springfield, IL), pp. 470–493.
- Henry, K. R., and Chole, R. A. (1980). "Genotypic differences in behavioral, physiological and anatomical expressions of age-related hearing loss in the laboratory mouse," *Audiology* **19**, 369–383.
- Hinchcliffe, R. (1962). "The anatomical locus of presbycusis," *J. Speech Hear. Disorder* **27**, 301–310.
- Hofstetter, K. M., and Ehret, G. (1992). "The auditory cortex of the mouse: Connections of the ultrasonic field," *J. Comp. Neurol.* **323**, 370–386.
- Ison, J. R., Gutierrez, E., Agrawal, P., Pak, J., and Vaughn, W. (1996). "Reflex inhibition by partially filled gaps in developing and aged mice. Ontogenetic changes in processing efficiency, not temporal acuity or audibility," *Soc. Neurosci. Abst.* **22**, 1822.
- Jerger, J. (1973). "Audiological findings in aging," *Adv. Oto-Rhino-Laryngol.* **20**, 115–124.
- Johnsson, L.-G. and Hawkins, Jr., J. E. (1972). "Vascular changes in the human inner ear associated with aging," *Ann. Otol. Rhinol. Laryngol.* **81**, 364–376.
- Kawamura, K. (1975). "The pontine projections from the inferior colliculus in the cat. An experimental study," *Brain Res.* **95**, 309–322.
- Kazee, A. M., Han, L. Y., Spongr, V. P., Walton, J. P., Salvi, R. J., and Flood, D. G. (1995). "Synaptic loss in the central nucleus of the inferior colliculus correlates with sensorineural hearing loss in the C57BL/6 mouse model of presbycusis," *Hearing Res.* **89**, 109–120.
- Lesser, H. D., O'Neill, W. E., Frisina, R. D., and Emerson, R. G. (1990). "On-off units in the mustached bat inferior colliculus are selective for transients resembling acoustic glint from fluttering insect targets," *Exp. Brain Res.* **82**, 137–148.
- Lynch-Armour, M. A., Frisina, R. D., McReynolds, E. E., Byrd, J., and Walton, J. P. (1995). "Cochlear nucleus inputs to a functionally-characterized region of the inferior colliculus decline with age for the CBA mouse model of presbycusis," *Assoc. Res. Otolaryngol. Abstr.* **18**, 71.
- Maurer, J. R., and Rupp, R. R. (1979). *Hearing and Aging* (Grune & Stratton, New York).

- McCroskey, R. L., and Kasten, R. N. (1982). "Temporal factors and the aging auditory system," *Ear Hear.* **3**, 124–127.
- Meininger, V., Pol, D., and Derer, P. (1986). "The inferior colliculus of the mouse. A nissl and golgi study," *Neurosci.* **17**, 1159–1179.
- Mesulam, M. M. (1982). *Tracing Neural Connections with Horseradish Peroxidase* (Wiley, New York).
- Moore, R. Y., and Goldberg, J. M. (1966). "Projections of the inferior colliculus in the monkey," *Exp. Neurol.* **14**, 429–438.
- Moore, J. K., Karapas, F., and Moore, R. Y. (1977). "Projections of the inferior colliculus in insectivores and primates," *Brain Behav. Evol.* **14**, 301–327.
- Morest, D. K. (1965a). "The laminar structure of the medial geniculate body of the cat," *J. Anat. (London)* **99**, 143–160.
- Morest, D. K. (1965b). "The lateral tegmental system of the midbrain and the medial geniculate body: Study with Golgi and Nauta methods in cat," *J. Anat. (London)* **99**, 611–634.
- Morest, D. K. and Oliver, D. L. (1984). "The neuronal architecture of the inferior colliculus in the cat: Defining the functional anatomy of the auditory midbrain," *J. Comp. Neurol.* **222**, 209–236.
- Morest, D. K., and Winer, J. A. (1986). "The comparative anatomy of neurons: Homologous neurons in the medial geniculate body of the opossum and cat," *Adv. Anat. Embryol. Cell. Biol.* **97**, 1–96.
- Oliver, D. L. (1984). "Neuron types in the central nucleus of the inferior colliculus that project to the medial geniculate body," *Neurosci.* **11**, 409–424.
- O'Neill, W. E., Frisina, R. D., and Gooler, D. M. (1989). "Functional organization of mustached bat inferior colliculus: I. Representation of FM frequency band important for target ranging revealed by ^{14}C -2-deoxyglucose autoradiography and single unit mapping," *J. Comp. Neurol.* **284**, 60–84.
- Pearson, J. D., Morrell, C. H., Gordon-Salant, S., Brant, L. J., Metter, E. J., Klein, L. L., and Fozard, J. L. (1995). "Gender differences in a longitudinal study of age-associated hearing loss," *J. Acoust. Soc. Am.* **97**, 1196–1205.
- Powell, E. W. and Hatton, J. B. (1969). "Projections of the inferior colliculus in the cat," *J. Comp. Neurol.* **136**, 183–192.
- Rasmussen, G. L. (1964). "Anatomic relationships of the ascending and descending auditory systems," in *Neurological Aspects of Auditory and Vestibular Disorders*, edited by W. S. Fields and W. S. Alford (Thomas, Springfield IL), pp. 5–19.
- Royster, L. H., and Thomas, W. G. (1979). "Age-effect hearing levels for a white nonindustrial noise-exposed population and their use in evaluating hearing conservation programs," *Am. Ind. Hyg. Assoc. J.* **40**, 504–511.
- Royster, L. H., Driscoll, D. P., Thomas, W. G., and Royster, J. D. (1980). "Age effect hearing levels on a black nonindustrial noise exposed population (NINEP)," *Am. Ind. Hyg. Assoc. J.* **41**, 113–119.
- Schmiedt, R. A. (1993). "Cochlear potentials in quiet-aged gerbils: Does the aging cochlea need a jump start?" in *Sensory Research, Multimodal Perspectives*, edited by R. T. Verrillo (Erlbaum, Hillsdale, NJ), pp. 91–103.
- Schuknecht, H. F., Watanuki, K., Takahashi, T., Belal, A., Kimura, R. S., Jones, D. D., and Ota, C. Y. (1974). "Atrophy of the stria vascularis, a common cause of hearing loss," *Laryngoscope* **84**, 1777–1821.
- Schuller, G., Covey, E., and Casseday, J. H. (1984). "Auditory pontine grey: connections and response properties in the horseshoe bat," *Eur. J. Neurosci.* **3**, 648–662.
- Schulte, B. A. and Schmiedt, R. A. (1992). "Lateral wall Na, -K-ATPase and endocochlear potentials decline with age in quiet-reared gerbils," *Hearing Res.* **61**, 35–46.
- Schweizer, H. (1981). "The connections of the inferior colliculus and the organization of the brainstem auditory system in the greater horseshoe bat (*Rhinolophus ferrumequinum*)," *J. Comp. Neurol.* **201**, 25–49.
- Shore, S., Helfert, S. C., Bledsoe, S. C., Altschuler, R. A., and Godfrey, D. A. (1991). "Descending projections to the dorsal and ventral divisions of the cochlear nucleus in guinea pig," *Hearing Res.* **52**, 255–268.
- Stiebler, I. and Ehret, G. (1985). "Inferior colliculus of the house mouse. I. A quantitative study of tonotopic organization, frequency representation, and tone-threshold distribution," *J. Comp. Neurol.* **238**, 65–76.
- Suga, N., Simmons, J., and Shimoza, T. (1974). "Neurophysiological studies on echolocation systems in awake bats producing CF-FM orientation sounds," *J. Exp. Biol.* **61**, 379–399.
- Walton, J. P., Frisina, R. D., Ison, J. R., and O'Neill, W. E. (in press). "Neural correlates of behavioral gap detection in the inferior colliculus of the young CBA mouse," *J. Comp. Physiol. A*.
- Walton, J. P., Wilson, W. W., Mees, A. C., O'Neill, W. E., and Frisina, R. D. (1995). "Age-related alteration in temporal processing of amplitude modulation and silent gaps in the inferior colliculus of the CBA mouse," *Assoc. Res. Otolaryngol. Abstr.* **18**, 71.
- Wenstrup, J. J. and Grose, C. D. (1995). "Inputs to combination-sensitive neurons in the medial geniculate body of the mustached bat: The missing fundamental," *J. Neurosci.* **15**, 4693–4711.
- Wenstrup, J. J., Larue, D. T., and Winer, J. A. (1994). "Projections of physiologically defined subdivisions of the inferior colliculus in the mustached bat: targets in the medial geniculate body and extrathalamic nuclei," *J. Comp. Neurol.* **346**, 207–236.
- Willard, F. H., and Ryugo, D. K. (1983). "Anatomy of the central auditory system," in *The Auditory Psychobiology of the Mouse*, edited by J. F. Willott (Thomas, Springfield, IL), pp. 201–304.
- Willott, J. F. (1986). "Effects of aging, hearing loss, and anatomical location on thresholds of inferior colliculus neurons in C57BL/6J and CBA/J mice," *J. Neurophysiol.* **56**, 391–408.
- Willott, J. F. (1991). *Aging and the Auditory System* (Singular Pub. Group, San Diego, CA).
- Willott, J. F. and Bross, L. S. (1990). "Morphology of the octopus cell area of the cochlear nucleus in young and aging C57BL/6J and CBA/J mice," *J. Comp. Neurol.* **300**, 61–81.
- Willott, J. F., Bross, L. S., and McFadden, S. L. (1992). "Morphology of the dorsal cochlear nucleus in C57BL/6J and CBA/J mice across the life span," *J. Comp. Neurol.* **321**, 666–678.
- Willott, J. F., Aitkin, L. M., and McFadden, S. L. (1993). "Plasticity of auditory cortex associated with sensorineural hearing loss in adult C57BL/6J mice," *J. Comp. Neurol.* **329**, 402–411.
- Willott, J. F., Parham, K., and Paris-Hunter, K. (1988). "Response properties of inferior colliculus neurons in young and very old CBA/J mice," *Hearing Res.* **37**, 1–14.
- Willott, J. F. and Urban, G. P. (1978). "Response properties of neurons in nuclei of the mouse inferior colliculus," *J. Comp. Physiol.* **127**, 175–184.
- Wilson, W. W., Walton, J. P., and Frisina, R. D. (1995). "Paradoxical gap detection thresholds of mouse inferior colliculus neurons in background noise," *Soc. Neurosci. Abstr.* **21**, 908.
- Zettl, M. L., O'Neill, W. E., and Frisina, R. D. (1995). "Calcium binding protein and GABA immunoreactivity in the inferior colliculus of the aged CBA/J mouse model of presbycusis," *Soc. Neurosci. Abstr.* **21**, 908.

Sound-pressure measurements in the cochlear vestibule of human-cadaver ears

Sunil Puria, William T. Peake, and John J. Rosowski

The Eaton-Peabody Laboratory of Auditory Physiology, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, Massachusetts 02114 and The Department of Otology and Laryngology, Harvard Medical School, and The Research Laboratory of Electronics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

(Received 3 June 1996; revised 30 December 1996; accepted 3 January 1997)

The middle-ear pressure gain for the 50-Hz to 12-kHz range was determined from the ratio of sound pressures measured in the vestibule and the ear canal of four human-cadaver ears. The magnitude of the middle-ear pressure gain is 20 dB for frequencies between 500 Hz and 2 kHz. Above 4 kHz, the gain changes as a function of frequency at a rate of approximately -8 dB/octave and below 400 Hz at 4 dB/octave. The standard error of the mean magnitude across the four ears is typically less than 3 dB. The phase angle of the pressure gain also changes with frequency. Interruption of the ossicular chain decreases the vestibule pressure by at least 20 dB. It is shown that air bubbles in the inner ear can diminish the vestibule pressure; procedures are used to remove bubbles. From these pressure measurements and previous measurements of stapes motion, the frequency dependence of behavioral thresholds for tones was tested to discover whether it corresponds to the constancy of a physiological variable at the cochlear input. Among pressure, power, or stapes-motion measures, the vestibule pressure is most nearly constant with frequency at the behavioral “minimum audible pressure.” © 1997 Acoustical Society of America. [S0001-4966(97)01405-7]

PACS numbers: 43.64.Ha, 43.66.Ba, 43.38.Pf [RDF]

INTRODUCTION

The middle ear couples the sound in the ear canal to the inner ear. The middle-ear pressure gain [the ratio of the vestibule pressure near the footplate to the ear-canal pressure near the tympanic membrane $P_V(\omega)/P_{EC}(\omega)$] has been measured *in vivo* in cat, guinea pig, and chinchilla ears (Nedzelitsky, 1980; Dancer and Franke, 1980; Franke and Dancer, 1985; Décor, 1989). Although it is not now practical to make pressure measurements in the inner ear of living humans, extracted human-cadaver temporal bones can be used to measure the pressure transformation.

Measurements have demonstrated (Merchant *et al.*, 1988; Rosowski *et al.*, 1990) that the middle-ear input impedance of temporal bone preparations (both human and guinea pig) is indistinguishable from ears of living subjects, at least at frequencies below 5 kHz. In support of this conclusion, comparison of measurements of motion of the umbo in live human ears and in temporal bones are similar for frequencies up to 10 kHz (Goode *et al.*, 1993). These results suggest that cadaver ears may be approximately acoustically equivalent to living ears when appropriate steps are taken. Thus, acousto-mechanical properties of the human middle and inner ear can be studied on human temporal bones. The goal of this study is to measure the middle-ear pressure gain (magnitude and phase) in human-cadaver ears as models of living ears. Other than our preliminary report (Puria *et al.*, 1993b) and recent measurements of the magnitude of the pressure gain in one ear by Hüttenbrink and Hudde (1994), direct measurements of the pressure gain have not been reported.

The pressure-gain measurements reported here can be compared to predictions of recent middle-ear models (Krin-

glebotn, 1988; Goode *et al.*, 1994). Although predictions and measurements agree to the extent that both have band-pass characteristics, substantial disagreements occur with the Goode *et al.* (1994) model for frequencies below 0.8 kHz and with both models at frequencies above 2–3 kHz. Thus modifications of the models are required to account for these new results.

To describe the acoustic input of the ossicular chain to the cochlea, knowledge of both stapes motion and inner-ear pressure is required. Measurements of the ratio of stapes velocity to ear-canal pressure in human temporal bones have been previously reported (Kringelbotn and Gundersen, 1985; Goode *et al.*, 1994). From the ear-canal to inner-ear pressure transfer function and the reported stapes-motion measurements, we can determine whether the frequency dependence of the human hearing threshold for tones can be described as a constant level of some physical variable at the cochlear input (e.g., sound pressure, volume velocity, volume acceleration, or power). We show that, at frequencies above 150 Hz, constant vestibule pressure and behavioral threshold curves have similar shapes, while below 150 Hz none of the variables tested describes the frequency dependence.

I. METHODS

A. Post-mortem materials

Measurements were made on fresh or thawed ears extracted from human cadavers. Most of the temporal-bone material was purchased from the National Disease Research Interchange (NDRI) of Philadelphia, PA which solicits post-mortem materials throughout the country from people who have donated their bodies to science. These temporal-bone

plugs are extracted at the time of autopsy, placed in dry ice, and shipped to the Massachusetts Eye and Ear Infirmary overnight.

Fifteen temporal bones were used in these experiments. Some were damaged on arrival, some were used to develop procedures, and a few were damaged during surgery and could not be used for measurements. Reliable pressure measurements were made in four temporal bone ears.¹

The donors of the temporal bones were females of ages 43, 42, and 82 for ears 2–4, respectively. The time between death and start of measurements were 15 months, 2.5 months, and 8 days for ears 2–4, respectively. This information was not available for ear 1.

One fresh bone (ear 4) was gathered at autopsy from the Massachusetts General Hospital by the methods which the Department of Otolaryngology uses for specimens used in surgical training. The fresh bone was prepared for measurements within a few days of autopsy without freezing. Measurements on this bone in the fresh state and after freezing suggest that freezing has little effect on the middle-ear pressure gain (Sec. II B 5).

B. Preparation of temporal bones

To make ear-canal and inner-ear pressure measurements, the temporal bones were prepared so as to place both a sound-source/microphone assembly in the ear canal near the tympanic membrane and a hydropressure transducer in the vestibule near the stapes footplate. The bony ear canal of each temporal bone was shortened by drilling bone away, and a brass ring was attached to the lateral side of the tympanic ring with carboxylate dental cement (ESPE, Durelon®, Norristown, PA). This arrangement produced a reversible and reproducible coupling of the custom-made sound-source/microphone assembly to the external meatus such that the microphone probe tube was within 3–5 mm of the umbo as shown in Fig. 1. A dental drill was used to enlarge the internal auditory canal and to drill a hole (approximately 1.6-mm diameter) through the saccular wall at the inferior vestibular nerve area, just anterior and superior to the singular canal (the duct for the nerve to the posterior semicircular canal). This hole into the inner ear allowed insertion of a pressure transducer.

As the inner ear manipulations produced the possibility of air entering the labyrinth, we provided for removal of air with two tubes (polyethylene) cemented into different regions of the perilymphatic space. The “inlet tube” was placed in the readily accessible superior part of the superior semicircular canal. The “outlet tube” was placed in one of two locations. In the early experiments the outlet tube was cemented into the apical turn of the cochlea. The apex was opened by drilling away the medial wall of the protympanum just posterior to the opening of the eustachian tube. Measurements on ear 1 were made with this configuration. For subsequent ears the outlet was located in the scala tympani near the round window; in theory this location enables more complete flushing of the cochlear spiral. To access the round window, the middle-ear cavity was opened widely by drilling into the hypotympanum through the jugular fossa, just anterior to the stylomastoid foramen; this approach allows access

to the basal end of the cochlea. The outlet tube was placed in the scala tympani just anterior to the round window. The tube inner diameters were between 0.34 and 0.61 mm with outer diameters between 0.91 and 1.5 mm. Tube lengths between 105 and 125 mm were used for these experiments.²

To support the bone during the measurements, a solid brass rod 5.5–7 cm long and 0.64 cm in diameter was cemented to the bone, typically in the carotid canal; this arrangement allowed the bone to be clamped to a pole mounted on the experimental table. The time required for the entire preparation was between 12 and 20 h for each temporal bone; in some cases the processes extended over two days.

All measurements reported here are with opened middle-ear cavities. Measurements indicate that the human middle-ear cavity plays a minor role in middle-ear sound transmission (McElveen *et al.*, 1982). We will return to this point in Sec. III.

C. Preparations for measurements

At the beginning of each measurement session the temporal bone was degassed by placing the bone in a beaker filled with degassed saline (9-gm NaCl per liter of H₂O) which was then placed in a glass vacuum chamber for approximately 30 min. The bone was positioned in the beaker so that the round-window outlet tube was in air while the inlet tube was immersed in saline. During degassing, bubbles were observed to escape through the outlet tube, indicating that air escaped while degassed saline entered the inner ear.

After degassing, a reservoir, consisting of a tube filled with degassed saline, was connected via a coupling tube (131610 25G, 0.5 mm, irrigating cystotome snipped at the tip; Steriseal, Worcestershire, England) to the inlet tube, providing a fluid source. The fluid was allowed to drain from the outlet tube (in some cases the inlet and outlet tubes were reversed by placing the reservoir on the round-window tube as described further in Sec. II). The inlet tube remained attached to the reservoir and the fluid level was maintained throughout the experiment.

The temporal bone was mounted on the experiment table in a sound-proofed chamber (Vér *et al.*, 1975). The bone was positioned so that the ear-canal coupling ring was in a plane perpendicular to the table. The sound-source/microphone assembly was placed on a micro-manipulator. A thin layer of clay between the brass ring on the bone and brass coupling ring on the sound delivery tube was used to form an acoustic seal between the two rings. The coupling rings helped reduce the session-to-session variations in the position of the microphone probe tube relative to the eardrum.

Changing the height of the fluid column attached to the superior semicircular canal (SSC) inlet tube can change the static pressure within the inner ear. Does this static pressure affect the middle-ear pressure gain? Control experiments show that changing the fluid reservoir level by less than ± 2 cm has effects less than ± 1 dB in magnitude and less than ± 7 deg in phase with smaller effects at most frequencies. In comparison, the effects of changes in the reservoir level were greater than typical changes observed between repeat mea-

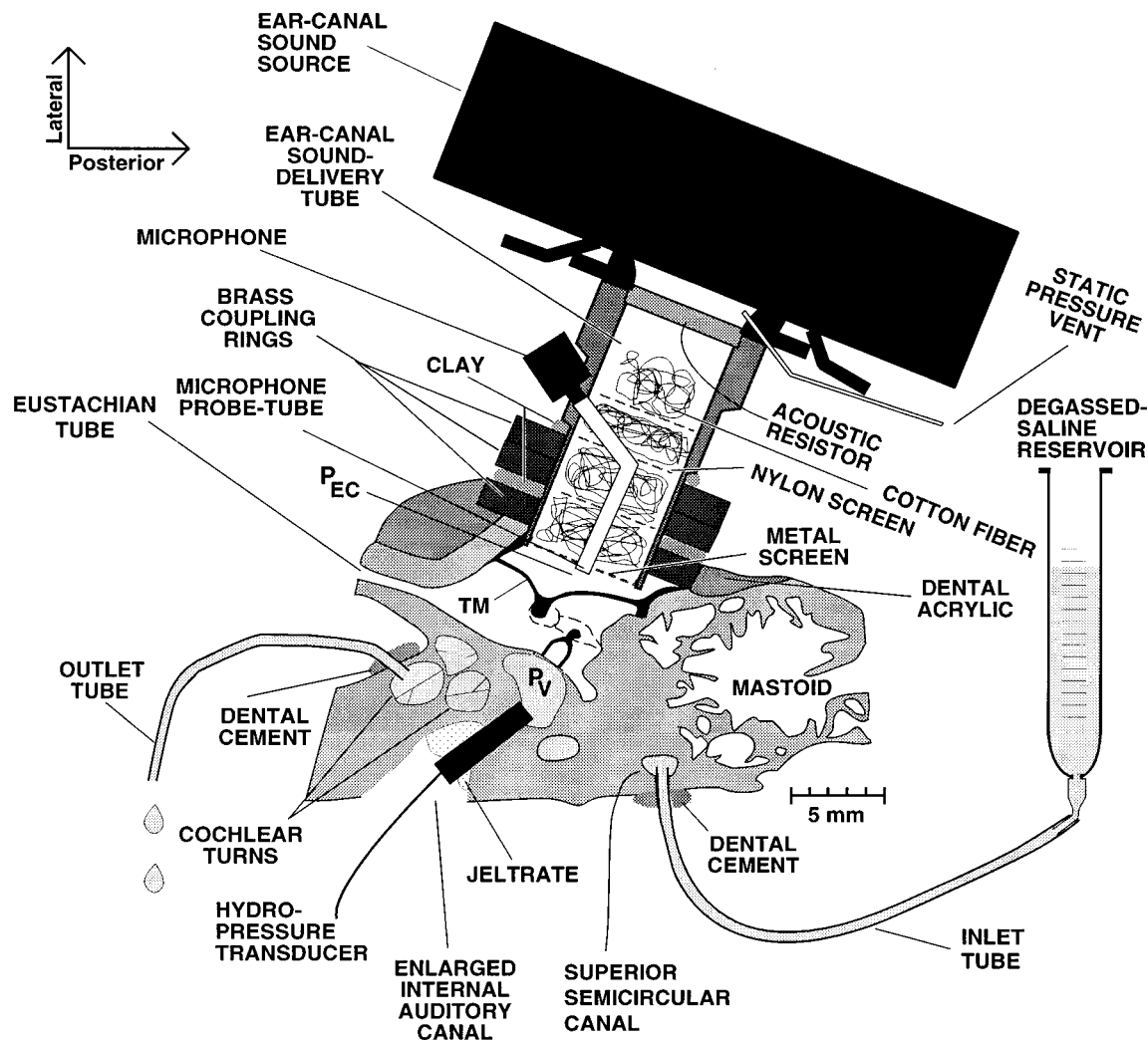


FIG. 1. This horizontal cross section of a human temporal bone (right ear) and the sound-source/microphone assembly is a simplified representation of the preparation. A sound source (DT-48) is coupled to the ear canal by the sound-delivery tube containing an acoustic resistor and layers of cotton and nylon screen to increase the impedance of the acoustic source. The ear-canal pressure P_{EC} was measured with a small electret microphone (EK-3133; Knowles Electronics, Ithaca, IL) attached to a probe tube. To measure the inner-ear pressure P_v a hydropressure transducer (EP IL-050; Entran Devices, Inc., Fairfield, NJ) was placed in the vestibule facing the footplate. To ensure that the cochlea remained fluid filled, an inlet tube was cemented into the superior semicircular canal and an outlet tube was cemented into the scala tympani near the round window. (The round window and the hypotympanum are at a location inferior to the cross section shown here.) The inlet was attached to a 12-cc reservoir of degassed saline. The temporal bone was cemented to a brass rod which was clamped to a pole mounted on a table. The sound-source/microphone assembly and the hydropressure transducer were placed on separate micromanipulators. The temporal bone is stationary while the transducers are maneuvered into the bone. The bodies of malleus and incus, visible in sections that are superior to that shown, are indicated by the thin dashed lines.

surements during the same measurement session, but smaller than typical changes between measurements on the same bone in different measurement sessions.

The hydropressure transducer was placed on a second micromanipulator and positioned in the vestibule. An alginate liquid impression material "Jeltrate" (L. D. Caulk Company, Division of Dentsply International Inc., Milford, DE) was applied between the internal auditory canal and the body of the hydropressure transducer. The impression material congeals within a minute into a mold that prevented fluid flow from the vestibule (Fig. 1).

The bones were kept moist by applying saline to the middle-ear cavity typically within 10 min just preceding each measurement. The application of fluid moistened the

tympanic membrane, ossicles, annular ligament, and the round-window membrane, supporting ligaments and surrounding structures.

D. Response measurement, calibration, and analysis

Stimulus generation and synchronous response averaging were done with a DSP-16+ board (Ariel Corp., Cranbury, NJ) with two channel D/A and A/D converters and a Digital Signal Processing chip. The SYSid[®] measurement and analysis system (SYSid Labs, Berkeley, CA) was used for all measurements (Puria and Allen, 1992; Puria *et al.*, 1993a). The stimulus delivered by the D/A was a flat spectrum chirp (swept tone) with quadratic phase (linear group

delay). The spectrum of the stimulus contained 512 frequency points with a constant frequency spacing of 48.7 Hz.

The output of the D/A converter was amplified (D-75; Crown International, Elkhart, IN) and fed to the sound source (DT-48; Beyerdynamic, Hicksville, NY) through a switch-selectable resistor for attenuation of the sound level. Simultaneously a response waveform was sampled by the A/D and recorded from either the probe-tube microphone or the hydropressure transducer. Typically, 500–2000 stimulus presentations were averaged.

Two basic constraints were applied to the sound-pressure levels used to make ear-canal and inner-ear pressure measurements. First, because of the low sensitivity of the hydropressure transducer the inner-ear pressure spectral level had to be greater than 100 dB SPL in order to maintain a signal-to-noise ratio of more than 30 dB (arbitrarily chosen). Second, ear-canal sound-pressure spectral levels larger than 80 dB SPL produce significant distortion in the ear-canal probe microphone output. These requirements precluded simultaneous measurements of the two pressures; thus the ear-canal pressure was measured at a lower stimulus level than the inner-ear pressure. The switch-selectable series resistor described above allowed quick switching between the different drive levels used in the two measurement types. For the ear-canal measurements the switch-selected resistor attenuated the output of the D/A converter by 70 dB, and for the inner-ear pressure measurements the attenuation was either 20 dB or 30 dB (harmonic distortion was typically less than 2%). Thus the two stimuli differed by as much as 50 dB. The earcanal microphone response was amplified (P5 pre-amplifier; Grass Instruments, Quincy, MA) by 40 dB and the hydropressure transducer output was amplified by 66 or 74 dB. These sound-source attenuations and amplifier gains were chosen to minimize harmonic distortion of the sound source, transducers, amplifier, and the DSP board (see below for linearity tests).

The ear-canal microphone and hydropressure transducer electric outputs were calibrated and the calibrations were used to convert the measured voltages to sound pressures. A pistonphone (type 4220, Brüel & Kjaer) of known sound-pressure level was used to calibrate a $\frac{1}{4}$ -in. reference microphone (type 4136, B&K) with a flat response (± 0.5 dB in the frequency range of interest). This reference microphone was used to calibrate the probe-tube microphone by measuring the response of both microphones in a small cavity where probe-tube and reference microphone were less than 1 mm apart. The calibrated probe-tube microphone was then used to calibrate the hydropressure transducer by measuring the responses of both when they were in the same cavity. Thus it was assumed that the hydropressure transducer calibration is the same in both fluid and air. This assumption is valid for the following reasons: (1) The impedance of the strain-gauge hydropressure transducer is much higher than the impedance of the cochlea; (2) the hydropressure transducer is sensitive to pressure and not fluid acceleration. Given these observations the hydropressure output is proportional to the pressure at its sensing end regardless of the impedance of the measured environment.

As a check on the linearity of the stimulus and measure-

ment system, the sound pressure produced by the sound source in a small brass cavity was measured by the hydro-pressure transducer at a high input level (20-dB attenuation) and by the probe-tube microphone at a lower level (70-dB attenuation). When normalized for drive level, the sound pressures measured by the hydropressure transducer and the ear-canal probe-tube microphones were within ± 0.5 dB in magnitude and $\pm 3.7^\circ$ in phase in the 0.05–11 kHz range and within ± 2.4 dB in magnitude and $\pm 16^\circ$ in phase in the 11–20 kHz range. The close agreement between the two suggests that at these levels the measurement system (sound source, microphones, and amplifier used in these experiments) is linear. The observed differences place a bound on possible calibration errors and nonlinearities in the measurement system.

The complex middle-ear transfer function was obtained from the ratio of the complex amplitudes of vestibule sound-pressure to ear-canal sound pressure. The magnitude of the middle-ear pressure gain $20 \log_{10} |P_V/P_{EC}|$ was obtained from the absolute value of the complex transfer function while the $\angle \{P_V/P_{EC}\}$ was obtained from its angle. All means and standard errors of the mean (SEM) were calculated from the magnitudes (in dB) and angles of the transfer functions. The total sound level reported is $10 \log_{10}$ of the sum of the squared pressure magnitudes (in Pascals) of all frequency components. All sound levels (in dB SPL) are referenced to 20 μ Pa.

Histological assessment showing that inner-ear structures such as the basilar membrane, Reissner's membrane, and organ of Corti were not damaged by the procedures outlined above was obtained from one temporal bone (ear 1) that was fixed, decalcified, embedded, sectioned, and stained by techniques similar to those described by Schuknecht (1993).

II. RESULTS

A. Ratio of vestibule pressure to ear-canal pressure

1. Pressure gain in normal middle ears

Figure 2 shows for the four ears the mean magnitude (panel A) and phase (panel B) of repeated pressure ratio P_V/P_{EC} measurements. These curves represent the pressure transformation for each ear. Also shown in Fig. 2 is the average of the four ears.

Because many measurements were made—interspersed with various manipulations—during each experimental session, criteria were established to select measurements to be included in each ear's mean. For each session we chose the measurement with the highest gain magnitude over the widest range of frequencies (i.e., maximum area under the curve). The justification for this criterion is our conclusion that most manipulations of the middle ear in its “normal state” tend to decrease the magnitude of the pressure gain in some frequency region. Measurements that lead to this conclusion are presented in subsequent sections.

The maximum gain magnitude (Fig. 2A) of the four-ear mean is 21 dB at 1.2 kHz. In the 500 Hz to 2 kHz frequency region, the mean gain of the four ears is approximately 20 dB. Below 400 Hz, the mean magnitude decreases with a

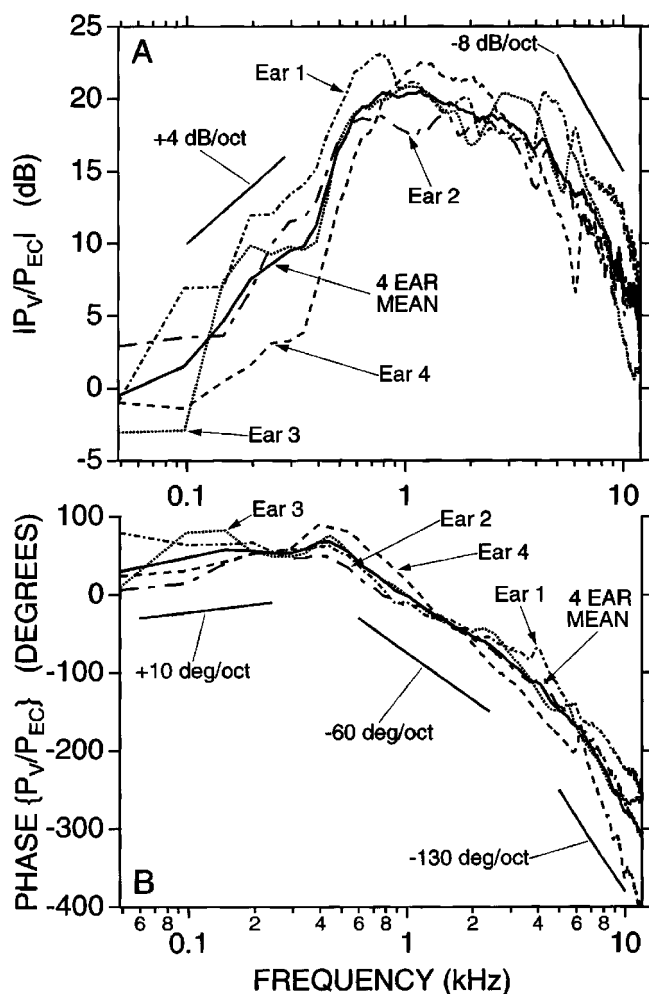


FIG. 2. Measured middle-ear pressure gain in four human ears. For each ear, means across measurement sessions are shown with magnitudes in (A) and phases in (B). Measurements made in separate experimental sessions (with one or more day between sessions) were included in the average. Ears 1–4 included 2, 4, 5, and 4 sessions, respectively. The mean for each ear was used to compute the “4 EAR MEAN” plots. The measurement frequency resolution is approximately 50 Hz.

slope of approximately 4 dB/oct; above 4 kHz the pressure decreases approximately -8 dB/oct. The mean phase of the pressure gain of the four ears (Fig. 2B) at frequencies below 400 Hz has a mean slope of 10 deg/oct and seems to approach 0 deg as the stimulus frequency decreases. Above 400 Hz the phase appears to have two different regions. In the 400 Hz to 4 kHz region the mean slope is approximately -60 deg/oct, whereas above 4 kHz the mean slope is approximately -130 deg/oct.

A measure of the session-to-session variability of each ear is the standard error of the mean³ (SEM) for all measurements used to calculate the mean. Figure 3 shows the SEM magnitude (panel A) and the SEM phase (panel B) for each temporal bone. To estimate the inter-ear variability the mean pressure gains from Fig. 2 were used to calculate the four-ear SEM shown in Fig. 3.

The SEM magnitude of the four ears is often greater than the SEM magnitude for each individual ear indicating more inter-ear variability than session-to-session variability in each ear. The exception to this is ear 1 which has a larger

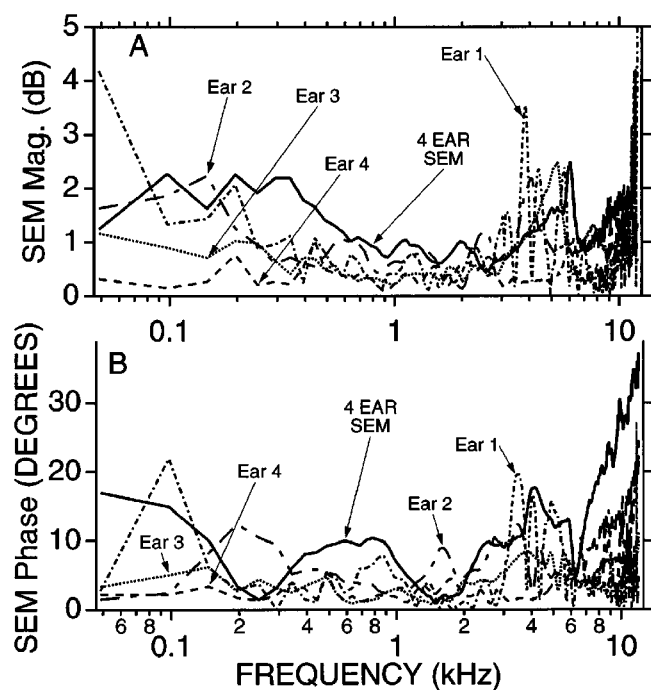


FIG. 3. The standard error of the mean (SEM) magnitude (in dB) is shown in panel A, while the standard error of the mean phase is shown in panel B for each of the four ears. Measurements used to compute the individual means in Fig. 2 were used to compute the SEM for each bone. To estimate the inter-ear variability the four mean curves in Fig. 2 were used for the “4 EAR SEM” plots.

SEM than the other ears, perhaps because measurements from only two sessions are available for that ear. The magnitude of the four-ear SEM is higher at frequencies below 0.5 and above 3 kHz either because there is more inter-ear variability and/or there is more variability due to a decrease in signal-to-noise ratio.

The SEM of the phase is shown in Fig. 3B. With some exceptions the four-ear SEM phase is larger than the mean session-to-session SEM, again indicating that inter-ear variations make a contribution to the variability. For frequencies greater than 6 kHz the inter-ear phase variation tends to increase systematically. One possible reason for the inter-ear variability in phase is variation in the location of the ear-canal microphone relative to the tympanic membrane (Fig. 1) among the different ears.

2. Test for mechanical coupling from earphone to inner ear

During measurements the sound-source/microphone assembly and the temporal bone touch each other and thus the two are mechanically coupled (see Fig. 1). It is therefore possible that some of the pressure in the inner ear is caused by vibration of the temporal bone produced by the sound source. One way to estimate the effect of the vibrations produced by the sound source is to make pressure measurements with the sound source mechanically isolated from the bone. If the pressure gain is unchanged with the source isolated the effect of mechanical coupling is insignificant.

Measurements made for three different “positions” of the sound-source/microphone assembly relative to the ear-canal coupling ring are shown in Fig. 4. Although the spec-

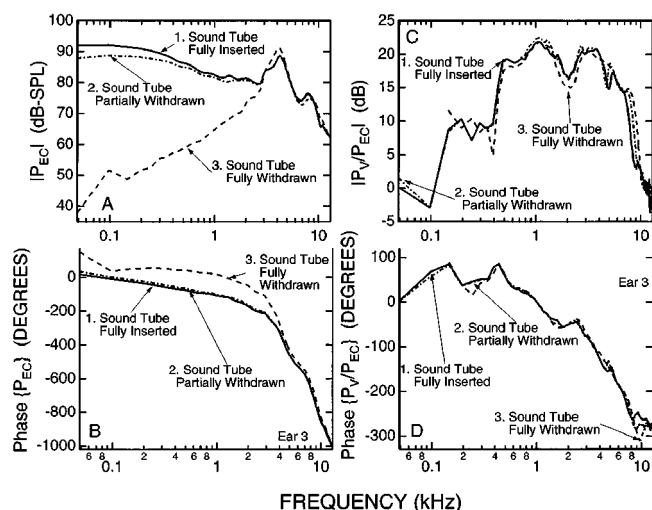


FIG. 4. Measurements with variations in mechanical connection of sound source to ear. The spectrum of the ear-canal pressure magnitude in dB SPL is shown in panel A, while the phase of the pressure is shown in panel B. The pressure transfer functions are shown in panels C and D. Three measurement conditions are shown in each panel. (1) With the sound-delivery (SD) tube fully inserted and thus the SD and bone are mechanically coupled (with clay between the ear canal and SD tube coupling rings). (2) The SD tube is partially withdrawn such that the tip of the SD tube is inside the ear-canal space but the brass rings attached to the bone and SD tube are not touching. (3) The SD tube is further withdrawn such that the tip of the SD tube is at the entrance of the bone brass ring. As shown in panel A, condition (3) significantly reduces the spectral level of the ear-canal pressure at low frequencies. However, the pressure reduction had little effect on the transfer function (panels C and D). In panels C and D the two lowest-frequency data points (0.05 and 0.1 kHz) for case 3 were removed because of low signal-to-noise ratio.

trum of the ear-canal sound level is different for each case, the total sound level is approximately 106 dB SPL for all three cases (Fig. 4A). Case 1 is a normal measurement in which the delivery tube is fully inserted into the coupling ring; a layer of clay is used to provide a good seal between the brass ring on the bone and the brass ring of the sound-delivery tube (see Fig. 1). For case 2, the sound-delivery assembly was withdrawn a few millimeters leaving a gap between the two brass rings, but the tip of the sound-delivery tube and the microphone probe tip were still inside the ear-canal brass ring. This configuration decreased the ear-canal pressure magnitude somewhat at low frequencies (Fig. 4A), but did not affect the middle-ear pressure gain (Fig. 4C and D). For case 3, the sound-delivery tube was further withdrawn such that its end was just at the ear-canal entrance. A large decrease (as much as 40 dB) in the pressure magnitude at the ear-canal entrance occurs at low frequencies (Fig. 4A), but the pressure gain (Fig. 4C and D) is essentially unchanged. This result indicates that during normal measurements mechanical coupling of sound-source vibrations has little effect on vestibular pressure.

3. Role of ossicular chain in pressure transformation

In this section we test whether the middle-ear pressure gain shown in Fig. 2 depends on ossicular sound conduction only or whether other mechanisms contribute significantly to the measured inner-ear pressure. In Fig. 5 panels A and B show the vestibule pressure measured before and after inter-

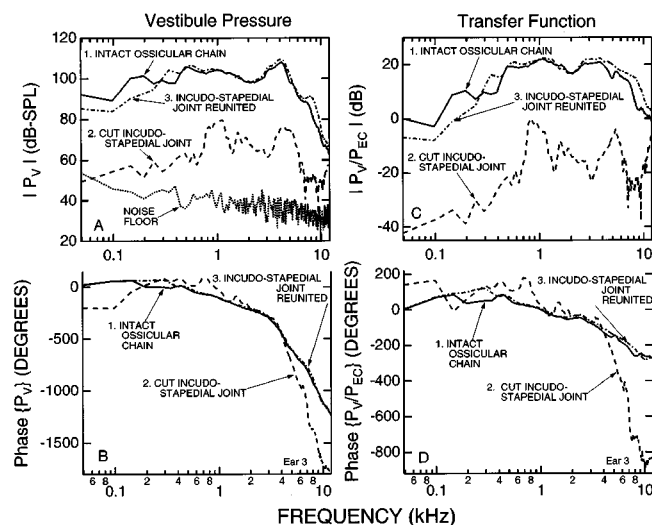


FIG. 5. The effect of interrupting the ossicular chain on the vestibule pressure and the corresponding transfer function. Three cases are shown: (1) intact ossicles; (2) after interrupting the ossicular chain by cutting the incudo-stapedial joint; and (3) after reuniting the ossicular chain by applying dental cement at the incudo-stapedial joint. The fourth curve in panel A labeled "noise floor" is the hydropressure transducer response without a stimulus, measured after interrupting the ossicular chain. The total sound level (panel A) for case 1 is 124 dB SPL while the noise floor is 65 dB SPL. Panels A and B show the magnitude and phase of the vestibule pressure, while panels C and D show the corresponding transfer-function magnitude and phase.

rupting the ossicles by cutting the incudo-stapedial joint and removing a part of the lenticular process so that there was a clear air gap between the ossicles. It is clear from panel A that the magnitude of the vestibule pressure after ossicular disarticulation decreases by more than 20 dB and is typically higher than the measured noise floor.

Panels C and D of Fig. 5 show the vestibule pressure normalized by the ear-canal pressure for the three conditions shown in panels A and B. While the transfer function is severely affected by ossicular disarticulation, disarticulation does not reduce the transfer function magnitude to 0 ($-\infty$ in dB). Indeed, there are frequency regions (between 0.8 and 1.3 kHz, 4 and 6 kHz, and 11 and 12 kHz) where the transfer function shows the sound-pressure in the vestibule is not much lower than the ear-canal pressure. The transfer of sound pressure to the cochlea with ossicular disarticulation indicates sound reaches the cochlea through mechanisms other than the ossicular chain.

The alternative coupling mechanisms are not understood, but some suggestions can be made. Measurements (not shown here) similar to those in Fig. 4 were made after interrupting the incudo-stapedial joint to test the possibility that the residual vestibule pressure is due to mechanical coupling from the sound source. These results showed that even with an interrupted incudo-stapedial joint, mechanical coupling does not contribute to the vestibule pressure at least for frequencies below 3 kHz. Another possibility is that the stapes and round window are directly stimulated by sound radiation from the tympanic membrane. In this scenario the pressure near the two windows is lower at lower frequencies than at higher frequencies due to the open middle-ear cavities, which is consistent with the measurements shown in

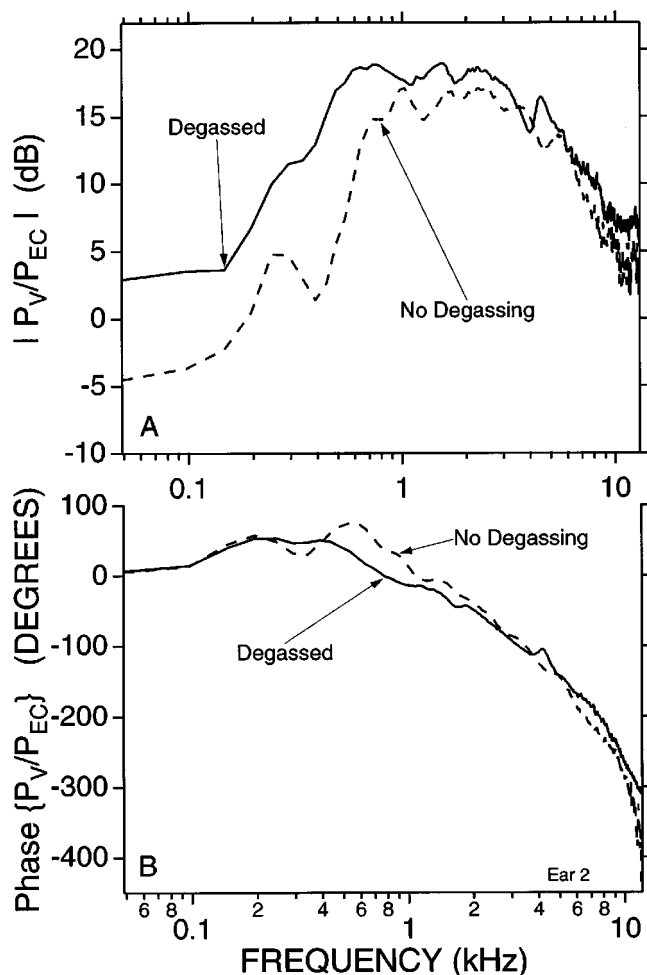


FIG. 6. Effect of degassing a temporal bone at the start of each experiment session. Middle-ear pressure gain measurements, in one ear, without any degassing at the beginning of the measurement session and with degassing at the beginning of the measurement session are shown. The bone was degassed for approximately 30 min in a vacuum chamber. The average of measurements made during four different experiment sessions for each of the two conditions is shown (see Fig. 3, ear 2, for SEM of the degassed condition).

Fig. 5. With this mechanism variations in the transfer function for frequencies above 1 kHz (Fig. 4C and D) could be due to resonances in the open middle-ear cavity space.

4. Dependence of middle-ear pressure gain on ear-canal sound-pressure level

The middle-ear pressure gain at different ear-canal sound-pressure levels was previously reported (Puria *et al.*, 1993b). For total sound levels in the ear canal of less than approximately 120 dB SPL, both the magnitude and phase of the pressure gain are independent of level while above 120 dB SPL the gain is level dependent. All measurements reported here are from the linear region of the middle-ear gain.

B. Control experiments: How do the measurements depend on our methods?

1. Middle-ear input impedance

We have previously demonstrated that opening the degassed inner ear, placing the hydrophone in the vestibule,

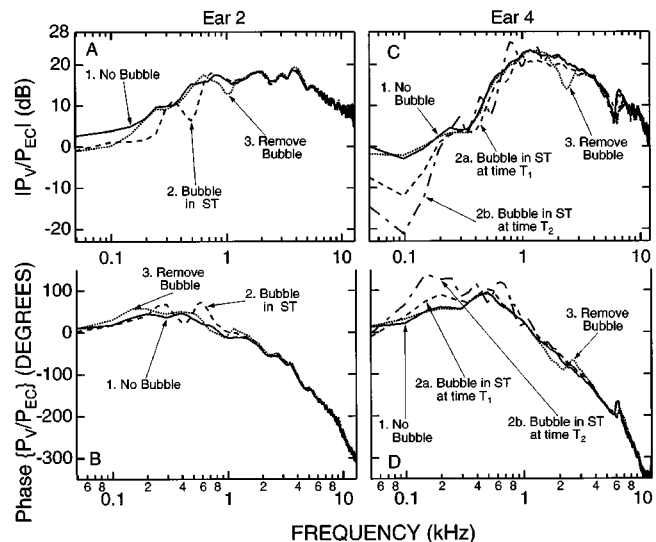


FIG. 7. Effects of air introduced into the scala tympani (ST) through the round-window tube on the pressure gain. The magnitude of the middle-ear pressure gain is shown in panels A and C, while the phase angle is shown in panels B and D. This figure shows measurements made on ear 2 (left panels) and ear 4 (right panels) during different states of the inner ear. (1) Before any air was introduced into the inner ear. For ear 2 (left panels), one measurement was made while the air was in the inner ear (2); for ear 4 (right panels), two measurements were made at time T_1 (2a) and time T_2 (2b) while the air was in the inner ear. (3) After air was removed from the inner ear. The difference in time T_1 and T_2 was 3 min and all measurements shown here were made within a 35-min time period. The volume of air introduced into ear 2 was approximately 8 μl , while for Ear 4 it was 20 μl .

and flushing the inner ear with degassed saline has little effect on the middle-ear input impedance in temporal bones (Puria *et al.*, 1993b). Although the impedance measurements are not very sensitive to minor changes they are sensitive to large changes, i.e., fixation of the stapes. Those results indicate that the acousto-mechanical properties of the ear are effectively unaltered by the manipulations we have used. Some of the effects of the procedures used in our preparation are illustrated below.

2. Degassing at the beginning of an experiment

What is the effect of degassing the temporal bones at the beginning of each experiment session? Averaged measurements with and without degassing show that degassing has a significant effect at frequencies below 1 kHz and little effect at higher frequencies (Fig. 6). However, it is not clear from these measurements where the air might be. To explore the effects of bubble location, measurements were made (in degassed bones) before and after air was injected into the inner ear. As intra-labyrinthine bubbles cannot be seen, these measurements serve as an indicator of when and possibly where bubbles might be present.

3. Air introduced into the scala tympani

To introduce bubbles into the scala tympani, air was intentionally allowed to enter the scala tympani through the round-window (RW) outlet tube. The saline reservoir was uncoupled from the superior semicircular canal (SSC) tube which was lowered so that fluid drained from this tube while air entered the RW tube. After air entered (approximately 8

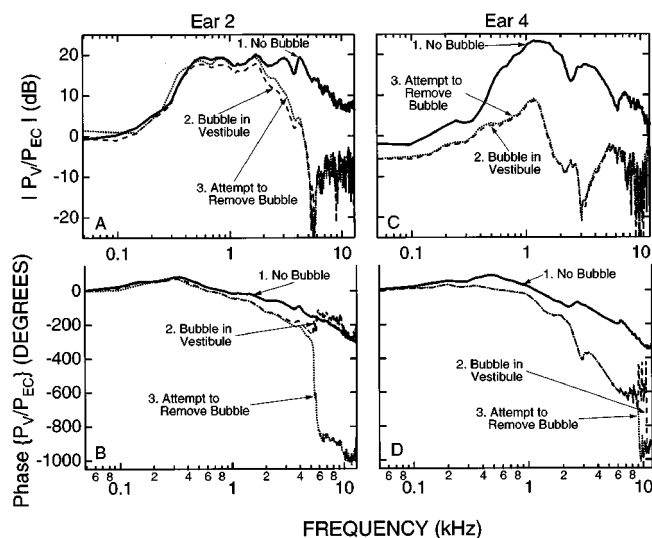


FIG. 8. Effect of air introduced into the vestibule. This figure shows measurements made on ear 2 (left panels) and ear 4 (right panels) during different states of the inner ear: (1) Before any air was introduced into the inner ear; (2) after air entered the inner ear through the superior semicircular canal tube; and (3) after attempts were made to remove the air from the inner ear. All measurements shown here were made within a 20-min time period. The volume of the injected air was not measured but calculations based on tube diameter and length suggest that the volume was less than $12 \mu\text{l}$ for ear 2 and less than $30 \mu\text{l}$ for ear 4 (the actual volume of air is likely to be less than $\frac{1}{2}$ of these estimates). Sudden changes in phase at high frequencies (panels B and D) are due to noisy estimates of phase resulting in 360° phase ambiguities (unwrapping errors).

μl for ear 2 and $20 \mu\text{l}$ for ear 4)⁴ the RW tube, the fluid reservoir was coupled to the RW tube, allowing saline to follow behind the trapped air into the RW tube and into the scala tympani. After the air entered the inner ear, several measurements were made. To verify that the measured effect was caused by the introduced air, bubbles were removed by reversing the fluid flow with the reservoir back on the SSC tube. In a successful flush of the trapped air, bubbles were seen to leave through the RW tube.

Figure 7 shows that air introduced into the inner ear through the scala tympani has an effect primarily at low frequencies. Similar observations were made on the other ears tested. Attempts to remove air were successful to the extent that the measured pressure gain after removal is closer to the pre-air situation. An exception to this is in the mid-frequency region (0.7–1.2 kHz for ear 2, and 2–3 kHz for ear 4) where removing the bubble causes an additional decrease in the measured pressure gain. The reason for this effect is unclear.

4. Air introduced into the vestibule

Air was allowed to enter the vestibule by uncoupling the saline reservoir and raising the level of the SSC tube (2–3 cm). As the inner ear fluid drained through the RW tube, air entered the SSC tube. When a volume of air was observed in the tube, the reservoir was reattached allowing saline to follow behind the trapped air and into the vestibule. After the air entered the vestibule several measurements were made. Figure 8 shows that air introduced into the vestibule affects the pressure transfer ratio primarily at high frequencies for ear 2 (panels A and B) and at both high and low frequencies for ear 4 (panels C and D).

Attempts were made to remove the air from the vestibule by putting the saline reservoir on the RW tube thereby allowing the contents of the labyrinth to be flushed. However, these attempts were all unsuccessful; bubbles never reappeared in the SSC tube. The lack of effect of this procedure on the measured pressure is consistent with the idea that the air introduced was the cause of this change (Fig. 8). Many other measurements made before and after injecting air into the vestibule suggest that small bubbles, which are likely to be trapped in the vestibule, have an effect primarily at high frequencies. These bubbles in the vestibule could be removed only by decoupling the bone from the measurement equipment and repeating the degassing procedure.

The empirical findings of the air injection experiments can now be used to interpret the effect of degassing at the beginning of the measurement session (Fig. 6). For all measurements the temporal bones are continuously flushed throughout the measurement session. This flushing procedure helps remove air bubbles that are in the path of the fluid flow. However, the flushing procedure does not necessarily remove bubbles that are trapped and not in the path of the fluid flow. For example, when air was intentionally injected into the scala vestibule it was difficult to remove the air by simply flushing it. This suggests that air injected through the SSC reaches a space which is not in the fluid path. Degassing helps remove at least some of these trapped bubbles. Degassing has an effect primarily at low frequencies (Fig. 6) suggesting that either prior to the degassing procedure air is trapped in the scala tympani, or subsequent flushing pushes air into the scala tympani. One place where air might get trapped is the space between the round window outlet tube and the round window. Thus the combination of flushing and degassing helps to remove air from the inner ear.

5. Effect of freezing

Most measurements reported in this study are from temporal bones that had been frozen and were thawed at the beginning of each measurement session. To determine the effect of freezing and thawing on the middle-ear pressure gain, measurements were made on a fresh (i.e., not previously frozen) temporal bone. The bone was stored in saline with a drop of Betadyne solution (topical antiseptic); the saline was changed daily. Measurements began eight days postmortem. Figure 9 shows measurements made on two separate days in the fresh state and on two subsequent days after freezing, thawing, refreezing, and rethawing the temporal bone. Because these measurements show no significant differences between fresh and thawed bones, we consider our results with thawed preparations to be representative of “fresh” ears.

III. DISCUSSION

A. A simple model for the effects of air bubbles on vestibule pressure

Our results show that vestibule pressure can be affected at low frequencies by air introduced into the scala tympani (Fig. 7), and at high frequencies when air is introduced into the vestibule (Fig. 8A and B). In both cases the magnitude of

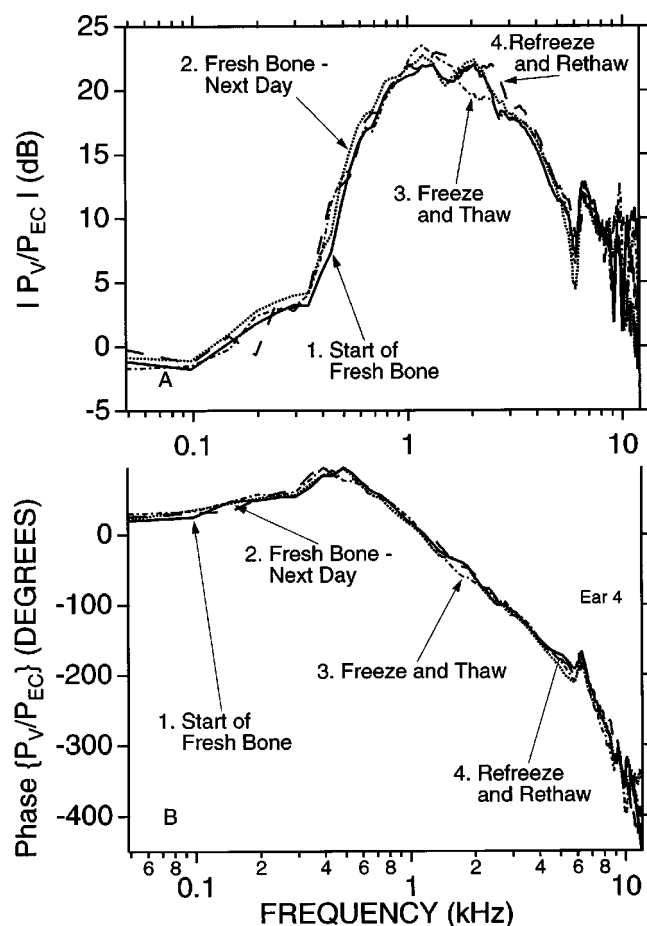


FIG. 9. Effect on the middle-ear pressure gain of repeated freezing and thawing of the temporal bone. Measurements for four conditions are shown: (1) First measurement of a fresh (not previously frozen) temporal bone; (2) second measurement of the fresh temporal bone on the next day; (3) measurement after freezing and thawing the bone; and (4) measurement after refreezing and rethawing.

the pressure typically decreases. These observations can be represented by the simple circuit model shown in Fig. 10, where the middle-ear acoustic input impedance looking out through the oval window (i.e., in the reverse direction) is represented by $Z_{oow}(\omega)$ and the equivalent middle-ear drive is represented by $U_{eq}(\omega)$ (also referred to as the Norton equivalent current). Under the assumption that the cochlea is incompressible (Wever and Lawrence, 1950; SHERA and ZWIG, 1992; KRINGLEBOTN, 1995; VOSS *et al.*, 1996) we represent the inner ear impedance by a one-port impedance Z_c in series with the round-window stiffness K_{rw} .

The effect of an air bubble in the scala tympani can be approximated by a stiffness K_{st} in parallel with the stiffness of the round window K_{rw} (Fig. 10). The magnitude of K_{st} is the ratio of the bulk modulus of air and the volume of the bubble: small bubbles have high impedance (BERANEK, 1954). The effect of K_{st} on P_V depends on the other circuit parameters. KRINGLEBOTN's (1988) model suggests that $|Z_{oow}|$ is considerably larger than $|Z_c|$ at frequencies below 1 kHz. This inequality implies that the middle ear acts as a volume velocity source, and that $|P_V|$ is proportional to the sum of the cochlear and round window impedances. LYNCH *et al.* (1982) and MERCHANT *et al.* (1996) indicate that the impedance of the

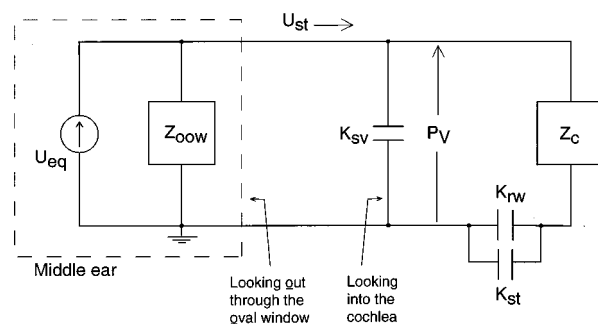


FIG. 10. A circuit representation for the effect of air bubbles on the vestibule pressure. The impedance of the footplate looking out through the oval window is represented by Z_{oow} and U_{eq} is the equivalent volume velocity of the middle ear. The input to the cochlea is the stapes volume velocity U_{st} . The cochlear input impedance is represented by the impedance Z_c in series with the round window impedance $K_{rw}/j\omega$. Elements K_{st} and K_{sv} represent the stiffness of air in the scala tympani and the scala vestibule, respectively. The effect of an air bubble in the scala tympani is represented by K_{st} with K_{sv} infinite, and the effect of an air bubble in the scala vestibule is represented by K_{sv} with K_{st} infinite. For the no bubble case K_{st} and K_{sv} are both infinite.

round window $K_{rw}/j\omega$ is generally of smaller magnitude than Z_c except at frequencies below approximately 100 Hz (see also Fig. 15A). This inequality implies a bubble-induced change in the impedance near the round window will have its biggest effect on the cochlear impedance and P_V at low frequencies which is the behavior we observed when we introduced bubbles into the scala tympani.

The effect of air in the vestibule at high frequencies can be explained by the bubble acting as a shunt stiffness K_{sv} from the vestibule to ground; that is, the bubble impedance appears in parallel with cochlear impedance and round window impedance (Fig. 10). At low frequencies the impedance of the bubble is large compared to the cochlear impedance and it has little effect. At high frequencies, the impedance of the bubble is smaller, such that significant volume velocity is lost compressing the air within the bubble. If the volume of the air is increased, then the cutoff frequency, where the effect of the bubble becomes significant, will decrease.

Air injected into the scala vestibule in ear 4 (Fig. 8C and D) decreases the gain magnitude at both low and high frequencies, while for ear 2 (Fig. 8A and B) the decrease is only at high frequencies. The volume of air injected into the scala vestibule was not measured. However, an upper bound on the air volume, based on inner diameter and length measurements of SSC tubes, can be calculated. For ear 4 the volume of air is less than 30 μl , while for ear 2 it is less than 12 μl . The relatively large volume of air⁴ injected into ear 4 suggests that perhaps part of the volume of air traveled from the scala vestibule, through the helicotrema and into the scala tympani. For this scenario the bubble can have an effect at both low and high frequencies.

B. Previous estimates of the magnitude of the middle-ear pressure gain

Other than our preliminary report (PURIA *et al.*, 1993b), direct measurements of both the magnitude and phase of the

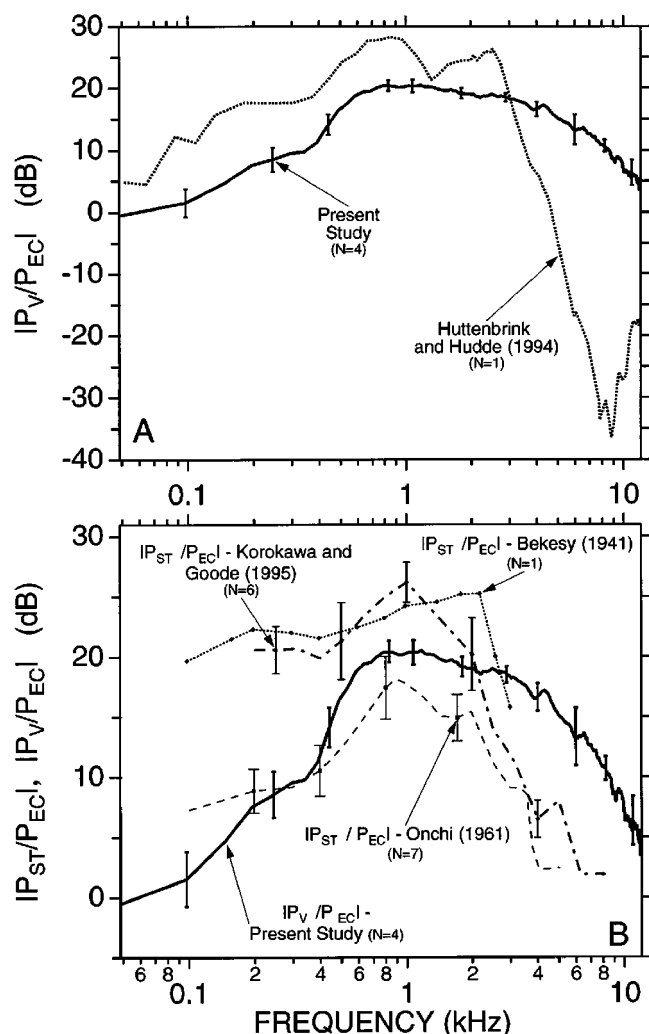


FIG. 11. Comparison with previous middle-ear-gain magnitude measurements. The vertical lines on each plot shows the standard error of the mean (SEM) for each measurement [with the exception of Hüttenbrink and Hudde (1994) where $N=1$ and thus the variability in measurement is unknown]. (A) Two estimates of middle-ear pressure gain from measurements of the ratio of the vestibule pressure and the ear-canal pressure $|P_V/P_{EC}|$. Hüttenbrink and Hudde (1994) measurements are not valid for frequencies above 6 kHz because of calibration errors. (B) The ear canal to stapes pressure gain $|P_{ST}/P_{EC}|$ estimated from measurements made before and after removing the middle ear (Onchi, 1961; Kurokawa and Goode, 1995) and Békésy's (1941, 1960) measurement for the immobilized stapes (see text) are compared with $|P_V/P_{EC}|$. Note that different vertical scales are used for panels A and B.

middle-ear pressure gain have not been previously reported for human ears. A direct measurement of the magnitude of the middle-ear pressure gain was reported by Hüttenbrink and Hudde (1994) in one ear. Their measurement is higher than the average of our measurements for frequencies below approximately 2 kHz (Fig. 11A). Above 3 kHz the measured gain decreases at a rate of approximately 30 dB/oct, much steeper than the 8 dB/oct decrease in the present measurements.

One possibility for the difference is intra-cochlear air as Hüttenbrink and Hudde's methods did not include degassing or flushing of the labyrinth. Although Hüttenbrink and Hudde attempted to prevent air from getting into the temporal bone by keeping the temporal bone under water while the bone was prepared for measurements, the results are consis-

tent with our results with air in the inner ear. A decrease in the measured gain at high frequencies can be expected if bubbles are trapped in the vestibule (Fig. 8A).

A different middle-ear pressure gain is the ear canal to stapes pressure gain P_{ST}/P_{EC} . The stapes pressure gain corresponds to the force of the incus on the head of the stapes divided by the footplate area and thus does not include the pressure drop due to the annular ligament or the stapes mass. The annular ligament has an effect primarily at low frequencies while the effect of the stapes mass is negligible (Lynch *et al.*, 1982; Merchant *et al.*, 1996). Thus the two pressure gains are expected to be similar in the middle and high stimulus frequencies but different at low frequencies.

Two estimates of the magnitude of the stapes pressure gain are compared with the present middle-ear pressure gain measurements in Fig. 11B. Onchi (1961) measured the ear-canal pressure relative to the pressure at the round window before removing the ossicles and the pressure at the stapes (on the middle-ear cavity side) relative to the round-window pressure after removing the TM, malleus, and incus. From these measurements he estimated the magnitude of the stapes pressure gain $|P_{ST}/P_{EC}|$, for the intact middle ear, in the 100 Hz to 5 kHz region. As with the present measurements, Onchi's measurements have a band-pass characteristic. However, the slope of Onchi's measurements is steeper at frequencies above 1 kHz, and shallower at frequencies below 1 kHz, than the present measurements. Kurokawa and Goode (1995) estimated the stapes pressure gain, in the 200 Hz to 9 kHz range, by comparing the stapes displacement to stimulus pressure ratio before and after removal of the TM, malleus, and incus. The Kurokawa and Goode measurements below 2 kHz are higher than the present measurements as expected since these measurements include the pressure used to overcome the impedance of the annular ligament. At frequencies above 2 kHz the Kurokawa and Goode (1995) estimates of the magnitude of the stapes pressure gain are lower than the present measurements. Again, one possibility is that air was trapped in the inner ear in these previous preparations.

Figure 11B also shows Békésy's (1941; 1960, pp. 95–101) measurement of the stapes pressure gain with an immobilized stapes, in the 100 Hz to 3 kHz region. Békésy's measurements were performed using a cancellation technique in which he determined the relative magnitudes of ear canal and intracochlear sound pressures required to cancel the motions produced by either pressure alone. For the range of frequencies where they overlap, the Kurokawa and Goode (1995) measurements yield results that are only slightly lower than Békésy's measurements. Békésy's measurements are consistent with model calculations which suggest that the maximum possible stapes pressure gain achievable is for the immobile stapes condition (Kringelbotn, 1988).

A study not involving pressure measurements but rather the impedance of the stapes and cochlea $Z_{sc}=Z_s+Z_c$ (where Z_s is the stapes and annular ligament impedance and Z_c is the cochlear input impedance) in a fresh (not previously frozen) human temporal bone preparation was recently reported (Merchant *et al.*, 1996). In that study the inner ear was never opened. An upper bound on the size of a bubble in the inner ear of that preparation was estimated to be 0.02 μl .

Air bubbles in the vestibule affect only the Z_c component (Fig. 10) and not the Z_s component. As shown in Merchant *et al.*, for frequencies below about 1 kHz Z_c is much less than Z_s and therefore in this range, measurements of Z_{sc} are not very sensitive to the presence of bubbles. For frequencies above 1 kHz, $|Z_{sc}|$ is comparable to $|Z_c|$ and thus measurements of Z_{sc} can reveal the presence of bubbles.

Indeed, the Merchant *et al.* measurements show that at frequencies above 5 kHz $|Z_{sc}|$ decreases while the angle approaches -90° suggesting the presence of a bubble in the inner ear (see Fig. 15). The impedance corresponding to a 0.02- μ l bubble becomes comparable to Z_c near 5 kHz. Because the bubble impedance appears in parallel with Z_c (Fig. 10), the effective cochlear impedance decreases as frequency increases and thus $|Z_{sc}|$ decreases. Thus in the Merchant *et al.* study, measurements of Z_{sc} and Z_c are likely affected, for frequencies above 5 kHz, by the presence of bubbles in the scala vestibule. For a given stapes velocity, a decrease in cochlear input impedance at high frequencies, due to air, has the effect of decreasing the vestibule pressure. As expected, the presence of a bubble in the vestibule decreases the cochlear input impedance which results in a decrease in $|Z_{sc}|$, and a decrease in the middle-ear pressure gain.

The analysis above suggests that air bubbles can enter the inner ear of even fresh temporal bones. Small cracks in the bone or the ripping of soft tissue during bone extraction can lead to the influx of air into the cochlea. Therefore, it is possible that any bone can have a significant air bubble in it. Thus, it is possible that in previous measurements the effect of air bubbles is significant, because they were not degassed nor were the bones flushed.

In summary, previous middle-ear pressure gain measurements (both $|P_V/P_{EC}|$ and $|P_{ST}/P_{EC}|$) are significantly lower than the present measurements for frequencies above 2–3 kHz. Below 1–2 kHz the present measurements are lower than the Kurokawa and Goode measurements; this is expected because of the added pressure drop across the annular ligament. Unexpectedly, measurements by Onchi are comparable to our measurements for frequencies below 1 kHz (rather than higher as in the Kurokawa and Goode measurements), whereas the one measurement by Hüttenbrink and Hudde is higher than our measurements.

C. Structural correlates of the pressure gain

The middle-ear pressure gain of Fig. 2 can be described in terms of the peak magnitude of the pressure gain and the frequency dependence. In simple theories of the middle ear the magnitude of the pressure gain is determined by dimensions of the middle-ear structures and is independent of frequency. The “ideal mechanical transformer model” proposes that the pressure gain is the product $(A_{tm}/A_{fp})(L_m/L_i)$ where the first term is the ratio of the tympanic-membrane area divided by the stapes-footplate area and the second term is the ratio of the lengths of lever arms of the malleus and incus (Wever and Lawrence, 1954, Chap. 7; Dallos, 1973, pp. 92–95; see also Rosowski, 1995). Mean anatomical measurements for humans (Wever and Lawrence, 1954) then predict a pressure-gain magnitude⁵ of 28 dB and zero phase. Our measurements are about 7 dB below this value in the

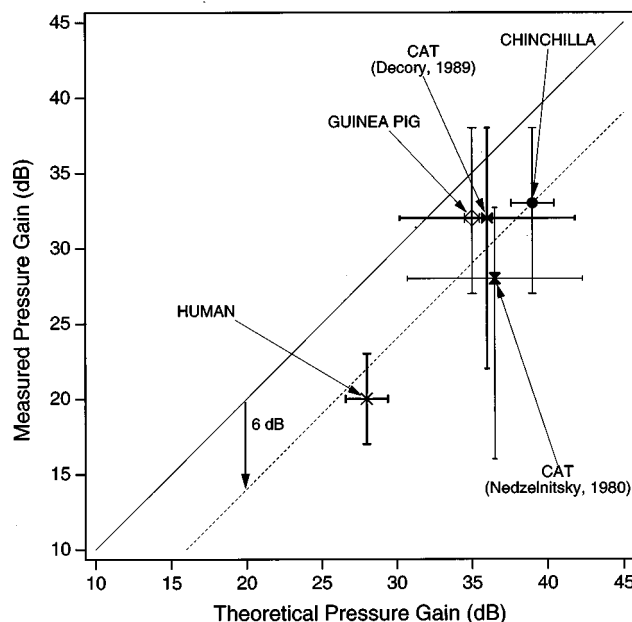


FIG. 12. Theoretical pressure gain, based on the product of (1) the ratio of the TM area (measured at the TM ring) to the footplate area and (2) ratio of malleus length to incus length, is plotted against the measured pressure gain. The symbols show the mean values. The mean of the theoretical gain was obtained from mean anatomical measurements reported by Wever and Lawrence (1954). [For clarity theoretical gain for cat (Nedzelitsky, 1980) was horizontally shifted by 0.5 dB.] The range of the theoretical gain, indicated by horizontal bars, was estimated from reported anatomical measurements (Wever and Lawrence, 1954; Vrettakos *et al.*, 1988; Hemilä *et al.*, 1995) under the assumption that dimensions co-vary. By using the largest and the smallest dimensions two estimates were obtained (see footnote 5 for an example) and the horizontal bars indicate the larger of the two extremes. The mean and range of the measured pressure was obtained at the frequency where the gain is maximum for each species (0.8–1.2 kHz). The vertical bars show the measured range. The guinea pig and chinchilla data are from Décorry (1989). The solid line indicates where the measured gain is equal to the theoretical gain and thus data points above and below indicate deviations of the measured gain relative to the theoretical gain. As shown by the dashed line, the five measured gains, on average, are 6 dB lower than the corresponding theoretical gains.

middle-frequency region, and even lower at frequencies below 400 Hz and above 4 kHz. In addition, the measured phase is not generally zero, and both magnitude and angle vary with frequency. Thus these first measurements of the magnitude and phase of the pressure gain contradict the transformer models. Nevertheless, it may be that the pressure gain in the mid-frequency range does depend on these structural features of the middle ear.

The mean and range of the pressure gain from “transformer theory” are compared with mean and range of the peak measured pressure from four species in Fig. 12. As the theoretical pressure gain increases the corresponding measured pressure gain in the four species also increases suggesting that the measured gain in different species is correlated with the dimensions of the middle-ear structures (Fig. 12). Although the measured middle-ear gain tends to increase with increasing predicted gain, the average of the difference between the five measured gains and corresponding predicted gain is approximately 6 dB (dashed line in Fig. 12).

Figure 13 shows the frequency dependence of the pres-

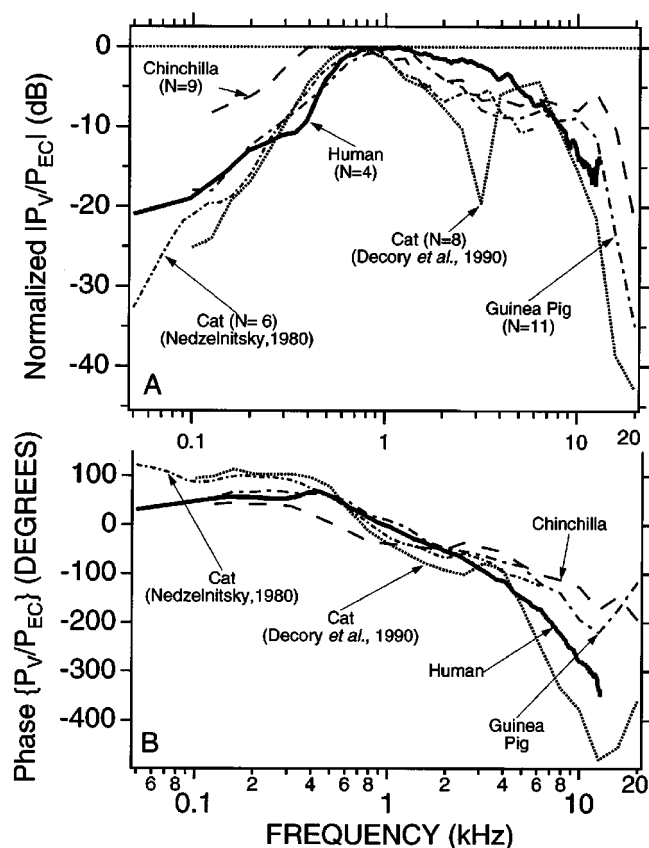


FIG. 13. Inter-species comparison of the average middle-ear pressure gain P_V/P_{EC} . Panel A shows the normalized magnitude, while panel B shows the phase. The average magnitudes were normalized by subtracting 20.52, 30.87, 27.52, 30.23, and 29.69 dB from human, cat (Décoré *et al.*, 1990; Nedzelitsky, 1980), chinchilla, and guinea pig middle-ear gains, respectively. The mean magnitude and phase for the human case is from Fig. 2. For the cat, Nedzelitsky's (1980) measurement is for cavities wide open, whereas the Décoré *et al.* (1990) measurement is with an open bulla cavity and intact septum [the pressure minimum near 3 kHz in the latter case is caused by the interaction of the remaining tympanic cavity and the foramen in the septum (Møller, 1965)]. Measurements in chinchilla and guinea pig are from Décoré *et al.* (1990).

sure gain in the same four species. In Fig. 13 the magnitude of the pressure gain has been normalized by the peak so that the frequency dependence of the four animals can be compared. In the frequency region where they overlap there are remarkable similarities in the frequency dependence of the measurements made in living animal ears and the human temporal bone ears. At low frequencies (below 700 Hz) the gain magnitude decreases at a rate faster than at mid-frequencies (1–6 kHz). As in the human ear the slope of the phase is different across different frequency regions.

In Fig. 12, why is there a tendency for the magnitude of the measured gain to be lower than the theoretical gain? And is this difference related to the frequency dependence of the pressure gain shown in Fig. 13? Network models of the middle ear (Møller, 1961; Zwislocki, 1962; Kringlebotn, 1988) predict deviations from the transformer theory (e.g., see Fig. 14). These deviations are largely attributed to the effect of other forces, e.g., from ligaments and other elastic structures, inertial forces, and slippages of the joints. For example, the pressure drop (or loss) due to middle-ear stiffness is consistent with the observation that the gain decreases

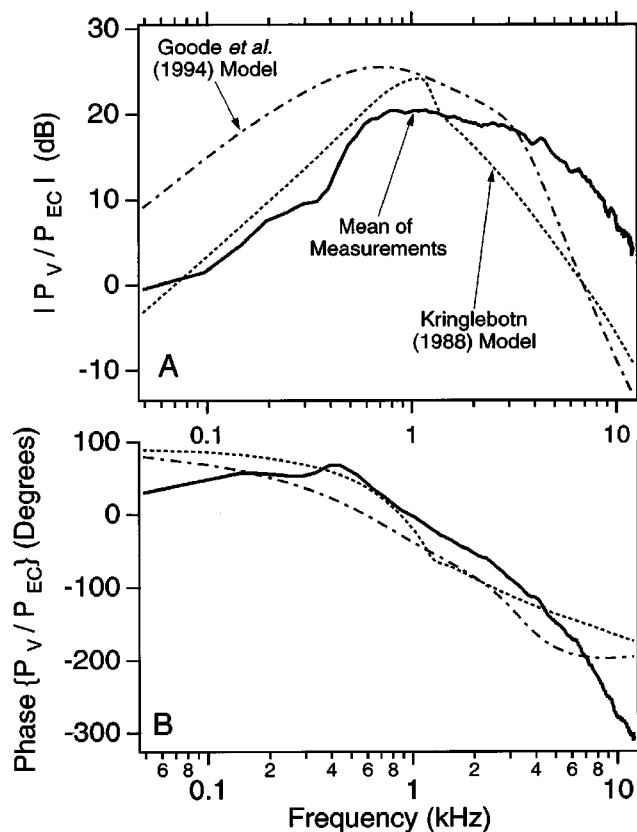


FIG. 14. Comparison of the average measured middle-ear pressure gain with model predictions of the pressure gain (middle-ear cavities wide open). The measured average is based on data of Fig. 2. For the Goode *et al.* (1994) model the inner-ear pressure corresponds to the pressure drop across the fluid mass and resistance of the cochlea. For the Kringlebotn (1988) model the inner-ear pressure corresponds to the pressure drop across the resistance of the cochlea.

as frequency decreases (Fig. 13). Evidence supporting the role of the middle-ear stiffness in determining gain at low frequencies comes from the observation that in human, cat, and guinea pig the middle-ear stiffness is similar, whereas in chinchilla it is lower (Rosowski, 1995, Table 6.3). Consequently, the chinchilla middle-ear pressure gain at low frequencies has a slope similar to the other species but the break frequency is lower (Fig. 13A).

The middle-ear gain also decreases at high frequencies (Fig. 13A). Since the impedance of the middle-ear stiffness is less significant at high frequencies, the decrease in pressure at the higher frequencies is not likely due to the stiffness of the middle ear. It is not clear what causes the decrease at the higher frequencies but one possibility is a decrease in sound transmission to the inner ear due to slippage in the incudo-malleolar joint or incudo-stapedial joint. Indeed, measurements and model results show that slippage between ossicles result in a decrease in stapes displacement (Guinan and Peake, 1967; Peake and Guinan, 1967; Gyo *et al.*, 1987; Goode *et al.*, 1994), which should result in decreased pressure at the higher frequencies.

An alternative explanation to pressure drop due to "other" forces comes from experimental observations that the human tympanic membrane does not move as a piston but rather breaks up into multiple modes (Tonndorf and

Khanna, 1972; Løkberg *et al.*, 1980). Figure 12 shows that the net effect of (1) multiple modes of the TM, and/or (2) forces due to elastic structures, inertia and joint slippage discussed above are on the average approximately 6 dB in the middle-frequency region. By separating the two effects one could address questions regarding the role of the multiple modes of the TM in sound transmission through the middle ear (i.e., does multi-moding of the TM counter the forces due to ligaments and inertia, or do the multiple modes further decrease the pressure gain?). More measurements are needed to separate these two effects.

D. Comparison with middle-ear models

Network models of the human middle ear date back to the early work of Onchi (1949, 1961), Møller (1961), Zwislöcki (1962, 1965), and Shaw and Stinson (1981, 1983) that explain acoustic measurements in terms of the interaction of mechanical and acoustic lumped elements. The predictions of the recent updates to these models (Kringelbott, 1988; Goode *et al.*, 1994) are compared to our measurements in Fig. 14. Although the three magnitude curves all have band-pass shapes, there are some clear differences between measurements and model predictions. The peak gain is approximately 25 dB in the Goode *et al.* model and 24 dB in the Kringelbott model. The maximum measured gain is approximately 5 dB lower. At frequencies below 1 kHz the Kringelbott model magnitude is in approximate agreement with the measured data, whereas in the Goode *et al.* model the magnitude is greater than the measured gain by as much as 12 dB. At low frequencies, both model phases approach 90° whereas the average of the measured phase is less positive and decreases with decreasing frequency.

Above 3 kHz the model calculations and these measurements are in disagreement. In both models the magnitude of the pressure gain is lower than the measured pressure gain. In addition there are differences between the measurements and the Goode *et al.* model in the slope of the gain magnitude. The slopes of the measured and model phases are also in disagreement. Further work is required to determine the reasons for these differences.

E. Cochlear input impedance

The pressure in the vestibule produced by the footplate volume velocity is proportional to the inner-ear input impedance. From the ratio of measured vestibule pressure and stapes volume velocity P_V/U_{ST} we obtain an empirical estimate of the impedance of the cochlea and the round window $Z_{crw}(\omega) = Z_c(\omega) + Z_{rw}(\omega)$.

The mean pressure (Fig. 2) was divided by a fit⁶ to measurements of the round-window volume velocity to ear-canal pressure ratio (Kringelbott and Gundersen, 1985) to obtain the empirical $Z_{crw}(\omega)$ shown in Fig. 15. Also shown in Fig. 15 is a measurement of Z_{crw} from one ear by Merchant *et al.* (1996). This measurement was obtained by taking the difference between the input impedance of the stapes (measured at the head of the stapes) before and after draining the cochlea. For frequencies up to 5 kHz, the phases of the two estimates (Fig. 15B) based on measurements differ by less than 30 deg.

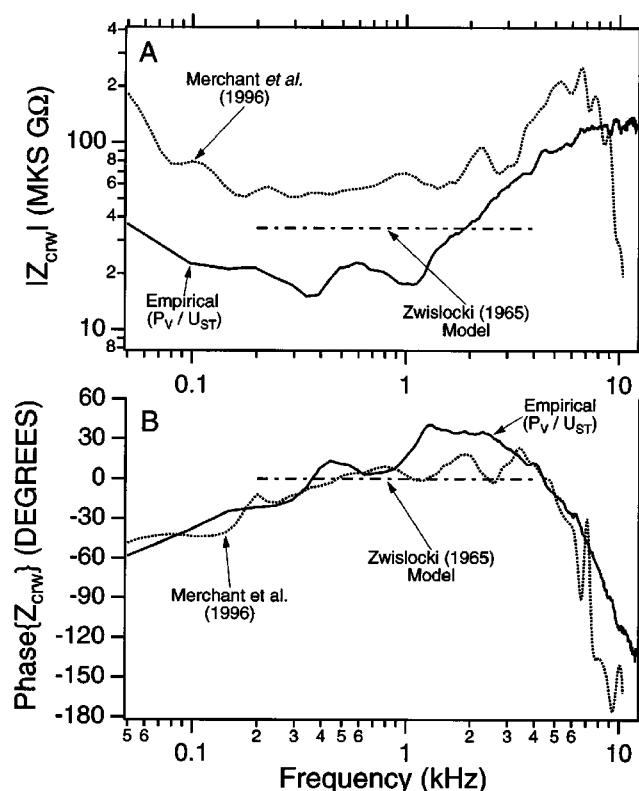


FIG. 15. An empirical estimate of the input impedance for the human cochlea $Z_{crw}(\omega)$ obtained from the ratio of the mean vestibule pressure (Fig. 2) and a model fit⁶ to mean round-window (RW) volume velocity data (Kringelbott and Gundersen, 1985). Measurement of Z_{crw} for one ear by Merchant *et al.* (1996) is also shown. Zwislöcki's (1965) estimate is based on theoretical considerations for the cochlear input impedance.

In the same frequency region, the shapes of the two magnitude curves are similar, but the absolute magnitude of our empirical Z_{crw} is lower than the Merchant *et al.* (1996) measurements by approximately a factor of 3. Both suggest that Z_{crw} is compliance dominated ($-90^\circ \leq \text{Phase}\{Z_{crw}\} \leq -30^\circ$) for frequencies below 0.2 kHz [possibly as a result of round window impedance Z_{rw} (Lynch *et al.*, 1982)], resistive in the 0.3–2 kHz range and above 2 kHz Z_{crw} does not seem to be describable by simple lumped-element circuits.

We have argued in a previous section that for frequencies above 5 kHz the decrease in magnitude of Z_{crw} measured by Merchant *et al.* (1996), and shown in Fig. 15, is likely due to air in the inner ear. Our empirical estimate of Z_{crw} shows no such decrease in magnitude, suggesting that if there were air bubbles present in the measurements presented in Fig. 2 then the volume of those bubbles is likely to be significantly less than the 0.02 μl Merchant *et al.* reported. Our estimates of the phase of Z_{crw} is less than -90° for frequencies above approximately 10 kHz whereas the Merchant *et al.* (1996) estimate of the phase is less than -90° for frequencies above approximately 8 kHz. Since driving point impedances cannot have angles less than -90° (Bode, 1945) there appears to be an error in both measurements at the highest frequencies. For the empirical case, either the phase of the measurements shown in Fig. 2B or the phase of the stapes volume velocity⁶ used to compute Z_{crw} , or both, are in error. Although we cannot tell which is the case, one possi-

bility is that the ear-canal delay in the Kringlebotn (1988) model fit was not explicitly taken into account.

The measurement-based estimates of Z_{crw} are also compared with a theoretical prediction of the cochlear input impedance for the human by Zwislocki (1948, 1965) in Fig. 15. Zwislocki's model prediction suggests that the cochlear input impedance is resistive (i.e., zero phase) and independent of frequency. The measurements suggest that in the 0.3–2 kHz region this is a reasonable approximation. In contrast to Zwislocki (1965), a model prediction of Z_c by Puria and Allen (1991) is a function of frequency and suggests that Z_c is not purely resistive. Differences in the two model predictions is the result of including reflections due to changes in the scalae area apical to the oval window in the Puria and Allen (1991) model, whereas in the Zwislocki model only the characteristic impedance at the base of the cochlea was considered.

One possible reason for the differences in the model predictions of the middle-ear pressure (Fig. 14) is the cochlear load for the middle ear (Puria and Allen, 1995). In the Kringlebotn model the cochlear input impedance is purely resistive, while in the Goode *et al.* model the cochlear load is resistive at frequencies below 5 kHz and mass dominated above approximately 10 kHz, with a transition region in between. Figure 15 suggests that the cochlear impedance is not simply resistive over a wide frequency range but is generally frequency dependent, suggesting that discrepancy between measurements and model results (Fig. 14) might be decreased by using a more realistic cochlear input impedance (Fig. 15).

F. Cochlear input at the threshold of hearing

From knowledge of both the pressure in the inner ear (present study) and the stapes volume velocity (Kringlebotn and Gundersen, 1985; Goode *et al.*, 1994) for a given sound-pressure level at the tympanic membrane it is possible to determine acoustic input to the cochlea at behavioral threshold. In particular, we ask “Do the frequency variations of the sound pressure at behavioral threshold correspond to either constant pressure in the vestibule, constant stapes velocity, constant stapes acceleration, or constant power into the inner ear?”

Different methods have been used to measure minimum audible pressure (MAP) at the TM as a function of frequency by several investigators. Killion (1978) combined various measurements of monaural MAPs into a curve which is shown in Fig. 16. The monaural MAP estimate is based on measurements both in the free field and with an earphone. Both types of measurements were corrected for the ear-canal effects and averaged. According to this estimate the absolute minimum threshold of hearing is approximately 8 dB SPL at the TM at a frequency of 1.3 kHz.

Also shown in Fig. 16 are four curves each of which shows the pressure at the TM required to maintain a different acoustic variable constant at the cochlear input. In each case the constant level of the variable has been chosen so that the minimum of the curve equals the minimum in Killion's threshold estimate. Thus similarities in the frequency dependence of any of these curves to the measured MAP supports

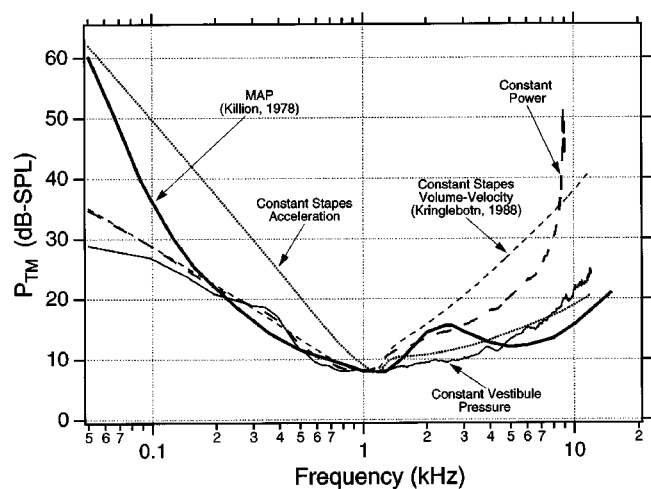


FIG. 16. Comparison of the minimum audible pressure (MAP) in man to four calculated curves each of which shows the TM pressure required to produce a constant magnitude of some acoustic quantity at the inner-ear input. The Killion (1978) curve is an estimate of the averaged MAP near the tympanic membrane. The mean middle-ear pressure gain (Fig. 2A) was used to compute $|P_{\text{TM}}|$ for a “constant vestibule pressure” of 28 dB SPL. A fit to the Kringlebotn and Gundersen (1985) data (see footnote 6) was used to compute a “constant stapes volume velocity” of $3 \times 10^{-5} \text{ mm}^3/\text{s}$. Equation (1) (Sec. III F 3), the cochlear input impedance of Fig. 15, the middle-ear pressure gain of Fig. 2, and $W_{\text{th}} = 15 \times 10^{-18} \text{ W}$ at the input to the cochlea, were all used to compute the constant power curve. A constant stapes acceleration of $65 \mu\text{m}/\text{s}^2$ was computed from the “constant stapes velocity” curve. Threshold calculations for all four acoustic quantities are set to a minimum threshold of 8 dB SPL.

the hypothesis that the cochlea acts as a detector of this variable. We next discuss each of these variables.

1. Constant pressure at the cochlear input

“Do the average pressure measurements of Fig. 2 have the same frequency dependence as behavioral thresholds?” In Fig. 16, the mean of the four ears shown in Fig. 2 is inverted to show the pressure at the tympanic membrane required to maintain a constant pressure of 28 dB SPL in the vestibule (corresponding to the lowest mean MAP at the TM). Comparison to the behavioral measurements indicates that there is reasonable agreement between pressure and behavioral threshold at frequencies above 150 Hz. This conclusion is consistent with observations by Lynch *et al.* (1982, Fig. 25) that shows agreement in behavioral measurements and constant inner-ear pressure in cat at stimulus frequencies above approximately 250 Hz.

At frequencies below 150 Hz the behavioral and constant vestibule pressures diverge. The average slope of the behavioral curve is more than -20 dB/oct , whereas the average slope of the constant pressure curves is approximately -4 dB/oct . Possible reasons for this divergence is that the cochlea actually responds to the pressure difference between the scala vestibule and scala tympani. At high frequencies the scala tympani pressure is much lower in magnitude than the scala vestibule pressure, while at low frequencies the scala tympani pressure can become comparable to the scala vestibule pressure due to the impedance of the round window (Nedzelitsky, 1980). Thus at low frequencies the sound-pressure difference across the cochlea is no longer simply

related to the pressure in the vestibule. The scala tympani pressure in humans is not available and thus it was estimated from the stapes velocity and round window impedance. Subtracting the estimated scala tympani pressure from the measured vestibule pressure increased the threshold by only 3–4 dB at 50 Hz and less at higher frequencies. Thus for frequencies below 150 Hz the hypothesis that at behavioral threshold the inner ear detects a constant pressure difference can be rejected.

2. Constant stapes volume velocity

Next we ask “Do the estimates of stapes volume velocity by Kringlebotn and Gundersen (1985)⁶ have the same frequency dependence as behavioral thresholds shown in Fig. 16?” Figure 16 shows that, in a manner similar to the pressure measurements, the constant stapes volume-velocity curve diverges from the measured MAP data at frequencies below 150 Hz because the two have average slopes that differ by more than 16 dB/oct. For frequencies above 1 kHz the two shapes also differ. The constant stapes volume velocity curve has a slope of approximately 9 dB/oct. MAP estimates vary by about 5 dB in the 1- and 5-kHz regions, and above 5 kHz the slope of the behavioral threshold is also about 9 dB/oct, but the threshold level is lower by as much as 25 dB at 10 kHz. Thus for frequencies above 1 kHz the overall shape of the stapes velocity measurement differs from the behavioral threshold measurements and the hypothesis that the inner ear detects a constant velocity at its input can be rejected.

3. Constant power into the cochlea

It has been previously proposed that the cochlea acts as a power detector at threshold (Khanna and Tonndorf, 1969; Rosowski *et al.*, 1986; Rosowski, 1991). We test this hypothesis by computing the pressure at the TM needed to maintain a constant power into the cochlea as follows:

$$|P_{TM}(\omega)| = \sqrt{\frac{W_{th}}{\text{Re}\{Y_c(\omega)\} |P_V(\omega)/P_{EC}(\omega)|^2}}, \quad (1)$$

where $\text{Re}\{Y_c\}$ is the real part of the reciprocal of the cochlear input impedance (curve marked empirical P_V/U_{ST} in Fig. 15), $|P_V/P_{EC}|^2$ is the square magnitude of the middle-ear pressure gain (Fig. 2), and W_{th} is a constant. [Although P_{TM} and P_{EC} , in Eq. (1), both correspond to ear-canal pressure, different symbols are used to distinguish between the measured variable P_{EC} and the calculated variable P_{TM} .] The TM pressure required to maintain a constant power W_{th} of 15×10^{-18} W into the cochlea is shown in Fig. 16. These calculations are consistent with previous estimates of power at the input to the inner ear⁷ (Rosowski, 1991). While there are similarities between the constant power and threshold curves in the 150 Hz to 3 kHz region, the shapes of the two curves are dissimilar for frequencies below 150 Hz and above 3 kHz, and the hypothesis that the inner ear detects a constant sound power at its input can be ruled out for frequencies below 150 Hz and above 3 kHz.

4. Constant stapes acceleration

In comparison to the constant velocity curve, the behavioral threshold is too high at frequencies below 150 Hz, while at frequencies above 2 kHz the threshold estimate is too shallow. This suggests that a “spectral tilt” in the constant stapes velocity curve could result in a better agreement between the two. One mechanism for such a tilt is the time derivative of the stapes velocity which is the stapes acceleration (in the frequency domain, stapes accelerations = $j\omega$ stapes velocity). Figure 16 shows the sound pressure required at the TM to maintain a constant stapes acceleration of $65 \mu\text{m/s}^2$. The constant acceleration curve is within the range of the MAP measurements for frequencies above 1 kHz and in the 50-Hz region, but in the 75 Hz to 0.7 kHz region, the hypothesis that the inner ear detects a constant acceleration at its input can be ruled out.

5. Behavioral threshold summary

Figure 16 shows that in the 1.5–3.5 kHz frequency region there is at first an increase in the behavioral threshold and then a decrease. None of the physiological measurements show a similar “hump.” Killion (1978) has hypothesized that this hump in sensitivity is due to the middle ear. The constant pressure curve in Fig. 16 is due to the ossicular components of the middle ear and since there is no hump seen in the pressure measurements, the hump in the behavioral measurements cannot be attributed to the ossicles. The possibility still exists that the hump is due to the middle-ear cavities. Indeed measurements and model calculations show a similar decrease in stapes motion when cavities are added to the ossicular chain (McElveen *et al.*, 1982). Thus if the middle-ear gain was measured with the middle-ear cavities closed then it is likely that the “hump” would also be measured in the constant pressure curve.

Guinan and Peake (1967) compared behavioral threshold estimates of MAP in cats to physiological measurements of constant stapes displacement, constant stapes velocity, and constant stapes acceleration. In their analysis the constant stapes acceleration curve is closer to the behavioral threshold measurements than constant displacement or velocity in the 60 Hz to 16 kHz frequency range.

In summary, constant inner-ear pressure is a superior broadband indicator of hearing threshold. At frequencies below 150 Hz, none of the measures tested seem to be correlated to behavioral threshold. Stapes acceleration is correlated to threshold for frequencies above 1 kHz whereas stapes velocity and constant power are not correlated. In the 150 Hz to 1 kHz frequency region constant pressure, constant stapes acceleration, and constant power are correlated with behavioral threshold. This analysis suggests that the inner ear is limited in sensitivity by either: (1) constant pressure for frequencies above 150 Hz; or (2) constant pressure, constant stapes velocity, or constant power for frequencies below 1 kHz and constant stapes acceleration or constant pressure for frequencies above 1 kHz.

G. Future measurements

We have extended the preparation shown in Fig. 1 to include measurement of reverse pressure gain in addition to the forward pressure gain measurements reported here. Preliminary results suggest that the reverse pressure gain is not a simple transformation of the forward pressure gain (Puria and Rosowski, 1996).

IV. SUMMARY

We have developed a method for measuring the middle-ear pressure gain in a human temporal bone preparation. The measurements made in four ears suggest:

- (1) The middle ear has a maximum pressure gain of approximately 20 dB in the 500 Hz to 2 kHz range. Above and below this range the gain decreases.
- (2) The measured pressure gain is due largely to the tympanic membrane and ossicles.
- (3) Air in the normally fluid-filled inner ear reduces the measured gain. Air bubbles injected into the scala tympani reduce the gain mostly at low frequencies, while air bubbles injected into the vestibule reduce the gain mostly at high frequencies. Representation of bubbles as compliances in a simple model approximately accounts for the measured effects.
- (4) Repeated freezing and thawing of the temporal bone has little effect on the measured pressure gain.
- (5) With some exceptions in the mid-frequency range, the pressure gain of existing network models of the human middle ear are inconsistent with the present measurements.
- (6) By taking the ratio of pressure measurements and previous measurements of stapes motion we obtain an estimate of the cochlear input impedance for humans. The estimate suggests that the cochlear plus round window input impedance is compliance dominated at frequencies below 200 Hz, resistive in the 0.3–2 kHz range, and a complicated function of frequency above 2 kHz.
- (7) For frequencies above 150 Hz, the shape of the curve corresponding to constant inner-ear pressure is in approximate agreement with behavioral threshold estimates of minimum audible pressure (MAP).

ACKNOWLEDGMENTS

We are grateful to the members of the Eaton-Peabody and Otopathology Laboratories for assistance. Special thanks to Diane Jones for obtaining the temporal bones and performing the histological work and to M. E. Ravicz for acoustic apparatus design and construction. We also thank Armand L. Dancer, Richard L. Goode, Nelson Y. S. Kiang, Saumil N. Merchant, Brian C. J. Moore, Michael E. Ravicz, Christopher Shera, and Susan Voss for comments and suggestions. Supported (in part) by research Grants Nos. F32 DC00073, R03 DC02677-01, and P01 DC00119 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

¹Measurements in ear 1 have been previously reported (Puria *et al.*, 1993b).

²Theoretical considerations, based on the tube dimensions, imply that the tube impedance is 300×10^9 MKS ohms at 50 Hz and increases proportionately with frequency (Beranek, 1954). The tube impedance is greater than the cochlear input impedance by more than a factor of 5 for frequencies above 80 Hz and at least a factor of 1.5 at 50 Hz (see Fig. 15A). Thus the inlet and outlet tubes are likely to have an insignificant effect on the pressure measurements reported here.

³The SEM is defined as STD/\sqrt{N} , where STD is the standard deviation and N is the number of measurements.

⁴In comparison, the total inner-ear volume based on measurements of cross-sectional areas of cochlear passages (Wever, 1949, p. 279) is approximately 100 μl .

⁵The details of the method used to compute the mean and range of the theoretical pressure gain $T_r = (A_{tm}/A_{fp}) \times (L_m/L_i)$ for all species shown in Fig. 12 are shown here with an example for the human ear. The mean and range of the tympanic membrane area A_{tm} , footplate area A_{fp} , total length of the malleus L_m , and the length along the long process of the incus L_i listed below are from Wever and Lawrence (1954, Appendix D).

	A_{tm} (mm ²)	A_{fp} (mm ²)	L_m (mm)	L_i (mm)	T_r (dB)
min	55.8	2.65	7.6	7	29.4
mean	66.3	3.15	8.35	7	28
max	85	3.75	9.1	7	27.2

Using the mean anatomical values we obtain a mean T_r of 28 dB. Two estimates of the range of the theoretical gain are obtained by assuming that the dimensions of the TM and ossicles increase or decrease together (but not necessarily in a proportional manner). Using the maximum values we obtain one estimate of the range to be 29.4 dB, while another range of 27.2 dB was obtained using the minimum values. The two ranges were subtracted from the mean and the larger of these two estimates (1.4 dB) used for Fig. 12.

⁶A modification to the Kringlebotn (1988) model fit to the Kringlebotn and Gundersen (1985) measurements of the round-window volume velocity was used. For the model fit Kringlebotn (1988) increased the measured mean magnitude of the round-window velocities by 5 dB at all frequencies. The increased velocities were thought to be a better descriptor of the velocities measured in the youngest bones. However, this increase is inconsistent with the mean measurements of stapes volume velocity magnitude by Goode *et al.* (1994). Thus we have used the fit to the mean data originally reported by Kringlebotn and Gundersen (1985) to calculate $Z_{crw}(\omega)$.

⁷Rosowski (1991) calculated the diffuse-field sound pressure required to maintain constant power into the human cochlea, which is similar in shape to the constant power curve shown in Fig. 16 for frequencies below 1 kHz and above 4 kHz. In the 1–4 kHz region, a difference exists between the constant power curve shown in Fig. 16 and Rosowski's (1991, Fig. 12) estimate of the constant power into the cochlea and this difference can be accounted for by external-ear resonance effects (e.g., Blauert, 1983). The power at threshold W_{th} by the two methods also differs. Rosowski (1991) estimated W_{th} to be 1×10^{-18} W, whereas for the present calculations, W_{th} is 15×10^{-18} W. Part of the reason for this difference is that the threshold pressure in Rosowski (1991) was -5.3 dB SPL corresponding to a minimum audible field (MAF), whereas the threshold pressure in Fig. 16 is 8 dB SPL corresponding to a MAP. For a threshold pressure of -5.3 dB SPL W_{th} is estimated to be 0.7×10^{-18} W for the present measurements. The remaining difference is due to differences between the cochlear load used for power threshold calculations (Fig. 15).

Békésy, G. von (1941). "Über die messung der schwingungsamplitude der gehörknöchelchen mittels einer kapazitiven sonde," *Akust. Zh.* **6**, 1–16.

Békésy, G. von (1960). *Experiments in Hearing* (McGraw-Hill, New York).

Beranek, L. L. (1954). *Acoustics* (McGraw-Hill, New York).

Blauert, J. (1983). *Spatial Hearing* (MIT, Cambridge, MA).

Bode, H. W. (1945). *Network Analysis and Feedback Amplifier Design* (Van Nostrand, New York).

Dallos, P. (1973). *The Auditory Periphery-Biophysics and Physiology* (Academic, New York).

Dancer, A., and Franke, R. (1980). "Intracochlear sound pressure measurements in guinea pigs," *Hearing Res.* **2**, 191–205.

Décory, L. (1989). "Origins of interspecific differences in susceptibility to noise," Ph.D. thesis, The University of Bordeaux II.

- Décory, L., Franke, R. B., and Dancer, A. L. (1990). "Measurement of eardrum acoustic impedance," in *The Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. Geisler, J. Matthews, M. Ruggero, and C. Steele (Springer-Verlag, New York), pp. 270–277.
- Franke, R., and Dancer, A. (1985). "Mesures de pression acoustique intracochléaire chez le cobaye," *Rev. Acoust.* **74**, 439–477.
- Goode, R., Ball, G., and Nishihara, S. (1993). "Measurement of umbo vibration in human subjects—method and possible clinical applications," *Am. J. Otol.* **14**, 217–251.
- Goode, R., Killion, M., Nakamura, K., and Nishihara, S. (1994). "New knowledge about the function of the human middle-ear: Development of an improved analog model," *Am. J. Otol.* **15**, 145–154.
- Guinan, J., and Peake, W. T. (1967). "Middle-ear characteristics of anesthetized cats," *J. Acoust. Soc. Am.* **41**, 1237–1261.
- Gyo, K., Aritomo, H., and Goode, R. L. (1987). "Measurement of the ossicular vibration ratio in human temporal bones by use of a video measuring system," *Acta Oto-Laryngol.* (Stockh.) **103**, 87–95.
- Hemilä, S., Nummela, S., and Reuter, T. (1995). "What middle ear parameters tell about impedance matching and high frequency hearing," *Hearing Res.* **85**, 31–44.
- Hüttenbrink, K., and Hudde, H. (1994). "Untersuchungen zur schalleitung durch das rekonstruierte mittelohr mit einem hydrophon," *HNO* **42**, 49–57.
- Khanna, S., and Tonndorf, J. (1969). "Middle ear power transfer," *Arch. Klin. exp. Ohr-, Nas.-u. Kehlk. Heilk.* **193**, 78–88.
- Killion, M. C. (1978). "Revised estimate of minimum audible pressure: Where is the missing 6 dB?," *J. Acoust. Soc. Am.* **63**, 1501–1508.
- Kringlebotn, M. (1988). "Network model for the human middle ear," *Scand. Audiol.* **17**, 75–85.
- Kringlebotn, M. (1995). "The equality of volume displacements in the inner ear windows," *J. Acoust. Soc. Am.* **98**, 192–196.
- Kringlebotn, M., and Gundersen, T. (1985). "Frequency characteristics of the middle ear," *J. Acoust. Soc. Am.* **77**, 159–164.
- Kurokawa, H., and Goode, R. (1995). "Sound pressure gain produced by the human middle ear," *Otolaryngology—Head and Neck Surgery* **113**, 349–355.
- Løkberg, O., Høgmoe, K., and Gundersen, T. (1980). "Vibration measurement in the human tympanic membrane—*in vivo*," *Acta Otolaryngol.* **89**, 37–42.
- Lynch, T. J. III, Nedzelnitsky, V., and Peake, W. T. (1982). "Input impedance of the cochlea in cat," *J. Acoust. Soc. Am.* **72**, 108–130.
- McElveen, J., Goode, R., Miller, C., and Falk, S. (1982). "Effect of mastoid cavity modification on middle ear sound transmission," *Ann. Otol. Rhinol. Laryngol.* **91**, 526–532.
- Merchant, S. N., Ravicz, M. E., and Rosowski, J. J. (1996). "Acoustic input impedance of the stapes and cochlea in human temporal bones," *Hearing Res.* **97**, 30–45.
- Merchant, S. N., Davis, P. J., Rosowski, J. J., and Coltrera, M. D. (1988). "Normality of the input immittance of middle ears from human cadavers," *Assn. Res. Oto-Laryngol. Abstr.* 211–212.
- Møller, A. R. (1961). "Network model of the middle ear," *J. Acoust. Soc. Am.* **33**, 168–176.
- Møller, A. R. (1965). "An experimental study of the acoustic impedance of the middle ear and its transmission properties," *Acta Oto-Laryngol.* **60**, 129–149.
- Nedzelnitsky, V. (1980). "Sound pressures in the basal turn of the cat cochlea," *J. Acoust. Soc. Am.* **68**, 1676–1689.
- Onchi, Y. (1949). "A study of the mechanism of the middle ear," *J. Acoust. Soc. Am.* **21**, 404–410.
- Onchi, Y. (1961). "Mechanism of the middle ear," *J. Acoust. Soc. Am.* **33**, 794–805.
- Peake, W., and Guinan, J. (1967). "Circuit model for the cat's middle ear," *MIT Res. Lab. Electron. Q. Prog. Rep.* **84**, 320–326.
- Puria, S., and Allen, J. B. (1991). "A parametric study of cochlear input impedance," *J. Acoust. Soc. Am.* **89**, 287–309.
- Puria, S., and Allen, J. B. (1992). "SYSid—Audio-band measurement and analysis system," *J. Acoust. Soc. Am.* **92**, 2469.
- Puria, S., Allen, J. B., Elko, G. W., Jeng, P. S., and Kirkegaard, D. L. (1993a). "Measurements and analysis with SYSid," *J. Acoust. Soc. Am.* **93**, 2356.
- Puria, S., Rosowski, J. J., and Peake, W. (1993b). "Middle-ear pressure gain in humans: Preliminary results," in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. Horst, P. van Dijk, and S. van Netten (World Scientific, Singapore), pp. 345–351.
- Puria, S., and Allen, J. B. (1995). "Model and measurements of the cat middle-ear: Evidence of tympanic membrane acoustic delay," *J. Acoust. Soc. Am.* (submitted).
- Puria, S., and Rosowski, J. J. (1996). "Measurement of reverse transmission in the human middle ear: Preliminary results," in *Diversity in Auditory Mechanics*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, and E. L. Hecht-Poinar (World Scientific, Singapore).
- Rosowski, J. J. (1991). "The effects of external- and middle ear filtering on auditory threshold and noise-induced hearing loss," *J. Acoust. Soc. Am.* **90**, 124–135.
- Rosowski, J. J. (1995). "Outer and middle ears," in *Comparative Hearing: Mammals*, edited by R. R. Fay and A. N. Popper (Springer-Verlag, New York).
- Rosowski, J. J., Carney, L. H., Lynch, T. J. III, and Peake, W. T. (1986). "The effectiveness of external and middle ears in coupling acoustic power into the cochlea," in *Peripheral Auditory Mechanisms*, edited by J. B. Allen, J. L. Hall, A. Hubbard, S. T. Neely, and A. Tubis (Springer-Verlag, New York), pp. 3–12.
- Rosowski, J. J., Davis, P. J., Merchant, S. N., Donahue, K. M., and Coltrera, M. D. (1990). "Cadaver middle ears as models for living ears: Comparisons of middle ear input impedance," *Ann. Otol. Rhinol. Laryngol.* **99**, 403–412.
- Schuknecht, H. F. (1993). *Pathology of the Ear* (Lea and Febiger, Philadelphia).
- Shaw, E. A. G., and Stinson, M. R. (1981). "Network concepts and energy flow in the human middle ear," *J. Acoust. Soc. Am.* **69**, S44.
- Shaw, E. A. G., and Stinson, M. R. (1983). "The Human External and Middle Ear: Models and Concepts," in *Mechanics of Hearing*, edited by E. de Boer and M. A. Viergever (Delft U.P., Delft, The Netherlands), pp. 3–10.
- Shera, C. A., and Zweig, G. (1992). "An empirical bound on the compressibility of the cochlea," *J. Acoust. Soc. Am.* **92**, 1382–1388.
- Tonndorf, J., and Khanna, S. M. (1972). "Tympanic-membrane vibrations in human cadaver ears by time-averaged holography," *J. Acoust. Soc. Am.* **52**, 1221–1233.
- Vér, I., Brown, R. M., and Kiang, N. (1975). "Low-noise chambers for auditory research," *J. Acoust. Soc. Am.* **58**, 392–398.
- Voss, S. E., Rosowski, J. J., and Peake, W. T. (1996). "Is the pressure difference between the oval and round window the effective stimulus for the cochlea?," *J. Acoust. Soc. Am.* **100**, 1602–1616.
- Vrettakos, P. A., Dear, S. P., and Saunders, J. C. (1988). "Middle ear structure in the chinchilla: A quantitative study," *Am. J. Otolaryngol.* **9**, 58–67.
- Wever, E. G. (1949). *Theory of Hearing* (Wiley, New York).
- Wever, E. G., and Lawrence, M. (1950). "The acoustic pathways to the cochlea," *J. Acoust. Soc. Am.* **22**, 460–467.
- Wever, E. G., and Lawrence, M. (1954). *Physiological Acoustics* (Princeton U.P., Princeton, NJ).
- Zwislocki, J. (1948). "Theorie der schneckenmechanik," *Acta Oto-Laryngol. Suppl.* **72**, 1–76.
- Zwislocki, J. (1962). "Analysis of the middle-ear function. Part I: Input impedance," *J. Acoust. Soc. Am.* **34**, 1514–1523.
- Zwislocki, J. (1965). "Analysis of some auditory characteristics," in *Handbook of Mathematical Psychology*, edited by R. D. Luce, R. R. Bush, and E. Galanter (Wiley, New York), Vol. 3, Chap. 1, p. 37.

Click-evoked otoacoustic emissions and the influence of high-frequency hearing losses in humans

Paul Avan^{a)}

Biophysics Laboratory, School of Medicine, University of Auvergne, PO Box 38, 63001 Clermont-Ferrand, France

Michel Elbez and Pierre Bonfils

Auditory Research Laboratory, UMR CNRS Neurobiologie de réseaux sensorimoteurs, University Paris V, 78 rue de la Convention, 75015 Paris, France

(Received 18 April 1996; revised 6 January 1997; accepted 9 January 1997)

Click-evoked otoacoustic emissions (cEOAEs) are thought to reflect the presence of highly tuned mechanisms involved in sound processing inside the cochlea. When the sensitivity and tuning of the inner ear are impaired in some frequency range, the spectral components of cEOAEs in the same frequency range are expected to be altered if the previous premise is correct. Although clinical experience does not contradict such an interpretation, fundamental aspects of cEOAE generation and propagation in the cochlea are not clear enough to preclude possible additional influences of remote cochlear places on cEOAE. In order to analyze this possibility, ultra-high-frequency hearing thresholds between 8 and 16 kHz were assessed in 43 human subjects that had clinically normal hearing thresholds in the frequency range of cEOAEs. The magnitude of their cEOAEs was found to be correlated to their average ultra-high-frequency hearing threshold, especially when ears presenting spontaneous otoacoustic emissions were not taken into account ($p=0.002$, $r^2=0.29$). Age and ultra-high-frequency hearing thresholds were correlated ($p<0.01$, $r^2=0.40$); thus it is not possible to exclude that aging was the primary cause of the observed trend. The contribution of ultra-high-frequency hearing status to cEOAE magnitude, perhaps in relation to age, seems to explain a significant part of the variance of "normative" emission data and may be of interest for early detection of high-frequency hearing impairments. © 1997 Acoustical Society of America. [S0001-4966(97)05805-0]

PACS numbers: 43.64.Jb, 43.66.Sr [RDF]

INTRODUCTION

Otoacoustic emissions, first described by Kemp (1978), have turned out to be a promising technique for hearing screening when fast and noninvasive objective measurements are needed. They can be separated into several classes: spontaneous otoacoustic emissions (SOAEs) that are present in the absence of external stimulus, evoked otoacoustic emissions (EOAEs), which can be further subdivided into stimulus frequency otoacoustic emissions (SFOAEs) evoked by continuous pure tones, click-evoked otoacoustic emissions (cEOAEs), and distortion-product otoacoustic emissions (DPOAEs) elicited by couples of pure tones with neighboring frequency. All EOAEs are sounds reemitted by the cochlea, that can be recorded in humans as well as a large number of animal species including mammals. They rapidly disappear when the cochlea is impaired as a result of genetic diseases (Horner *et al.*, 1985), noise exposure (Anderson and Kemp, 1979), ototoxicity (McFadden and Plattsmier, 1984; Brown *et al.*, 1989), or less specific pathological states such as those associated with aging (Bonfils *et al.*, 1988). Thus it is generally accepted that EOAEs are related to the same sensory processes as those responsible for normal hearing sensitivity and tuning, which presumably involve motile cochlear outer hair cells.

Although the correlations between EOAEs and tuning processes make it tempting to assume that all types of EOAEs are frequency-specific indicators of cochlear function, no direct proof of this hypothesis has been produced thus far. Nevertheless, the effect of a continuous pure tone of varying frequency superimposed on the stimulus has been described on SFOAEs (Kemp and Chum, 1980) and DPOAEs (e.g., Brown and Kemp, 1984; Harris *et al.*, 1992). Very sharp suppression tuning curves have been reported, suggesting that these EOAEs present a large degree of frequency specificity. Gorga *et al.* (1993a) applied decision theory to determine the optimal combinations of sensitivity and specificity when using human DPOAEs as predictors of hearing thresholds around the frequencies of eliciting tones; they showed that DPOAEs were indeed suitable predictors. Many clinical studies have confirmed that good correlations can be found between the shapes of audiograms and the frequency spectrum of DPOAEs (Gaskill and Brown, 1990; Avan and Bonfils, 1993; Moulin *et al.*, 1994). There is some evidence that a decrease or lack of cEOAEs in some frequency range indicates an impairment of cochlear function in the part coding for the same frequency range (Norton, 1993). Gorga *et al.* (1993b) reported that TEOAEs analyzed into octave bands performed as well as DPOAEs to distinguish normal from impaired ears and even better than DPOAEs around 1 kHz.

Some experimental data are not fully consistent with

^{a)}Electronic mail: paul.avan@u-clermont1.fr

such an interpretation and perhaps reveal that EOAE behavior is more complex when the cochlea is partly impaired. First, relationships between EOAE level and aging were described by several authors; some of them included elderly subjects whose audiogram was normal for their category of age and therefore exhibited some degree of hearing loss in the range of EOAEs (e.g., Collet *et al.*, 1990, with cEOAEs); it was not really surprising to find that EOAE level decreased with age. Others included subjects with clinically normal hearing in the frequency range of emissions (LePage and Murray, 1991, with cEOAEs; Lonsbury-Martin *et al.*, 1991, with DPOAEs) and found that despite a normal audiogram, EOAE levels kept decreasing with age e.g., according to LePage and Murray, the emitted power of cEOAEs lost an average 8.6 dB/decade for the first two decades, then 4.2 dB/decade above 45 years of age. However, Stover and Norton (1993) measured all types of EOAEs in normally hearing aging subjects and argued that hearing sensitivity should be included as a controlled variable for studies of intrinsic aging processes to be valid. Second, Avan *et al.* (1995) used a guinea pig model of acoustic overexposure to high-frequency tones and showed that although hearing thresholds remained unaffected at EOAE frequencies, cEOAEs were affected in the presence of hearing threshold shifts at much higher frequencies. This finding suggests that EOAEs may also reveal remote changes in cochlear sensitivity in addition to their well-documented sensitivity to changes at the places tuned to their frequencies. An alternative explanation is that EOAEs detect essentially local changes in cochlear function, that these changes may remain “infraclinical” on a pure-tone audiogram and be revealed by more sensitive EOAE measurements.

The goal of the present study was to evaluate a relationship between cEOAEs and high-frequency hearing in adult human subjects. Aging also is accompanied by a large variety of audiometric profiles and even if hearing thresholds remain within normal range between 0.5 and 4 kHz, i.e., the frequency interval in which cEOAEs are normally found, it is well acknowledged that the high-frequency part of the audiogram, namely the 8- to 16-kHz range, provides an early reflection of aging processes (Osterhammel and Osterhammel, 1979; review in Borchgrevink *et al.*, 1996). A group of human adult subjects ranging from 24 to 50 years of age took part to the present work. The inclusion criterion was that audiometric thresholds remained clinically normal at and below 4 kHz, and their ultra-high-frequency hearing was systematically evaluated. Click-evoked otoacoustic emissions were measured and possible correlations were examined between their levels and high-frequency hearing status.

I. METHODS

A. Selection of subjects

Forty-three adult subjects (20 males, 23 females) volunteered to participate in recording sessions. Their age ranged from 24 to 50 years, with a mean of 34.5 yr [standard deviation (s.d.) 7.5]. These subjects had to pass preliminary tests in order to exclude any past or present hearing pathology. An otoscopy was systematically performed using a microscope

and was requested to be strictly normal, in other words, subjects presenting thickened eardrums, scars, or pellucid areas were discarded. Tympanometric peaks had to be centered (± 50 daPa) and stapedius reflex thresholds had to be within normal range. Last, a pure-tone audiogram was performed in a soundproof room at conventional frequencies ranging from 0.25–8 kHz (half-oct steps above 1 kHz). The criterion for normal hearing at and below 4 kHz was that whatever the audiometric frequency f between 0.25 and 4 kHz, pure-tone auditory thresholds $AT(f)$ had to be better than or equal to 15 dB HL. When the two ears of a subject met all these criteria, one of them was chosen at random for further testing. Otherwise, the normal ear was tested.

B. Experiment

High-frequency audiometry was performed between 8 and 16 kHz (1 kHz steps) in the same soundproof room as conventional audiometry, using an Inter Acoustics AS10HF device with KOSS HV/1A earphone. This type of equipment has been used by Dreschler *et al.* (1985) to establish normative data concerning high-frequency hearing in young adults. For every ear, two measurements of hearing thresholds were available at 8 kHz, one from conventional and one from high-frequency audiometer. They always coincided within ± 5 dB.

The presence or absence of SOAEs was assessed using a miniature probe microphone Etymotic ER10B inserted in the external auditory meatus. The output of its preamplifier was sent into a frequency analyzer Hewlett–Packard HP 3561A and analyzed in successive 1-kHz-wide frequency bands from 0.5 to 8.5 kHz (resolution 2.5 Hz). With such a resolution, the noise floor varied from -15 dB sound-pressure level (SPL) at 1 kHz, to $-20/-25$ dB SPL above 2 kHz. Two standard criteria (Probst *et al.*, 1991) had to be met in order to ensure that a SOAE was present; firstly, a narrow-band signal (total width less than 5 Hz) had to be present at least 5 dB above noise floor; secondly, it had to be repeatable with minor frequency changes when a second measurement was performed 5 min later. Only presence or absence of SOAE was taken into account for further processing.

Click-evoked otoacoustic emissions (cEOAEs) were recorded with ILO88 Otodynamics equipment. Its functions have been described extensively (e.g., in Bray and Kemp, 1987). Briefly, the probe of this equipment contains a miniature stimulator and microphone, connected to a ISA computer board that generates stimuli for the earphone and performs data acquisition from the microphone. The probe tip was tightly adapted to the external auditory meatus with a rubber sealer of appropriate size fitted to the probe tip before insertion. Special care was taken that the rubber probe tip was flush with the end of the probe. The stimuli were clicks with a peak-equivalent sound-pressure level of 71 ± 1.9 dB SPL as controlled *in situ* during recordings by the microphone of the probe. In every ear, this level was obtained by applying a constant electrical drive to the stimulator. The frequency spectrum of clicks was computed and was requested to be flat (within 3 dB), from 0.5 to 4 kHz. The most frequent deviation from this condition was due to stimulus ringing which was clearly visible in the time domain and

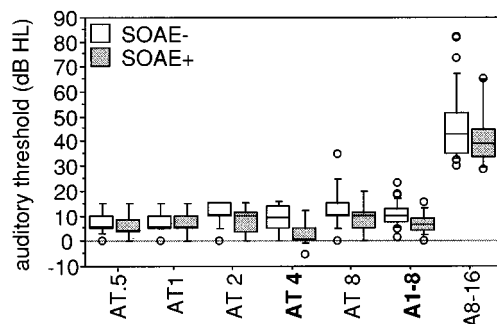


FIG. 1. Boxplots of auditory thresholds at octave frequencies from 0.5–8 kHz are represented for the two categories of subjects (SOAE+, shaded boxes, SOAE–, open boxes). The boxplots of average thresholds in the ranges 1–8 and 8–16 kHz are also depicted. On this figure, the box of a boxplot extends from the 25th to the 75th percentile and the horizontal bar inside the box represents the median. When median and 25th percentile coincide, they are represented with a bold line. The upper whisker extends from the box to the highest data value not exceeding 75th percentile + 1.5 × (75th percentile – 25th percentile). The lower whisker extends from the box to the lowest data value not below 25th percentile – 1.5 × (75th percentile – 25th percentile). Any data beyond these limits is plotted with an open circle. Threshold distributions significantly depend on SOAE presence only for AT4 and A1–8 (bold, $p < 0.05$).

resulted in a resonance peak of the stimulus spectrum around 3 to 4 kHz. A flat click-spectrum was found in almost every ear when the probe tip was correctly set, i.e., flush with the rubber adaptor. In other words, this condition ensured that no stimulus ringing was present; hence, the risk of contamination of the early part of the cEOAE by stimulus ringing was minimum. If not, the rubber adaptor had to be changed until a satisfactory stimulus was obtained. As shown later on, all these conditions were found to be of prominent importance for test/retest consistency.

Clicks were repeated at a rate of 50/s and a temporal averaging of acoustical responses was carried out. Noisy epochs were rejected when the acoustic level within the recording time window exceeded 2.5 mPa: This is a more stringent rejection level than the default setting, i.e., 4.6 mPa. We collected and averaged 260 noise-free samples of data within a time window ranging from 2.5 to 20 ms after click onset. Broadband cEOAE magnitude was computed within this time window. This procedure corresponds to the so-called “linear mode” of the ILO software, more extensively described in Bray and Kemp (1987). In addition, the ILO software allowed control of stimulus level stability every 3 s: It never varied by more than 15% of its initial value. For every subject, cEOAE recordings were done twice at one week intervals with the same electrical drive to the probe’s earphone.

The whole experiment complied with the requirements of French regulations concerning human experiments and with the guidelines of the declaration of Helsinki.

C. Data analysis

The subjects were selected in such a way that all hearing thresholds were normal according to the clinical definition up to 4 kHz (see median hearing thresholds, quartiles, and extreme values in Fig. 1). Nevertheless, in an attempt to

TABLE I. Mean auditory thresholds (s.d. in brackets) in dB HL at the tested frequencies. An asterisk indicates a significant difference between SOAE+ and SOAE– ($p < 0.05$).

Frequency (kHz)	0.5	1	2	4	8	A1–8	A8–16
Total sample	7.8 (4.0)	8.7 (4.1)	10.1 (5.3)	7.6 (6.4)	12.8 (11.0)	9.5 (4.7)	44.8 (13.5)
SOAE–	8.1 (4.1)	8.8 (3.9)	11.0 (4.6)	9.0 (6.1)	14.3 (12.2)	10.5 (4.6)	46.2 (14.0)
SOAE+	7.7 (4.0)	8.5 (4.7)	8.1 (6.3)	4.2* (6.1)	9.2 (6.4)	7.1* (4.3)	41.5 (11.9)

better characterize the audiometric status within this frequency range, the area $A(1,8)$ of every audiogram, i.e., delimited by actual audiogram curve and x axis between 1 and 8 kHz, was computed using dB HL and octave as axis units. The area of the high-frequency audiogram was also computed, according to the formula $A(8,16) = \sum AT(f)/8$ for every f ranging from 8–16 kHz (1-kHz step). Actually this formula is slightly approximative because thresholds should be sampled every 1/8th oct instead of every kHz. Then the two obtained values were normalized to a 1-oct width according to the relationships $A1–8 = A(1,8)/3$ and $A8–16 = A(8,16)$. The unit for A8–16 is the same as for A1–8, i.e., dB HL/octave. Subjects were divided into two subgroups depending on the presence (SOAE+, $n = 13$) or absence (SOAE–, $n = 30$) of spontaneous otoacoustic emissions. The main reason for this subdivision is that SOAE+ ears were expected to give rise to larger cEOAEs than SOAE– ears, because a large fraction of cEOAE power was influenced by synchronized spontaneous emission components. The overall magnitude of cEOAEs, i.e., cEOAEmag including the whole frequency spectrum was evaluated in dB. Pearson’s correlation coefficients and least-squares-fit lines were computed between A1–8, A8–16, age, and cEOAEmag, for the total sample of ears and for the two subgroups SOAE+ and SOAE–.

II. RESULTS

A. Audiometric data

At 8 kHz, only 9 subjects out of 43 had abnormal thresholds, i.e., 20 dB HL in five cases, 25 dB HL in two cases and 35 dB HL in two cases. Accordingly, A1–8 had a small mean value of 9.5 dB/oct (s.d. 4.7), with a minimum of 0.0 and a maximum of 23.3. After splitting the subjects into categories SOAE+ ($n = 13$) and SOAE– ($n = 30$), Student’s t tests (unpaired comparisons) revealed a small but significant difference between categories for mean auditory thresholds at 4 kHz and for A1–8 (Fig. 1 and Table I). No significant influence of category was found for other mean hearing thresholds.

The distribution of A8–16 describing the ultrahigh-frequency hearing status between 8 and 16 kHz was much more scattered with a mean value of 44.8 dB/oct (s.d. 13.5) and extreme values of 28.7 and 82.5 dB/oct. There was no significant difference between SOAE+ and SOAE– subgroups (Table I). In contrast with A1–8 ($r^2 = 0.03$, p

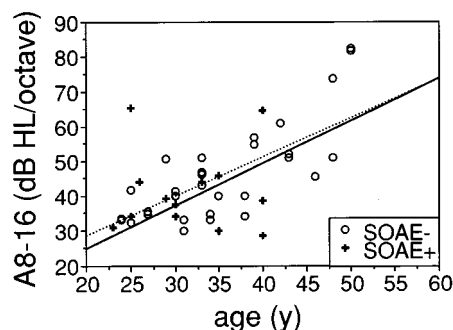


FIG. 2. The averaged ultra-high-frequency hearing threshold A8–16 is plotted against age. The regression line for SOAE– ears is significant (solid line, $p < 0.01$; $r^2 = 0.40$) (SOAE– ears, circles; SOAE+ ears, crosses; dashed line: regression computed from all ears, SOAE+ and SOAE–).

≥ 0.05), A8–16 significantly depended on age (Fig. 2) according to the following linear regression: $A8-16 = 6.2 + 1.12 \times \text{age}$ ($r^2 = 0.40$, $p < 0.01$).

B. Characteristics of cEOAEs, relation to audiometric data

Click-evoked otoacoustic emissions were found in every selected ear with numerous frequency components spreading throughout an interval (f_{\min}, f_{\max}), with f_{\min} ranging from 0.62 to 0.92 kHz, and f_{\max} from 3.22 to 5.50 kHz. Spontaneous otoacoustic emissions were found in 13 ears at one or several frequencies between 1 and 5 kHz and resulted in long-lasting cEOAE components at the same frequencies that could be considered as synchronized spontaneous emissions. In SOAE– ears, most of cEOAE energy was found in the frequency interval 0.6–2 kHz.

When measurements were repeated at 1-week intervals with a strictly controlled probe tip position, test–retest differences never exceeded 1 dB for the peak level of clicks and 0.45 dB for the overall magnitude of cEOAE. When the probe tip was moved from its ideal position, i.e., flush with the probe adaptor, to a deeper position, stimulus ringing occurred together with obvious spectral changes around 3 to 4 kHz. The measured stimulus peak levels might vary by as much as 12 dB, with a mean of 4.6 dB when the adaptor was 6 mm ahead of the probe tip. In contrast, cEOAEs were little affected by probe tip position provided the electrical drive to the probe earphone was set constant, and concomitant emission magnitude changes never exceeded 0.8 dB.

The mean magnitude of cEOAEs was 6.5 dB (s.d. 6.6 dB, range 25 dB). It varied significantly with category, with a mean value of 4.2 dB (s.d. 5.7) in SOAE– subjects and 12.1 (s.d. 5.0) in SOAE+ ears (significant mean difference of 7.9 dB; $p = 0.0001$). As warranted by the selection of subjects, no systematic relationship was observed between cEOAE magnitude and audiometric status below 8 kHz. None of the Pearson's correlation coefficients between cEOAE level and any of the auditory thresholds from 0.25 to 8 kHz, nor A1–8, was significant ($p > 0.05$). Similar results were found within each SOAE category. These results are illustrated by the scatterplot in Fig. 3(a).

A significant linear regression was found between cEOAE magnitude and A8–16 [Fig. 3(b), cEOAEmag

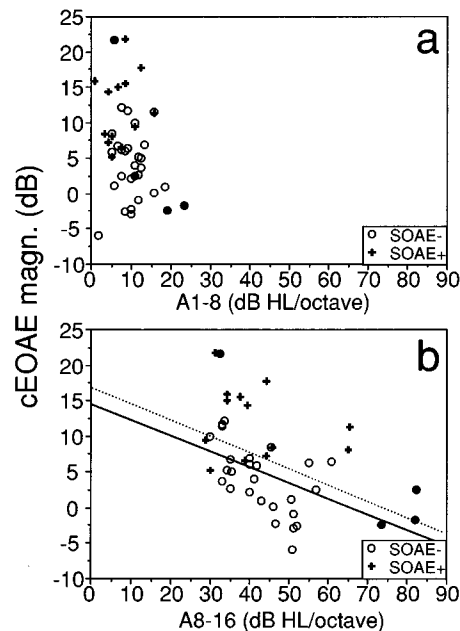


FIG. 3. (a) Broadband cEOAE magnitudes are plotted against averaged hearing thresholds between 1 and 8 kHz (SOAE– ears, circles; SOAE+ ears, crosses). A1–8 values seldom exceed 15 dB HL/octave. Mean cEOAE magnitude is larger for SOAE+ than SOAE– ears. Four outlying data collected in SOAE– ears are pinpointed using closed circles. (b) The same data as in (a) (same symbols) are plotted against ultra-high-frequency hearing status; the solid line describes the negative correlation between cEOAE magnitude of SOAE– ears and A8–16 ($p < 0.0001$, $r^2 = 0.29$). The dashed line is obtained when all ears are included regardless of spontaneous emission presence or absence ($r^2 = 0.23$).

$= 16.9 - 0.23 \times A8-16$, $n = 43$, $r^2 = 0.23$, $p = 0.001$], whereas the other correlations between cEOAEmag and audiometric data were not significant. When subjects were split into SOAE categories, the above mentioned relationship remained almost unchanged for SOAE– ears [Fig. 3(b), slope -0.22 , $n = 30$, $r^2 = 0.29$, $p = 0.002$]. The magnitude of cEOAE was significantly correlated not only with A8–16, but also with age, with a slope of -0.53 in the complete sample ($r^2 = 0.37$, $n = 43$, $p = 0.001$), and -0.30 for SOAE– ears ($r^2 = 0.25$, $n = 30$, $p = 0.01$, Fig. 4). No significant correlation could be found in the smaller sample of SOAE+ ears ($n = 13$, $r^2 = 0.05$, $p = 0.44$ for cEOAEmag versus A8–16).

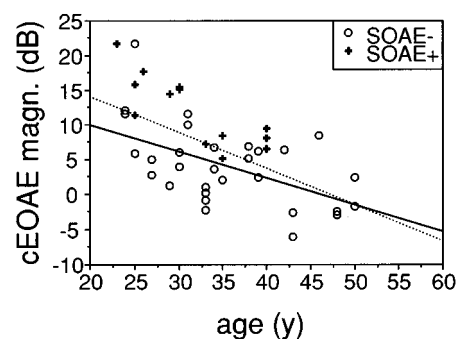


FIG. 4. When plotted against age, cEOAE magnitude can also be described with a significant regression line (solid line, $p = 0.001$; $r^2 = 0.37$) (same symbols as in Fig. 2).

C. Control of the possible effect of outlying data

Figure 3(b) shows that the distribution of datapoints in the scatterplot was quite even except for a few ears represented with closed circles: One of them had a very large cEOAE magnitude and three other ones had unusually poor high-frequency hearing thresholds. Removing them from the computation of regression allowed us to check that they had little influence on the general trend for cEOAEs versus A8–16 which remained significant, i.e., $r^2=0.31$ ($n=26$, $p=0.002$).

III. DISCUSSION

A. General

The present experiment demonstrates that the magnitude of cEOAEs exhibits a significant correlation with the audiometric status of the part of the cochlea tuned to ultra-high frequencies, and also with age. In healthy human adult subjects with clinically normal auditory thresholds up to 4 kHz, i.e., in the frequency interval of most cEOAE components, the overall magnitude of cEOAEs shows a fairly large variance. Twenty-three percent of this variance relates to the values of hearing thresholds in the remote ultra-high-frequency interval 8–16 kHz. Some of the residual variance seems to be associated with the presence of synchronized spontaneous emissions within cEOAE responses, so that the amount of variance of cEOAE magnitude that can be taken into account by the variations of high-frequency auditory thresholds increases up to 29% when only SOAE- ears are considered. As it is well-known that ultra-high-frequency thresholds actually represent early gauges of cochlear aging (Osterhammel and Osterhammel, 1979), it is not surprising to find that cEOAE magnitude is also correlated with age. However, it is noteworthy that this correlation exists in the absence of any age-related change in auditory threshold at and below 4 kHz. The linear relationship disclosed between cEOAE magnitude and ultrahigh-frequency thresholds implies a mean decrease of about 2 dB SPL per 10 dB HL increase of auditory thresholds.

B. Technical considerations (probe position, recording mode)

Although statistically significant, the observed trend is small thus special caution in the measurement technique was needed to observe it. Several questions had to be addressed; namely, (1) how to adjust the probe adaptor to obtain a repeatable stimulus with flat spectrum and constant magnitude, (2) what voltage to send to the earphone to observe a constant response that could be considered as characteristic of the measured ear, and (3) what recording mode would be best suitable to reveal the properties of cEOAE responses.

First, the ideal condition to insert the probe in the external auditory meatus is known to be the one with a probe adaptor flush with probe tip (Kemp *et al.*, 1990). It ensured a repetitive stimulus magnitude with a flat frequency spectrum between 0.5 and 4 kHz. Second, when a constant voltage was sent to the earphone, both in-the-ear stimulus amplitude and cEOAE magnitude proved to remain stable at 1-week inter-

vals. In contrast, when the probe tip was not ideally placed, stimulus ringing was observed probably in relation to resonance phenomena (Kemp *et al.*, 1990). Ringing biased *in situ* stimulus measurements and more than 12-dB changes in measured stimulus levels could be observed in the same ear from one test to another. In contrast, the magnitude of cEOAEs was very little affected by changes in probe tip versus adaptor tip positions provided the voltage drive to the probe earphone was kept constant. This indicates that the input level into the cochlea was also constant, hence a constant voltage level seems to guarantee a constant SPL at the tympanic membrane. Clinical procedures involving a correction of the voltage input to the probe earphone in order to obtain a constant “target” click level at the probe microphone (review in Whitehead *et al.*, 1995) may be incorrect if the probe tip position is not appropriate: Such unduly performed corrections may lead to an increase in the variance of cEOAE magnitude data.

In the absence of stimulus ringing, cEOAEs were recorded using the so-called linear mode (Kemp *et al.*, 1990) and average click levels of 71.9 dB p.eq SPL. This paradigm is simpler than standard clinical testing procedures. Recall that standard stimulus levels are about 10 dB louder; thus the risk that residual stimulus ringing influences the early part of cEOAE responses increases. It makes it necessary to use the so-called nonlinear mode. Briefly, in such conditions, cEOAE amplitudes no longer increase linearly with stimulus level; their nonlinearly growing part, which is free from linearly growing stimulus ringing, is derived from a computation involving cEOAEs elicited by two different stimulus levels. The obtained “cEOAE” response actually is the result of this nonlinear derivation. It is not straightforward to interpret small changes in the magnitude of derived cEOAEs (Grandori and Ravazzani, 1993): It is the reason why we chose to work with directly recorded cEOAEs.

C. Multiple influences on cEOAEs: Spontaneous emissions, age, ultra-high-frequency hearing status

Subject selection was deliberately restricted to clinically normal pure-tone audiograms up to 4 kHz with normal middle-ear function. Auditory thresholds in the range of normal cEOAEs were not expected to be confounding factors. Indeed, no relationship was found neither between average auditory threshold A1–8 in dB/octave and cEOAE magnitude, nor between A1–8 and age. On the other hand, auditory thresholds at frequencies >4 kHz were let free to vary from normal limits to more or less severe impairment. Actually, apart from a few exceptions, these impairments were restricted to the ultra-high-frequency range. One goal of subject selection was to obtain a homogeneous age distribution among the final selection of subjects, i.e., an almost constant number of subjects in every 5-yr-wide class of age; meeting strict audiometric criteria was increasingly difficult when age exceeded 40 and this explains the relatively small size of the final sample. Despite this size, a clear and significant correlation appeared between magnitude of cEOAEs and ultra-high-audiometric status.

The percentage of SOAE+ ears in this sample was 30% in good agreement with many other reports in normally hear-

ing adults (see review in Probst *et al.*, 1991). Presence of SOAEs in one ear is thought to be the reflect of a particularly healthy cochlea (Bonfils, 1988). Accordingly, a mean difference of 3 dB has been reported by McFadden and Mishra (1993) between hearing thresholds in SOAE+ and SOAE- ears. In the present work, a small but significant difference in the mean A1-8 was found between the two subgroups, and its amount was 3.4 dB in good agreement with McFadden and Mishra. The presence of SOAE led to peculiar cEOAE recordings, because synchronization of SOAEs by click stimuli systematically resulted in the presence of long-lasting components whose duration exceeded that of the time window of analysis. As already pointed out by Kulawiec and Orlando (1995), their contribution to the overall cEOAE magnitude was often prominent, and the average cEOAE magnitude was much larger for SOAE+ than SOAE- ears, by almost 8 dB. For this reason, it seemed more appropriate to carry out a separate analysis of SOAE+ and SOAE- ears.

Owing to the small number of subjects in subgroup SOAE+, it is difficult to derive any conclusion as to the influence of A8-16 on their cEOAE magnitude. An important point is that the same trend was found in the SOAE- subgroup as for the total sample. In this subgroup, it is therefore possible to conclude that although cEOAEs do not contain significant energy above 4 kHz, and although their magnitude is dominated by frequency components below 2 kHz, they reflect to some extent the ultrahigh-frequency hearing impairment that probably arises from a progressive impairment of auditory receptors in the basalmost part of the cochlea. Surprisingly, more than two octaves separate the places tuned to the frequencies of emissions from the early impaired cochlear base.

Figure 2(a) shows that such impairment tends to gradually appear with increasing age, so that age and A8-16 are not independent variables. Thus it does not seem possible to decide whether the observed decrease of cEOAEs is primarily due to ultra-high-frequency impairment or to aging. Aging by itself may not be a well-defined univariant process and other factors, such as general deterioration of the cochlear function, noise exposure throughout life, genetic background, etc. may also be important. There is also evidence from non-noise-exposed populations of remote countries, that aging alone is not necessarily a factor of audiometric deterioration (review in Borchgrevink *et al.*, 1996). In a recent observation based on two groups of 20 subjects aged between 19 and 29 and 40 and 61 yr with carefully matched hearing thresholds, Prieve and Falter (1995) reported that the age dependence of cEOAE magnitude was not significant. Unfortunately they did not mention ultrahigh-frequency hearing thresholds. A similar study on larger samples and with a systematic evaluation of high-frequency hearing sensitivity, exposure to noise, general physiological status, etc., might allow to separate age contribution from other factors. An alternative complementary experiment might consist of constraining even more the range of age of subjects in order to limit the possibility that data variance may be influenced by aging, and check the influence of ultra-high-frequency hearing status alone.

D. Possible consequences

The present results may contribute to a tentative explanation of part of the large variance that is commonly observed in normative cEOAE data (Smurzynski and Kim, 1992). The amount of spread that can be accounted for by ultra-high-frequency hearing status is given by the determination coefficients, i.e., about 30%; thus it is clear that a number of other variables are needed to account for a greater portion of the variance. Two alternative hypotheses can be proposed to account for the influence of ultrahigh-frequency hearing status. First, it is well acknowledged that on average, there is a baso-apical gradient of age-induced hair cell lesions inside the cochlea. It is possible to speculate that pure-tone audiometry may not be very sensitive to beginning lesions of outer hair cells, so that a significant percentage of basal hair cells may be impaired before an elevation of ultra-high-frequency hearing thresholds is detected. Audiometric impairment would be restricted to high-frequencies even in the presence of damage to more apical hearing receptors. If cEOAEs provide an earlier functional reflect of this damage at places tuned to cEOAE frequencies, they may apparently anticipate standard audiometry. If this premise is correct, the present data would mean that cEOAEs are much more sensitive than pure-tone audiometry to age-related loss in sensitivity, as there is a shift of more than 2 oct between the progressions of these two tests.

Alternatively, cEOAEs might partly depend on the function of basalmost cochlear places. Although the present data are insufficient to support this hypothesis, there is some experimental evidence (Avan *et al.*, 1995) that in guinea pigs exposed to loud high-frequency pure tones, temporary high-frequency threshold shifts are observed between 5 and 20 kHz, and cEOAE amplitudes at frequencies below 5 kHz are temporarily decreased in the absence of any threshold elevation at these frequencies. It is important to note that all these data are purposely restricted to cochleas with normal hearing thresholds in the frequency range of cEOAEs. Obviously they do not preclude a good correspondence between cochlear damage at a given place and cEOAE components at corresponding frequencies, as reported in several clinical studies (Gorga *et al.*, 1993b; Norton, 1993). No inference can be drawn from such studies about possible influences of cochlear base because auditory thresholds above 8 kHz were not mentioned in these reports.

A recent abstract by Arnold *et al.* (1996) shows that a similar trend is present with another category of otoacoustic emissions, i.e., DPOAEs. They are thought to be a highly frequency-specific probe of the cochlea because they are generated around the places tuned to primary tones. However, DPOAEs elicited by primary tones between 4-8 kHz and around 2 kHz appeared to be influenced by ultra-high-frequency thresholds in a sample of 50 human subjects whose audiometric characteristics were similar to ours [linear relationship, slope 0.18 to 0.25, Lonsbury-Martin (personal communication) to be compared to 0.22 in the present work, r^2 determination coefficients 0.16-0.18 vs 0.23-0.29 for cEOAEs]. Hence, the described results do not seem specific of one particular category of otoacoustic emissions.

The reported findings that the magnitude of cEOAEs

reflects partly the influence of high-frequency hearing losses may be of interest in situations such as follow-up of antimitotic or aminoglycoside treatments (Hotz *et al.*, 1994). Their early ototoxic effect involves outer hair cells and primarily those tuned to high frequencies (Henley and Rybak, 1995). However, the weakness of the expected effect makes it necessary to design careful monitoring paradigms. Further modeling work will be needed to account for these properties of otoacoustic emission generators.

ACKNOWLEDGMENTS

Parts of this work were funded by Grant No. 4AIC04 from CNAMTS-INSERM and by "Fondation pour la Recherche Médicale."

- Anderson, S. D., and Kemp, D. T. (1979). "The evoked cochlear mechanical response in laboratory primates," *Arch. Otolaryngol.* **224**, 47–54.
- Arnold, D. J., Lonsbury-Martin, B. L., Martin, G. K., and Stagner, B. B. (1996). "Influence of sensitivity of ultra-high frequency hearing on distortion-product otoacoustic emission levels in humans," *Abstr. Assoc. Res. Otolaryngol.*, Vol. 19, edited by G. R. Popelka, 25 pp.
- Avan, P., and Bonfils, P. (1993). "Frequency specificity of human distortion product otoacoustic emissions," *Audiology* **32**, 12–26.
- Avan, P., Bonfils, P., Loth, D., Erminy, M., and Elbez, M. (1995). "Transient-evoked otoacoustic emissions and high-frequency acoustic trauma in the guinea pig," *J. Acoust. Soc. Am.* **97**, 3012–3020.
- Bonfils, P. (1988). "Spontaneous otoacoustic emissions: Clinical interest," *Laryngoscope* **99**, 752–756.
- Bonfils, P., Piron, J. P., Uziel, A., and Pujol, R. (1988). "A correlative study of evoked otoacoustic emission properties and audiometric thresholds," *Arch. Otolaryngol.* **245**, 53–56.
- Borchgrevink, H. M., Hallmo, P., and Mair, I. W. S. (1996). "Extended high-frequency hearing loss from noise exposure," in *Scientific Basis of Noise-Induced Hearing Loss*, edited by A. Axelsson, H. Borchgrevink, R. P. Hamernik, P. A. Hellstrom, D. Henderson, and R. J. Salvi (Thieme, New York), pp. 299–312.
- Bray, P., and Kemp, D. T. (1987). "An advanced cochlear echo technique suitable for infant screening," *Brit. J. Audiol.* **21**, 191–204.
- Brown, A. M., and Kemp, D. T. (1984). "Suppressibility of the $2f_1-f_2$ stimulated acoustic emissions in gerbil and man," *Hearing Res.* **13**, 29–37.
- Brown, A. M., McDowell, B., and Forge, A. (1989). "Acoustic distortion products can be used to monitor the effects of chronic gentamicin treatment," *Hearing Res.* **42**, 143–156.
- Collet, L., Gartner, M., Moulin, A., and Morgon, A. (1990). "Age-related changes in evoked otoacoustic emissions," *Ann. Otol. Rhinol. Laryngol.* **99**, 993–997.
- Dreschler, W. A., van den Hulst, R. J., Tange, R. A., and Urbanus, N. A. (1985). "The role of high-frequency audiometry in early detection of ototoxicity," *Audiology* **24**, 387–395.
- Gaskill, S. A., and Brown, A. M. (1990). "The behavior of the acoustic distortion product, $2f_1-f_2$, from the human ear and its relation to auditory sensitivity," *J. Acoust. Soc. Am.* **88**, 821–839.
- Gorga, M. P., Neely, S. T., Bergman, B. M., Beauchaine, K. L., Kaminski, J. R., Peters, J., and Jesteadt, W. (1993a). "Otoacoustic emissions from normal-hearing and hearing-impaired subjects: distortion product responses," *J. Acoust. Soc. Am.* **93**, 2050–2060.
- Gorga, M. P., Neely, S. T., Bergman, B. M., Beauchaine, K. L., Kaminski, J. R., Peters, J., Schulte, L., and Jesteadt, W. (1993b). "A comparison of transient evoked and distortion product otoacoustic emissions in normal-hearing and hearing-impaired subjects," *J. Acoust. Soc. Am.* **94**, 2639–2648.
- Grandori, F., and Ravazzani, P. (1993). "Nonlinearities of click-evoked otoacoustic emissions and the derived nonlinear technique," *Br. J. Audiol.* **27**, 97–102.
- Harris, F. P., Probst, R., and Xu, L. (1992). "Suppression of the $2f_1-f_2$ otoacoustic emissions in humans," *Hearing Res.* **64**, 133–141.
- Henley, C. M., and Rybak, L. P. (1995). "Ototoxicity in developing mammals," *Brain Res. Rev.* **20**, 68–90.
- Horner, K. C., Lenoir, M., and Bock, G. R. (1985). "Distortion-product otoacoustic emissions in hearing-impaired mutant mice," *J. Acoust. Soc. Am.* **78**, 1603–1611.
- Hotz, M. A., Harris, F. P., and Probst, R. (1994). "Otoacoustic emissions: an approach for monitoring aminoglycoside-induced ototoxicity," *Laryngoscope* **104**, 1130–1134.
- Kemp, D. T. (1978). "Stimulated acoustic emissions from within the human auditory system," *J. Acoust. Soc. Am.* **64**, 1386–1391.
- Kemp, D. T., and Chum, R. (1980). "Properties of the generator of stimulated acoustic emissions," *Hearing Res.* **2**, 213–232.
- Kemp, D. T., Ryan, S., and Bray, P. (1990). "A guide to the effective use of otoacoustic emissions," *Ear Hear.* **11**, 93–105.
- Kulawiec, J. T., and Orlando, M. S. (1995). "The contribution of spontaneous otoacoustic emissions to the click-evoked otoacoustic emissions," *Ear Hear.* **16**, 515–520.
- LePage, E. L., and Murray, N. M. (1991). "Normative otoacoustic emission data using the ILO88 system," in *InterNoise 91, Proceedings Vol. 1*, edited by A. Lawrence (Noise Control Foundation, Poughkeepsie, New York), pp. 217–220.
- Lonsbury-Martin, B. L. (personal communication).
- Lonsbury-Martin, B. L., Cutler, W. M., and Martin, G. K. (1991). "Evidence for the influence of aging on distortion-product otoacoustic emissions," *J. Acoust. Soc. Am.* **89**, 1749–1759.
- McFadden, D., and Mishra, R. (1993). "On the relation between hearing sensitivity and otoacoustic emissions," *Hearing Res.* **71**, 208–213.
- McFadden, D., and Plattsmier, H. S. (1984). "Aspirin abolishes spontaneous otoacoustic emissions," *J. Acoust. Soc. Am.* **76**, 443–448.
- Moulin, A., Bera, J. C., and Collet, L. (1994). "Distortion product otoacoustic emissions and sensorineural hearing loss," *Audiology* **33**, 305–326.
- Norton, S. J. (1993). "Applications of transient-evoked otoacoustic emissions to paediatric populations," *Ear Hear.* **14**, 64–73.
- Osterhammel, D., and Osterhammel, P. (1979). "High-frequency audiometry: age and sex variations," *Scand. Audiol.* **8**, 73–81.
- Prieve, B. A., and Falter, S. R. (1995). "COAEs and SSOAEs in adults with increased age," *Ear Hear.* **16**, 521–528.
- Probst, R., Lonsbury-Martin, B. L., and Martin, G. K. (1991). "A review of otoacoustic emissions," *J. Acoust. Soc. Am.* **89**, 2027–2067.
- Smurzynski, J., and Kim, D. O. (1992). "Distortion-product and click-evoked otoacoustic emissions of normally-hearing adults," *Hearing Res.* **58**, 227–240.
- Stover, L., and Norton, S. J. (1993). "The effects of aging on otoacoustic emissions," *J. Acoust. Soc. Am.* **94**, 2670–2681.
- Whitehead, M. L., Stagner, B. B., Lonsbury-Martin, B. L., and Martin, G. K. (1995). "Effects of ear-canal standing waves on measurements of distortion-product otoacoustic emissions," *J. Acoust. Soc. Am.* **98**, 3200–3210.

Relations between notched-noise suppressed TEOAE and the psychoacoustical critical bandwidth

Joachim Neumann, Stefan Uppenkamp, and Birger Kollmeier^{a)}
AG Medizinische Physik, Universität Oldenburg, D-26111 Oldenburg, Germany

(Received 30 May 1996; revised 22 November 1996; accepted 23 November 1996)

Narrow-band transitory evoked otoacoustic emissions (TEOAE) were recorded for nine normal hearing subjects in the presence of a broadband tone complex suppressor. Introducing a spectral notch at the frequency of the narrow-band stimulus causes the suppression effect to decrease, the more so the wider the notch. This decrease in suppression permits an estimate of the size of one critical band. One advantage of this approach is that no active participation of the subjects is required. The estimated critical bandwidth is then compared with independent estimates based on a simultaneous masking experiment, using the same stimuli. The two measures of the critical bandwidth coincide well for those six subjects with spontaneous otoacoustic emissions. However, the bandwidth estimate based on the OAE measurements is too large for the other three subjects without spontaneous emissions. Simulations of the suppression effect with a driven van der Pol oscillator with moderate undamping produce critical bandwidth estimates consistent with those observed in the psychoacoustical experiments. This allows an estimate of the “effective” amount of undamping on the basilar membrane that is required to produce the critical bandwidth observable in psychoacoustic experiments. © 1997 Acoustical Society of America. [S0001-4966(97)01105-3]

PACS numbers: 43.64.Jb, 43.66.Dc [RDF]

INTRODUCTION

The recording of otoacoustic emissions allows one to obtain data on the peripheral hearing system without any active participation of the subject. Clinical interest in otoacoustic emissions is typically focused on the determination of hearing thresholds. The spectrum of transitory evoked otoacoustic emissions (TEOAE) or the distortion product otoacoustic emissions (DPOAE, “DPgram”) usually serve as an estimate of the audiogram. However, otoacoustic emissions are mostly measured with stimuli well above the threshold of hearing so that they might relate better to suprathreshold phenomena than to the audiogram. Therefore, otoacoustic emissions might help in determining functional parameters of the inner ear that relate to parameters derived from suprathreshold psychoacoustical tests.

One of the most important parameters of this kind is the critical bandwidth (CBW). It describes the width of the frequency band within which spectral energy of a masker is integrated (Fletcher, 1940; Greenwood, 1961; Zwicker and Feldtkeller, 1967). The size of one critical band also has great importance for experiments that study the interaction of tones within the auditory system. For example, the level of a perceived cubic difference tone decreases for a ratio of the primaries f_2 and f_1 larger than 1.2 (Goldstein, 1967; Hall, 1972; Smoorenburg, 1972; Weber and Mellert, 1975). Similarly, the level of combination tones measured as distortion product otoacoustic emissions (DPOAE) in the ear canal vary with the spectral distance of the primaries f_2 and f_1 (Harris *et al.*, 1989; Gaskill and Brown, 1990).

Brown *et al.* (1993) quantitatively compared this “characteristic of the DPOAE-filters” with the psychoacoustical

critical bandwidth expressed as equivalent rectangular bandwidth (ERB) for each of a set of subjects. The CBW was determined in a forward masking experiment using noise maskers with differing spectral notchwidth according to Patterson (1976). Brown *et al.* concluded that the DPOAE-tuning curve may serve as an estimate for the size of one critical band. One problem with these OAE experiments is that the stimuli and procedures to estimate the critical bandwidth differ considerably from those utilized in the psychoacoustical experiment. Since experimental paradigms as well as the assumed shape of the auditory filter significantly influence the estimates of the CBW (Kollmeier and Holube, 1992), a quantitative comparison between the CBW based on DPOAE and the CBW based on masking experiments seems to be difficult for the experiments described so far.

The experiments presented here therefore use the same stimuli for measurements of the suppression of narrow-band TEOAE and for psychoacoustical CBW measurements. The OAE experiments are based on the observation that TEOAE can be synchronized by additional tones (Kemp, 1979; Kemp and Chum, 1980; Wilson, 1980; Wit *et al.*, 1981; Zwicker and Schloth, 1984; Long *et al.*, 1988). In the case of TEOAE or synchronized spontaneous otoacoustic emissions (SOAE), the effect of an additional sinusoid decreases with increasing spectral distance to the suppressed emission component. The variation of this distance allows the recording of characteristic “tuning curves” based on the level of the suppressed TEOAE (Uppenkamp and Kollmeier, 1994). This tuning curve exhibits a bandwidth that approximates one critical band, with Q_3 varying between 3 and 8 for subjects with SOAE and Q_3 varying between 1 and 3 for subjects without SOAE. However, the relation between this effect and the critical bandwidth measured with psychoacoustical methods is not yet completely understood.

^{a)}Corresponding author.

In contrast to the experiments described by Uppenkamp and Kollmeier (1994), a broadband tone complex with a variable spectral notch is used in this study. This tone complex serves as suppressor in the TEOAE recordings and as masker in the notched-noise masking experiments. In both experiments, the width of the notch in the tone complex is varied. In addition, the same tone pulse is used in both experiments. In the TEOAE experiments the tone is used to evoke the otoacoustic emission, whereas in the psychoacoustical masking experiment it serves as the signal that the subject is requested to detect.

I. METHODS

A. Experimental setup for OAE measurements

Otoacoustic emissions are recorded in an IAC403-A sound-insulated chamber. The acoustic stimulation of the ear is carried out with an insert ear phone (Etymotic Research ER-2), which has a flat frequency response up to 10 kHz. The acoustic signal is recorded in the sealed ear canal with a miniature electret microphone (Knowles EA 1843). The microphone sensitivity, including a pre-amplifier with a gain of 46 dB, is 1.55 V per Pa at 1000 Hz. The output of the pre-amplifier is connected to a custom-designed amplifier with a gain of 20 dB. The signal is then passed through a butterworth high-pass filter with a cut-off frequency of 200 Hz to reduce low-frequency noise. The signal is digitized using a 16-bit A/D converter on a signal-processing board (Ariel corporation DSP-32C) and recorded in two separate memory buffers.

The digital signal processor is used to calculate the root-mean-square of the signal in real time. Noise reduction is carried out by an averaging technique that uses the inverse of the rms value of the response to the signal as a weighting factor. These segments have a duration of 46 ms, yielding a stimulus rate of 21.6 Hz. Segments with high rms values are rejected and segments with little noise receive a high weight. Furthermore, the cross-Fourier-transform of the two buffers is calculated concurrently. The real part of this cross-spectrum is summed for all frequencies to serve as an estimate of the level of the otoacoustic emission. The noise level is estimated by the rms of the difference of the two buffers. The time signals and TEOAE spectra are displayed on the host PC throughout the recording session.

B. Subjects

Nine normal hearing subjects, aged from 23 to 34 years, 5 male and 4 female, participated voluntarily in this study. They all exhibit normal hearing, as indicated by ear inspection and routine audiometry. Six subjects show spontaneous otoacoustic emissions (SOAE). For three subjects the level of the SOAE is more than 14 dB above the noise floor which exhibits a spectral power density of approximately $14 \mu\text{Pa}/\sqrt{\text{Hz}}$ (i.e., -3 dB SPL/Hz).

C. OAE experiments

The experimental procedure can be subdivided into four steps:

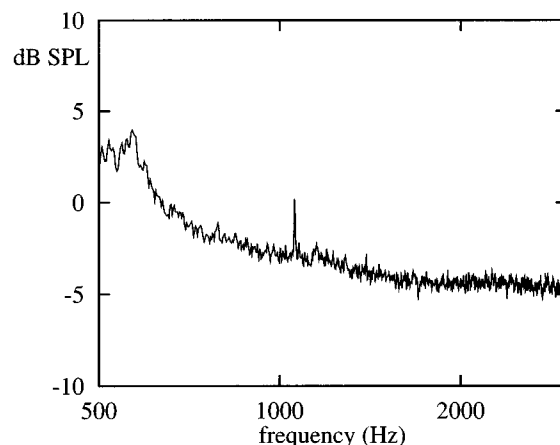


FIG. 1. Spontaneous otoacoustic emission (SOAE) from the left ear of normal hearing subject BG. The ordinate gives the spectral power density in dB SPL. The component at 1058 Hz is 3.2 dB above the noise floor.

- (1) Broadband stimulation of TEOAE to select a prominent spectral component.
- (2) Recording of narrow-band-evoked TEOAE at a low stimulus level at the frequency of a prominent spectral component.
- (3) Suppression of the narrow-band TEOAE with suppressors of variable notchwidths.
- (4) Evaluation of the CBW from changes in the suppression effect.

These four steps are described in detail in the following sections and illustrated by exemplary measurements for a normal hearing subject (BG). This subject has a spontaneous otoacoustic emission (SOAE) at 1058 Hz (cf. Fig. 1).

1. Broadband stimulation of TEOAE

In the first step, a broadband TEOAE is recorded in nonlinear averaging mode according to Bray and Kemp (1987) at a stimulus level of 40 dB SPL peak equivalent. The stimulation utilizes a short chirplet signal with spectral power in the range of 500 Hz to 6000 Hz (cf. Neumann *et al.*, 1994). In contrast to click stimuli, chirplet signals allow an optimal stimulation of any frequency range, narrow band as well as broadband. In addition, chirplet signals contain more energy than a click stimulus with the same maximum amplitude. Figure 2 shows the chirplet-evoked TEOAE

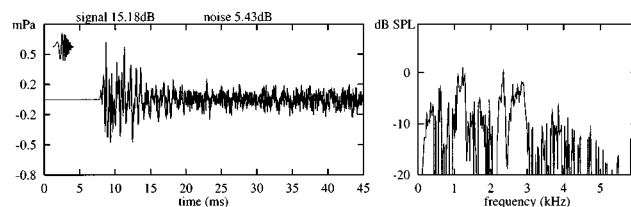


FIG. 2. Otoacoustic emissions from normal hearing subject BG. Left: time signal of chirplet-evoked TEOAE, right: spectrum of the TEOAE. The insert at the left panel shows the broadband chirplet stimulus employed. The TEOAE spectrum of this subject shows a typically peaked structure with major components between 500 Hz and 3000 Hz. The component at 1058 Hz is selected for narrow-band stimulation.

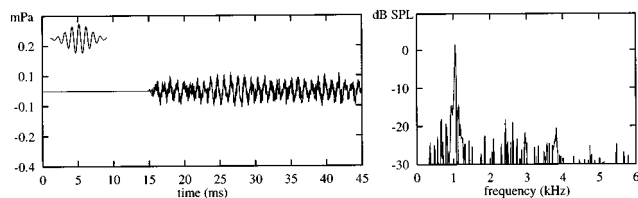


FIG. 3. Narrow-band-evoked TEOAE from subject BG at low stimulus level (same data representation as in Fig. 2). The insert at the left panel shows the tone pulse stimulus employed. The carrier frequency of the tone pulse is 1058 Hz. The SOAE clearly contributes to the transitory evoked otoacoustic emission in this case.

from subject BG and a sketch of the broadband chirplet stimulus as an insert at the left panel.

2. Narrow-band stimulation of TEOAE

In the second step, a prominent component is selected from the spectrum of the recorded broadband emission. In some cases this component is a synchronized spontaneous otoacoustic emission. For example, the SOAE at 1058 Hz of subject BG is visible as one major peak in the TEOAE spectrum. This component is selected for the subsequent narrow-band stimulation.

The stimulation with a tone pulse is always a compromise between the limited maximal duration of the stimulus and the concentration of the spectral power. An optimal tone pulse with a Gaussian envelope and a constant relative bandwidth of $\Delta f_{3\text{ dB}}/f_0 = 0.17$ is employed. These tone pulses are theoretically described by Strube (1989) and were utilized for OAE recordings by Uppenkamp and Kollmeier (1994). The time signal of such a tone pulse is given by

$$s(t) = e^{-0.005(\omega t)^2} \cos(\omega t). \quad (1)$$

The duration of this tone pulse varies with the center frequency. For example, a 1000 Hz tone pulse has an amplitude above 1% of the maximum for a duration of 9.7 ms. The recording of the TEOAE is performed in linear averaging mode. The stimulus level is successively reduced until the emission disappears in the background noise. The output level of the acoustic transducer is then set 10 dB above this level. As an example, Fig. 3 shows the narrow-band-evoked emission for subject BG where a stimulus frequency of 1058 Hz and a stimulus level of 18 dB SPL peak equivalent was used.

3. Suppression of the narrow-band TEOAE

In the third step, the TEOAE is suppressed with a broadband tone complex. This suppressor is designed to cancel out during the averaging procedure in order not to interfere with the recording procedure of the TEOAE. Therefore, a complex of continuous tones is generated at frequencies that are odd harmonics of half the stimulus rate $f_r/2 = 10.8$ Hz:

$$\text{tone complex}(t) = \sum_n \sin\left(2\pi(2n+1)\frac{f_r}{2}t + \varphi_n\right). \quad (2)$$

Each individual component of the tone complex shows a phase shift of π at the beginning of successive averaging

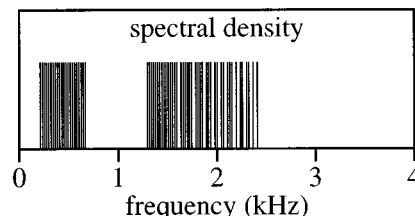


FIG. 4. Power spectrum of a tone complex with an incomplete harmonic series (10 components per critical band in two octaves) that is used as suppressor signal. A spectral notch of 300 Hz is placed at 1000 Hz. The omitted frequency components are added at the spectral boundaries. The harmonics of the resulting tone complex extend from 380 Hz to 2180 Hz.

frames. In order to obtain a signal with a noise-like waveform within each period, the starting phase φ_n is randomly chosen for all frequency components. If the averaging procedure is based on pairs of successive frames, the suppressor is canceled out in the resulting signal. The weighted averaging for noise reduction is also based on the rms value of pairs of successive frames. Thus, the suppressor has no influence on the weighted averaging and the noise reduction thus achieved. Due to the logarithmic place-frequency transformation in the cochlea the tone complex described so far would concentrate most energy in the basal part of the basilar membrane. In order to provide a uniform excitation by the suppressor, only an incomplete harmonic series is employed where the spacing of the harmonics is varied according to the bark scale (Zwicker and Terhardt, 1980). With this distribution of the harmonics, approximately the same power falls within each critical band. The tone complexes utilized in the experiments exhibit a spacing of 0.1 Bark (10 components per critical band). The complexes extend across a minimum spectral range of two octaves centered around the probe tone frequency. A spectral notch with variable bandwidth is placed at the frequency of the tone pulse. In order to keep the total power of the suppressor constant for different values of the notch, the spectral extent of the suppressor is varied. The same number of frequency components that is omitted in the region of the notch is added symmetrically both at the upper and lower spectral boundary of the original tone complex to keep the signal power constant. The width of the spectral notch was varied in the experiments from 0 Hz to 400 Hz in 10 steps of increasing size. An example of the power spectrum of such a notched tone complex is given in Fig. 4. To obtain a strong suppression effect, the level of the suppressor is set higher than the level of the tone pulse. Figure 5 shows the narrow-band TEOAE in the presence of a suppressor without notch for subject BG. As in Fig. 3, the stimulus frequency is 1058 Hz and the stimulus level is 18 dB SPL peak equivalent. The level of the suppressor was set to 44 dB SPL in this case.

4. Bandwidth determination from otoacoustic emissions (CBW_{OAE})

In the fourth and last step, the influence of the notch-width on the suppressed TEOAE is used to evaluate the critical bandwidth (CBW_{OAE}). The parameter observed is the energy of the TEOAE which is calculated within an octave

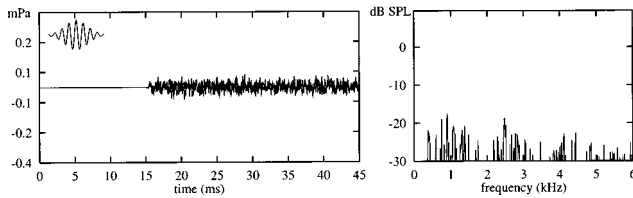


FIG. 5. Narrow-band-evoked TEOAE from subject BG in the presence of a suppressor tone complex with a spectral range of two octaves without spectral notch (same data representation as in Figs. 2 and 3). Note that the suppressor itself is not visible in the averaged time signal. In comparison with Fig. 3 the level of the emission is reduced.

band centered at the tone pulse frequency. The TEOAE level decreases in the presence of a suppressor without a spectral notch, but recovers with increasing notchwidth. The CBW_{OAE} is estimated from this data in a manner similar to that well-known from psychoacoustics. For this purpose the suppression effect is calculated as the difference between the level of the unsuppressed TEOAE and the level of the TEOAE in the respective suppressed condition. Figure 6 shows the decrease of the suppression effect with growing notchwidth for subject BG (rhombi). The filter describing the influence of the notched suppressor on the narrow-band-evoked TEOAE is assumed to be a symmetrical rounded exponential filter $roex(f, f_m, a)$ centered at the frequency f_m (cf. Glasberg and Moore, 1990). The prediction of the suppression effect is based on the assumption that the suppression effect SE is proportional to the energy of the suppressor in the auditory filter:

$$SE \sim \int_{-\infty}^{\infty} roex(f, f_m, a) \cdot S_{sup}(f) df. \quad (3)$$

Here, $S_{sup}(f)$ is the spectrum of the employed tone complex suppressor and the $roex$ -filter is defined as

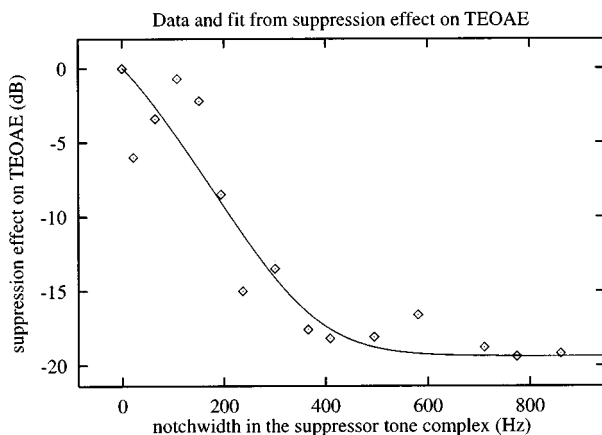


FIG. 6. Suppression effect of different tone complexes on the narrow-band TEOAE for subject BG. The suppression effect is expressed as the decrease of the OAE level due to the suppressor within an octave band centered around the stimulus frequency. Abscissa: width of the spectral notch in the tone complex. Ordinate: magnitude of the suppression effect in relation to the maximal observed suppression, normalized to 0 dB. The suppression effect decreases with increasing notchwidth. The solid line shows the prediction of the suppression effect based on a $roex$ -filter with parameter $a=62.5$. This value of a corresponds to a CBW_{OAE} of $4a=250$ Hz.

$$roex(f, f_m, a)$$

$$= \begin{cases} \frac{1}{2a} \left(1 - \frac{f-f_m}{a} \right) \exp\left(\frac{f-f_m}{a}\right), & \text{for } f < f_m, \\ \frac{1}{2a} \left(1 - \frac{f_m-f}{a} \right) \exp\left(\frac{f_m-f}{a}\right), & \text{for } f \geq f_m. \end{cases} \quad (4)$$

The $roex$ -filter can be described by a single parameter a . This parameter a can be determined by fitting the $roex$ -filter-based suppression prediction to the experimental data. For this purpose, a modified least-squares fit using a Lorentz error distribution is used which is more tolerant towards extremely deviating values than the standard least-squares fit. In the following, the CBW_{OAE} is characterized by the value of $4a$. This is the bandwidth of a rectangular filter with the same total power (ERB). For the concept of the equivalent rectangular bandwidth cf. Moore (1993) and Kollmeier and Holube (1992). The solid line in Fig. 6 shows the $roex$ -filter-based suppression prediction for subject BG. An optimal fit is achieved for $a=62.5$, which corresponds to a CBW_{OAE} of $4a=250$ Hz.

D. Psychoacoustical experiments

In order to quantitatively compare the individual CBW_{OAE} with the psychoacoustical critical bandwidth (CBW_{PSY}), simultaneous masking experiments were performed that resemble the “classical” notched-noise experiments (Patterson, 1976). The acoustic stimuli are the same as those used in the TEOAE experiments. The stimuli are transformed to analog signals by a 16-bit D/A converter at a sampling rate of 22050 Hz. They were low-pass-filtered, adjusted in level and monaurally presented to the subjects via a headphone (Sennheiser HDA200) in a soundproof booth. The timing, stimulus presentation and the recording of the responses was computer-controlled by a Sun workstation. The subject’s task is to detect the probe tone in one out of three intervals in each trial (3-IFC paradigm). Subject responses were given via a computer keyboard.

The same harmonic tone complex that was used as suppressor in the OAE experiments serves as masker, and the same tone pulse stimulus serves as probe tone in these experiments. The masker is set to a level of 30 dB above subjective threshold (this threshold was determined in a pilot experiment by three normal hearing subjects using the method of adjustment). As in the OAE experiments, the spectral notch is centered at the frequency of the tone pulse. In contrast to the OAE experiments, the tone pulse is placed 60 ms after the start of the masker to avoid an overshoot effect (Zwicker, 1965). The level of the probe tone is changed in an one-up-two-down paradigm. The initial step size of 8 dB was reduced by a factor of 2 after each upper reversal during the initial phase of the track, with a minimal step size of 1 dB. The average level for the last six reversals in each adaptive track was used as threshold estimate. The threshold estimation is made three times for each of ten different notchwidths.

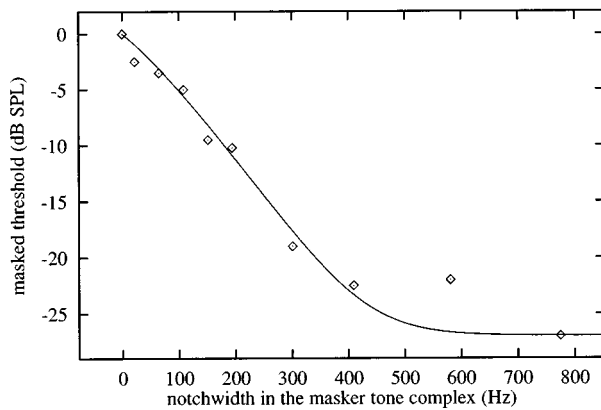


FIG. 7. Decrease in masked threshold of a 1058 Hz tone pulse presented 60 ms after the onset of a notched tone complex with varying notchwidth, subject BG, right ear. The masking effect is normalized to the condition with the maximum effect (=0 dB), i.e., suppressor without spectral notch. The masking effect decreases with increasing notchwidth. The solid line represents the roex-filter-based fit of the data which is optimal for $a=54.5$, corresponding to a CBW_{PSY} of $4a=218$ Hz.

E. Bandwidth determination from psychoacoustical experiments (CBW_{PSY})

Figure 7 shows the masking effect of the tone complex on the detection of the test tone for subject BG. The masked threshold plotted on the ordinate is normalized to the condition with the highest masking effect (=0 dB). As expected, the masked threshold strongly depends on the width of the notch in the suppressor. The determination of the CBW_{PSY} is based on this dependence. The data from Fig. 7 are used as input data to a roex-filter-based masking prediction. As described in the previous section, the prediction of the CBW_{PSY} is based on a Lorentz-fit to the experimental data. The parameter of the roex filter ($a=54.5$) corresponds to a CBW_{PSY} of $4a=218$ Hz in Fig. 7. The quality of the roex-filter-based fit is a useful value for determining how appropriate the model assumptions are in predicting the measured data. For this purpose, the nonlinear deviation measure B_{nl} was employed (Schach and Schäfer, 1978; Press *et al.*, 1992). It is defined for N measured data points y_i with mean \bar{y} and the respective predicted values \hat{y}_i :

$$B_{nl} = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}. \quad (5)$$

In the case of a perfect fit $B_{nl}=1$. B_{nl} is zero if the mean \bar{y} is used as prediction \hat{y}_i for all data points. If the prediction of the measured data is worse, B_{nl} can be negative. The prediction of the masking effects displayed in Fig. 7 exhibits a value of $B_{nl}=0.997$. Thus, the model assumptions based on the roex-filter seems to provide a good description of the data.

II. RESULTS

Figure 8 gives the results of both experiments for all subjects. In contrast to Figs. 6 and 7, the “raw” data are shown in this figure to allow an estimate of the interindividual differences. The three left panels show the levels of

the TEOAE in presence of the different tone complexes as a function of the width of the spectral notch in the suppressor tone complex. The three right panels show the corresponding psychoacoustic results, that is, the masked threshold of the test tone relative p.e. to the masker fixed at a sensation level of 30 dB. The subjects are divided into three groups: subjects without SOAE (upper panels), subjects with moderate SOAE (middle panels) and subjects with strong SOAE (one or more SOAE components more than 14 dB above the noise floor, lower panel).

In the three left panels, the suppression effect decreases with increasing notchwidth for all subjects. The effect is stronger for subjects with strong or moderate SOAE than for subjects without SOAE (upper left panel). Furthermore, the decrease of the suppression effect is not monotonic. Some subjects show a local minimum of the suppression effect below 200 Hz. The three panels on the right show the individual masking effects in the psychoacoustical experiment. As expected from the literature on notched-noise masking experiments, all subjects exhibit a decrease of the masked threshold with increasing notchwidth. For eight out of the nine subjects, the interquartile ranges are smaller than 3 dB in most conditions. One subject with strong SOAE had large interquartile ranges (subject TB, lower right panel). This subject reported difficulties in detecting the probe tone in a reliable way. The level of the SOAE at 1553 Hz is 13.6 dB SPL for this subject. This is in the range of the masked threshold, because the masker level employed in the psychoacoustical experiments is 30 dB SL. The difficulties in detecting the probe tone might thus originate from an interaction of the SOAE with the perceived probe tone.

As described in the previous section, estimates of the CBW_{OAE} and CBW_{PSY} are derived from the data displayed in Fig. 8. Table I lists these values for all subjects. The prediction of the suppression and masking effects displayed in Fig. 8 exhibits values of B_{nl} in the range between 0.96 and 1. Thus, the theoretical curves based on a roex-filter shape yield an accurate description of both sets of data.

Table I also lists the relative critical bandwidth (CBW_{PSY}/f or CBW_{OAE}/f). These values are plotted as a scatter diagram in Fig. 9. For all subjects the CBW_{PSY}/f corresponds well with the value of 0.2 that is reported in the literature (cf. Glasberg and Moore, 1990). However, the CBW_{OAE}/f varies substantially across subjects: While six subjects with SOAE (BG, AP, HG, HH, AS, and TB) also show a CBW_{OAE}/f value close to 0.2, the three subjects without SOAE (SU, TW, and ML) exhibit a substantially larger CBW_{OAE}/f of 0.4 to 0.5.

III. SIMULATIONS WITH A DRIVEN VAN DER POL OSCILLATOR

The generation of otoacoustic emissions can be modeled with simulations of basilar membrane mechanics including active mechanisms (Davis, 1983; Koshigoe and Tubis, 1983; Duifhuis *et al.*, 1986; Zwicker, 1986; Talmadge *et al.*, 1990; van den Raadt and Duifhuis, 1990; Neely and Stover, 1993; Kanis and de Boer, 1993). These models, however, have many free parameters and predict the detailed generation of OAE by a variety of different mechanisms. Under the sim-

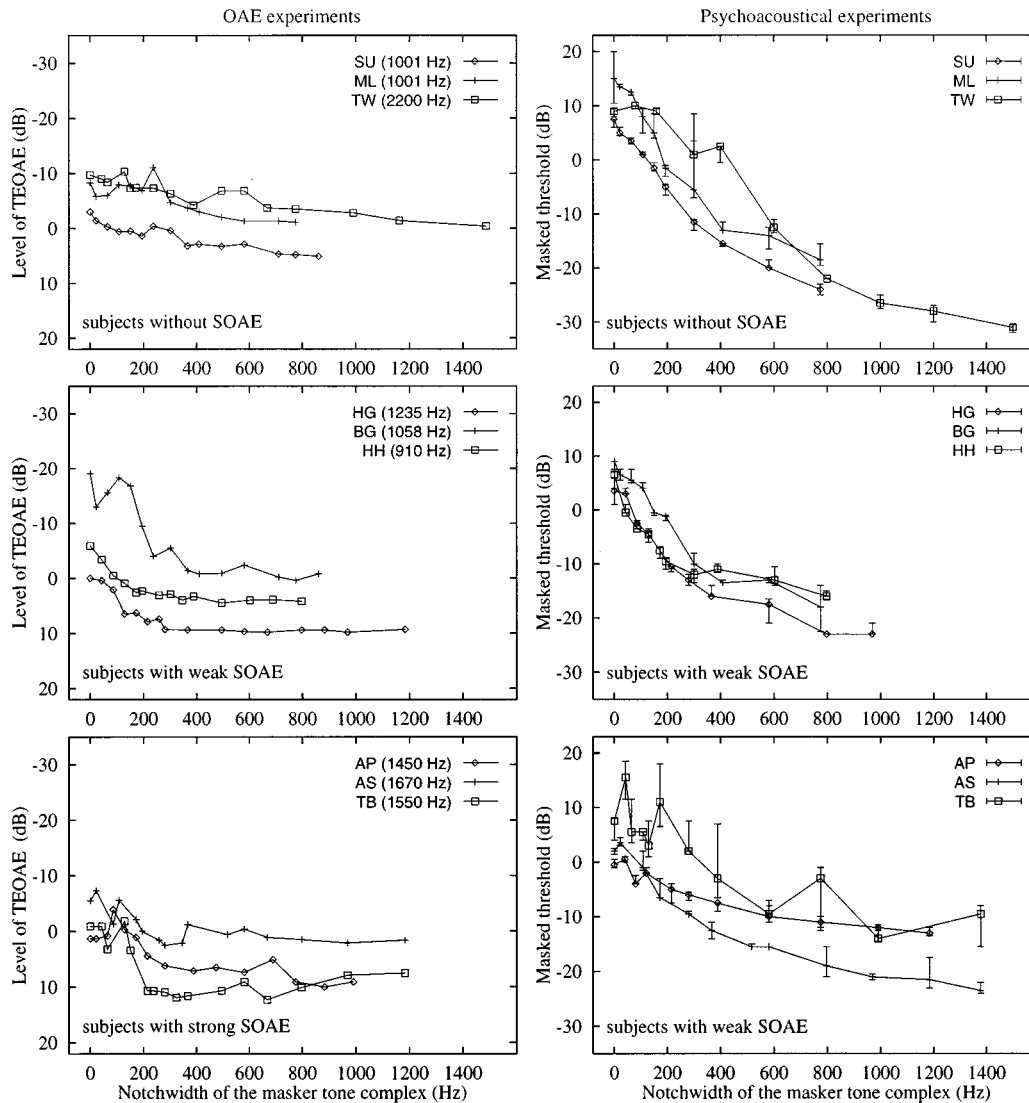


FIG. 8. Results from both experiments for three sets of subjects: subjects without SOAE (upper panel), subjects with weak SOAE (middle panel) and subjects with strong SOAE (more than 14 dB above the noise floor, bottom panel). Left: sound pressure level of the narrow-band TEOAE in dependence on the width of the notch in the suppressor. Please note the reversed ordinate in the left column, that is, higher levels are plotted pointing downwards. Right: results of notched-noise masking experiments. Threshold is given here relative p.e. to the masker set to a sensation level of 30 dB.

TABLE I. Values of the critical bandwidth in Hz determined from OAE-measurements (CBW_{OAE}) and from psychoacoustic experiments (CBW_{PSY}) for all subjects. For comparison, the relative bandwidth is also given (CBW_{PSY}/f and CBW_{OAE}/f). The goodness of fit is expressed by the nonlinear correlation coefficient B_{nl} that ranges between -1 and 1 (for perfect fit). The fifth and sixth column give the sound pressure levels of the stimuli applied in the OAE experiment L_{Puls} : level of evoking tone pulse, L_{Sup} : level of suppressor tone complex). The seventh column (ΔL_{OAE}) gives the maximum suppression effect in dB. The masker in the psychoacoustical task was set to a sensation level of 30 dB for all subjects.

Subject	Side	SOAE (dB)	Frequency	L_{Puls} (dB)	L_{Sup} (dB)	ΔL_{OAE}	CBW_{OAE}	CBW_{OAE}/f	B_{nl}	CBW_{PSY}	CBW_{PSY}/f	B_{nl}
ML	right	none	1000	26	47	9.9	408	0.41	0.963	192	0.19	1.000
SU	right	none	1000	30	51	8.1	502	0.50	0.961	219	0.22	0.998
TW	left	none	2200	20	45	9.9	938	0.43	0.975	368	0.17	0.999
BG	left	3.2	1058	18	44	18.7	250	0.24	0.993	218	0.21	0.997
HH	left	5.1	910	25	47	10.4	160	0.18	0.997	150	0.17	0.984
HG	right	11.1	1235	30	52	9.8	232	0.19	0.994	212	0.17	0.999
AS	left	14.1	1666	30	51	9.8	234	0.14	0.966	292	0.18	0.994
AP	left	15.1	1459	25	48	13.9	334	0.23	0.971	298	0.20	0.998
TB	left	18.0	1553	28	50	14.1	182	0.12	0.989	334	0.22	0.992

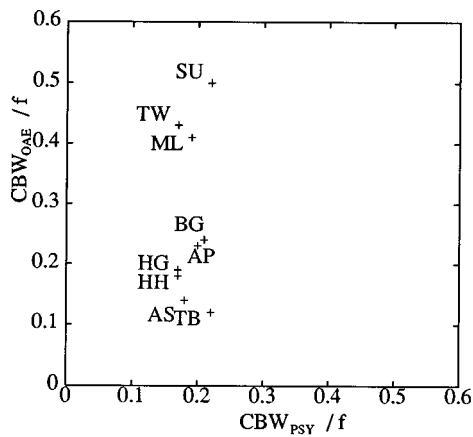


FIG. 9. Comparison of CBW_{OAE} and CBW_{PSY} . Subjects with SOAE (BG, HG, TB, AP, HH, AS) exhibit similar values of CBW_{PSY}/f and CBW_{OAE}/f , whereas subjects without SOAE (SU, ML, TW) exhibit a CBW_{OAE}/f that is twice as high as the CBW_{PSY}/f .

plifying assumption that the generation of otoacoustic emissions is a local oscillation process on the basilar membrane, including some “negative damping,” a single van der Pol oscillator may be used to model the main physical principle of OAE generation.

A. The van der Pol oscillator as a model for OAE

The van der Pol oscillator equation is the simplest example of a nonlinear self-sustained oscillator. If $x(t)$ denotes the time-dependent elongation of the oscillator which is driven by an external force $E(t)$, the van der Pol equation can be written as,

$$\ddot{x} + (-d_1 + d_2 x^2)\dot{x} + \omega_0^2 x = E(t), \quad d_1, d_2 \geq 0. \quad (6)$$

In Eq. (6), the parameter d_1 denotes a constant undamping term and represents the “active” properties of the oscillator. The parameter d_2 determines the nonlinear damping which becomes dominant for large elongations. The parameter ω_0 is the characteristic circular frequency of the oscillator without damping. Depending on the choice of the parameters d_1 , d_2 and the force $E(t)$, the oscillator may produce a stationary sinusoidal oscillation, or even behave like a chaotic strange attractor (Parlitz and Lauterborn, 1987).

The single van der Pol oscillator has been shown to be a suitable model for some properties of spontaneous otoacoustic emissions, including suppression tuning curves and entrainment to external tones (e.g., van Dijk and Wit, 1990a; Long *et al.*, 1988, 1991), and several time constants determining the relaxation dynamics (Talmadge *et al.*, 1990; Murphy *et al.*, 1995). A detailed analysis of amplitude and frequency fluctuations of spontaneous emissions illustrates that a linear stiffness oscillator, as given in Eq. (6), cannot account completely for the experimental findings (van Dijk and Wit, 1990b). Nevertheless, it has been shown (Talmadge and Tubis, 1993) that a cochlear model with distributed damping of the van der Pol type can account for even more complex properties of evoked and spontaneous otoacoustic emissions, such as the approximate 0.4 Bark frequency periodicity.

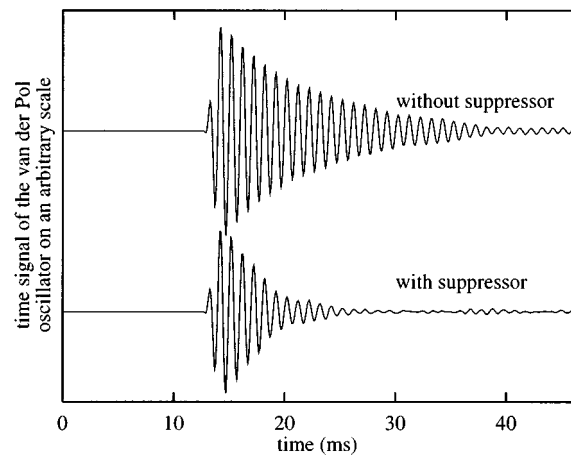


FIG. 10. Averaged output of a van der Pol oscillator with $\omega_0/2\pi=1000$ Hz, evoked by a 1000 Hz tone pulse. The value of the damping term d_2 was set to 10 000 and the undamping term d_1 was set to 100. The external force includes a tone pulse with an amplitude of $6.33\omega_0^2$ and a noise term with an amplitude of $0.32\omega_0^2$. Upper curve: simulated “emission” on an arbitrary scale. Lower curve: simulated suppressed “emission.” The time signal is set to zero during the evoking tone pulse.

As shown before (Uppenkamp and Kollmeier, 1994), the single van der Pol oscillator can be utilized to model some experimental findings in the interaction of narrow-band transitory evoked otoacoustic emissions with additional continuous tones as well. In those simulations, the external force $E(t)$ consisted of the evoking stimulus tone pulse and one continuous sinusoid that served as suppressor and canceled out in successive averaging frames. The power of the simulated emission showed a decline if the frequency of the additional tone was near the circular frequency of the oscillator ω_0 . Hence, the synchronization of the emission with the original stimulus is reduced and the “response” of the system to the original stimulus is attenuated.

In analogy to the experiments in Sec. II, simulations of narrow-band-evoked otoacoustic emissions in presence of tone complexes have now been performed using the simple model of a single driven van der Pol oscillator.

B. Numerical results

Since the single van der Pol oscillator does not include the function of the middle ear and the wave propagation along the cochlear partition, the time function $x(t)$ of the driven oscillator has to be interpreted in terms of movement of the basilar membrane at a certain place, characterized by its best frequency. This signal is segmented into sections of 46 ms (the stimulus repetition rate). The signal following each tone pulse stimulus is interpreted as evoked otoacoustic emission. Thus, the time delay between the generation of the OAE and the signal at the recording microphone is neglected. The temporal development of the system was computed using a numerical integration procedure (fourth order Runge-Kutta, cf. Press *et al.*, 1992). Figure 10 gives an example of a simulated narrow-band TEOAE with and without the suppressor tone complex. During the temporal extent of the stimulus the time signal is set to zero. The simulated otoacoustic emission is calculated for 17 different values of

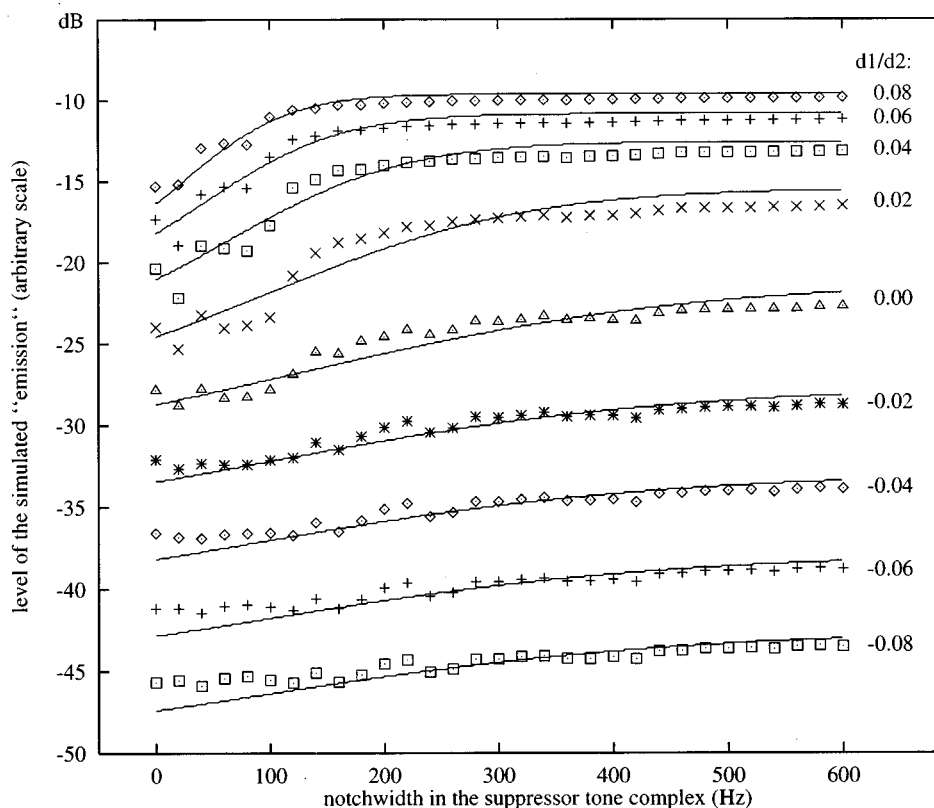


FIG. 11. Level of the simulated “emissions” in the presence of suppressor tone complexes with varying notchwidth (symbols) and the appropriate roex-filter-based fit (solid lines). The parameter d_1 varies from a ratio of $d_1/d_2=0.08$ to $d_1/d_2=-0.08$. For positive undamping (upper curves) a local minimum of the suppression effect can be observed below 100 Hz. The dynamic range of the suppression effect as well as the level of the “emission” is reduced in the case of undamping (lower curves).

the undamping parameter d_1 ranging from -800 to 800 and for 31 different notchwidths of the suppressor tone complex. As shown in Fig. 11, the reduction of the level of the simulated emission depends on the width of the spectral notch in the suppressor tone complex and on the value of d_1 . The dynamic range of the suppression effect is least for large negative values of the undamping parameter d_1 . For positive values of d_1 , a minimum of the “emission level” can be observed for notchwidths in the suppressor tone complex ranging between 50 Hz and 150 Hz. This might correspond to the local minima found in the OAE-data for subjects with a strong SOAE (cf. lower left panel of Fig. 8).

Similar to the method described in Sec. II, the level of the suppressed emissions can serve as input for a roex-filter-based prediction. The estimates of the simulated critical bandwidth (CBW_{SIM}) is based on the predictions shown as solid lines in Fig. 11. Figure 12 gives the resulting values of CBW_{SIM} as a function of d_1/d_2 , i.e., the ratio of the undamping parameter d_1 and the nonlinear damping parameter d_2 . Apparently, CBW_{SIM} decreases for positive values of d_1 (undamping), whereas the CBW_{SIM} is larger for negative values (damping). In this aspect, the van der Pol oscillator behaves as expected from a linear resonator. The local maxima of the CBW_{SIM} near a value of $d_1 = 0$ is due to the local minimum of the emission level for small notchwidth (see above). This causes a reduced slope of the roex-filter-based fit (solid lines

in Fig. 11) for values of d_1/d_2 between 0 and 0.04 and thus results in a larger value of CBW_{SIM} .

IV. DISCUSSION

The major results of this study can be summarized as follows:

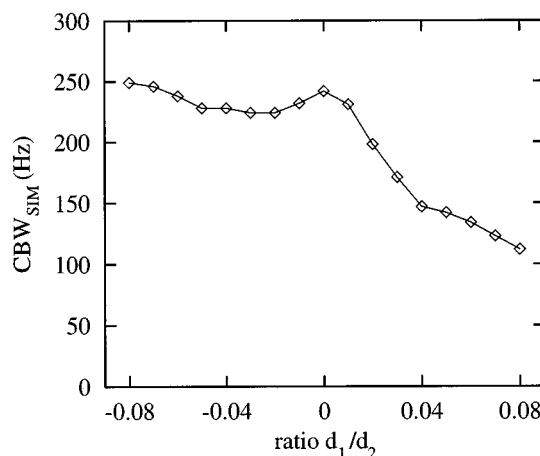


FIG. 12. Values of CBW_{SIM} for different ratios d_1/d_2 . The grade of undamping is changed while the parameter d_2 is kept at a constant value of 10 000.

- (1) The level of narrow-band TEOAE is reduced in the presence of a suppressor tone complex.
- (2) The decline of this suppression effect with increasing notchwidth in the suppressor allows one to estimate the width of one critical band. A similar bandwidth can be obtained from simulations of the suppression effect using a single driven van der Pol oscillator.
- (3) The values of CBW_{PSY} and CBW_{OAE} differ significantly for subjects without SOAE.

With respect to the first point it should be noted that the reduction of the TEOAE level in the presence of a suppressor can be explained as a synchronization effect (Neumann *et al.*, 1997). The suppressor tone complex exhibits a phase difference of π in successive segments and the time segments are averaged in pairs of two. Thus, the suppressor and the synchronized portion of the otoacoustic emission completely cancels in the averaged signal. The strength of this synchronization effect strongly depends on the spectral distance between the suppressed components of the emission and the components of the suppressor. In most cases, the maximum suppression effect is achieved for a tone complex without spectral notch (left panels in Fig. 8). In some cases, however, an additional local minimum of the suppression effect occurs for a notchwidth between 40 Hz and 300 Hz (e.g., subjects ML, AS, TB, and BG). Subjects with SOAE exhibit a strong suppression effect of 14 to 23 dB. The suppression effect levels off for notchwidths greater than about 300 Hz. Subjects without spontaneous otoacoustic emissions, on the other hand, exhibit only a small reduction of the emission level of 8 to 10 dB. Although the upper bandwidth limit for the suppression effect is less pronounced for these subjects, the general dependence of the suppression effect on the notchwidth of the suppressor is comparable.

With respect to the second point it should be noted that the decline of the suppression effect with increasing notchwidth shows the same general shape as the decline of the masked threshold in the psychoacoustical experiments. Both experiments use the same acoustical stimuli and depend on the interaction of energy in a localized region on the basilar membrane. Both experiments are compatible with the concept of auditory filters. It can be assumed that the spectral range within which an additional tone can suppress an otoacoustic emission is related to the range within which masking energy is integrated across frequency. The lower cluster in Fig. 9 shows that values of CBW_{OAE} and CBW_{PSY} coincide well for subjects with SOAE. Both estimates of the critical bandwidth also agree with the value of $CBW_{PSY}/f \approx 0.2$ given in the literature (Glasberg and Moore, 1990; Zwicker, 1982). However, within this cluster, the variations of CBW_{OAE} and CBW_{PSY} seem to be independent. In addition, the lack of a coincidence for subjects without SOAE also shows that no strict relation can be found between CBW_{OAE} and CBW_{PSY} (see below).

The simulation of the suppression effect with a driven van der Pol oscillator can repeat most findings of the OAE experiments. With an appropriate choice of the amount of undamping, the results of the suppression experiments can be simulated for subjects with and without SOAE. This finding

is not surprising in view of the previous work by Long *et al.* (1988, 1991), and Uppenkamp and Kollmeier (1994), who showed that the van der Pol equation with an appropriate undamping yields a frequency-dependent suppression effect which resembles the well-known critical bandwidth effect. However, these authors did not derive critical bandwidth estimates in the same way as performed here (that is motivated by psychoacoustical bandwidth estimation procedures). The suppression effect simulated here has the same order of magnitude as that in the OAE experiments and depends on the size of the spectral notch in the suppressor tone complex (cf. Fig. 11). This figure also shows that the total range of the suppression effect is large for positive values of the undamping parameter d_1 , and limited for large negative values of d_1 . This corresponds with the observation that the suppression effect is greater for subjects with SOAE whereas subjects without SOAE exhibit shallow slopes (left panels of Fig. 8). A local minimum similar to the minimum of the suppression effect in subjects without SOAE can be observed in the simulations for positive undamping (cf. upper traces of Fig. 11). The size of the critical bands in the simulations (cf. Fig. 12) is in the same range as found in the actual OAE experiments. The rate of undamping is the most important parameter for the value of CBW_{SIM} . The bandwidth estimate is large for positive damping ($d_1 < 0$) and decreases in the case of positive undamping ($d_1 > 0$). The value that corresponds to the critical bandwidth of a normal hearing subject ($CBW_{SIM}/f \approx 0.2$) can be observed for $d_1/d_2 \approx 0.02$. Thus, the results from the simulations of a cochlear amplifier with moderate undamping (i.e., amplification gain just above one) are in good agreement with the data from normal hearing subjects. It may even be argued that this agreement at small positive values provides an estimate of the “effective” mechanical undamping at the basilar membrane level that is required for a normal function of the auditory system.

With respect to the third point it should be noted that the CBW_{OAE} is larger for subjects without SOAE than for subjects with SOAE (cf. Fig. 9). This result is visible in the shallow slopes of the suppression effect in the upper left panel of Fig. 8. The CBW_{OAE} does not coincide with the CBW_{PSY} , since the CBW_{PSY} is approximately the same for all subjects. For those subjects that do not exhibit a SOAE close to the test frequency, the obtained CBW_{OAE} value overestimates the actual CBW_{PSY} value. There are several possible explanations of the divergence of CBW_{OAE} and CBW_{PSY} values. In the first place, the level of narrow-band TEOAE is comparably low in the absence of SOAE. As a consequence, the maximal achievable suppression effect is limited and the spread of the data might be too large to derive a valid estimate of the critical bandwidth. Nevertheless, this can not explain the observed systematic divergence between CBW_{OAE} and CBW_{PSY} .

The observation of greater interindividual variability in the OAE data (left panels of Fig. 8) that cannot be found in the masked threshold data (right panels of Fig. 8) supports the conjecture that both methods test a different subset of the properties of the auditory system. The psychoacoustical detection task might involve effects like “off-frequency listen-

ing” or central processes that do not primarily reflect cochlear mechanisms and cannot be tested with otoacoustic emissions. These effects might cause a psychoacoustical critical bandwidth that is smaller than expected from OAE experiments. On the other hand, the recording of otoacoustic emissions involves properties of the auditory system that do not directly contribute to sound perception. For example, the propagation of the emission from the place of its origin to the apex, in the middle ear, and into the recording system might suppress and substantially alter the signal. Thus, the OAE as they are generated in the inner ear might not be entirely represented by the signal recorded in the ear canal. These processes might also influence the apparent arbitrariness of the occurrence of SOAE. Apart from involving a different subset of the properties of the auditory system, the frequency range that contributes to either CBW_{OAE} or CBW_{PSY} might differ in principle. The psychoacoustical task is based on the global excitation pattern on the cochlea whereas the analysis of the suppression effect evaluates the level of a single frequency component and does not account for level changes at other frequencies. Possibly, this difference is less pronounced in the presence of a spontaneous otoacoustic emission. It is known that a “leading” SOAE oscillation is able to synchronize nearby oscillators (van Hengel and Maat, 1993). As a consequence, the SOAE oscillation might concentrate most OAE energy at a single frequency whereas multiple oscillators are involved in the generation of TEOAE. This might explain why the values of CBW_{OAE} and CBW_{PSY} are in agreement for subjects with SOAE only.

V. CONCLUSION

The present OAE experiments as well as the masked threshold experiments depend on the interaction of energy in a localized region on the basilar membrane. The prediction of the size of a critical band from an OAE experiment succeeds for those subjects with spontaneous otoacoustic emissions (six out of the nine tested). For the remaining subjects without spontaneous emissions, the critical bandwidths from the OAE experiment were larger than in the psychoacoustical experiment. The OAE experiment could be modeled with a single driven van der Pol oscillator that produced critical bandwidth estimates consistent with those observed in the psychoacoustical experiment if a moderate undamping was chosen. Therefore, the “effective” amount of undamping at the basilar membrane level can be estimated which is required to provide the critical bandwidth observed in psychoacoustic experiments.

ACKNOWLEDGMENTS

This study was supported by Deutsche Forschungsgemeinschaft (Ko 942/11-1). The helpful comments and criticism by the reviewers, B. Scharf and A. Tubis, are gratefully acknowledged.

Bray, P., and Kemp, D. T. (1987). “An advanced cochlear echo technique suitable for infant screening,” *Br. J. Audiol.* **11**, 191–204.

- Brown, A. M., Gaskell, S. A., Carylton, R. P., and Williams, D. M. (1993). “Acoustic distortion as a measure of frequency selectivity: relation to psychophysical equivalent rectangular bandwidth,” *J. Acoust. Soc. Am.* **93**, 3291–3297.
- Davis, H. (1983). “An active process in cochlear mechanics,” *Hearing Res.* **9**, 79–91.
- van Dijk, P., and Wit, H. P. (1990a). “Synchronization of spontaneous otoacoustic emissions to a $2f_1 - f_2$ distortion product,” *J. Acoust. Soc. Am.* **88**, 850–856.
- van Dijk, P., and Wit, H. P. (1990b). “Amplitude and frequency fluctuations of spontaneous otoacoustic emissions,” *J. Acoust. Soc. Am.* **88**, 1779–1793.
- Duifhuis, H., Hoogstraten, H. W., van Netten, H. W., Diependaal, R. J., and Bialek, W. (1986). “Modelling the cochlear partition with coupled van der Pol-Oscillators,” in *Peripheral Auditory Mechanisms*, edited by J. B. Allen, J. L. Hall, A. E. Hubbard, S. T. Neely, and A. Tubis (Springer-Verlag, Berlin), Lecture Notes in Biomathematics **64**, pp. 290–297.
- Fletcher, H. (1940). “Auditory patterns,” *Rev. Mod. Phys.* **12**, 47–65.
- Gaskell, S. A., and Brown, A. M. (1990). “The behavior of the acoustic distortion product, $2f_1 - f_2$, from the human ear and its relation to auditory sensitivity,” *J. Acoust. Soc. Am.* **88**, 821–839.
- 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. (1967). “Auditory nonlinearity,” *J. Acoust. Soc. Am.* **41**, 676–689.
- Greenwood, D. D. (1961). “Critical bandwidth and the frequency coordinates of the basilar membrane,” *J. Acoust. Soc. Am.* **33**, 1344–1356.
- Hall, J. L. (1972). “Auditory distortion products $f_2 - f_1$ and $2f_1 - f_2$,” *J. Acoust. Soc. Am.* **51**, 1863–1871.
- Harris, F. P., Lonsbury-Martin, B. L., Stagner, B. B., Coats, A. C., and Martin, G. K. (1989). “Acoustic distortion products in humans: systematic changes in amplitude as a function of f_2/f_1 ratio,” *J. Acoust. Soc. Am.* **85**, 220–229.
- van Hengel, P., and Maat, A. (1993). “Periodicity in frequency spectra of click evoked and spontaneous OAE, theory meets experiment,” in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 47–53.
- Kanis, L. J., and de Boer, E. (1993). “The emperor’s new clothes: DP emissions in a locally active nonlinear model of the cochlea,” in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 304–314.
- Kemp, D. T. (1979). “Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea,” *Arch. Otorhinol.* **224**, 37–45.
- Kemp, D. T., and Chum, R. (1980). “Properties of the generator of stimulated acoustic emissions,” *Hearing Res.* **2**, 213–232.
- Kollmeier, B., and Holube, I. (1992). “Auditory filter bandwidths in binaural and monaural listening conditions,” *J. Acoust. Soc. Am.* **92**, 1889–1901.
- Koshigoe, S., and Tubis, A. (1983). “A nonlinear feedback model for outer hair-cell stereocilia and its implications for the response of the auditory periphery,” in *Mechanics of Hearing*, edited by E. de Boer and M. A. Viergever (Delft U.P., Delft), pp. 153–160.
- Long, G. R., Tubis, A., and Jones, K. L. (1991). “Modeling synchronization and suppression of spontaneous otoacoustic emissions using Van der Pol oscillators: Effects of aspirin administration,” *J. Acoust. Soc. Am.* **89**, 1201–1212.
- Long, G. R., Tubis, A., Jones, K. L., and Sivaramakrishnan, S. (1988). “Modification of the external-tone synchronization and statistical properties of spontaneous otoacoustic emissions by aspirin consumption,” in *Basic Issues in Hearing*, edited by H. Duifhuis, J. W. Horst, and H. P. Wit (Academic, London), pp. 93–100.
- Moore, B. C. J. (1993). “Frequency analysis and pitch perception,” in *Human Psychophysics*, edited by W. A. Yost, A. N. Popper, and R. R. Fay (Springer-Verlag, Berlin), pp. 56–115.
- Murphy, W. J., Talmadge, C. R., Tubis, A., and Long, G. R. (1995). “Relaxation dynamics of spontaneous otoacoustic emissions perturbed by external tones: I. Response to pulsed single-tone suppressors,” *J. Acoust. Soc. Am.* **97**, 3702–3710.
- Neely, S. T., and Stover, L. J. (1993). “Otoacoustic emissions from a nonlinear, active model of cochlear mechanics,” in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 64–71.

- Neumann, J., Uppenkamp, S., and Kollmeier, B. (1994). "Chirp evoked otoacoustic emissions," *Hearing Res.* **79**, 17–25.
- Neumann, J., Uppenkamp, S., and Kollmeier, B. (1997). "Interaction of otoacoustic emissions with additional tones: Suppression or synchronization?," *Hearing Res.* **103**, 19–27.
- Parlitz, U., and Lauterborn, W. (1987). "Period-doubling cascades and devil's staircases of the Driven van der Pol oscillator," *Phys. Rev. A* **36**, 1428–1434.
- Patterson, R. D. (1987). "Auditory filter shapes derived with noise stimulation," *J. Acoust. Soc. Am.* **59**, 640–654.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical Recipes in C* (Cambridge U. P., Cambridge).
- van den Raadt, M. P. M. G., and Duifhuis, H. (1990). "A generalized Van der Pol-oscillator cochlea model," in *The Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer-Verlag, Berlin), Lecture Notes in Biomathematics **87**, pp. 227–234.
- Schach, S., and Schäfer, T. (1978). *Regressions und Varianzanalyse: Eine Einführung* (Springer-Verlag, Berlin).
- Smootenburg, G. F. (1972). "Audibility region of combination tones," *J. Acoust. Soc. Am.* **52**, 603–614.
- Strube, H. W. (1989). "Evoked otoacoustic emissions as cochlear Bragg reflections," *Hearing Res.* **38**, 35–45.
- Talmadge, C. L., and Tubis, A. (1993). "On modeling the connection between spontaneous and evoked otoacoustic emissions," in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 25–32.
- Talmadge, C. L., Long, G. R., Murphy, W. J., and Tubis, A. (1990). "Quantitative evaluation of limit-cycle oscillator models of spontaneous otoacoustic emissions," in *The Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer-Verlag, Berlin), Lecture Notes in Biomathematics **87**, pp. 235–242.
- Uppenkamp, S., and Kollmeier, B. (1994). "Narrow band stimulation and synchronization of otoacoustic emissions," *Hearing Res.* **78**, 210–220.
- Weber, R., and Mellert, V. (1975). "On the nonmonotonic behavior of cubic distortion products in the human ear," *J. Acoust. Soc. Am.* **57**, 207–214.
- Wilson, J. P. (1980). "Evidence for a cochlear origin for acoustic reemissions threshold fine-structure and tonal tinnitus," *Hearing Res.* **2**, 233–252.
- Wit, H. P., Langevoort, J. C., and Ritsma, R. J. (1981). "Frequency spectra of cochlear acoustic emissions ('Kemp-Echoes')," *J. Acoust. Soc. Am.* **70**, 437–445.
- Zwicker, E. (1965). "Temporal effects in simultaneous masking by white-noise bursts," *J. Acoust. Soc. Am.* **37**, 653–663.
- Zwicker, E. (1982). *Psychoakustik* (Springer-Verlag, Berlin).
- Zwicker, E. (1986). "'Otoacoustic emissions' in a nonlinear cochlear hardware model with feedback," *J. Acoust. Soc. Am.* **80**, 154–162.
- Zwicker, E., and Feldtkeller, R. (1967). *Das Ohr als Nachrichtenempfänger* (Hirzel, Stuttgart).
- Zwicker, E., and Schloth, E. (1984). "Interrelation of different otoacoustic emissions," *J. Acoust. Soc. Am.* **75**, 1148–1154.
- Zwicker, E., and Terhardt, E. (1980). "Analytical expression for critical-band rate and critical bandwidth as a function of frequency," *J. Acoust. Soc. Am.* **68**, 1523–1525.

A psychoacoustic model for the noise masking of plosive bursts^{a)}

James J. Hant, Brian P. Strobe, and Abeer A. Alwan

Speech Processing and Auditory Perception Laboratory, Department of Electrical Engineering, School of Engineering and Applied Sciences, UCLA, 405 Hilgard Avenue, Los Angeles, California 90095

(Received 26 March 1996; revised 18 November 1996; accepted 19 November 1996)

A model for predicting the masked thresholds of the voiceless plosive bursts /k,t,p/ in background noise is proposed. Because plosive bursts are brief, are generated by a noise source, and have different spectral characteristics, the modeling approach accounts for duration, center frequency, and signal bandwidth. Noise-in-noise masking experiments are conducted using a broadband masker and bandpass noise signals of varying bandwidth (100–5483 Hz), duration (10–300 ms), and center frequency (0.4–4 kHz). Data from these experiments are used to parametrize an auditory filter model in which the effective bandwidth and the signal-to-noise ratio at threshold for each filter are duration dependent. The duration-dependent filter model is then used to predict the thresholds of synthetic and naturally spoken plosive bursts in background noise. Finally, results from pilot notched-noise experiments are presented which support duration-dependent frequency selectivity.

© 1997 Acoustical Society of America. [S0001-4966(97)02204-2]

PACS numbers: 43.66.Ba, 43.71.An, 43.66.Dc [WJ]

INTRODUCTION

Plosives are often confused in noisy environments but there is no quantitative model to predict these perceptual confusions. In this paper, a psychoacoustic model to predict the noise masking of plosive bursts is developed. The model is a step toward a more general quantitative model of human speech perception in noisy environments.

Plosive bursts are brief signals that are generated by a noise source. To understand how they are masked by background noise, the durational effects on the masking of noise signals need to be studied. In addition, because the bursts have different spectral shapes, the effects of signal frequency and bandwidth should also be assessed. In this study, a series of psychoacoustic experiments using bandpass-noise signals of varying duration, bandwidth, and center frequency are conducted. Data from these experiments are used to derive a psychoacoustic model for noise-in-noise masking. The model is then used to predict the noise-masked thresholds of natural and synthetic plosive bursts. Finally, notched-noise experiments at short durations are conducted to assess implications of the model.

I. BACKGROUND AND MOTIVATION

A. Background and motivation for the psychoacoustic experiments

Proposed mechanisms that explain the masking of short-duration tones and bandpass noises in background noise include: reduced frequency selectivity, temporal integration, and other mechanisms specifically sensitive to signal tran-

sients. The relative importance of each of these mechanisms and the interplay between them is not completely understood.

Scharf (1970) suggested changing frequency selectivity with duration and attributed it to wider critical bandwidths at short durations. Scharf cites Scholl's (1962) study on noise-masked thresholds of bandpass noise signals. At long durations (500 ms) intensity thresholds increased when the signal bandwidth became larger than a critical band but at very short durations (3 ms), thresholds remained constant as bandwidth increased, suggesting widened auditory filters.

Several other investigations have suggested decreased frequency selectivity at shorter durations. Van den Brink and Houtgast (1990) described thresholds of long-duration tone complexes (100 ms) by a $5 \log(n)$ rule, where n is the number of 1/3 octave spaced frequency components. Thresholds of short-duration signals (10 ms) were described by an $8 \log(n)$ rule. More efficient spectral integration at shorter durations was used to explain these trends. Similarly, profile analysis experiments using tone complexes of varying frequency spacing and duration showed thresholds greatly dependent on duration and frequency ratio (Dai and Green, 1993). Dai and Green proposed a mechanism of broadening auditory filters to describe these results.

Moore *et al.* (1987) and Wright and Dai (1994) explored durational effects by conducting notched-noise experiments using short-duration tones. Moore *et al.* found that filter shapes did not significantly vary as a function of the temporal position of the signal with respect to the masker, suggesting that auditory filters do not develop over time. Wright and Dai supported these results by measuring similar filter shapes for both short- (5-ms) and long-duration (295-ms) tones. However, in both studies, filters were measured using a long-duration masker; perhaps measured filter shapes were related to the long-duration notch rather than the short-duration signal. To explain their findings, Wright and Dai (1994) suggested that at short durations, instead of widening filters,

^{a)}Parts of this work were presented at the 130th Meeting of the Acoustical Society of America, St. Louis, MO [Hant *et al.*, J. Acoust. Soc. Am. **98**, 2908 (A) (1995)] and the 4th Meeting of the International Conference on Spoken Language Processing, Oct. 1996, Philadelphia, PA [Hant *et al.*, ICSLP, (1), 570–574, 1996].

subjects sum energy across several filter outputs. The effect of summing energy from several auditory filters at short durations is similar to “increased spectral integration” or “broadening auditory filters” since the three mechanisms imply decreased frequency selectivity at short durations.

Duration also affects frequency discrimination; Moore (1973) measured frequency difference limens for 10-ms tones to be as much as an order of magnitude higher than those of 200-ms tones.

Temporal integration has also been used to describe the masking of short-duration stimuli. Hughes (1946) quantified temporal loudness summation by measuring the audibility thresholds of tones in quiet as a function of duration (t). He found that his results can be modeled by the following equation:

$$(I - I_0)t = I_0\tau = \text{const.}$$

This model assumes that the auditory system integrates stimulus intensity above a minimal intensity (I_0), over a time constant (τ), and that this integration is linear with time. Other investigators such as Garner and Miller (1947), and Plomp and Bouman (1959) conducted similar threshold experiments and calculated time constants on the order of 100 ms.

If energy integration of approximately 100 ms were the only mechanism involved with the change in tone thresholds as a function of duration, temporal resolution might be limited by this integration time. To address this discrepancy, Viemiester and Wakefield (1991) proposed a “multi-look” mechanism that separates the process of loudness summation from temporal resolution. This mechanism implies that sounds are sampled over a temporal window less than 10 ms long. These samples, or looks, are then stored in memory and can be accessed or processed separately. This mechanism allows for loudness summation of signals over long durations without reducing temporal resolution.

There have also been several studies on the noise masking of noise signals. However, data from these studies are not sufficient to develop a model for predicting the masking of noisy speech segments, such as plosive bursts. Green (1960) measured the masked thresholds of bandpass noises as functions of signal duration and center frequency but did not significantly vary signal bandwidth. Schacknow and Raab (1976) measured the masked thresholds of bandpass noises as functions of both signal and masker bandwidth but did not vary signal duration. Scholl (1962) tested the effects of both bandwidth and duration on the thresholds of bandpass noises in noise but only at a center frequency of 2 kHz. The psychoacoustic experiments reported in this paper measure the masked thresholds of noise signals at four bandwidths, four durations, and five center frequencies. By measuring the masked thresholds of noise over a range of bandwidths, durations, and center frequencies, a model for predicting the noise masking of plosive bursts can be developed.

B. Background for the speech perception experiments

Previous perceptual experiments have shown that plosives are often confused in noisy environments. Miller and

TABLE I. Bandpass noise signals used in the experiments. The CB values were chosen from values provided by Zwicker and Terhardt (1980).

Center freq. (in Hz)	1 CB bandwidth (in Hz)	2 CB bandwidth (in Hz)	4 CB bandwidth (in Hz)	8 CB bandwidth (in Hz)
400	100	200	400	800
1000	162	324	648	1296
2000	301	602	1204	2408
3000	479	958	1916	3832
4000	685	1371	2742	5483

Nicely (1955) showed that noise masking impairs listeners’ ability to identify the plosive’s place of articulation. It is believed that a plosive’s place of articulation is signaled by several acoustic cues such as the plosive’s spectrum at the consonant release and direction and rate of formant transitions (e.g., Blumstein *et al.*, 1977; Blumstein and Stevens, 1980; Walley and Carell, 1983).

Few studies have examined the perceptual importance of different acoustic cues in noise. Using stationary broadband noises with spectral shapes resembling those of the plosives, Farar *et al.* (1987) measured the discrimination thresholds for different plosive burst pairs in noise as a function of the bursts’ duration. Discrimination thresholds decreased by 20 dB as the bursts’ duration increased from 10 to 300 ms. To predict these data, a model that filtered the signal with non-overlapping critical bands, averaged the power via temporal integration, and added Gaussian noise was used. However, this model did not predict the durational effects on thresholds. In another study, Alwan (1992) conducted identification experiments with synthetic and burstless /bV/ and /dV/ stimuli whose F_2 trajectories were masked with a bandpass noise masker. The results suggest that cues other than the bursts and formant transitions, such as the amplitude differences in F_3 and F_4 , can be used as place cues in noise.

II. EXPERIMENT 1: MASKING OF BANDPASS NOISES

In this section, a series of psychoacoustic experiments using bandpass noises of varying center frequency, duration, and bandwidth are described. Data from these experiments will be used to parameterize a model that predicts the masking of natural and synthetic plosives.

A. Subjects

Four subjects (two male, two female) with normal hearing and tested tone thresholds below 20 dB HL participated in the experiments. All were native talkers of American English and ranged in age from 24 to 30 years. Two of the subjects were experienced with psychoacoustic experiments and not paid for their participation, the other two were paid.

B. Stimuli

Table I summarizes the bandwidths and center frequencies tested. The bandpass noises (signals) had center frequencies ranging from 400 to 4000 Hz, and bandwidths between 100 and 5483 Hz. The bandwidths correspond to 1–8 critical bands (CB) for each center frequency tested. Each of these bandpass noise signals were played for the listeners at four

durations: 10, 30, 100, and 300 ms. Signals were turned on and off using a raised-cosine window with a rise–fall time of 1 ms. All signals were generated by digitally filtering one of 32 Gaussian noise sources.

The masker was broadband noise at a spectrum level of 36 dB/Hz and a duration of 750 ms. Bandpass noise signals were centered in time with respect to the masker. The masker was randomly selected from 32 Gaussian noise sources and turned on and off using a raised cosine window with a rise–fall time of 20 ms.

C. Methods

Stimuli were presented binaurally to listeners in a soundproof room via Telephonics TDH49P headphones. Computer software generated the test tokens as 16 bit/16 kHz digital sequences. An Ariel ProPort 656 board performed digital-to-analog conversion. The resulting analog waveforms were amplified using the preamp of a Sony 59ES DAT recorder which was connected to the headphones. The entire system was calibrated within ± 0.5 dB before each experiment using a Larson Davis 800B sound-level meter. Masked thresholds were determined using an adaptive 2AFC paradigm (Levitt, 1971). Three correct responses determined a successful subtrial while one incorrect response determined an incorrect subtrial. Thresholds, therefore, are defined to be the 79% correct point. Step sizes were initially set to 4 dB, then reduced to 2 dB after the first reversal, and finally to 1 dB after the third reversal. From a total of nine reversals, the average of the last six determined threshold.

D. Results

The results of the bandpass noise experiments are shown in Fig. 1. At each center frequency tested, the masked thresholds for the bandpass noises are plotted as a function of signal duration with signal bandwidth as a parameter. The data are expressed in this manner to reveal bandwidth effects. Thresholds are reported both as intensity thresholds [dB sound-pressure level (SPL)], plotted on the left side of the figure, and spectrum-level thresholds (dB/Hz), plotted on the right side of the figure. Thresholds are averaged across subjects with standard deviations expressed by the error bars.

Previous experiments have shown that the masked thresholds of steady-state tones in broadband noise occurs when the signal-to-noise ratio (SNR) in a critical band is equal to -4 dB (Hawkins and Stevens, 1950). Because the bandpass noise signals had bandwidths greater than or equal to a critical band, this value corresponds directly to a spectrum level SNR. Thus, for a broadband masker of 36 dB/Hz, previous tone-masking experiments predict spectrum level thresholds of $36 - 4 = 32$ dB/Hz. This threshold is shown as a dotted line in the spectrum level plots. The change in thresholds with duration for the 1 CB noises is similar to that measured for tones by Plomp and Bouman (1959).

E. Discussion

At short durations the intensity thresholds are similar across bandwidth, while at long durations, spectrum level thresholds are similar across bandwidth. This may imply re-

duced frequency selectivity at short durations. If at short durations subjects are summing signal energies over a wide frequency region, intensity thresholds will be similar across bandwidth. If at long durations subjects are summing signal energies over a narrower frequency region, then spectrum level thresholds will be similar across bandwidth. Figure 1 shows that these bandwidth effects are consistent for all center frequencies tested. At 4000 Hz, however, these effects are less pronounced.

The observed bandwidth effects cannot be the result of spectral splatter. A 10-ms, 100-Hz bandwidth signal spreads to a 3-dB equivalent bandwidth of only 150 Hz.

Van den Brink and Houtgast (1990) found similar bandwidth effects with tone complexes and described the phenomena as more efficient “spectral integration” at short durations. More efficient spectral integration at short durations implies reduced frequency selectivity and can therefore explain the bandwidth trends in the current data. Dai and Green (1993) found similar results conducting profile analysis experiments and attributed their results to broadening auditory filters with decreasing duration. Similarly, broadening auditory filters with decreasing duration suggests reduced frequency selectivity. The data can also be described by a summation of energy across filter outputs as proposed by Wright and Dai (1994). Summation of energy across filter outputs, again implies reduced frequency selectivity.

Figure 1 shows that spectrum level thresholds for wider bandwidth signals do not decrease at durations greater than 30 ms. Spectrum level thresholds for wider bandwidth signals are least affected by changing frequency selectivity. Therefore, mechanisms not related to the signal’s bandwidth, such as temporal integration, might occur over a time period equal to or less than 30 ms. If these noise signals were integrated over time constants of order 100 ms as proposed by Plomp and Bouman (1959), then the thresholds for the wider bandwidth noises would decrease past 30 ms; not a general trend observed in the data.

In Fig. 1 there is also a noticeable kink at 30 ms, especially for the 8 CB thresholds. This may reflect mechanisms sensitive to signal transients. The transients of a 30-ms noise signal may be more detectable than those of a longer duration noise, which has onsets and offsets that are more spaced apart in time.

III. A MODEL OF DURATION-DEPENDENT AUDITORY FILTERS

The durational dependence of noise-masked thresholds may be the result of several mechanisms that include changing frequency selectivity, temporal integration, and sensitivity to signal transients. In this section, a model of duration-dependent auditory filters is proposed. The model divides the durational mechanisms into those that are dependent on signal bandwidth and those that are not. The proposed model is described by two duration- and frequency-dependent parameters: the processing efficiency, K , and a frequency-selectivity parameter, p .

As stated previously, duration-dependent filters are one of several mechanisms to describe changing frequency selectivity. Both spectral integration and summation of energy

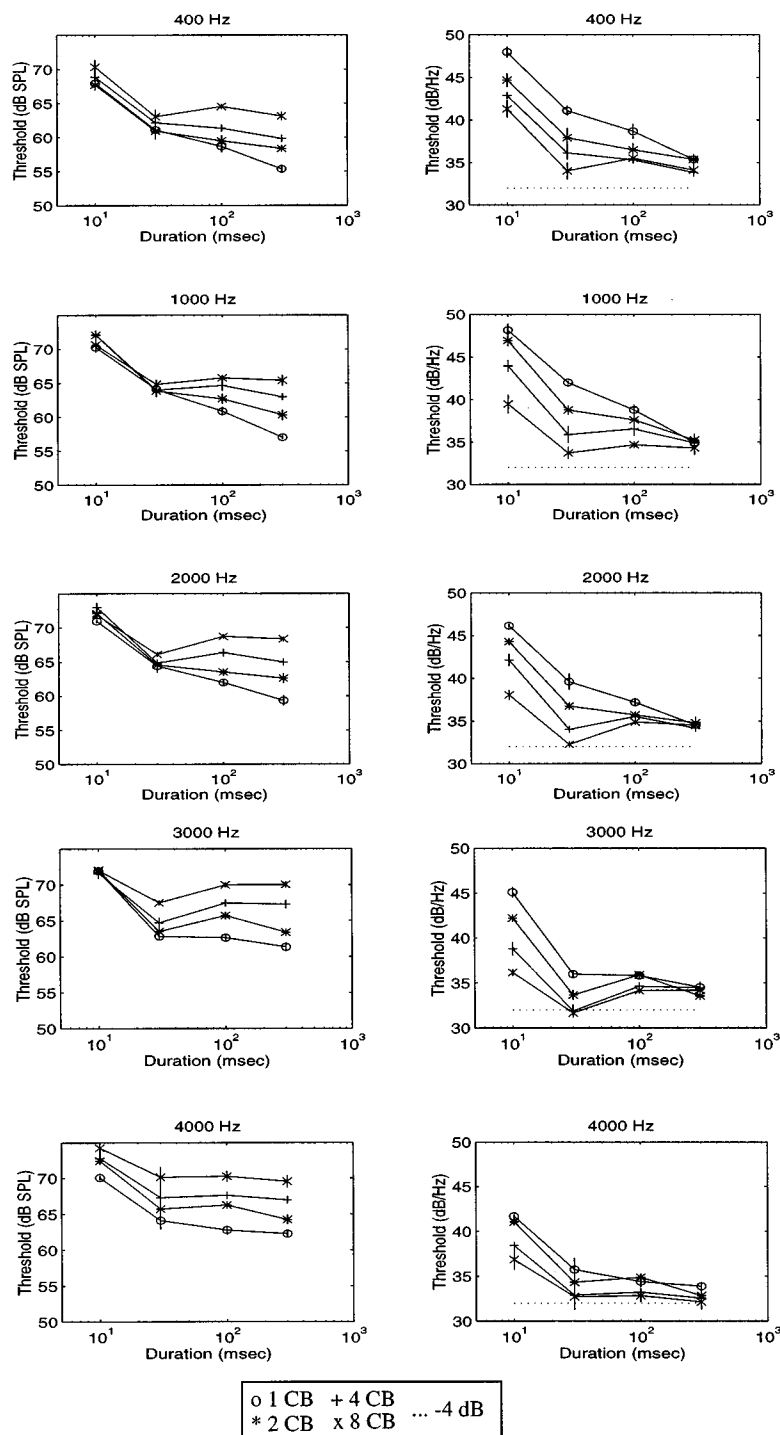


FIG. 1. Average bandpass noise thresholds in a flat noise masker. The data shown are the thresholds of bandpass noises centered at 400 Hz, 1 kHz, 2 kHz, 3 kHz, and 4 kHz in a flat noise masker with a spectrum level of 36 dB/Hz. Thresholds are expressed as total intensity thresholds (dB SPL) in the left column and spectrum level thresholds (dB/Hz) in the right column. Contours represent the different bandwidths tested. Dotted lines in the right column represent the thresholds as predicted from long duration tone-masking experiments. Standard deviations across subjects are expressed by the error bars.

across filter outputs can describe duration-dependent frequency selectivity, but, to predict speech masking thresholds, they add unnecessary complications. “Spectral integration,” as defined by the $N \log(n)$ rule (Van den Brink and Houtgast, 1990), predicts thresholds only for spectrally flat signals and hence, cannot be used to predict the thresholds of spectrally shaped signals such as plosive bursts. Summation of energy across filter outputs requires quantifying a two stage process, an initial filter followed by higher-level sum-

mation. Duration-dependent auditory filters provides a direct, single-stage, model of changing frequency selectivity.

A. Model background and description

Critical bands, a measure of auditory frequency selectivity, were first introduced by Fletcher (1940). Patterson (1976) and several other investigators later used notched-noise experiments to quantify the precise shapes of auditory

filters. The roex function has been used to describe the shape of the auditory filter (Glasberg and Moore, 1990):

$$W(g) = (1 + pg)e^{-pg}, \quad (1)$$

where g is the normalized frequency deviation from the filter's center frequency, and p is the slope of the filter skirts.

At threshold, the signal-to-noise ratio at the output of the filter, defined as the filtered signal intensity divided by the filtered noise intensity, is equal to a constant (K). The bandwidth of the filter is determined by the parameter p in the roex function.

In the duration-dependent filter model, the parameters p and K are functions of signal duration. The parameter p determines how the effective bandwidth of the auditory filter changes with duration, while the parameter K determines how the processing efficiency in each filter changes with duration. p quantifies changing frequency selectivity, and K quantifies duration-dependent mechanisms that are independent of signal bandwidth. Below is a list of assumptions used in the model.

(1) Masked thresholds are determined by the auditory filter with the highest SNR. The wide-bandwidth noises excite several adjacent filters. Our model assumes that redundant subthreshold responses are not combined to provide a net superthreshold response. All filters provide detection information, but filters with the maximum SNR provide detection information at the lowest signal intensity, and therefore, determine threshold.

With the wide-bandwidth noise signals, several filters have a similar, maximum SNR. To simplify the model parametrization, we assume detection thresholds are determined by the filter at the signal's center frequency. This assumption is incorrect if either of the two parameters in the model vary with center frequency in a manner that makes the detection of wide-bandwidth signals more advantageous away from the center channel. With a wide-bandwidth constant spectrum-level signal, the filter bandwidth does not affect SNR at the filter output: a filter bandwidth increase adds as much signal power as noise power. Therefore, the simplifying assumption only requires processing efficiency not to vary significantly with frequency.

In previous notched-noise experiments, processing efficiency changes by 2–4 dB for frequencies ranging between 124 and 4000 Hz (Patterson, 1976; Patterson *et al.*, 1982; Fidell *et al.*, 1983). In this study, thresholds of long-duration narrow-band noise signals imply that processing efficiency changes by 1.5 dB over a similar frequency range. Therefore, our assumption that processing efficiency does not vary significantly with center frequency increases the model's wide-bandwidth threshold estimations by at most 1.5 dB.

(2) Auditory filters have the shape of a symmetric roex function where the parameter p determines the slope of the filter skirts. Asymmetries of auditory filters are ignored.

(3) Parameters p and K vary as functions of signal duration and center frequency.

(4) At 300 ms, filters have reached steady state. The p values at 300 ms are determined directly from previous notched-noise data.

B. Model derivation for bandpass noises

Substituting $g = |f - c_f|/c_f$ into Eq. (1), an expression for $W(f)$ is obtained

$$W(f) = \left(1 + \frac{p|f - c_f|}{c_f}\right) e^{-(p|f - c_f|/c_f)}, \quad (2)$$

where c_f is the filter's center frequency.

The parameter K defines the SNR of the filter at threshold and is denoted by the following equation

$$K = \frac{P_s}{P_n} = \frac{\int_{-\infty}^{\infty} S(f)W(f)df}{\int_{-\infty}^{\infty} N(f)W(f)df}, \quad (3)$$

where P_s and P_n are the power of the signal and masker through the auditory filter. $S(f)$ and $N(f)$ are the power spectra of the signal and the noise masker, respectively.

To apply this model to the bandpass noise data, $S(f)$ is a bandpass noise of bandwidth b and spectrum level S (power/Hz). $N(f)$ is a broadband noise of spectrum level N over frequencies ranging between 0 and 8 kHz. Therefore, $S(f)$ and $N(f)$ can be taken out of their respective integrals. Since the roex function is symmetric about its center frequency, c_f , K can be expressed as

$$K = \frac{S \int_{c_f}^{c_f + b/2} W(f)df}{N \int_{c_f}^{8 \text{ kHz}} W(f)df}. \quad (4)$$

By integrating the roex function, and solving for S , we can express the model predictions as a function of the parameters K , N , c_f , b , and p ,

$$S_{\text{model}} = \frac{KN(2c_f/p - (8 \text{ kHz} - c_f + 2c_f/p)e^{-(p(8 \text{ kHz} - c_f)/c_f)})}{[2c_f/p - (b/2 + (2c_f/p)e^{-(pb/2c_f)})]}, \quad (5)$$

where S_{model} is the model prediction for the signal threshold in power/Hz. N , c_f and b are known for each signal duration, bandwidth, and center frequency.

C. Determining K and p values

Both the processing efficiency, K , and filter-shape parameter, p , vary with center frequency and duration [$K(d, c_f)$ and $p(d, c_f)$]. Once specified, these parameters determine the model's prediction of the masking thresholds. The filter-shape parameters at long durations were set to those reported by Glasberg and Moore (1990). An iterative procedure was used to estimate the remaining parameters by minimizing the MSE of the model's prediction of the masking thresholds. Unconstrained MSE minimization, however, leads to discontinuous K and p functions of center frequency and duration.

Therefore, a second iterative estimation was used that minimizes a cost function equal to the sum of the MSE of the prediction and the MSE of the linear fit of $K(c_f, d)$ and $p(c_f, d)$. That is, parameters were chosen to predict the noise masking data and to vary smoothly as a function of center frequency and duration, using a soft constraint. The linear constraints used $K(c_f, d)$ expressed in dB, and $p(c_f, d)$ expressed in ERB (equivalent rectangular bandwidth).

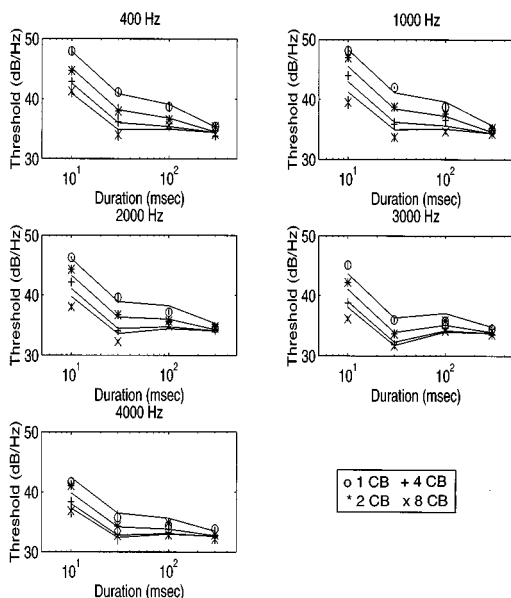


FIG. 2. Model fit to bandpass noise data. Symbols represent the experimental thresholds (in dB/Hz) for the 400–4000 Hz bandpass noise data. Solid lines represent the model fits.

D. Results

Figure 2 shows the model prediction of the bandpass noise thresholds. Model errors range between 0 and 2 dB. The parameter $K(d, c_f)$ traces the general change in thresholds as a function of duration. The parameter $p(d, c_f)$ determines the width of the auditory filter, and therefore, specifies the separation of thresholds with signal bandwidth.

Figure 3 shows the model estimates for changing filter

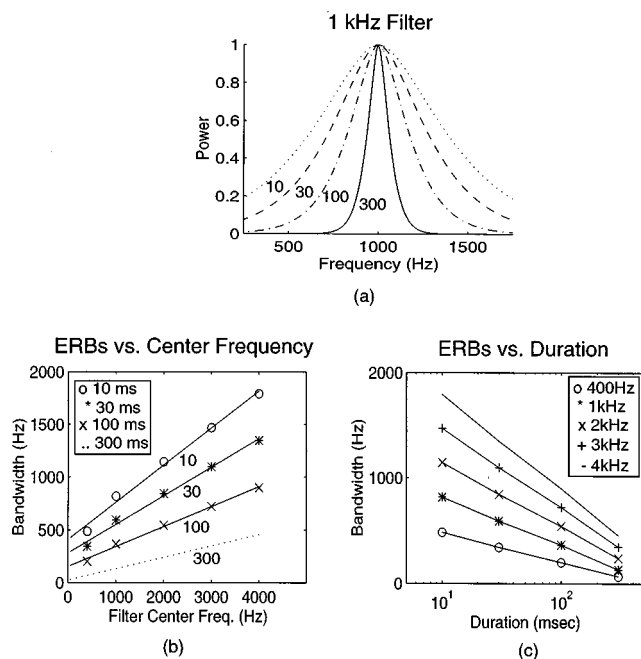


FIG. 3. Model estimates of filter bandwidths. (a) Estimated filter shapes for the 1 kHz filter at different durations. (b) ERBs vs center frequency with duration as a parameter. (c) ERBs vs duration with center frequency as a parameter.

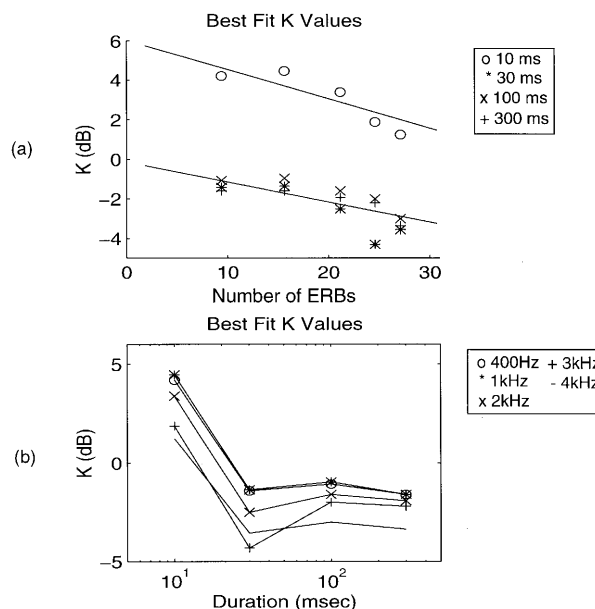


FIG. 4. Best fit K values. (a) K values vs number of ERBs with duration as a parameter. (b) K values vs duration with center frequency as a parameter.

bandwidths as a function of duration. Figure 3(a) plots the predicted filter shape for the 1-kHz filter at 10, 30, 100, and 300 ms. At 1 kHz, the model predicts filter bandwidths increase by a factor of 6 from 300 to 10 ms. This number is similar to the results of Dai and Green (1993) which approximated a bandwidth increase of 7.

Figure 3(b) plots ERBs as functions of filter center frequency with duration as a parameter. ERBs are greater with increasing center frequency and decreasing duration. Figure 3(c) plots the same data showing filter ERBs as functions of duration with center frequency as a parameter. The fractional change in bandwidth appears to be greater for filters with a lower center frequency. From 300 to 10 ms, the ERBs for the 400 and 4000 Hz filter change by a factor of 7 and 4, respectively.

The model's K values are shown in Fig. 4. Figure 4(a) graphs K values, in dB, as a function of the filter's ERB number with duration as a parameter. Linear K values at 10 ms and 30–300 ms were fit to exponential functions of center frequency. Figure 4(a) shows these functions on a dB scale. Figure 4(b) plots the same K values as functions of duration with center frequency as a parameter.

Figure 4 shows that there is a large difference between K values at 10 ms and those at the longer durations. Processing efficiency is reduced at short durations. The difference between K values at 10 ms and steady state could be interpreted as an indirect measure of integration time constants. If sounds are integrated over time constants near 30 ms, then the difference between K values at 10 ms and steady state should approach a factor of 3, or about 5 dB, as seen in the figure. Figure 4(b) shows a kink in the K values at 30 ms, reflecting the threshold changes for bandpass noises in Fig. 1.

E. Discussion

It is also possible to model the bandpass noise data using a model of *fixed* frequency selectivity, temporal integration,

and a cross-channel mechanism. With this model, sounds are first filtered through a fixed filter bank. Each filter output is then integrated in time and processed through a detection device. The outputs of the detection devices from several frequency channels are used to detect the signal. Assuming statistically independent channels which are combined optimally, this cross-channel mechanism leads to a signal detectability d' that increases with the square root of the number of frequency channels which provide redundant information (Green and Swets, 1966). For a noise signal eight times wider than a critical band, there are approximately eight adjacent filters with redundant information and the spectrum level threshold decreases by a factor of $\sqrt{8}$, or 4.5 dB. Thus, the cross-channel d' mechanism predicts a decrease in spectrum level thresholds with increasing signal bandwidth.

For the bandpass noise data, this spread in thresholds decreases with increasing duration. Therefore, the cross-channel (d') mechanism has to be *less* efficient at long durations. Although this is possible, it seems unlikely that our ability to use information across filter outputs would become less efficient with increasing duration. If a cross-channel d' mechanism were significant at long durations, there would be a large spread in spectrum level thresholds across bandwidth; not a general trend observed in the data. In addition, the magnitude of the spread in thresholds at short durations appears to be too large to be described solely by a cross-channel d' mechanism. At 10 ms, the 1 and 8 CB noises show a difference in spectrum level thresholds of almost 8 dB, more than twice as large as that predicted by a \sqrt{n} cross-channel d' mechanism. If *energy*, rather than *information*, were summed over the different filter outputs, as implied by decreasing frequency selectivity, the predicted difference in thresholds would equal 9 dB. Finally, in Sec. V we present results from a set of notched noise experiments that directly measure reduced frequency selectivity at short durations. These results cannot be explained by a cross-channel d' mechanism.

It is our working hypothesis that frequency selectivity is the result of at least two processes: low-level basilar filtering, and higher-level neural processing of the filter outputs (Licklider, 1951; Zwicker, 1965; Lyon, 1984; Patterson *et al.*, 1992; Wang and Shamma, 1994). While the first may be nearly instantaneous, the second may require time to evolve. Although perceptually parametrized models are influenced by both processes, a static critical band model neglects the possible temporal evolution of the second process. By allowing *effective* filter bandwidths to change with duration, we can approximate the evolution of the second process.

IV. SPEECH EXPERIMENTS

In the previous section, a duration-dependent auditory filter model was derived to predict noise-masked thresholds of bandpass noises. In this section, the model is used to predict the noise masking of plosive bursts. Both natural and synthetic voiceless plosives are used as stimuli in masking experiments and their thresholds are compared to model predictions.

A. Speech samples

Three talkers (two male and one female) were recorded. All were native talkers of American English and had no prior phonetic training. Words with initial plosives were recorded in the middle of the carrier phrase "Say — again." /k/'s were recorded preceding four different vowels ranging from back to front in the words "coup," "cot," "cat," and "key." /t/'s and /p/'s were recorded preceding three different vowels ranging from back to front in the words "too," "tat," and "tea" and "pooh," "pat," and "pea," respectively. Each of the three talkers recorded the different speech samples twice. The speech signals were recorded digitally (16 bit/32 kHz) using a calibrated microphone [Beyerdynamic M 101 N(C)] with a relatively flat frequency response. The microphone, with a windscreen, was placed at a distance of approximately 10 in. and at an angle of 45° from the talker to limit any distortion caused by the burst of the plosive.

The word containing the plosive was manually isolated from its carrier phrase. Seven ms of silence was retained in front of the word, to ensure that the burst was not cut out of the waveform. Any dc offset was removed and the signal was downsampled to 16 kHz. To isolate the plosive from the word, the beginning of each speech signal was segmented using a raised cosine window with a rise-fall time of 1 ms and a duration of either 17 or 37 ms (10 or 30 ms without silence).

B. Analysis of speech waveforms

Figure 5 shows examples of burst spectra for plosives in each of the different carrier words. These DFT spectra are of the first 10 ms of the plosive and were computed using a Hamming window. The title on each graph corresponds to the specific talker, either "jh," "bps," or "sc," and the carrier word from which the plosive was spoken.

The spectral shapes for the plosive bursts are generally consistent with what has been reported in the literature. The /k/'s, /t/'s, and /p/'s appear to fall into the respective compact, diffuse rising, and diffuse falling templates originally described by Blumstein and Stevens (1980).

The /k/ bursts in the words "coup" and "cot" have a prominent spectral peak in the range of 1.7–2 kHz with a secondary peak at around 4 kHz; these are classified as back /k/ bursts. The /k/ bursts in the words "cat" and "key," on the other hand, have a broader spectral peak in the range of 2.4–3 kHz and a less prominent peak near 4 kHz; these are classified as front /k/ bursts. The spectral peaks of both the back and front /k/'s remain prominent through a majority of their aspiration noise. However, these peaks become broader as the vowel is approached. The time waveforms show that the back and front /k/'s can have one to three bursts in their first 10 to 30 ms.

The /t/ bursts have a broad spectral shape in the range of 5–6 kHz which is relatively independent of vowel context. At 30 ms this peak is still prominent. The typical /t/ time waveform consists of one burst with significant aspiration lasting for about 100 ms.

The /p/ bursts are the most variable. The only common

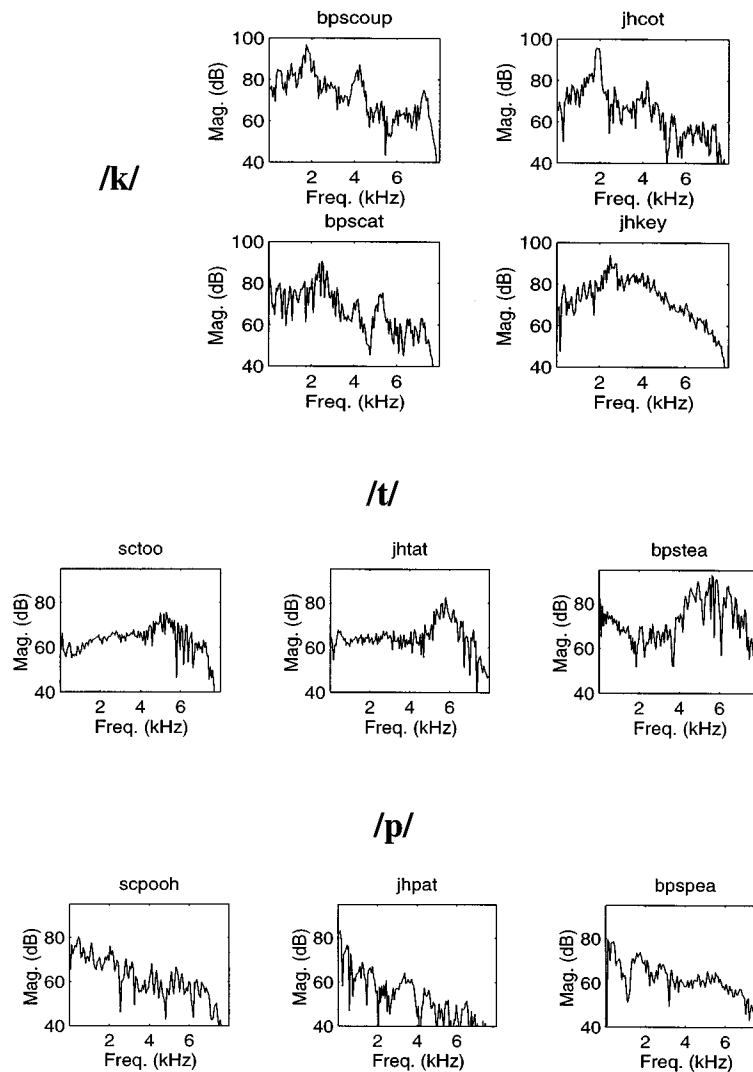


FIG. 5. Examples of spectral shapes for the different speech tokens. Each token is denoted by its talker and its carrier word.

characteristic among the /p/ bursts is a “diffuse falling” spectrum.

The female subject, “sc,” showed slightly higher resonance peaks than the two other male talkers for both back and front /k/’s. For the /t/ and /p/, however, no considerable differences between the female and male talkers were observed.

C. Synthesis of speech signals

Noise bursts were synthesized to resemble the spectral shapes of the natural back /k/, front /k/, /t/, and /p/. For the

TABLE II. Klatt synthesizer parameters for back /k/, front /k/, and /t/. The parallel branch of the synthesizer was used. Resonances refer to the synthesizer parameters $F2$ and $F3$, bandwidths, to parameters $B2F$ and $B3F$, and gains, to parameters $A2F$ and $A3F$.

Speech sound	Resonances (in kHz)	Bandwidths (in Hz)	Gain (in dB)
back /k/	1.7, 3.5	50, 300	55, 48
front /k/	2.7	600	45
/t/	5.5	800	58

front and back /k/, and for /t/, the bursts were synthesized by exciting the parallel branch of the Klatt formant synthesizer (KLSYN88-Klatt, 1980) with the amplitude of the frication source (AF) set to 60 dB. The exact center frequencies, bandwidths, and levels of the resonance peaks for each of the bursts were set using the parameters $F2, F3, \dots, B2F, B3F, \dots$, and $A2F, A3F, \dots$, respectively; parameter values are shown in Table II. The /p/ burst was synthesized to have a “diffuse falling” spectral shape and no resonance peaks. Thus, flat noise was filtered to have a falling slope of 30 dB/8000 Hz.

All signals were synthesized at 300 ms. Using a raised-cosine window with a rise–fall time of 1 ms, the bursts were then segmented into durations of either 10, 30, 100, or 300 ms. The spectral shapes for the synthetic bursts are shown in Fig. 6.

D. Model Predictions of Speech Thresholds

In Sec. III, the model parameters K and p were calculated at discrete durations of 10, 30, 100, and 300 ms and center frequencies of 400, 1000, 2000, 3000, and 4000 Hz. To predict the thresholds of signals with arbitrary spectral shape, K and p values need to be interpolated over a con-

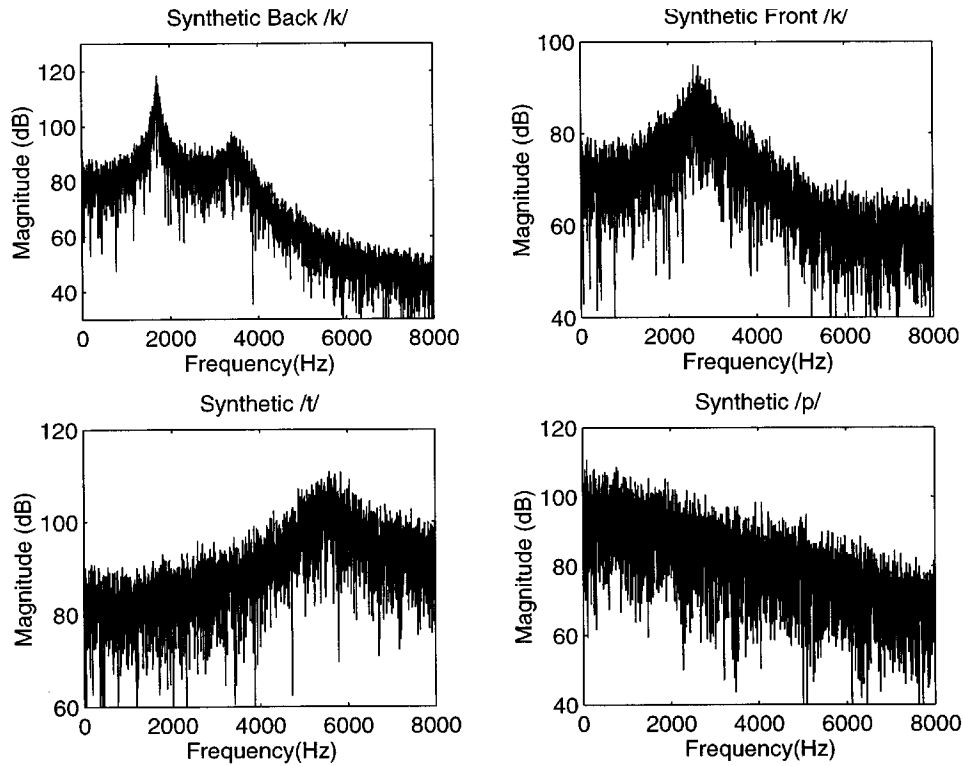


FIG. 6. Spectral shapes of the synthetic plosive bursts.

tinuous frequency range. The interpolation of these parameters is shown by the curves in Figs. 3 and 4.

To predict speech thresholds, the noise masker and plosive burst are first processed through a bank of filters characterized by the interpolated p values. Adjacent filters in the filter bank have center frequencies ranging between 200 and 6500 Hz, separated by 50 Hz increments. The SNR corresponding to each filter is determined by dividing the filtered signal intensity by the filtered noise intensity. The filter with the maximum SNR, denoted by the frequency c'_f , is used to determine the signal threshold.

Recall that the parameter K is defined as the filtered signal power, P_s , divided by the filtered noise power, P_n .

$$K(d, c'_f) = \frac{P_s(d, c'_f)}{P_n(d, c'_f)}. \quad (6)$$

The filtered signal power at threshold, P_{s_0} , is determined by solving for P_s at c'_f .

$$P_{s_0}(d) = K(d, c'_f) P_n(d, c'_f). \quad (7)$$

The total signal power at threshold, S_0 , is P_{s_0} scaled by the ratio of the total signal power, S , to the filtered signal power, P_s :

$$\begin{aligned} S_0(d) &= P_{s_0}(d) \frac{S}{P_s(d, c'_f)} \\ &= K(d, c'_f) P_n(d, c'_f) \frac{S}{P_s(d, c'_f)}. \end{aligned} \quad (8)$$

$K(d, c'_f)$ is a parameter of the model. $P_n(d, c'_f)$ is determined by multiplying the power spectrum of the masker with

the power spectrum of the highest SNR filter and summing across frequency. The ratio $S/P_s(d, c'_f)$ is the total signal power divided by signal power through the highest SNR filter. (This ratio is only a function of the spectral shape of the signal and filter.)

Examples of processing the plosive bursts through the duration-dependent filter bank are shown in Fig. 7. The figure plots the output energies of the filter bank at each duration for the synthetic plosives back /k/, /t/, and /p/, respectively. The output intensity for each filter, denoted by circles on the graphs, is normalized to the total intensity of the signal. At long durations, auditory filters are narrow and hence, the filter bank preserves the original spectral shape of the burst. At shorter durations, however, auditory filters become wider and there is spectral smoothing. At 10 ms, this smoothing is so severe, that the /k/ burst no longer appears to have a second resonance peak and its once sharp spectral peak at 1.7 kHz is now broader. The spectral smoothing has less effect on the /p/ and /t/ bursts which have broader spectral peaks.

E. Methods for synthetic and natural speech experiments

The natural and synthetic bursts generated were used as signals in two sets of masking experiments. Thresholds were measured for these signals in a noise masker. Three of the four original subjects (2 male and 1 female) participated in the synthetic speech experiments, and a different three out of the four subjects (1 male and 2 female) participated in the natural speech experiments.

For the synthetic speech experiments, four plosive bursts

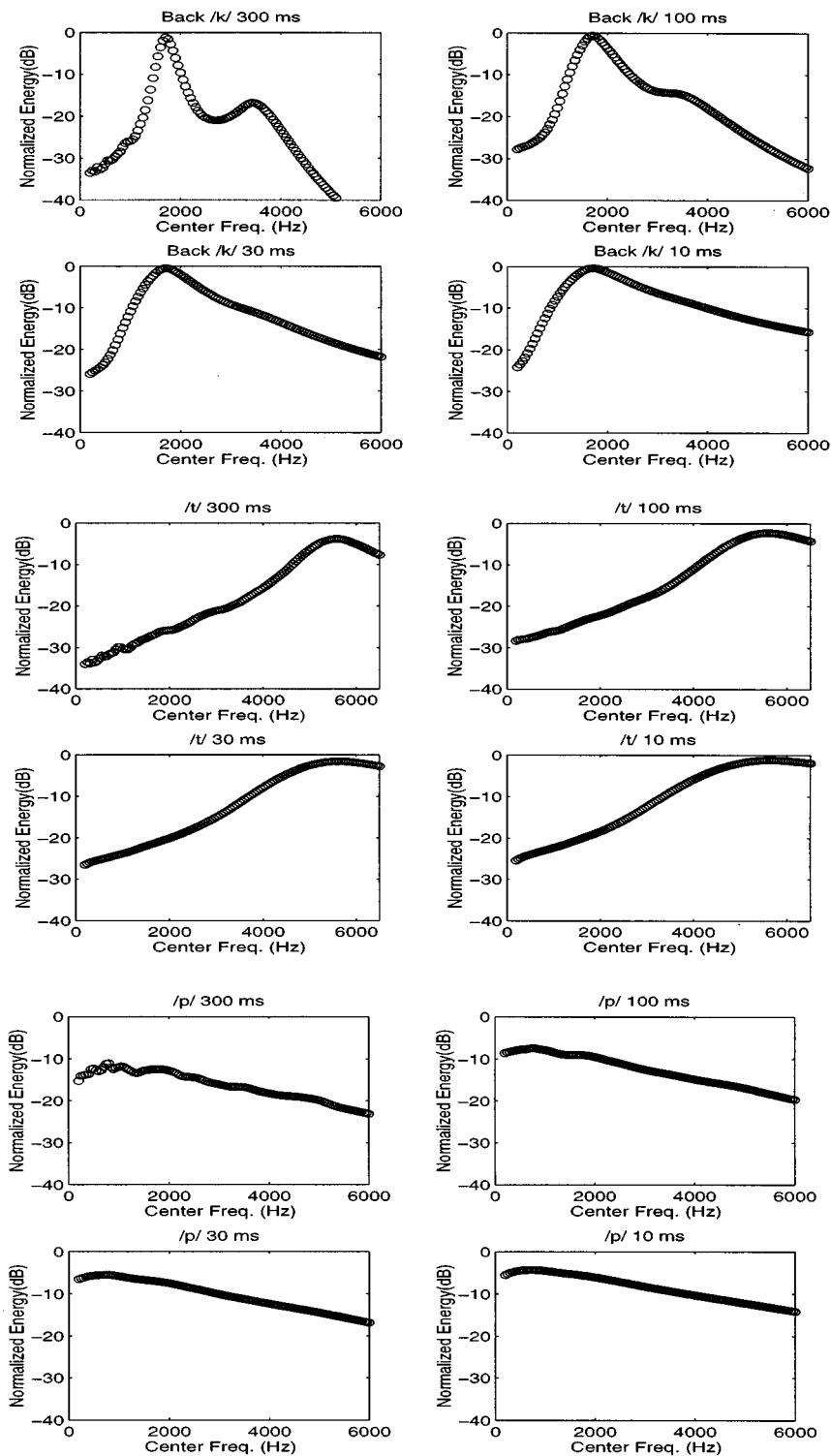


FIG. 7. Predicted spectral smoothing for synthetic /k,t,p/ bursts at different durations. The output energies for each filter, denoted by circles, are normalized to the total energy of the burst.

(back /k/, front /k/, /t/, and /p/) with durations of 10, 30, 100, and 300 ms were used, resulting in a total of 16 tokens. In the natural speech experiments, one plosive example from each of the carrier words and talkers, was used. Each was windowed at durations of 10 and 30 ms, yielding a total of 60 tokens.

The noise masker in both the synthetic and natural speech experiments had equal energy per critical band. The

masker was at a level of 56 dB/CB and had a duration of 750 ms. A DFT spectrum of this noise is plotted in Fig. 8.

Intensity thresholds for both the synthetic and natural speech experiments were defined as in the bandpass noise experiments. Spectrum level thresholds were not calculated for the plosive bursts because a single “spectrum level” does not characterize signals which have nonflat spectral shapes.

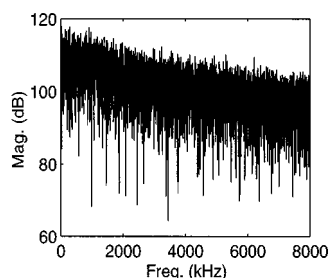


FIG. 8. Spectrum of the perceptually flat masker.

F. Model prediction of synthetic speech thresholds

Figure 9 shows the thresholds for the synthetic speech and the corresponding model predictions. The figure shows intensity thresholds (denoted by asterisks) versus duration for the synthetic plosive bursts back /k/, front /k/, /t/, and /p/. Thresholds are averaged across subjects with standard deviations expressed by the error bars. Intersubject differences are only evident for /p/ thresholds. Superimposed on these graphs are the model predictions for the thresholds (denoted by circles).

Figure 9 shows that the model predicts the thresholds well at all durations with only slight errors at 30 ms. The model predicts the large durational dependence of thresholds for the back and front /k/, and the smaller changes in threshold for /t/ and /p/. These durational dependencies are as expected. The front and back /k/ have relatively narrow spectral peaks and hence their intensity thresholds are more affected by widening auditory filters. These bursts are more accurately modeled by the narrow bandpass noises whose thresholds vary significantly with duration. The /t/ and /p/ bursts, on the other hand, have broader spectral peaks and therefore their thresholds are less affected by changing filter

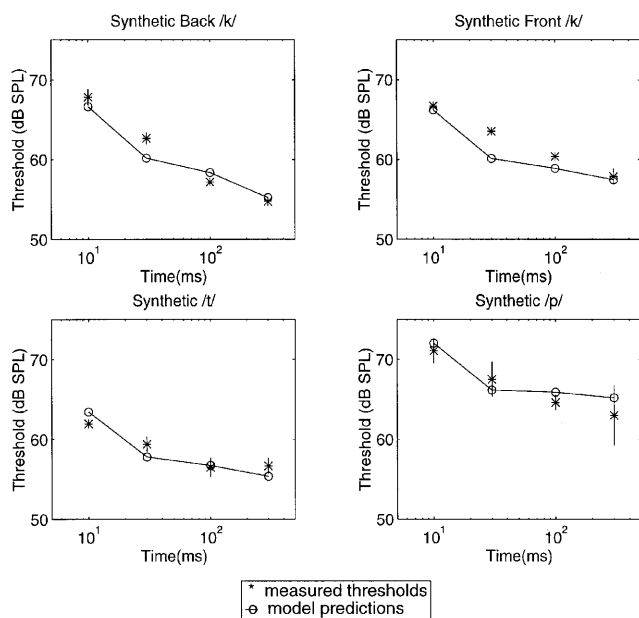


FIG. 9. Model predictions for the synthetic bursts. Asterisks denote perceptual thresholds of synthetic plosive bursts back /k/, front /k/, /t/, and /p/ in a *p*-flat masker (56 dB/CB). Error bars represent the standard deviations between subjects. The model predictions are denoted by the circles.

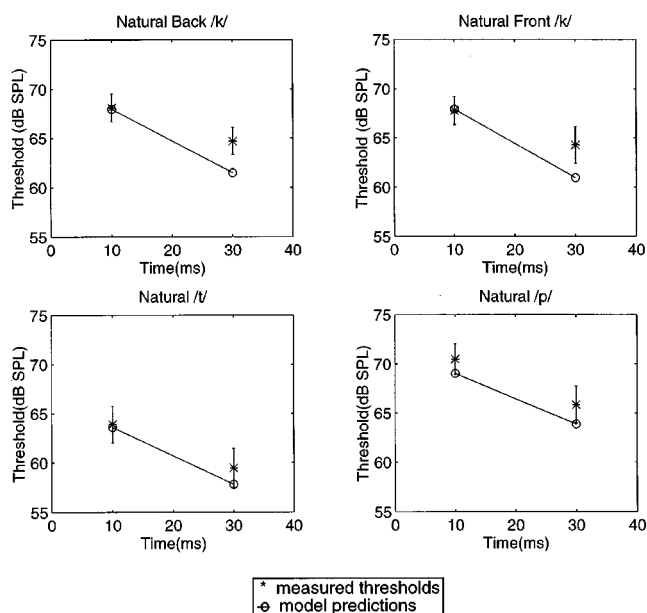


FIG. 10. Model predictions for the natural bursts. Asterisks denote the absolute thresholds of natural plosive bursts back /k/, front /k/, /t/, and /p/ in a *p*-flat masker (56 dB/CB). Thresholds are averaged over both the corresponding speech tokens and the three listeners, with error bars representing the standard deviations. The model predictions are denoted by the circles and lines.

shapes. These bursts are more accurately modeled by broad bandpass noises whose thresholds are less dependent on duration.

Figure 9 shows that the model slightly underestimates the thresholds for the front and back /k/ at 30 ms. These plosives do not show a kink in their thresholds as seen in the bandpass noise data. The kinks reduced the model's *K* values at 30 ms, and led to the underestimation of the /k/ thresholds. Perhaps, the transient effects for the synthetic /k/'s are less prominent than those of bandpass noises in the same frequency range.

G. Model predictions of natural speech thresholds

Model predictions for the natural plosive bursts are shown in Fig. 10. In this figure, thresholds for the natural plosives, denoted by the asterisks, are plotted as functions of duration. The intensity thresholds (dB SPL) are averaged over both the corresponding speech tokens and the three listeners. Thus, standard deviations, expressed by error bars, account for both talker and listener variations. The model predictions for these bursts are expressed by circles and lines.

Figure 10 shows that the duration-dependent filter model is successful in predicting thresholds at 10 ms. At 10 ms, the model predictions for all the plosives fall within the standard deviations of the data. The model errors for the 30 ms data are slightly larger, especially for the front and back /k/, with average errors between 2 and 3 dB. These model errors are similar to those seen for the synthetic stimuli.

The nonstationarity of the natural plosives may be another source of error. The filter model assumes constant spectral shape throughout the signal's duration. Despite the

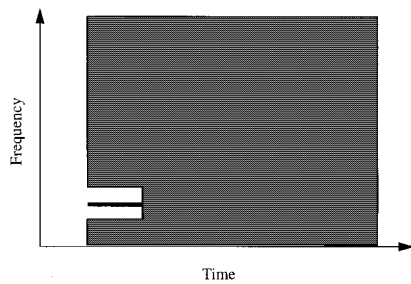


FIG. 11. Schematic spectrogram of the stimuli for the notched-noise experiment.

nonstationarity of the natural plosives, however, their thresholds are similar to the stationary synthetic stimuli at 10 and 30 ms.

H. Speech perception implications

A duration-dependent filter interpretation of the auditory system smooths the speech spectrum at short durations. This smoothing could have a large effect on what cues are used to perceive plosive bursts. The detailed spectral shape of the plosive burst may not be necessary for identifying the plosive's place of articulation. Figure 7 shows that broad spectral cues, like the concentration of energy at high or low frequencies, are retained at 10 ms and could play a role in identifying place. More subtle cues, however, like the bandwidth of spectral peaks, are perhaps smoothed at these short durations.

V. NOTCHED-NOISE EXPERIMENTS

In the previous sections, a duration-dependent filter model was derived to describe the noise masking of band-pass noises. At short durations, the model's filter shapes increase in bandwidth by a factor of 4–7. In this section, effective filter shapes at a short duration are directly measured by conducting a pilot set of notched-noise experiments. The data support the hypothesis that frequency selectivity is reduced at short durations.

Previous notched-noise experiments using short duration tones have shown filter shapes independent of both signal duration and temporal position (Moore *et al.*, 1987; Wright and Dai, 1994). However, in both studies, the notched-noise masker had a long duration. A decaying image of the tone (e.g., Patterson *et al.*, 1992) could have been compared to the long duration masker and used as a cue for detection. In this experiment, the decaying image of the tone is masked by filling in the notch with a bandpass-noise masker immediately after the tone is played.

A. Methods

Two male subjects with normal hearing participated in the experiments. Both subjects were experienced with psychoacoustic tasks. Notched-noise experiments were conducted at a center frequency of 1 kHz, for two tone durations, 20 and 300 ms. Notch widths varied between 0 and 1296 Hz (0 to 8 CB).

To reduce the effect of spectral splatter, all stimuli were turned on and off using a raised-cosine window with a rise–

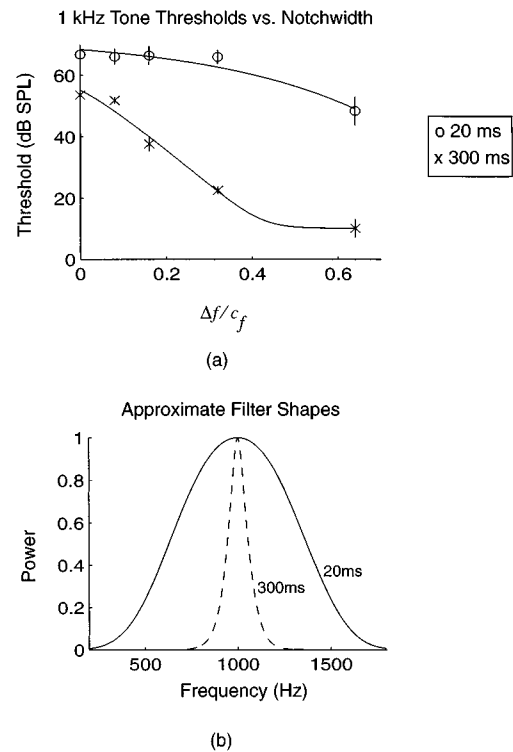


FIG. 12. Results of the notched-noise experiments. (a) 1 kHz tone thresholds versus bandwidth. Notch widths are expressed as the fractional deviation from the center frequency ($\Delta f/c_f$) with signal duration as a parameter. (b) Measured filter shapes for both 20 and 300 ms durations.

fall time of 10 ms. Signal duration was defined from the beginning of the onset to the end of the offset. A schematic spectrogram of the stimuli is shown in Fig. 11. The notched noise and tone were ramped on together. The tone was then ramped off while the notched noise continued for 280 ms past the offset of the tone. To fill in the notch, a bandpass noise masker was ramped on immediately after the tone. This bandpass noise continued through the remainder of the stimulus and was ramped off with the notched noise. The notched-noise and bandpass noise maskers had a constant spectrum level of 36 dB/Hz.

Thresholds were determined using an adaptive two-pair 2AFC procedure. Subjects chose between two pairs of notched noises where one of the pairs had a tone in one of the notches. This two-pair procedure was used to help subjects detect the short duration stimuli. Detection thresholds were measured three times for each subject, using the adaptive procedure described previously.

B. Results and discussion

The results of the notched-noise experiments are shown in Fig. 12. Figure 12(a) plots tone thresholds as a function of notch bandwidth, defined as the fractional deviation from center frequency ($\Delta f/c_f$), with signal duration as a parameter. Thresholds are averaged across the two subjects and three measurements with standard deviations expressed by the error bars. The 20-ms data is fit to an exponential of the form $a + 4.34[1 - \exp(-pg)]$, where g is the fractional deviation from the center frequency, and a and p are parameters. The value of 4.34 is chosen to ensure that the corresponding

filter shape does not have values greater than 1. From this fit, an approximate filter shape is calculated using the method described by Patterson (1976). The 300-ms thresholds are fit corresponding to a roex filter as described by Patterson *et al.* (1982). A roex filter is too steep to properly fit the 20-ms data. The resulting filters, for each duration, are shown in Fig. 12(b).

At short durations, tone thresholds only significantly decrease for notched widths greater than 640 Hz, or 4 CB, while at long durations, tone thresholds decrease continuously for notch bandwidths greater than 160 Hz, or 1 CB. Derived filter shapes are much broader at the short duration, reflecting reduced frequency selectivity. The ERB for the 20-ms filter is 745 Hz, while the 300-ms filter has an ERB of 114 Hz. These changes in filter bandwidth are comparable to those approximated by the duration-dependent filter model derived in Sec. III.

These effects cannot be the result of backward masking. Backward-masking thresholds for both subjects, using a 20-ms tone followed by a broadband masker, were only 18 dB SPL (57 dB below the intensity of the masker). In addition, the results cannot be due to an overshoot effect. Moore *et al.* (1987) found filter shapes independent of the temporal position of a tone with respect to the masker. The difference between this experiment and Moore's is that the notched noise is filled in with a bandpass noise after the tone is played.

In conclusion, the results of this experiment suggest that frequency selectivity is reduced at short durations, contrary to previous results. These results support the interpretation of the bandpass noise data as decreased frequency selectivity rather than a cross-channel d' mechanism. A more complete set of measurements, including a range of center frequencies and durations, will be presented in a future paper.

VI. CONCLUSIONS

The three conclusions of this study are:

- (1) The durational effects on noise masking of bandpass noises can be accurately described by an auditory filter model in which the effective bandwidth and processing efficiency of each filter are duration dependent. This study suggests that frequency selectivity decreases with decreasing duration.
- (2) The model is successful in predicting the noise masking of synthetic and natural plosive bursts. The model slightly underestimates the 30-ms data, especially for front and back /k/.
- (3) Notched-noise experiments at short durations measure a similar reduction in frequency selectivity.

ACKNOWLEDGMENTS

We thank our subjects for their cooperation. We would also like to thank W. Jesteadt, C. Turner, and an anonymous reviewer for their useful comments. This work was supported in part by NIH-NIDCD Grant No. 1 R29 DC 02033-01A1, The Whitaker Foundation, and NSF.

- Alwan, A. (1992). "The Role of F_3 and F_4 in Identifying the Place of Articulation for Stop Consonants," Proc. Intl. Conf. Spoken Lang. Proc., Vol. 2, 1063–1066, Banff, Canada.
- Blumstein, S. E., Stevens, K. N., and Nigro, G. N. (1977). "Property Detectors for Bursts and Transitions in Speech Perception," J. Acoust. Soc. Am. **61**, 1301–1313.
- Blumstein, S. E., and Stevens, K. N. (1980). "Perceptual Invariance and Onset Spectra for Stop Consonants in Different Vowel Environments," J. Acoust. Soc. Am. **67**, 648–662.
- Dai, H., and Green, D. M. (1993). "Discrimination of Spectral Shape as a Function of Stimulus Duration," J. Acoust. Soc. Am. **93**, 957–965.
- Farar, C., Reed, C. M., Ito, Y., Durlach, N. I., Delhorne, L. A., Zurek, P. M., and Braida, L. D. (1987). "Spectral-Shape Discrimination. I. Results from Normal-Hearing Listeners for Stationary Broadband Noises," J. Acoust. Soc. Am. **81**, 1085–1092.
- Fidell, S., Horonjeff, R., Teffeteller, S., and Green, D. M. (1983). "Effective Masking Bandwidths at Low Frequencies," J. Acoust. Soc. Am. **73**, 628–638.
- Fletcher, H. (1940). "Auditory Patterns," Rev. Mod. Phys. **12**, 47–65.
- Garner, W. R., and Miller, G. A. (1947). "The Masked Thresholds of Pure Tones as a Function of Duration," J. Exp. Psychol. **37**, 293–303.
- Glasberg, B. R., and Moore, B. C. (1990). "Derivation of Auditory Filter Shapes from Notched Noise Data," Hear. Res. **47**, 103–138.
- Green, D. M. (1960). "Auditory Detection of a Noise Signal," J. Acoust. Soc. Am. **32**, 121–131.
- Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics* (Wiley, New York).
- Hawkins, J. E., and Stevens, S. S. (1950). "The Masking of Pure Tones and of Speech by White Noise," J. Acoust. Soc. Am. **22**, 6–13.
- Hughes, J. W. (1946). "The Threshold of Audition for Short Periods of Stimulation," Proc. R. Soc. London, Ser. B **133**, 486–490.
- Klatt, D. H. (1980). "Software for a Cascade/Parallel Formant Synthesizer," J. Acoust. Soc. Am. **67**, 971–995.
- Levitt, H. (1971). "Transformed Up-Down Methods in Psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.
- Licklider, J. C. R. (1951). "A Duplex Theory of Pitch Perception," *Experientia* **7**, 128–134.
- Lyon, R. F. (1984). "Computational Models of Neural Auditory Processing," in *Proceedings of the IEEE ICASSP*, San Diego (IEEE, New York), pp. 36.1.1–36.1.4.
- Miller, G. A., and Nicely, P. E. (1955). "An Analysis of Perceptual Confusions Among Some English Consonants," J. Acoust. Soc. Am. **27**, 338–352.
- Moore, B. C. (1973). "Frequency Difference Limens for Short-Duration Tones," J. Acoust. Soc. Am. **54**, 610–619.
- Moore, B. C., Poon, P. W., Bacon, S., and Glasberg, B. (1987). "The Temporal Course of Masking and the Auditory Filter Shape," J. Acoust. Soc. Am. **81**, 1873–1880.
- Patterson, R. D. (1976). "Auditory Filter Shapes Derived by Noise Stimuli," J. Acoust. Soc. Am. **59**, 640–654.
- 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 the Speech Threshold," J. Acoust. Soc. Am. **72**, 1788–1803.
- 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; Proceedings of the 9th International Symposium on Hearing June 1991*, edited by Y. Cazals and K. Horner (Pergamon, Oxford), pp. 429–446.
- Plomp, R., and Bouman, M. A. (1959). "Relation Between Hearing Threshold and Duration for Tone Pulses," J. Acoust. Soc. Am. **31**, 749–758.
- Schacknow, P. N., and Raab, D. H. (1976). "Noise-Intensity Discrimination: Effects of Bandwidth Conditions and Mode of Masker Presentation," J. Acoust. Soc. Am. **60**, 893–906.
- Scharf, B. (1970). "Critical Bands," in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), Vol. 1.
- Scholl, H. (1962). "Über die Bildung der Hörschwellen und Mithörschwellen von Impulsen," *Acustica* **62**, 91–101.

- Van den Brink, W. A., and Houtgast, T. (1990). "Efficient Across-Frequency Integration in Short-Signal Detection," *J. Acoust. Soc. Am.* **87**, 284–291.
- Viemeister, N. F., and Wakefield, G. H. (1991). "Temporal Integration and Multiple Looks," *J. Acoust. Soc. Am.* **90**, 858–865.
- Walley, A. C., and Carrel, T. D. (1983). "Onset Spectra and Formant Transitions in the Adult's and Child's Perception of Place of Articulation in Stop Consonants," *J. Acoust. Soc. Am.* **73**, 1011–1022.
- Wang, K., and Shamma, S. (1994). "Self-Normalization and Noise-Robustness in Early Auditory Representations," *IEEE Trans. Speech Audio Process.* **2**, 412–435.
- Wright, B., and Dai, H. (1994). "Detection of Unexpected Tones with Short and Long Durations," *J. Acoust. Soc. Am.* **95**, 931–938.
- Zwicker, E. (1965). "Temporal Effects in Simultaneous Masking by White-Noise Bursts," *J. Acoust. Soc. Am.* **37**, 653–663.
- Zwicker, E., and Terhardt, E. (1980). "Analytical Expressions for Critical-Band Rate and Critical Bandwidth as a Function of Frequency," *J. Acoust. Soc. Am.* **68**, 1523–1525.

Spectral weights in level discrimination by preschool children: Synthetic listening conditions

Melodie S. Willihnganz, Mark A. Stellmack,^{a)} Robert A. Lutfi,^{b)}
and Frederic L. Wightman^{c)}

Waisman Center, University of Wisconsin–Madison, Madison, Wisconsin 53705

(Received 26 February 1996; revised 18 November 1996; accepted 19 November 1996)

On most auditory discrimination and detection tasks young children perform more poorly than adults. The current experiment applies a technique which potentially can reveal the extent to which the adult–child performance difference results from suboptimal attentional strategies or simply greater internal noise in the children. In this experiment preschool children and adults were asked to discriminate between complex tones comprised of three random-amplitude sinusoidal components. A trial-by-trial correlational analysis [R. A. Lutfi, *J. Acoust. Soc. Am.* **97**, 1333–1334 (1995)] provided an estimate of the weight listeners placed on the level information from individual spectral components in making the discrimination. The patterns of weights were interpreted as measures of “attentional strategy.” Both children and adults produced reliable patterns of weights. This is an especially important result since measuring a single weighting pattern requires large numbers of trials and hence multiple sessions with the children. While individual weighting patterns were reliable, weighting patterns differed both within and across groups. Moreover, neither the children nor the adults produced weighting patterns that would maximize percent correct in the task. A substantial proportion of the responses from both children and adults could be predicted from their weighting patterns even when performance was near chance. However, differences in overall performance between children and adults could not be accounted for by differences in their weighting functions. © 1997 Acoustical Society of America. [S0001-4966(97)02004-3]

PACS numbers: 43.66.Ba, 43.66.Fe [WJ]

INTRODUCTION

It is well documented that young children perform poorly compared to adults in auditory detection and discrimination tasks. Their detection thresholds are higher and their discrimination jnd's are larger. For example, data suggest that frequency resolution is poorer in young children than in adults (Allen *et al.*, 1989; Jensen and Neff, 1993), as is spectral pattern discrimination (Allen and Wightman, 1992), intensity discrimination (Jensen and Neff, 1993) and duration or gap discrimination (Elfenbein *et al.*, 1993; Irwin *et al.*, 1985; Jensen and Neff, 1993; Wightman *et al.*, 1989; Trehub *et al.*, 1995). Although adult-child differences in these auditory tasks appear to diminish by the early school years (Fior, 1972; Irwin *et al.*, 1985; Jensen and Neff, 1993; Maxon and Hochberg, 1982), the differences for some children and some tasks may persist (e.g., children labeled as having an auditory processing disorder). Persistent auditory skill deficits can be quite serious, as they are sometimes manifested in language (Bloom and Lahey, 1978) and learning disabilities (Katz and Wilde, 1994), reading problems (Sawyer, 1981), and failure in school (Cherry, 1992).

Some complex auditory tasks require listeners to integrate information across frequencies (e.g., Allen and Wightman, 1992) or to ignore irrelevant information at nonsignal frequencies (e.g., Allen and Wightman, 1995). Because of

the importance of audition for normal development, it is important to understand how and why performance on complex auditory tasks such as these is different in young children and adults. Some studies have interpreted poorer performance in children in terms of differences in sensitivity between children and adults. For example, higher auditory thresholds at some frequencies (Schneider *et al.*, 1986), and higher ear-canal resonant frequency (Kruger, 1987) have been suggested as factors in adult-child performance differences. Both of these factors potentially could render some stimulus components inaudible for children but not for adults.

Lower motivation and higher incidence of guessing have also been considered as explanations for children's poorer performance on auditory tasks. To address the motivation issue Trehub *et al.* (1981) compared young children's auditory thresholds with and without reinforcement with the view that the reinforcement condition would produce higher motivation. Their findings suggest that motivation significantly affected threshold estimates, with lower thresholds obtained in conditions in which reinforcement was present. In a computer simulation of possible guessing effects, Wightman and Allen (1992) concluded that a high rate of guessing or forgetting (50%) could account for most, if not all, of their observed adult-child performance differences. Bargones *et al.* (1995) estimated a guessing rate of 30% for infants based upon the upper asymptotes of their psychometric functions. However, Schneider and Trehub's (1992) assessment of infant psychometric functions suggests an average guess-

^{a)}Current address: University of Minnesota, N218 Elliott Hall, 75 East River Rd., Minneapolis, MN 55455.

^{b)}Also Department of Communicative Disorders.

^{c)}Also Department of Psychology.

ing rate of only 14%, which would have little effect on auditory performance.

It also may be possible to account for the observed auditory performance differences between adults and children in terms of the specific ways auditory information is used by the listener, that is, the listener's attentional strategy. For our purpose here "attentional strategy" will refer to the relative extent to which a listener "attends to" or "weights," in the context of some perceptual decision, the component parts or dimensions of a complex sound. For example, attentional strategy might refer to the extent to which a listener pays attention to different frequency regions in a wideband sound or to the loudness as opposed to the duration of a sound. There is some reason to believe that children may have different attentional strategies than adults. For example, given that high-frequency auditory thresholds appear to mature earlier than low-frequency thresholds (Schneider *et al.*, 1986), young children might attend to high-frequency information more than to low-frequency information in performing certain auditory tasks. An attentional strategy that focuses on more audible stimulus components may work well for some tasks (e.g., an analytic listening task that requires selective attention to a single high-frequency component), but may be less effective for other tasks (e.g., a task in which the relevant information is at low frequencies, or a task that requires synthesis of information across frequencies).

Note that the concept of an inappropriate attentional strategy is to be distinguished from the type of inattention that was simulated by Wightman *et al.* (1989). Wightman *et al.* simulated inattention to the stimulus as a whole, which results in random responses and chance performance on a certain proportion of trials. In contrast, when a listener uses an inappropriate weighting strategy, the listener's responses are based on an irrelevant component of the auditory information (or, in some cases, an inappropriate combination of information across frequencies). This does not produce random responses. Rather, responses are correlated with some aspect of the stimulus, albeit one which is irrelevant with respect to the task at hand. An inappropriate listening strategy of this sort will be reflected in the weighting function. It is this latter form of inattention that is the focus of the present experiment.

Procedures have been developed recently to quantify a listener's attentional strategy by measuring the weight individuals place on different elements of a complex auditory stimulus when making a discrimination (Berg, 1989; Richards and Zhu, 1994; Lutfi, 1995). In one such procedure (Lutfi, 1995), the weights are computed from the correlation between the values of stimulus component parameters and the listener's response from trial to trial. For example, Doherty and Lutfi (1996) used this procedure to estimate how normally hearing and hearing-impaired adult listeners weighted individual frequency components in a level discrimination task. Listeners identified which of two multitone complexes had a greater overall level. The levels of components in each complex were randomly and independently varied on each trial. Doherty and Lutfi (1996) reported that normally hearing listeners used a wide variety of attentional strategies, as revealed by the weighting patterns, even though

they all performed equally well in the task. The patterns of weights for individual listeners were also quite reliable across sessions.

The primary goals of the study reported in this article were to demonstrate the feasibility of quantifying the attentional strategies of young children by measuring weighting functions, to assess the reliability of those weighting functions, and to compare the weighting functions with those obtained in comparable conditions from adult listeners. In a companion paper (Stellmack *et al.*, 1997), weighting functions are used to examine adult-child differences in listening strategies used in various analytic listening tasks.

I. METHOD

A. Subjects

The listeners were 9 normally hearing children, 4 females and 5 males, who ranged in age from 4 years 3 months to 5 years 6 months (Mean=4 years 10 months), and 6 normally hearing adults. The children were students of the Wisconsin Early Childhood Program who had received their parents' permission to participate and who demonstrated a willingness to perform the listening task. Teachers reported that none of the children demonstrated behaviors suggestive of developmental delay or learning problems. The children were rewarded at the end of each session with small toys. The adults were University of Wisconsin students who were paid an hourly rate to participate, except for SNB (the second author).

Of the children who received parental permission and were willing to accompany the experimenters, only those children who were able to perform a basic listening task were retained as listeners for the main experiment. In the basic listening task, a 1000-Hz pure tone was presented at a random level with a range of as large as 39.5–84.5 dB sound-pressure level (SPL) during each of two listening intervals. The child was asked to select the "louder" interval. One child was unable to consistently perform this task above the chance level and was omitted from the experiment.

Additional criteria for inclusion in the study included pure-tone thresholds of 15 dB hearing level (HL) (ANSI, 1989) or better in the test ear at octave frequencies from 250 to 4000 Hz. Pure-tone thresholds were obtained from all listeners using a diagnostic audiometer (Grason–Stadler GSI-10), calibrated to ANSI specifications (ANSI, 1989). Assessment of each listener's middle-ear function was needed as middle-ear problems (common in young children) can cause fluctuations in hearing that may increase a listener's detection threshold by as much as 20–30 dB (Kokko, 1974). Therefore, tympanometry was performed immediately prior to each test session using a screening tympanometer (Grason–Stadler, GSI-27A Auto-Tymp), calibrated to ANSI specifications (ANSI, 1987). Peak-compensated static acoustic admittance results were within the 90% normal range (ASHA, 1990) for all listeners.

B. Stimuli

As illustrated in Fig. 1, each trial consisted of the presentation of two stimuli. Each stimulus was made up of three

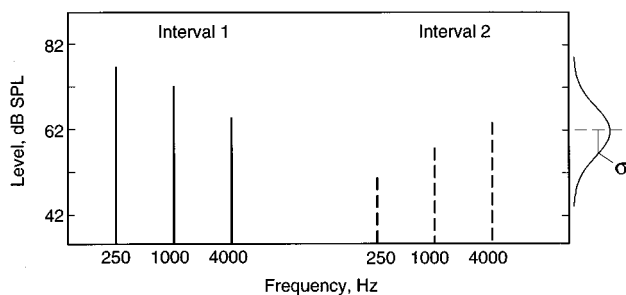


FIG. 1. An example of a single experimental trial is shown. The general procedure involves a two-interval, forced-choice situation in which two multitone complexes comprised of three fixed frequency components, 250, 1000, and 4000 Hz are presented. The listener must select the interval that has the greater overall level. Tone levels are randomly and independently sampled from one normal distribution with a mean level of 62 dB SPL. Task difficulty is determined by the standard deviation of the distribution.

sinusoidal components, 250, 1000, and 4000 Hz. The level of each component was randomly and independently selected for each interval from a Gaussian distribution of levels (in dB) with a mean level of 62 dB SPL, and a standard deviation (σ) which ranged from 1 to 9 dB. Component levels were limited to $\pm 2.5 \sigma$ from the mean. A correct response was defined as selection of the interval in which the sum of the component levels in dB SPL was larger (which we refer to as having the greater overall level). The multitone complex with the greater overall level appeared randomly in interval 1 or 2.

The σ of the component level distributions determines the variability of overall stimulus level. If the range of overall levels is large, the mean difference in overall level between the two stimuli presented on each trial is also large, simplifying the listener's task (identification of the interval in which overall level is greater). When the range of overall levels is small, the listener's task is more difficult. Thus, the σ used for each trial block determines the difficulty of the task for that block.

In this experiment the σ was selected for each listener to produce a performance level of approximately 70%–80% correct on practice trials. During practice, listeners ran blocks of trials that were identical to those in the actual experiment. The experimenter adjusted the σ between blocks until the listener achieved approximately 70%–80% correct. This was done so that weighting functions could be assessed at levels of performance that were approximately equal for all listeners.

The stimuli were generated digitally (50-kHz sampling rate), and converted to analog signals using an IBM-compatible PC interfaced with 16-bit digital-to-analog converters (DA3-2, Tucker–Davis Technologies). The signals were passed unfiltered through an amplifier (HB5, Tucker–Davis Technologies) to the listener's headphones. The two stimuli on each trial were 300 ms in duration, including 16.7-ms \cos^2 on-off ramps and were separated by 300 ms of silence. They were presented monaurally over TDH-49 headphones in GS-001 supra-aural cushions.

C. Procedure

Listeners were tested individually in a double-walled sound-attenuating chamber (Audio Suttle Corporation, AS-216). Two researchers accompanied the children for the safety of both the child and the experimenters (so that the child was not alone in a sound-attenuating chamber with a single adult). One adult initiated stimulus presentations when the child appeared quiet, attentive, and ready for the next trial. The second adult merely sat and observed. The adult listeners either sat alone and controlled stimulus presentations, or they were accompanied by a researcher who controlled the stimulus presentations. (For some of the adult listeners, rather than instructing them in how to start the computer programs and enter the necessary information, the experimenter performed these tasks for them.) Listeners sat in front of a 17-in. touch-sensitive color monitor (Energraphics Intellitouch) that was used to provide visual cues to the occurrence of trial intervals and to provide feedback.

D. Experimental task

An important requirement of developmental research is that the experimental task be applicable over a wide age range (Schneider and Trehub, 1992). The experimental task used here is simple, requires limited linguistic sophistication, and is generally considered to be within the cognitive abilities of normally developing children as young as 3½ years old (Boys Town NICD, 1983; Wadsworth, 1989). Specifically, children (and adults) were asked to "Listen to the two sounds and point to the one that is loud(er)." Two sounds were presented in a standard two-interval, forced-choice (2IFC) paradigm involving multiple blocks of approximately 20–30 trials (about 5 min per block for children). Stimulus intervals were cued by the presentation of box symbols on the touch-screen computer monitor. Listeners responded by touching the symbol that appeared on the monitor coincident with the presentation of the stimulus that was perceived to be louder.

Typically, 100–200 trials per session were obtained from children until a set of 500 trials was obtained, while 500 trials were obtained from adults in a single session. A second set of 500 trials was obtained from each subject at least one week later. Additionally, four children produced a third set of 500 trials which were obtained in a single session. The four children were selected on the basis of their willingness to participate for the amount of time required to collect 500 trials in one day.

E. Feedback

Feedback was provided after each response using one of three formats per block of trials—a picture puzzle, a car race, or a target game. For the picture puzzle format, whenever a correct response was made a portion of a digitized picture appeared on the monitor. For an incorrect response, the box symbol corresponding to the correct stimulus brightened and flashed three times. For the race car format, whenever a correct response was made, a computer-generated picture of a race car designated as the listener's car advanced toward a finish line. For an incorrect response, a competitor's race car advanced. For the target game format, a computer-generated

picture of 25 darts and a target appeared on the screen. Correct responses caused a dart to be projected to the center of the target, whereas incorrect responses caused a dart to be projected to the right or left of the target.

F. Correlational method of computing weights

For the data reported below, the relative weight given to the level information of each component was estimated by computing the correlation between the level change (across intervals of a trial) of each component and the listener's binary responses (Lutfi, 1995; Richards and Zhu, 1994). We assume that the observer's binary response is based on the value of a decision variable D , which is the weighted sum of the level differences between intervals of the components, as follows:

$$D = \sum_{i=1}^N w_i \Delta L_i + \epsilon, \quad (1)$$

where w_i is the observer weight for the i th component, ΔL_i is the level change across intervals of the i th component, and ϵ is additive error encompassing all variability unaccounted for by the weighted level differences. Responses are assumed to be based upon the value of D such that listeners respond

$$\begin{aligned} \text{"1"} & \text{ if } D > 0; \\ \text{"2"} & \text{ if } D < 0; \\ \text{"1"} \text{ or "2"} & \text{ randomly if } D = 0. \end{aligned} \quad (2)$$

The relative weights are then given by

$$w_i = r_{R(\Delta L_i)} \frac{\sigma_R}{\sigma_{\Delta L_i}}, \quad (3)$$

where R is the correlation between the listener's response and the level change of the i th component, σ_R is the standard deviation of the listener's binary response, and $\sigma_{\Delta L_i}$ is the standard deviation of the distribution of level changes across intervals of the i th component. In the present experiment, the levels for the three components were selected from distributions with equal standard deviations. Because Eq. (3) merely scales the correlation coefficients by a common factor when $\sigma_{\Delta L_i}$ is equal for all i (i.e., for all components), it was not necessary to apply Eq. (3) to the correlation coefficients in the present experiment. However, to facilitate comparisons, the correlation coefficients were normalized such that their magnitudes sum to 1.00.²

The value of the noise term (ϵ) in Eq. (1) cannot be computed directly. As an estimate of the relative influence of the noise term, the percentage of responses that can be predicted by the weights was computed. In order to do so, the computed weights were inserted into Eq. (1) and ϵ was set to zero. Simulated responses were obtained by inserting into Eq. (1) the actual level changes that were presented to the listener on each trial and using Eq. (2) to generate a response. The percentage of simulated responses that agreed with the listener's actual responses (the percent of responses predicted by the weights) was then computed. This value estimates the relative contribution of ϵ in the decision pro-

cess. If ϵ is zero, then 100% of the responses can be predicted by the weights. As ϵ becomes large, the percent of responses predicted by the weights will drop to a lower limit of 50%.

An estimate of the 95% confidence interval was constructed about each weight to test its significance. The confidence intervals were used to test whether the weights were significantly different from 0.00 and to test for differences between weights within listeners in order to determine if weights were statistically reliable across experimental sessions. Although the standard error decreases as the correlation coefficient increases in magnitude, for convenience we use $1/\sqrt{n}$ (the standard error of σ around 0.00, n = number of trials) as the standard error in constructing confidence intervals about each correlation coefficient. As a result, our test for significant differences between weights is in most cases more conservative than the 0.05 level.

A measure of the efficiency of the listener's attentional strategy was calculated from the root-mean-square rms deviation of the listener weights from the ideal. Because in this experiment each component provides an equal amount of information about the task, the ideal listening strategy is to weight all components equally (i.e., $w_1 = w_2 = w_3 = 0.33$). Thus, the rms deviation of the weights about the ideal for each weighting function is simply the rms of the normalized weights about 0.33. The smaller this rms value, the closer the observed weighting strategy is to the ideal, that is, the more efficient the weighting strategy.

II. RESULTS

A. Weighting functions

Weighting functions, consisting of the estimated weights for each of the three components of the stimulus, define the attentional strategy for each listener. Figures 2 and 3 show the weighting functions obtained in this study from the children and adults. The weighting functions in each figure are based on approximately 500 trials collected at a fixed standard deviation level. Each of the first two weighting functions obtained from the children represent trials collected over several days, whereas the third weighting function represents trials collected on one day.

In the present experiment, an ideal observer would weight all three components equally because all three components are equally informative. In terms of the normalized weights shown in Figs. 2 and 3, an ideal observer would give weights of 0.33 to all three components. Note that in most cases the actual weighting functions for both children and adults vary from the ideal to various extents. When considered as groups, the children and the adults vary from the ideal roughly to the same extent.

B. Reliability of the weighting functions

Confidence intervals were computed for each listener's weights and are shown as error bars in Figs. 2 and 3 only for weights that were significantly different (0.05 level) across sets of trials. Eight of the children produced weighting functions that were reliable (Fig. 2). The children's third weighting functions, which were collected on one day, are generally

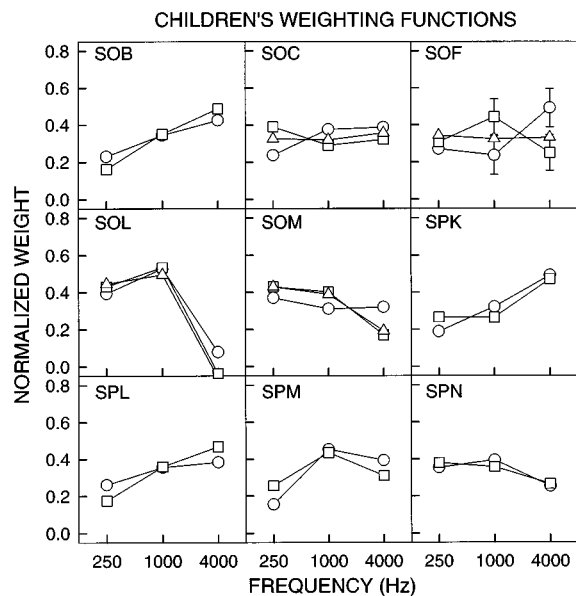


FIG. 2. Children's individual weighting functions are shown, with each panel containing weighting functions for a single child. Each weighting function (1st, 2nd, or 3rd, in the order in which they were obtained) represents one set of approximately 500 trials. Error bars are shown only for weights that are statistically different at the 0.05 level of significance. Circles: 1st weighting function; squares: 2nd weighting function; triangles: 3rd weighting function.

consistent with the first two weighting functions, which were collected over several days. Contrary to what might be expected, a smaller proportion of adults (three of six) produced reliable weighting functions. However, for each adult whose weighting functions were not reliable, the weight of only one component varied over sessions.

C. Child and adult weighting functions

Statistically similar weights (as defined by overlapping 95% confidence intervals), were averaged across all trial sets for each listener so that weighting functions could be com-

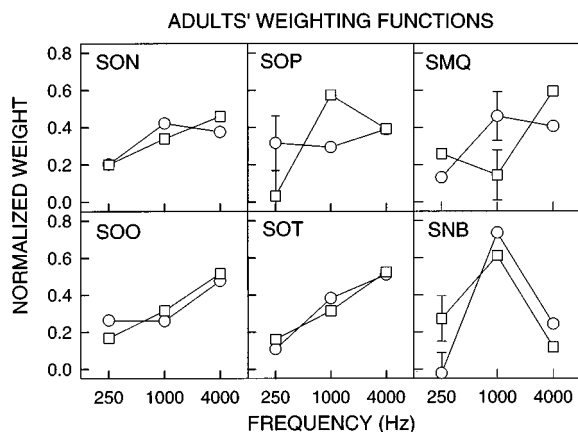


FIG. 3. Adults' individual weighting functions are shown, with each panel containing weighting functions for a single adult. Each weighting function (1st or 2nd, in the order in which they were obtained) represents one set of approximately 500 trials. Error bars are shown only for weights that are statistically different at the 0.05 level of significance. Circles: 1st weighting function; squares: 2nd weighting function.

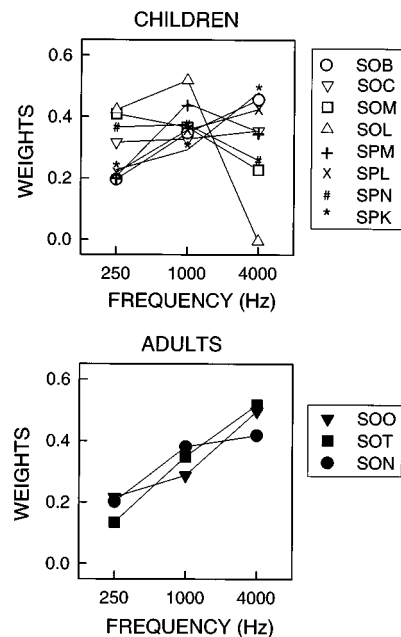


FIG. 4. Individuals' reliable weighting functions were averaged together and the averaged weighting functions are shown here to demonstrate general weighting patterns. Children and adults are represented in separate panels.

pared across listeners. The averaged weighting functions are shown in the top and bottom panels of Fig. 4 for children and adults, respectively. In general, the children's weighting functions are different from one another. In contrast, the adults weighted the low-frequency component least and the high-frequency component most.

D. Performance related to the weighting functions

Figure 5(a)–(c) displays listeners' percent correct scores as a function of rms of the obtained weights about the ideal (0.33), standard deviation, and percent agreement between responses obtained and responses predicted from the weighting model. Standard t tests were used to assess the statistical significance of differences in these indices. Results revealed that only the standard deviation required to achieve 70%–75% correct was significantly different between the child and adult groups ($t = -5.04$, $p < 0.00003$). Children required a larger standard deviation than adults to perform the listening task equally well. No statistically significant differences (at the 0.05 level) were found between the child and adult groups for rms of weights about the ideal, or for percent agreement between obtained and predicted responses.

III. DISCUSSION

The aims of this study were to evaluate the feasibility of obtaining weighting functions from young normally hearing children, to determine the reliability of these functions, and to compare them to weighting functions obtained from adults. A number of points may be made to summarize the results. First, reliable weighting functions were obtained from children as young as 4 years 3 months old, and eight of the nine children and three of the six adults produced weighting functions that were reliable. Second, there were signifi-

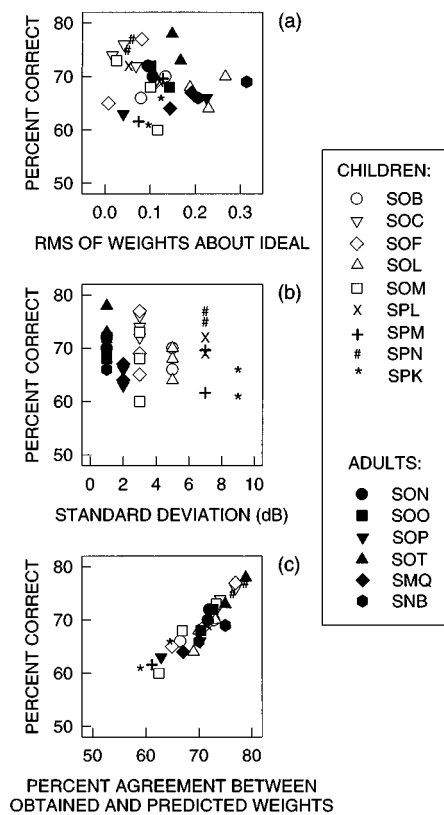


FIG. 5. Performance differences between children and adults are shown. Percent correct is shown as a function of (a) rms of weights about the ideal, (b) standard deviation of the level distribution, and (c) percent agreement between obtained and predicted weights. Note: These graphs contain multiple data points per listener.

cant differences in the standard deviation required to reach criterion performance. Children required a larger standard deviation than adults. The rms error between obtained and ideal weights was similar for both groups of listeners, suggesting that differences in weighting functions were not responsible for the differences in standard deviation required for criterion performance. This conclusion is supported by the fact that the two groups did not differ significantly in percent agreement between obtained trial-by-trial responses and those predicted by the weighting model.

The fact that reliable weighting functions can be obtained from young children is an important result. Weighting functions provide a quantitative measure of an individual's use of information, the "attentional strategy," which gives insight into how a task is performed rather than simply the level of accuracy at which it is performed. In the present experiment the children produced large individual differences in weighting functions in spite of their similar levels of performance (percent correct scores).

The finding that eight of the children produced weighting functions that were reliable is consistent with the observations of Allen and Wightman (1994) with respect to the reliability of psychometric functions. Although children typically show considerable short-term variability in their performance on psychoacoustic tasks (Allen *et al.*, 1989; Wightman *et al.*, 1989), it appears that reliable estimates of weighting functions can be obtained over large numbers of

trials. Further, it is important to note that the children's third weighting functions, which were collected on one day, were generally consistent with their first two weighting functions, which were collected over several days. This is important because it is not always possible to collect many trials from a child in one day. However, while reliability was found within child listeners, variability was evident across these listeners as seen in the diversity of their weighting functions. This diversity in weighting functions is consistent with studies of normally hearing adults (Doherty and Lutfi, 1996).

The weighting functions indicate that neither the children nor the adults in this study used auditory information optimally. In this experiment, optimal listening would be holistic (Smith and Evans, 1989), attending equally to the three components of the stimuli. In fact, most listeners in this study appeared to attend in a differentiated way, in which one or more components are weighted to a greater extent.

Berg (1990) has reported that higher level components (in terms of SPL) are generally given greater weight regardless of their reliability. In the present experiment, the unequal weighting of the three components might be accounted for by differences in hearing level, sensation level (SL), or loudness. For example, given that the mean level of all components was 62 dB SPL, the mean hearing levels of the 250-, 1000-, and 4000-Hz components were 37, 55, and 53 dB HL, respectively (Weissler, 1968). This is consistent with the observation that the majority of listeners gave least weight to the 250-Hz component. However, it remains the case that an ideal observer would give equal weight to all three components in the present task because the model described in Eq. (1) is based upon the level *change* of each component across intervals rather than the absolute level. The model does not predict differences in weights due to differences in component mean levels. Whether differences in HL, SL, or loudness affect weighting functions in synthetic listening tasks is a question to be addressed in future research.

It was not surprising that overall performance was unrelated to the shape of the weighting function, as Doherty and Lutfi (1992, 1996) reported similar findings for their adult normally hearing listeners. Although the children produced more varied weighting functions than the adults, the children's weighting functions appeared no less efficient than adults' on measures of rms from the ideal and percent agreement between obtained and predicted responses. It is true of weighting functions such as those measured in this paper that as the rms deviation of the actual weights about the ideal increases, percent correct decreases. Many different weighting functions have the same rms deviation from the ideal, therefore it is to be expected that performance will not be related to the specific shape of the weighting function.

The relationship between the weighting functions and overall performance may be more obvious with a different listening task. In this study the frequency components provided equal information regarding overall loudness. Thus, listeners could perform this task without attending selectively to any one frequency. Even though optimal performance would be achieved only by attending equally to all three components, better-than-chance performance would be obtained with any pattern of weights. If all of the information

necessary to perform the task was carried by only one of the three frequencies in the complex then performance would be more adversely affected by an inefficient weighting strategy. This topic is pursued in the companion paper (Stellmack *et al.*, 1997).

The single measure to suggest that children were poorer performers than adults on this auditory attention task is the standard deviation required for criterion performance, 70%–80% correct. Children consistently required a larger standard deviation, producing a larger between-interval difference in overall level, to perform at the same percent correct level as adults. In the present experiment, no consistent adult-child differences in weighting functions were observed when the standard deviations of the level distributions were adjusted to hold performance constant across listeners. Generally, it is possible that consistent differences between adults and children would be seen if the level standard deviations were equal for adults and children. However, in the task described here, if σ had been equated for all listeners, it would have resulted in either the children performing too poorly or the adults performing too well for the purposes of this study. If σ is set too small, performance would be at chance, the weights would not be significant, and the weights would account for a chance proportion of responses (50%). If σ is set too large, the listener is likely to perform so well that the weighting functions are ideal or nearly so. The focus of the present study was on the weighting strategies employed by children and adults when performance is imperfect but at levels greater than chance. As a result, σ was adjusted for each listener to bring performance into a criterion range. Future studies will be needed to determine whether weighting functions of children and adults differ when σ is equated across listeners.

Future studies also will investigate the weighting patterns that young children use to perform other types of listening tasks. The companion paper, for example (Stellmack *et al.*, 1997), describes the results of an experiment in which children are required to attend selectively to one component of a complex and to ignore the others. Selective auditory attention is an important skill, and the method described here, involving estimation of attentional weights in a forced-choice discrimination paradigm, offers a potentially useful means of quantifying it.

ACKNOWLEDGMENTS

Supported by research Grant No. HD23333-07 from the National Institutes of Health.

¹Psychophysically, the louder sound may not always correspond to the sound having the greater overall level inasmuch as loudness varies with frequency. However, these instructions were used in order to simplify the description of the experimental task, particularly for the children.

²The *magnitudes* (absolute values) of the correlation coefficients were normalized to sum to 1.00 rather than the actual correlation coefficients in order to deal with situations in which negative correlations were obtained. In these cases, normalizing the actual correlation coefficients occasionally can produce values beyond the range of -1.00 to 1.00 . Normalizing the magnitudes of the correlation coefficients more accurately reflects the relative weight given to each component across conditions. The *sign* of each weight then simply indicates whether the listener's responses were consis-

tent with (+) or opposite to (–) the direction of the level change of that component.

- ANSI (1989). ANSI S3.9-1989, "American National Standard specification for audiometers" (American National Standards Institute, New York).
- ANSI (1987). ANSI S3.39-1987, "American National Standard specification for instruments to measure aural acoustic impedance and admittance (aural acoustic immittance)" (American National Standards Institute, New York).
- American Speech-Language-Hearing Association (1990). "Guidelines for screening for hearing impairments and middle ear disorders," ASHA 32 (Suppl. 2), 17–24.
- Allen, P., and Wightman, F. (1992). "Spectral pattern discrimination by children," J. Speech Hear. Res. 35, 222–233.
- Allen, P., and Wightman, F. (1994). "Psychometric functions for children's detection of tones in noise," J. Speech Hear. Res. 37, 205–215.
- Allen, P., and Wightman, F. (1995). "Effects of signal and masker uncertainty on children's detection," J. Speech Hear. Res. 38, 503–511.
- Allen, P., Wightman, F., Kistler, D., and Dolan, T. (1989). "Frequency resolution in children," J. Speech Hear. Res. 32, 317–322.
- Bargones, J. Y., Werner, L. A., and Marean, G. C. (1995). "Infant psychometric functions for detection: Mechanisms of immature sensitivity," J. Acoust. Soc. Am. 98, 99–111.
- Berg, B. G. (1989). "Analysis of weights in multiple observation tasks," J. Acoust. Soc. Am. 86, 1743–1746.
- Berg, B. G. (1990). "Observer efficiency and weights in a multiple observation task," J. Acoust. Soc. Am. 88, 149–158.
- Bloom, L., and Lahey, M. (1978). "Normal Language Development," in *Language Development and Language Disorders* (Wiley, New York), pp. 67–266.
- Boys Town National Institute for Communication Disorders in Children (1983). "Tina Bangs: Vocabulary Comprehension Scale Summary Sheet for Language Age" (Form 429).
- Cherry, R. (1992). "Screening and evaluation of central auditory processing disorders in young children," in *Central Auditory Processing: A Transdisciplinary View*, edited by Katz, Stecker, and Henderson (Mosby-Year Book, St. Louis, MO), pp. 129–140.
- Doherty, K. A., and Lutfi, R. A. (1992). "Pure-tone infograms measured in listeners with sensorineural hearing loss," *Abstracts of the 15th Midwinter Research Meeting, Association for Research in Otolaryngology*, edited by D. J. Lim, 213, 69.
- Doherty, K. A., and Lutfi, R. A. (1996). "Spectral weights for overall level discrimination in listeners with sensorineural hearing loss," J. Acoust. Soc. Am. 99, 1053–1058.
- Elfenbein, J. L., Small, A. M., and Davis, J. M. (1993). "Developmental patterns of duration discrimination," J. Speech Hear. Res. 36, 842–849.
- Fior, R. (1972). "Physiological maturation of auditory function between 3 and 13 years of age," *Audiology* 11, 317–321.
- Irwin, R., Ball, A., Kay, N., Stillman, J., and Rosser, J. (1985). "The development of auditory temporal acuity in children," *Child Development* 56, 614–620.
- Jensen, J. K., and Neff, D. L. (1993). "Development of basic auditory discrimination in preschool children," *Psychol. Sci.* 4, 104–107.
- Katz, J., and Wilde, L. (1994). "Auditory processing disorders," in *Handbook of Clinical Audiology*, edited by J. Katz (Williams and Wilkins, Baltimore), 4th ed., pp. 490–502.
- Kokko, E. (1974). "Chronic secretory otitis media in children," *Acta Otolaryngol.* 372 (Suppl.), 7–44.
- Kruger, B. (1987). "An update on the external ear resonance in infants and young children," *Ear Hear.* 8, 333–336.
- Lutfi, R. A. (1995). "Correlation coefficients and correlation ratios as estimates of observer weights in multiple-observation tasks," J. Acoust. Soc. Am. 97, 1333–1334.
- Maxon, A. B., and Hochberg, I. (1982). "Development of psychoacoustic behavior: Sensitivity and discrimination," *Ear Hear.* 3, 301–308.
- Richards, V. M., and Zhu, S. (1994). "Relative estimates of combination weights, decision criteria, and internal noise based on correlation coefficients," J. Acoust. Soc. Am. 95, 423–434.
- Sawyer, D. J. (1981). "The relationship between selected auditory abilities and beginning reading achievement," *Language, Speech, and Hearing Services in the Schools* 12, 95–99.
- Schneider, B. A., and Trehub, S. E. (1992). "Sources of developmental change in auditory sensitivity," in *Developmental Psychoacoustics*, edited by L. A. Werner and E. W. Rubel (American Psychological Association, Washington, DC), pp. 3–46.

- Schneider, B. A., Trehub, S. E., Morrongiello, B. A., and Thorpe, L. A. (1986). "Auditory sensitivity in preschool children," *J. Acoust. Soc. Am.* **79**, 447–452.
- Smith, L. B., and Evans, P. (1989). "Similarity, identity, and dimensions: Perceptual classifications in children and adults," in *Object Perception Structure & Process*, edited by Shepp and Ballesteros (Erlbaum, Hillsdale, NJ), pp. 325–356.
- Stellmack, M. A., Willihnganz, M. S., Wightman, F. L., and Lutfi, R. A. (1997). "Spectral weights in level discrimination by preschool children: Analytic listening conditions," *J. Acoust. Soc. Am.* **101**, 2811–2821.
- Trehub, S. E., Schneider, B. A., and Bull, D. (1981). "Effect of reinforcement on infants' performance in an auditory detection task," *Developmental Psychol.* **17**, 872–877.
- Trehub, S. E., Schneider, B. A., and Henderson, J. L. (1995). "Gap detection in infants, children, and adults," *J. Acoust. Soc. Am.* **98**, 2532–2541.
- Wadsworth, B. J. (1989). "The Stage of Preoperational Thought," in *Piaget's Theory of Cognitive and Affective Development* (Longman, New York), pp. 59–93.
- Weissler, P. G. (1968). "International standard reference zero for audiometers," *J. Acoust. Soc. Am.* **44**, 264–275.
- Wightman, F., and Allen, P. (1992). "Individual differences in auditory capability among preschool children," in *Developmental Psychoacoustics*, edited by L. A. Werner and E. W. Rubel (American Psychological Association, Washington, DC), pp. 113–133.
- Wightman, F., Allen, P., Dolan, T., Kistler, D., and Jamieson, D. (1989). "Temporal resolution in children," *Child Development* **60**, 611–624.

Spectral weights in level discrimination by preschool children: Analytic listening conditions

Mark A. Stellmack,^{a)} Melodie S. Willihnganz, Frederic L. Wightman,^{b)} and Robert A. Lutfi^{c)}

Waisman Center, University of Wisconsin–Madison, Madison, Wisconsin 53705

(Received 26 February 1996; revised 18 August 1996; accepted 19 August 1996)

In this series of experiments, adult and child listeners were required to attend to a target tone in the presence of two distracters and to indicate in which of two intervals the target tone had the higher level. The attentional weight listeners placed on each component was estimated by computing the correlation between the level change of each component across intervals and the listener's response. In the first experiment, weights were obtained as a function of the mean level of the distracters (250 and 4000 Hz) for a 1000-Hz target. No consistent differences between the weighting functions of children and adults were observed. In a second experiment, weights were obtained as a function of the harmonic relationship between the distracters (250 and 4000 Hz, or 270 and 4320 Hz) and the 1000-Hz target. No difference was observed between the weighting functions computed with harmonic and inharmonic complexes. In the final experiment, each component of the complex (250, 1000, and 4000 Hz) was identified as the target in separate blocks of trials. In general, adults were able to weight the target component appropriately regardless of its frequency, while children tended to weight all components equally. The results suggest that preschool listeners may exhibit poorer attentional selectivity than adults. © 1997 Acoustical Society of America.

[S0001-4966(97)02104-8]

PACS numbers: 43.66.Ba, 43.66.Fe [WJ]

INTRODUCTION

Children's detection and discrimination thresholds are generally higher than those of adults (Allen and Wightman, 1994; Allen *et al.*, 1989; Elliott and Katz, 1980; Olsho *et al.*, 1988; Schneider *et al.*, 1986). Recent data suggest that these performance differences between children and adults might be a result of children's inability to focus attention on relevant components of a complex signal. For example, Allen and Wightman (1995) examined the effects of uncertainty in signal and masker frequencies on the detection psychometric functions of children and adults. When the signal frequency was 501 or 2818 Hz, chosen randomly on each trial, the majority of the adult listeners showed a decrement in performance relative to the fixed signal case. The decrement could be predicted by assuming that the adult listeners monitored both frequencies simultaneously. Many of the children showed no decrement in performance (but poorer performance with the signal frequency fixed), suggesting that their attention was not focussed at the signal frequency even in the fixed signal case. In a second experiment, in which a distracter tone was presented at a random frequency on each trial, thresholds for most of the adults and some of the children increased by about 10 dB relative to a no distracter condition. For most of the children, thresholds increased by 20 dB or more and, in many cases, the psychometric functions were so flat that thresholds were unmeasurable. As the authors observed, their results were consistent with the no-

tion that the children used a roving attention band when the signal was uncertain or distracters were added. In other words, they used a nonoptimal listening strategy in which attentional weight was given to the irrelevant information of the distracters. Bargones and Werner (1994) measured listening bands in infants and adults, and they concluded that while adults appeared to focus attention on an expected signal frequency, infants either listened through a misplaced or variably-placed attentional filter or through no filter at all. Upon comparing psychometric functions for detection of tones in quiet and in noise, Bargones *et al.* (1995) found that the psychometric functions of infants had reduced upper asymptotes, shallower slopes, and poorer thresholds than those of adults. Bargones *et al.* observed that, as was the case for Allen and Wightman (1995), some of their results might be attributable to differences between the listening strategies employed by infants and adults.

The accompanying paper (Willihnganz *et al.*, 1997) describes a method of measuring weighting functions that quantifies the way in which listeners use level information across frequencies. The relative weight that is given to the level information in each component of a complex indicates the extent to which a listener's responses depend upon the level information at each frequency. The weights are taken as a measure of the relative attention given to each component of a complex. Willihnganz *et al.* showed that although large individual differences exist across listeners, reliable weighting functions could be measured both for adults and for preschool children. The present paper uses that method to obtain weighting functions that quantify the behavior of children and adults in conditions that require selective listening. In these experiments, three tones of different frequency are pre-

^{a)}Current address: University of Minnesota, N218 Elliott Hall, 75 East River Rd., Minneapolis, MN 55455.

^{b)}Also Department of Psychology.

^{c)}Also Department of Communicative Disorders.

sented simultaneously to listeners during two intervals of each trial. Listeners are instructed to listen to one component (the target) while ignoring the other two (the distracters) and to indicate in which interval the target component is louder. The resulting weighting functions describe the relative influence of the level information of the three components on the listener's responses. In each of the following experiments, the optimal listening strategy (that which maximizes percent correct) is one in which greatest weight is given to the target and no weight is given to the distracters. If preschool listeners use a roving attention band, as suggested by Allen and Wightman (1995) and Bargones *et al.* (1995), a nonoptimal weighting function in which the distracter weights are equal to that of the target will result.

In the first experiment, weighting functions are measured for a fixed target frequency (1000 Hz) as a function of the mean level of the distracters (250 and 4000 Hz). In the second experiment, the harmonic relationship of the target and distracters is varied. In separate conditions of the second experiment, the 1000-Hz target is either harmonically related to the distracters (250 and 4000 Hz) or inharmonically related to them (270 and 4320 Hz). In the final experiment, for the same three-component complex (250, 1000, and 4000 Hz) with the same stimulus parameters, the component that is identified as the target is varied in different conditions. These experiments were conducted after the experiments described in the companion paper (Willihnganz *et al.*, 1997) and they were conducted in the order in which they are presented here. Some of the listeners in these experiments also performed as listeners in the companion paper.

I. EXPERIMENT 1: EFFECT OF MEAN DISTRACTER LEVEL

This experiment examined the weighting functions of children and adults in a task that required them to attend to a target component while ignoring two distracters. Weighting functions were obtained for different mean distracter levels [30, 60, and 70 dB sound-pressure level (SPL)], with the target mean level fixed at 70 dB SPL.

A. Methods

1. Subjects

The listeners were 5 normally hearing children, 2 females and 3 males, who ranged in age from 4 years 7 months to 5 years 7 months (Mean=5 years 2 months), and five normal-hearing adults. The children were students of the Waisman Early Childhood Program whose parents gave permission for them to participate in the study. All of the children had participated in previous listening experiments (some in Willihnganz *et al.*, 1997) and were chosen to participate in the present series of experiments on the basis of their willingness and availability to do so. The adult listeners included the first author and four volunteers who were paid to participate in the study.

Additional criteria for inclusion in the study included pure-tone thresholds of 15 dB hearing level (HL) (ANSI, 1989) or better in the test ear at octave frequencies from 250 to 4000 Hz. Pure-tone thresholds were obtained on all subjects using a diagnostic audiometer (Grason–Stadler GSI-

10), calibrated to ANSI specifications (ANSI, 1989). In order to screen the child listeners for middle-ear problems, tympanometry was performed immediately prior to each test session using a screening tympanometer (Grason–Stadler, GSI-27A Auto-Tymp), calibrated to ANSI specifications (ANSI, 1987). Peak-compensated static acoustic admittance results were within the 90% normal range (ASHA, 1990) for all subjects. Stimuli were presented monaurally to the left ear of all listeners except listener SPL, who exhibited a unilateral sensorineural high-frequency hearing loss in the left ear. As a result, stimuli were presented to SPL's right ear.

2. Stimuli

A three-component complex tone (250, 1000, and 4000 Hz) was presented to the listener during the two intervals of each trial. The level of each component (in dB SPL) was selected randomly and independently from a Gaussian distribution with a mean and standard deviation specified by the experimenter. The listeners were instructed to attend to the middle (target) component while ignoring the other two components (distracters), and to indicate in which interval the target component was louder. For the purposes of giving feedback, the "correct" interval was defined as the one in which the level of the target component was higher.

In this paradigm, because the level of the target component in each interval is selected randomly, the level difference for the target between intervals, the "signal" level, in effect, will vary from trial to trial. The signal level is determined by the standard deviation that is selected for the target level distribution. If the standard deviation is relatively large, the mean target level difference across intervals will be large, those level differences will be detectable on a high proportion of trials, and a relatively high percent correct will be obtained. As the standard deviation of the target level distribution is made smaller, the mean target level difference across intervals will become smaller, those level differences will be undetectable on a larger proportion of trials, and performance will drop. Thus, the level of performance is partially determined by the standard deviation of the target level distribution.

The mean of the distribution of target levels was 70 dB SPL. The standard deviation (σ) of the target level distribution (in dB) was chosen by the experimenter such that the listener achieved 70%–80% correct in practice trials when no distracters were present. (The experimenter adjusted the σ of the level distribution in successive blocks of trials until performance fell within the criterion range.) The largest σ chosen for any listener was 7 dB. The mean of the distribution of distracter levels was 30, 60, or 70 dB SPL in separate experimental conditions. The σ of the distracter level distribution was fixed at 3 dB. The target and distracter levels were limited to $\pm 2.5 \sigma$ from the mean in order to limit signal levels to a safe range.

The stimulus presented in each interval was 300 ms in duration, including 16.7-ms \cos^2 on-off ramps. The two intervals were separated by 300 ms of silence. The stimuli were generated digitally at a 50-kHz sampling rate and converted to analog signals on an IBM-compatible PC interfaced with 16-bit digital-to-analog converters (DA3-2, Tucker–

Davis Technologies). The signals were passed unfiltered through an amplifier (HB5, Tucker–Davis Technologies) to the listener's headphone (TDH-49). The headphones were fitted with supra-aural GS-001 ear cushions. Listeners wore earphones over both ears even though stimuli were presented to only one ear.

3. Procedure

Listeners were tested individually in a double-walled sound-attenuating chamber (AS-216, Audio Suttle Corporation). Two adult experimenters sat with the child listeners, with one experimenter initiating stimulus presentation when the child appeared quiet, attentive, and ready for the next trial. The second experimenter was present for security purposes (so that the child and adult experimenter were not alone in the sound-attenuating chamber). The adult listeners sat with an experimenter who controlled stimulus presentations, except for the first author who controlled stimulus presentations for himself.

Stimuli were presented to listeners in the context of several video games on a 17-in. touch-sensitive color monitor (Energraphics Intellitouch). In order to familiarize listeners with the stimuli, the individual components were associated with different characters on the screen. A practice game was presented in which listeners could touch one of the three characters (frog, happy face, and bird) and hear a pure tone at its corresponding frequency (250, 1000, and 4000 Hz, respectively) at a fixed level of 70 dB SPL. While the component sounded, its corresponding character opened its mouth. When the listener touched a box that was also on the screen, all three characters opened their mouths and a complex containing all three components was presented. In this way, listeners could be instructed to “listen to the happy face and ignore the frog and bird.” Listeners were allowed to manipulate the characters for a brief time, until they indicated that they understood the task.

Each experimental block consisted of 20–30 trials. During the trials, two boxes appeared on the screen, one of which would flash during each stimulus interval. The frog and bird, representing the distracters, also appeared on the screen and opened their mouths during each stimulus interval. Listeners indicated which interval contained the louder target component by touching one of the two boxes. After the listener made a response, the interval markers on the screen were replaced briefly by large and small happy faces which marked the correct and incorrect intervals, respectively. The animated characters appeared on the screen with one of the video games described in the accompanying paper (Willihnganz *et al.*, 1997). Correct responses were rewarded with appropriate visual reinforcement in the context of the particular video game that the listener was playing.

4. Computation of weights

Observer weights were computed as described in the accompanying paper (Willihnganz *et al.*, 1997). Essentially, computation of weights is derived from a model in which the observer's binary response is determined by the weighted sum of the level changes in dB of the three components

across the intervals of each trial. The weight for each component is the correlation between the listener's trial-by-trial response and the level change by the value of a decision variable D , which is the weighted sum of the level differences between intervals of the components, as follows

$$D = \sum_{i=1}^N w_i \Delta L_i + \epsilon, \quad (1)$$

where w_i is the observer weight for the i th component, ΔL_i is the level change across intervals of the i th component, and ϵ is additive error encompassing all variability unaccounted for by the weighted level differences. Responses are assumed to be based upon the value of D such that listeners respond

$$\begin{aligned} & \text{“1”} && \text{if } D > 0; \\ & \text{“2”} && \text{if } D < 0, \\ & \text{“1” or “2” randomly} && \text{if } D = 0. \end{aligned} \quad (2)$$

The relative weights are then given by

$$w_i = r_{R(\Delta L_i)} \frac{\sigma_R}{\sigma_{\Delta L_i}}, \quad (3)$$

where the R is the correlation between the listener's response and the level change of the i th component, σ_R is the standard deviation of the listener's binary response, and $\sigma_{\Delta L_i}$ is the standard deviation of the distribution of level changes across intervals of the i th component. In the present experiment, the levels for the three components in most cases are selected from distributions with unequal standard deviations. As a result, Eq. (3) was used to compute the actual weights from the correlation coefficients. The weights then were normalized such that the magnitudes of the weights summed to 1.0.¹

As in the accompanying paper (Willihnganz *et al.*, 1997) the percentage of responses that can be predicted by the weights was also computed as an estimate of the relative influence of the noise term [ϵ in Eq. (1)]. In order to do so, the computed weights were inserted into Eq. (1) and ϵ was set to zero. Simulated responses were obtained by inserting into Eq. (1) the actual level changes that were presented to the listener in each trial and using Eq. (2) to generate a response. The percentage of simulated responses that agreed with the listener's actual responses (the percent of responses predicted by the weights) was then computed. The smaller this value is, the larger the influence of ϵ in the decision process.

Given that the standard error is $1/\sqrt{n}$ for a correlation coefficient of 0.00 (where n is the number of trials), an estimate of the 95% confidence interval was constructed about 0.00 to test the significance of each weight. In the figures, filled symbols represent weights that are significantly different from 0.00 ($p < 0.05$).

As in the accompanying paper (Willihnganz *et al.*, 1997), the efficiency of the weighting strategy represented by each weighting function is assessed by computing the root-mean-square (rms) deviation of the observed weights from the ideal. In the experiments described below, correct responses are based upon level changes in only one component. Thus the ideal weighting strategy in each case is that in

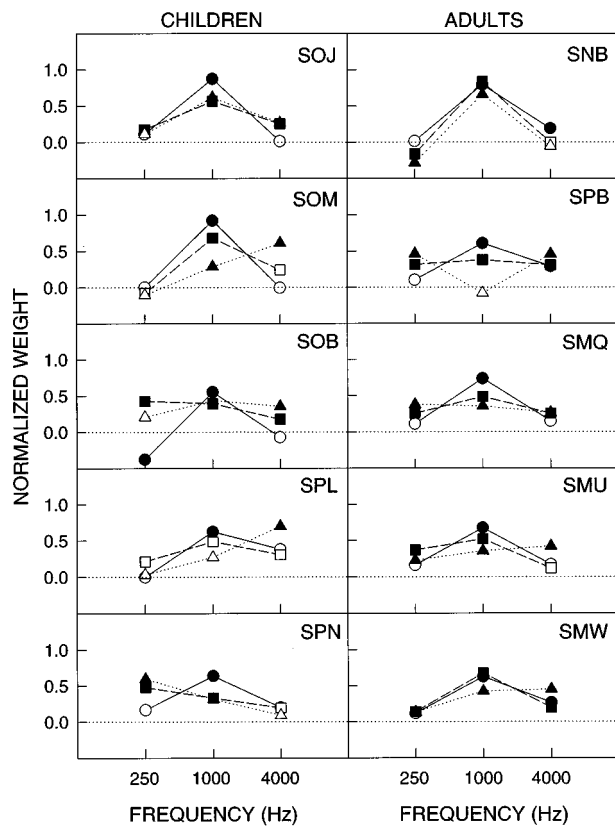


FIG. 1. Results of experiment 1, weighting functions for five children and five adults. Weights are plotted as a function of component frequency. Each symbol represents a different mean distracter level: circles: 30 dB SPL; squares: 60 dB SPL; triangles: 70 dB SPL. Filled symbols represent weights that are significantly different from zero ($p < 0.05$). Mean level of the target (1000-Hz component) was 70 dB SPL in all conditions. Each panel shows results for a single listener.

which the target level is given a weight of 1.00 and the distracter levels are given weights of 0.00 (e.g., $w_1 = 0.00$, $w_2 = 1.00$, $w_3 = 0.00$ when the center component, $i = 2$, is identified as the target). The rms deviation of the weights about the ideal then is the rms of the measured weighting function about the ideal weighting function for the particular experimental condition. Small rms values are associated with weighting strategies that more closely approximate the ideal. If a listener adopts the ideal weighting strategy, this rms value will be 0.00.

B. Results and discussion

The results for the five preschool and five adult listeners are shown in Fig. 1. In these figures, the normalized weights are plotted for each component with each symbol type representing a different mean distracter level. Filled symbols represent weights that are significantly different from 0.00 ($p < 0.05$). The weight for each component was obtained from Eq. (3) and the weights for the three components were normalized such that their magnitudes summed to 1.0. Each weighting function is based upon at least 100 trials. Table I shows the percent correct and percent of responses predicted by the weights for each subject in experiment 1. The standard deviation of the target level distribution for each subject is also shown in Table I.

TABLE I. Target level standard deviation, percent correct, and the percent of responses predicted by the weights for each subject in each condition of experiment 1.

	Target σ (dB)	Distracter mean level (dB SPL)	% Correct	% Responses predicted by weights
Children				
SOJ	2	30	86	88
		60	80	83
		70	78	84
SOM	4	30	86	85
		60	70	75
		70	57	72
SOB	5	30	84	87
		60	78	77
		70	74	80
SPL	7	30	74	74
		60	59	58
		70	63	67
SPN	7	30	91	91
		60	71	80
		70	71	79
Adults				
SNB	2	30	81	82
		60	80	86
		70	77	88
SPB	2	30	74	83
		60	66	77
		70	47	71
SMQ	3	30	73	73
		60	61	68
		70	65	72
SMU	3	30	74	81
		60	70	75
		70	64	77
SMW	2	30	72	78
		60	76	85
		70	69	81

Some general observations can be made regarding the results from both children and adults. When the mean distracter level is low (30 dB SPL), greatest weight is given to the middle (target) component, as is appropriate. As the mean distracter level increases, increasing weight is given to the distracters. When the target and distracter mean levels are equal (70 dB SPL), some listeners give greatest weight to one or more of the distracters, e.g., children SOM and SPL, and adult SPB. For both children and adults, a substantial proportion of responses were predicted by the weights even when performance (in terms of percent correct) was near chance.

As in the companion study (Willihnganz *et al.*, 1997), there are no consistent differences in the form of the weighting functions of children and adults. This fact becomes more apparent when one considers the efficiency of the weighting strategies adopted by children and adults. Figure 2 shows percent correct as a function of the rms deviation of the weights about the ideal. No difference between children (open symbols) and adults (filled symbols) is evident in terms of this measure of weighting efficiency. When considered as groups, children and adults cannot be distinguished

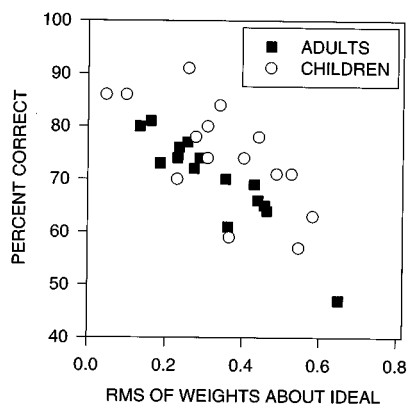


FIG. 2. Percent correct as a function of the rms of the weights about the ideal weights for all conditions of experiment 1. Open symbols represent children's data, filled symbols represent adults' data. The upper limit of the abscissa is the largest rms value that can be obtained for a non-negative target weight (when one of the distracters is given a normalized weight of 1.00).

on the basis of the rms deviation of their weights about the ideal weights.

The main difference between children and adults in this experiment is the fact that most children required a larger standard deviation for the target level distribution (shown in Table I) in order to reach the criterion level of performance (one child, SOJ, performed at a level comparable to that of the adults). Large individual differences such as these are commonly observed when comparing the listening behavior of children and adults (e.g., Allen and Wightman, 1994).

These data show that the weighting functions provide a description of behavior that is not available in performance measures such as percent correct. In many cases, the weights predict a relatively large proportion of responses even when performance in terms of percent correct is near chance (e.g., SOM, 70 dB SPL distracter mean; SPB, 70 dB SPL distracter mean). In these cases, poor performance can be partially attributed to inappropriate weighting of level information across frequencies.

II. EXPERIMENT 2: EFFECT OF HARMONIC RELATIONSHIP BETWEEN COMPONENTS

In experiment 1, as the distracter mean level was increased, weighting functions became less optimal. When the target and distracter mean levels were equal, every listener gave significant weight to at least one distracter component, and all but one listener (SPL) gave significant weight to at least two components. The stimuli were harmonic three-component complexes, and the listeners were asked to attend to one component while ignoring the others. Harmonically related components tend to be perceptually fused while inharmonically related components can be more easily segregated (Bregman, 1990; Yost, 1991). Consequently, one might expect that listeners would weight level information in a target component more appropriately if that component is inharmonically related to concurrent distracters. This experiment examines the effect of the harmonic relationship between components on performance and weighting functions in the level discrimination task of experiment 1.

A. Methods

Experiment 2 was the same as experiment 1 with the following exceptions. (1) the harmonic relationship between the components was varied between blocks of trials, and (2) the distracter mean level was always 70 dB SPL.

1. Subjects

The listeners were the same five children and five adults as in experiment 1.

2. Stimuli

A three-component complex was presented to the listener during the two intervals of each trial. The three components were either harmonically related (250, 1000, and 4000 Hz) or inharmonically related (270, 1000, and 4320 Hz). In both cases, listeners were instructed to indicate in which interval the 1000-Hz (target) component was louder. The mean level of the target and distracters was 70 dB SPL. (The mean level of the distracters was not varied as it was in experiment 1.) The standard deviations of the target and distracter levels for each listener were the same as in experiment 1, as was the equipment.

3. Procedure

The procedure was the same as in experiment 1. Stimuli were presented in the context of video games, with the individual stimulus components associated with different animated characters. Blocks of harmonic and inharmonic trials (approximately 20–30 trials each) were alternated to control for practice effects. Thus, it was necessary to repeat the harmonic condition of experiment 1. Before each block of trials, listeners were allowed to manipulate the characters in the practice game using the harmonic or inharmonic components in order to familiarize them with the stimuli that would be used in the experimental block that followed.

B. Results and discussion

The results are shown in Fig. 3. The circles represent the weighting functions for the harmonically related components and the squares represent weighting functions for inharmonic stimuli. Table II shows percent correct and the percentage of responses predicted by the weights.

There are no significant differences for any listener between any component weight in the harmonic case and the corresponding component weight in the inharmonic case. Nor are there any consistent differences between the percent correct values or the percentage of responses predicted by the weights in the two conditions. Some listeners displayed better performance (in terms of percent correct) with the inharmonic stimuli (e.g., adult SMU) while other listeners performed better with the harmonic stimuli (e.g., child SPN).

Figure 4 shows percent correct as a function of the rms deviation of the weights about the ideal for all listeners in this experiment. There appears to be little difference between children and adults in terms of the efficiency of their weighting strategies. In comparing Figs. 1 and 3, it can be seen that there was little change in the weighting patterns used by individual listeners in experiments 1 and 2.

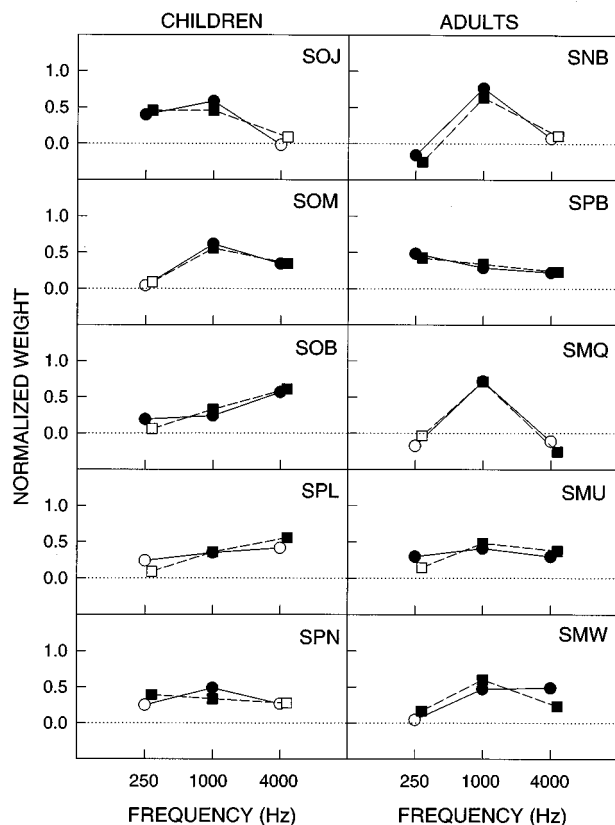


FIG. 3. Results of experiment 2. Weights are plotted as a function of component frequency. Each symbol represents a different condition: circles, harmonic stimuli; squares, inharmonic stimuli. Filled symbols represent weights that are significantly different from zero ($p < 0.05$). Mean target and distracter levels were 70 dB SPL. The middle component was the target in both conditions. Each panel shows results for a single listener.

Given that no consistent differences were found between the results for harmonic and inharmonic stimuli, we chose to use harmonic stimuli in experiment 3.

III. EXPERIMENT 3: EFFECT OF TARGET FREQUENCY

Of importance in processing complex auditory stimuli is the ability to attend to relevant information while ignoring simultaneous interfering information. Furthermore, a listener must have the ability to attend to different spectral portions of the auditory stimulus because useful information might be present in any frequency region. For example, speech discrimination often requires listeners to make discriminations on the basis of information in one region of the spectrum while ignoring information at other frequencies. The specific discrimination to be made determines which spectral region contains relevant information. The following experiment compares the ability of children and adults to attend to different spectral regions by examining the weighting functions of preschool and adult listeners when the component that is identified as the target varies across conditions.

Three-component stimuli are presented to listeners and, in separate conditions, the low-, middle-, or high-frequency component is defined as the target. As in the previous experiments, the listener is instructed to identify in which of two

TABLE II. Percent correct and the percent of responses predicted by the weights for each subject in each condition of experiment 2.

		% Correct	% Responses predicted by weights
Children			
SOJ	Inharmonic	65	75
	Harmonic	67	69
SOM	Inharmonic	71	73
	Harmonic	74	72
SOB	Inharmonic	71	75
	Harmonic	65	72
SPL	Inharmonic	63	58
	Harmonic	68	69
SPN	Inharmonic	71	70
	Harmonic	80	85
Adults			
SNB	Inharmonic	76	86
	Harmonic	71	73
SPB	Inharmonic	64	77
	Harmonic	64	78
SMQ	Inharmonic	78	82
	Harmonic	73	73
SMU	Inharmonic	74	75
	Harmonic	60	59
SMW	Inharmonic	71	81
	Harmonic	69	82

intervals the target is louder while ignoring the remaining components (distracters).

A. Methods

Experiment 3 was the same as experiment 1 with the following exceptions (1) target frequency was varied across conditions, (2) distracter mean level was 30 or 70 dB SPL, as opposed to 30, 60, or 70 dB SPL in experiment 1, and (3) the standard deviation of the distracter level distribution was equal to that of the target level distribution, while in experiment 1, it was fixed at 3 dB.

1. Subjects

The listeners were the same five children and five adults as in experiments 1 and 2.

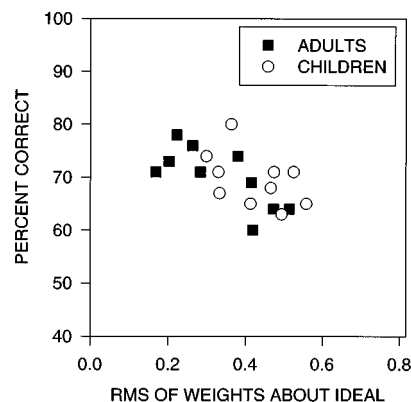


FIG. 4. Percent correct as a function of the rms of the weights about the ideal weights for all conditions of experiment 2. Open symbols represent children's data, filled symbols represent adults' data.

2. Stimuli

A three-component complex (250, 1000, and 4000 Hz) was presented to the listener during the two intervals of each trial. One of the components was identified as the target (see Sec. III A 3 below) and listeners were instructed to indicate in which interval the target component was louder. As in experiment 1, component levels were selected randomly and independently for each interval from a Gaussian distribution. The mean of the level distribution for the target was 70 dB SPL, and that of the distracters was either 30 or 70 dB SPL. The standard deviation of the target level distribution for each listener was the same as that used in experiment 1. The standard deviation of the distracter level distribution was equal to that of the target level distribution. As a result, the stimulus parameters remained constant for each listener regardless of which component was the target. All remaining stimulus parameters and the equipment were the same as those used in experiment 1.

3. Procedure

The procedure was largely the same as in experiment 1. Stimuli were presented in the context of video games, with the individual stimulus components associated with different animated characters. Whenever a different component was used as the target, listeners were allowed to manipulate the characters in the practice game as a way of familiarizing them with the experimental task. For example, when the 4000-Hz component was to be the target, listeners were told to "listen to the bird while ignoring the happy face and frog." At least 100 trials were run in each condition. All of the trials with a particular target were presented before the target was changed to another component. All of the trials with the distracters at a mean of 30 dB SPL were presented before the mean distracter level was raised to 70 dB SPL.

B. Results and discussion

The results for the child listeners are shown in Fig. 5 and those for the adults are shown in Fig. 6. The three panels in each row represent data for the low-, middle-, and high-frequency target conditions, respectively, for a given listener. Within each panel, circles represent weights for the 30 dB SPL mean distracter level, and the squares represent weights for the 70 dB SPL mean distracter level. Weights that are significantly different from 0.00 ($p < 0.05$) are represented by filled symbols. Table III shows percent correct and percentage of responses predicted by the weights for the children, while Table IV shows these same data for the adults.

Figure 7 shows average weighting functions, in which each weight is the average across all listeners for the children and adults in each condition. Within each panel, the average percent correct and average percentage of responses predicted by the weights for all child and adult listeners are shown.

At the lower distracter mean level, the results are consistent with those of experiment 1. On average, both children and adults give greatest weight to the target component when the mean level of the distracters is 30 dB SPL (circles in Fig. 7). However, unlike experiment 1, differences between the

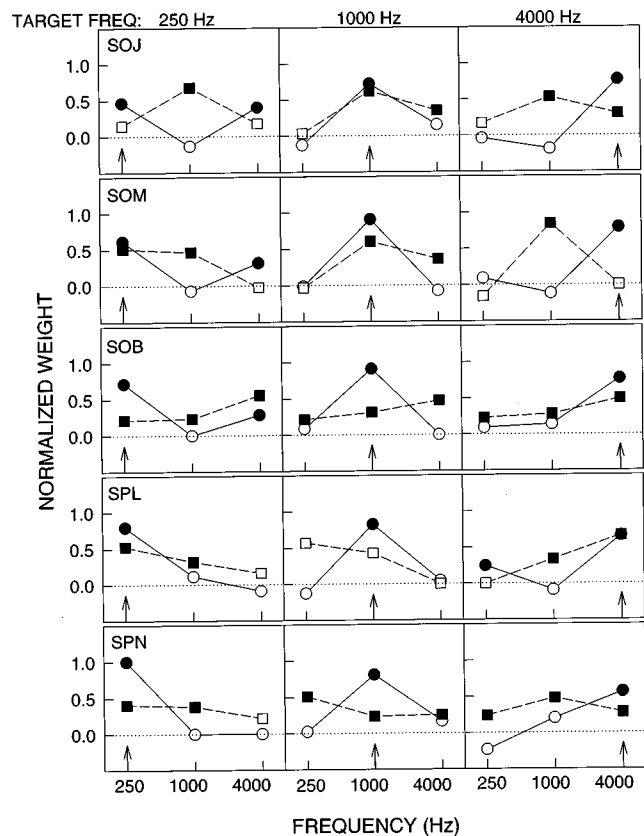


FIG. 5. Results of experiment 3, weighting functions for the preschool listeners. Each row of panels shows data from a single listener. Each column of panels represents conditions in which a different component was the target as shown at the top of the figure and as indicated by the vertical arrow at the bottom of each panel. Within each panel, the different symbol types represent different mean distracter levels: circles: 30 dB SPL, squares: 70 dB SPL. The target mean level was 70 dB SPL. Filled symbols represent weights that are significantly different from zero ($p < 0.05$).

behavior of children and adults become evident when the target and distracter mean levels are equal. The average weighting functions for children and adults differ when the mean distracter level is equal to that of the target (70 dB SPL, squares in Fig. 7). The adults, on average, continue to give greatest weight to the target component, although slightly less weight than when the mean distracter level is 30 dB SPL. The average weighting functions for the children suggest that nearly equal weight was given to all three components when the component mean levels were equal regardless of which component was the target. It appears that the children were less able than the adults to selectively listen to the target component when the distracters were at a comparable level.

The differences between the weighting strategies of the children and adults are evident in the rms of the weights about the ideal for the two different distracter mean levels. Figure 8 shows percent correct as a function of the rms deviation of the weights about the ideal for all conditions of this experiment in which the distracter mean level was 30 dB SPL. As in the previous experiments, there is little difference between children and adults in terms of the efficiency measure. However, it can be seen that adults tended to adopt a more efficient weighting strategy than children when the dis-

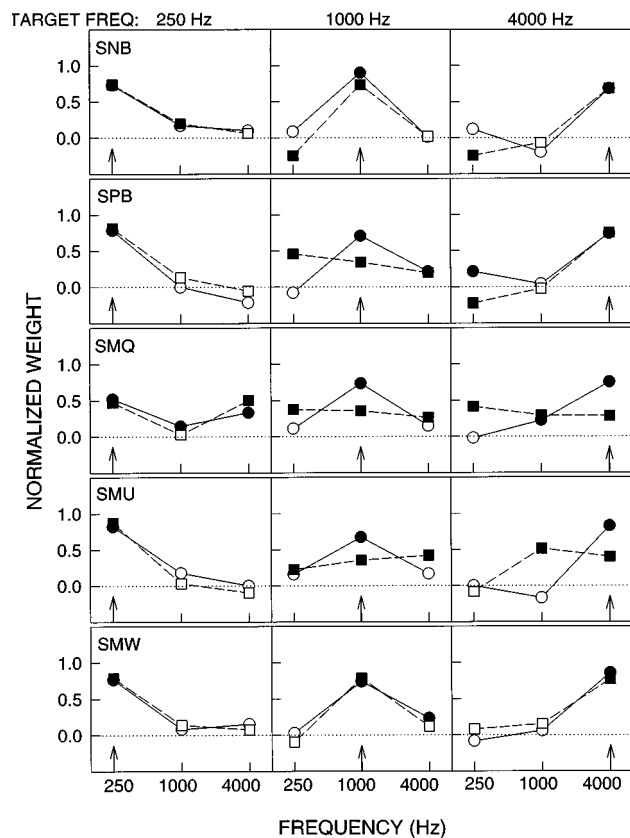


FIG. 6. Results of experiment 3, weighting functions for the adult listeners in the same format as Fig. 5. Within each panel, the different symbol types represent different mean distracter levels: circles: 30 dB SPL, squares: 70 dB SPL. The target mean level was 70 dB SPL. Filled symbols represent weights that are significantly different from zero ($p < 0.05$).

tracter mean level was 70 dB SPL (Fig. 9). The children's rms values in Fig. 9 generally are larger than those of the adults, indicating less efficient weighting strategies.

As in the previous experiments, substantial individual differences exist in these data. With respect to the children's weighting functions (Fig. 5), those of SOJ and SOM can be distinguished from those of the remaining children. SOJ and SOM participated in a great number of prior experiments in which the center component was always the target. When the mean distracter level was 70 dB SPL, these two listeners gave greatest weight to the middle component independent of which component was the target. The remaining three children (SOB, SPL, and SPN) gave roughly equal weight to all three components when the distracters had the same mean level as the target.

The differences between the data of SOJ and SOM and those of the other children may be due to differences in experience. For approximately one month prior to collection of data in the present experiments, SOJ and SOM participated in a number of pilot experiments in which a 1000-Hz component was always the target and 250- and 4000-Hz components were always distracters. Approximately 2000–3000 trials were run in these pilot conditions. During this time, none of the other child listeners provided data or listened to stimuli. From the data shown in Fig. 5, one might speculate that all of the children became generally confused when the

TABLE III. Percent correct and the percent of responses predicted by the weights for each preschool listener in each condition of experiment 3.

	Target Freq.	Distracter mean level (dB SPL)	% Correct	% Responses predicted by weights
SOJ	250	30	60	59
		70	52	74
	1000	30	76	77
		70	75	82
	4000	30	75	76
SOM	250	30	66	66
		70	63	67
	1000	30	80	77
		70	69	74
	4000	30	66	70
SOB	250	30	79	83
		70	58	75
	1000	30	77	80
		70	66	81
	4000	30	78	82
SPL	250	30	81	81
		70	66	70
	1000	30	63	65
		70	49	62
	4000	30	82	79
SPN	250	30	81	81
		70	64	70
	1000	30	80	80
		70	55	71
	4000	30	65	66
		70	64	88

experimental task required them to attend to different components in different conditions. The highly experienced child listeners (SOJ and SOM) appeared to revert to a previously overlearned listening strategy while the remaining children weighted all of the information equally.

The adult listeners also showed large individual differences in their weighting functions (Fig. 6). Some adults (SNB and SMW) gave greatest weight to the target component when the mean component levels were equal. The remaining adult listeners displayed various weighting strategies which approached the ideal (greatest weight given to the target, no weight given to the distracters) to different degrees.

While the average weighting functions for the adults suggested that greatest weight was given to the target component, adult weighting functions that resembled those of the children (e.g., SMQ) were observed. The general differences between adults and children that were described must be qualified by the fact that large individual variation was present.

IV. GENERAL DISCUSSION

A. Relevance of weighting functions

As described in the companion paper (Willihnganz *et al.*, 1997), the weighting functions provide information

TABLE IV. Percent correct and the percent of responses predicted by the weights for each adult listener in each condition of experiment 3.

	Target freq.	Distracter mean level (dB SPL)	% Correct	% Responses predicted by weights
SNB	250	30	81	83
		70	84	81
	1000	30	82	82
		70	78	82
	4000	30	86	88
		70	85	83
SPB	250	30	75	74
		70	72	74
	1000	30	74	79
		70	65	78
	4000	30	79	77
		70	84	84
SMQ	250	30	71	83
		70	60	64
	1000	30	73	73
		70	65	72
	4000	30	86	83
		70	59	75
SMU	250	30	71	72
		70	70	69
	1000	30	74	81
		70	64	77
	4000	30	74	74
		70	65	73
SMW	250	30	81	82
		70	72	74
	1000	30	82	88
		70	66	67
	4000	30	82	83
		70	76	76

about a listener's behavior that is not available in simple performance measures such as percent correct. In the present experiments, inappropriate weighting of level information in the various components results in poor performance in terms of percent correct. Even when low percent correct scores are obtained, the weighting functions provide information about the listener's behavior, as can be seen when one considers the percent of responses predicted by the weights. Weighting functions also provide a means of describing one form of individual differences in listening behavior.

Figure 10 shows the percent of responses predicted by the weights as a function of percent correct for all data in the present series of experiments. The weights generally predict a substantial percentage of the listener's responses even when performance in terms of percent correct approaches chance. Thus, the weighting functions can show the contribution of the listener's weighting strategy to the observed percent correct scores.

B. Behavioral differences between children and adults

The present series of experiments examined selective listening behavior in children and adults in terms of the relative weight given to the information in a target component

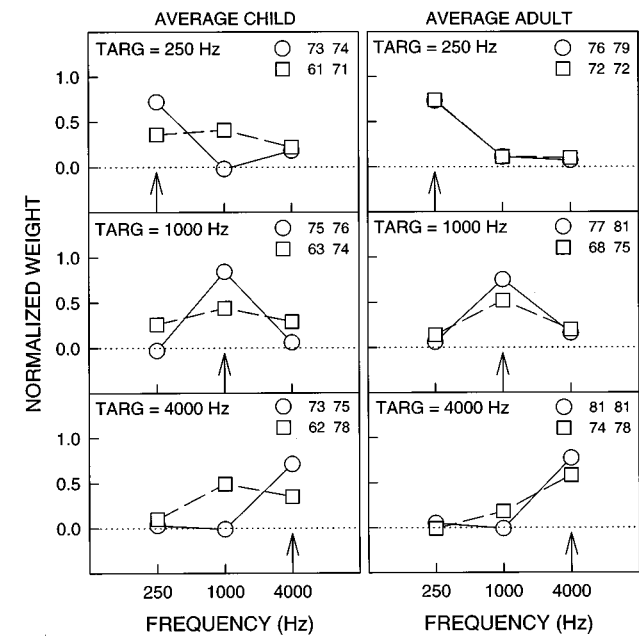


FIG. 7. Average weighting functions measured for children and adults in experiment 3. The leftmost column shows weighting functions averaged across the five children, the rightmost column shows average weighting functions for the five adults. Each row of panels shows average weighting functions for a different target frequency, as indicated in the panels. The different symbol types represent different mean distracter levels: circles: 30 dB SPL; squares: 70 dB SPL. The numbers shown in the upper portion of each panel indicate, respectively, the average percent correct and average percent of responses predicted by the weights for each symbol type.

and distracter components. In experiment 1, when the levels of the distracters had a standard deviation of 3 dB, children required a larger standard deviation on the level of the target in order to achieve percent correct scores that were comparable to those of the adults. However, there was no consistent difference between the way in which children and adults weighted the level information of the individual components. In experiment 2, no difference was observed between the weighting of level information of components in harmonic and inharmonic complexes for either children or adults. In experiment 3, in which the target and distracter levels had

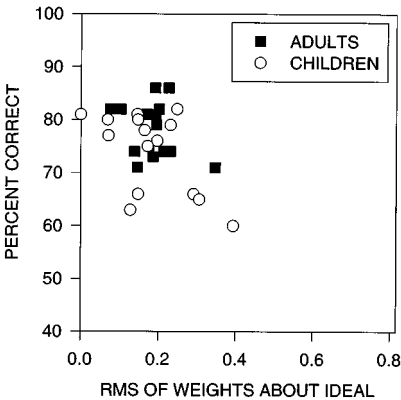


FIG. 8. Percent correct as a function of the rms of the weights about the ideal weights for the conditions of experiment 3 in which the distracter mean level was 30 dB SPL. Open symbols represent children's data, filled symbols represent adults' data.

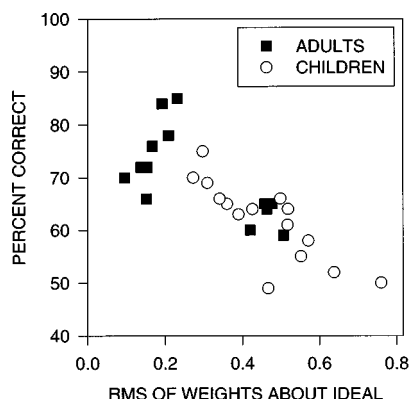


FIG. 9. Percent correct as a function of the rms of the weights about the ideal weights for the conditions of experiment 3 in which the distracter mean level was 70 dB SPL. Open symbols represent children's data, filled symbols represent adults' data.

equal standard deviations, either the low-, middle-, or high-frequency component was identified as the target. When the weighting functions were averaged across listeners, the adults generally weighted the proper component while children tended to weight all components roughly equally when the mean levels of the components were equal.

In relatively simple listening tasks in which the level information of all components in a complex is relevant (Willihnganz *et al.*, 1997) or in which one component is defined as the target for the duration of the experiment (experiments 1 and 2), children cannot be distinguished from adults on the basis of their weighting functions. Large individual differences in weighting functions exist, but no consistent pattern emerges among children or adults. In a more complex listening task in which the component identified as the target is varied across conditions of the experiment (experiment 3), it seems that children are less able than adults to redistribute the weight given to components when those components have relatively equal salience. As a result, children tend to weight all components equally over a block of trials while adults generally weight the target component more heavily.

Furthermore, it is apparent that performance in the ana-

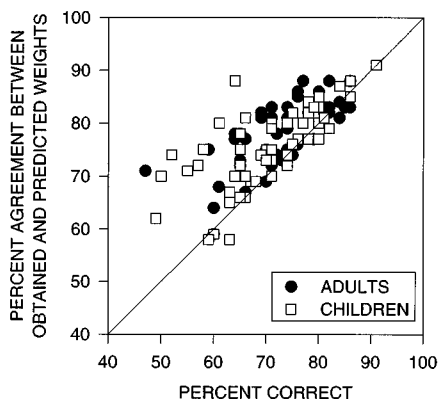


FIG. 10. Percentage of responses predicted by the weights (percent agreement between obtained and predicted weights) as a function of percent correct for all data in this series of experiments. The different symbol types represent data for children and adults, as shown in the figure.

lytic listening tasks of the present study, where listeners were required to attend to a single component, is more dependent upon the shape of the weighting functions than is performance in the synthetic listening tasks of the companion paper (Willihnganz *et al.*, 1997), where listeners were required to attend to all components equally. When feedback is consistent with the level change of only a single component, as in the present experiments, performance will suffer to the extent that listeners weight an irrelevant component, that is, to the extent that a nonoptimal weighting strategy is used. This can be seen in the figures showing percent correct as a function of the rms of the observer weights about the ideal (particularly Figs. 2 and 9). Percent correct decreases as the rms of the weights about the ideal increases (as efficiency decreases). When all components provide equal information in performing a task, as in Willihnganz *et al.* (1997), performance is less dependent upon the specific weighting strategy used. In those experiments, performance was unrelated to the specific shape of the weighting functions.

As stated in the Introduction, Allen and Wightman (1995) and Bargones *et al.* (1995) pointed out that the detection psychometric functions they observed would result from the use of nonoptimal listening strategies. The present results describe some of the differences between the listening strategies of children and adults. Although adults are more sensitive to level differences than children (based on the signal levels required to achieve a criterion level of performance), adults display superior weighting strategies only under certain conditions. Adults show no superiority over children in using the level information in a target component and ignoring interfering components as long as the target frequency remains constant for a fairly long period of time (in the present case, across several experimental sessions). However, when the task is varied in such a way that listeners are required to ignore information that was previously relevant to the task, children are less able than adults to make the necessary change in weighting the available information. In the present study, two of the children appeared to revert to a particular overlearned strategy when the task increased in complexity. This raises the possibility that the differences observed between the listening behaviors of children and adults stem from the speed with which the task is learned or, more likely, unlearned. Perhaps adults are simply able to learn a new task more quickly and as a result can be more flexible in the weighting strategies they adopt. In contrast, children may have a greater tendency to learn to perform a listening task in a specific way and then retain that strategy even when the demands of the task change. Future studies are needed to explore this possibility.

Although the results of the present study and those of previous studies suggest that children may use a roving attention band, Greenberg *et al.* (1970) found that both children and adults demonstrate frequency-selective behavior in a signal-detection task. It may be significant that their child listeners ranged in age from 6 years 4 months to 8 years 6 months, while the oldest child in the present study was 5 years 7 months. The child listeners of Allen and Wightman (1995), Bargones *et al.* (1995), and Bargones and Werner (1994) were all 5 years of age or younger. Perhaps these

studies identify a critical age when a more adult-like selective listening ability develops.

In general, the present results and those of Allen and Wightman (1995) indicate that children are less able to focus their attention than adults in complex listening tasks. When the task is well-defined over a relatively long period of time, such as when the target frequency or distracter frequencies are fixed, children and adults cannot be distinguished on the basis of their weighting functions (i.e., listening strategies). However, if stimulus uncertainty is increased, such as when a distracter is redefined as a target or if the distracter frequencies are randomized, the performance of children in terms of percent correct is more adversely affected than that of adults. To some extent, the decrement in performance can be attributed to children's lower degree of attentional selectivity, as evidenced by their nonoptimal weighting functions.

ACKNOWLEDGMENTS

Supported by research Grant No. HD23333-07 from the National Institutes of Health.

¹The *magnitudes* (absolute values) of the correlation coefficients were normalized to sum to 1.00 rather than the actual correlation coefficients in order to deal with situations in which negative correlations were obtained. In these cases, normalizing the actual correlation coefficients occasionally can produce values beyond the range of -1.00 to 1.00 . Normalizing the magnitudes of the correlation coefficients more accurately reflects the relative weight given to each component across conditions. The *sign* of each weight then simply indicates whether the listener's responses were consistent with (+) or opposite to (−) the direction of the level change of that component.

- Allen, P., and Wightman, F. (1994). "Psychometric functions for Children's detection of tones in noise," *J. Speech Hear. Res.* **37**, 205–215.
- Allen, P., and Wightman, F. (1995). "Effects of signal and masker uncertainty on children's detection," *J. Speech Hear. Res.* **38**, 503–511.
- Allen, P., Wightman, F. L., Kistler, D. J., and Dolan, T. R. (1989). "Frequency resolution in children," *J. Speech Hear. Res.* **32**, 317–322.
- ANSI (1987). ANSI S3.39-1987, "American National Standard specification for instruments to measure aural acoustic impedance and admittance (aural acoustic immittance)" (American National Standards Institute, New York).
- ANSI (1989). ANSI S3.9-1989, "American National Standard specification for audiometers" (American National Standards Institute, New York).
- American Speech-Language-Hearing Association (1990). "Guidelines for screening for hearing impairments and middle ear disorders," *ASHA* **32** (Suppl. 2), 17–24.
- Bargones, J. Y., and Werner, L. A. (1994). "Adults listen selectively; infants do not," *Psychol. Sci.* **5**, 170–174.
- Bargones, J. Y., Werner, L. A., and Marean, G. C. (1995). "Infant psychometric functions for detection. Mechanisms of immature sensitivity," *J. Acoust. Soc. Am.* **98**, 99–111.
- Bregman, A. S. (1990). *Auditory Scene Analysis* (MIT, Cambridge, MA).
- Elliott, L. L., and Katz, D. R. (1980). "Children's pure tone detection," *J. Acoust. Soc. Am.* **67**, 343–344.
- Greenberg, G. Z., Bray, N. W., and Beasley, D. S. (1970). "Children's frequency-selective detection of signals in noise," *Percept. Psychophys.* **8**, 173–175.
- Olsho, L. W., Koch, E. G., Carter, E. A., Halpin, C. F., and Spetner, N. B. (1988). "Pure-tone sensitivity of human infants," *J. Acoust. Soc. Am.* **84**, 1316–1324.
- Schneider, B. A., Trehub, S. E., Morrongiello, B. A., and Thorpe, L. A. (1986). "Auditory sensitivity in preschool children," *J. Acoust. Soc. Am.* **79**, 447–452.
- Willihnganz, M. S., Stellmack, M. A., Lutfi, R. A., and Wightman, F. L. (1997). "Spectral weights in level discrimination by preschool children. Synthetic listening conditions," *J. Acoust. Soc. Am.* **101**, 2803–2810.
- Yost, W. A. (1991). "Auditory image perception and analysis: The basis for hearing," *Hear. Res.* **56**, 8–18.

Formant transition duration and speech recognition in normal and hearing-impaired listeners

Christopher W. Turner

Department of Speech Pathology and Audiology, The University of Iowa, Iowa City, Iowa 52242

Sarah J. Smith, Patricia L. Aldridge, and Suzanne L. Stewart

Syracuse University, Communication Sciences and Disorders, Syracuse, New York 13244

(Received 8 July 1996; accepted for publication 19 December 1996)

Listeners with sensorineural hearing loss often have difficulty discriminating stop consonants even when the speech signals are presented at high levels. One possible explanation for this deficit is that hearing-impaired listeners cannot use the information contained in the rapid formant transitions as well as normal-hearing listeners. If this is the case, then perhaps slowing the rate of frequency change in formant transitions might assist their ability to perceive these speech sounds. In the present study, sets of consonant plus vowel (CV) syllables were synthesized corresponding to /ba, da, ga/ with formant transitions for each set ranging from 5 to 160 ms in duration. The listener's task was to identify the consonant in a three-alternative, closed-set response task. The results for normal-hearing listeners showed nearly perfect performance for transitions of 20 ms and longer, whereas the shortest transitions yielded poorer performance. A group of eight hearing-impaired listeners pure-tone averages (PTAs) ranging from 30 to 62 dB HL) was also tested. The hearing-impaired listeners tended to show poorer performance than the normals for transitions of all durations; however, the performance of a few hearing-impaired subjects was equal to that of normals for the shortest-duration transitions. A strong inverse relation was observed between degree of hearing loss and improvement in score as a function of transition duration. These results suggest that increasing the duration of formant transitions for listeners with more severe hearing losses may not provide a helpful solution to their speech recognition difficulties. © 1997 Acoustical Society of America. [S0001-4966(97)01805-5]

PACS numbers: 43.66.Lj, 43.66.Sr, 43.66.Ts, 43.71.Es [JWH]

INTRODUCTION

It is well-known that listeners with sensorineural hearing impairment make errors identifying the place of articulation for stop consonants (e.g., Owens *et al.*, 1972). Even after deficits in audibility, or the effects of elevated sensitivity thresholds, are accounted for, hearing-impaired listeners often show poorer recognition for these speech sounds (Turner and Robb, 1987; Turner *et al.*, 1992). One possible explanation for this difficulty is that the listener with sensorineural hearing loss has difficulty in using the information contained in the rapid formant transitions which can signal the place of articulation (e.g., Dorman *et al.*, 1985; Summerfield *et al.*, 1985). Zeng and Turner (1990) showed that listeners with sensorineural hearing loss had difficulty using formant transitions presented at audible levels in fricative-vowel speech syllables. The durations of these transitions, however, were slightly longer (about 40–50 ms) than those commonly observed for formant transitions in stop consonants (15–30 ms). Nonetheless, it remains a distinct possibility that the rapid formant transitions in stop consonants are a particularly difficult cue for the impaired auditory system to utilize.

If listeners with sensorineural hearing loss do have problems using short-duration glide information, then increasing the duration of formant transitions by slowing down the rate of frequency change might serve to improve their performance. Ochs *et al.* (1989) looked at the ability of normal-hearing and hearing-impaired listeners to identify waveform-

edited stop consonants when the durations of the formant transitions were systematically increased. They found that listeners with hearing loss, as well as those with normal hearing, showed improved performance when the duration of the presented formant transitions was increased. In their study however, increasing the duration of the stimuli meant that the frequency extent of the transitions also increased.

In a recent psychoacoustic investigation, Summers and Leek (1995) measured the ability of normal-hearing and hearing-impaired subjects to discriminate pure-tone frequency glides in the spectral region of typical second formants (F_2) for speech sounds. Glide duration, rate of frequency change, and the extent of frequency change were varied in their study, in order to investigate the hypothesis that the shortest and most rapidly changing formant transitions would be particularly troublesome for the listeners with hearing loss. Contrary to the predictions of the original hypothesis, the hearing-impaired listeners and the normal-hearing listeners performed equally well in discriminating brief (30 ms) tone glides in which frequency changed rapidly over ranges of 200 Hz or greater. The hearing-impaired listeners showed only a small improvement in discrimination performance as the rate of frequency change was decreased, whereas the normal-hearing listeners showed a large improvement in performance. These results might argue against a signal-processing strategy for improving speech recognition in hearing-impaired listeners that changes the duration of short formant transitions (where listeners with hearing

TABLE I. Thresholds for the hearing-impaired listeners (in dB HL).

Subject	Frequency (in kHz)						
	0.25	0.5	1.0	2.0	3.0	4.0	8.0
BP	25	45	50	40	45	45	85
DB	10	20	35	55	50	45	20
JH	15	25	25	40	45	50	60
SG	15	20	50	45	50	50	65
LC	20	25	35	40	55	60	65
WS	10	10	25	70	65	70	60
BB	60	65	60	60	65	80	95
JP	20	25	60	55	55	55	65

loss tend to perform as well as normal-hearing listeners) to longer and slower transitions (which would yield little or no improvement in discriminability).

In this report, we describe the results of an experiment designed to investigate in a more direct fashion (1) the comparison between the ability of normal and hearing-impaired listeners to identify the place of articulation of stop consonants as a function of formant transition duration and, (2) whether slowing the rate of frequency transition by increasing the glide durations in speech sounds might serve to improve the speech discrimination abilities of listeners with hearing loss. Rather than using a pure-tone glide discrimination task as in the Summers and Leek study, we used a simple identification task with synthetic speech stimuli. We asked subjects with normal and impaired hearing to discriminate between three synthetic “stop-consonant” stimuli, corresponding to the bilabial, alveolar, and velar places of articulation. A score (in percent correct) was obtained for formant transition durations of 5, 10, 20, 40, 80, and 160 ms. As this report shows, our results show that sensorineural hearing loss does not specifically affect the perception of syllables with rapidly changing formant transitions. Our results also show that slowing transition rate for these listeners does not consistently improve their performance, and does not bring performance to the levels achieved by normals listening to normal-duration speech stimuli.

I. METHODS

A. Subjects

Five normal hearing and eight hearing-impaired listeners participated in these experiments. The subjects with normal hearing had thresholds at or better than 20 dB HL for the audiometric test frequencies from 250 to 8000 Hz. The hearing-impaired subjects were classified as having hearing loss of sensorineural origin on the basis of an absence of an audiometric air-bone gap, normal middle-ear immittance results, and word discrimination scores consistent with their audiograms. Three-frequency (0.5, 1.0, and 2.0 kHz) pure-tone average hearing losses ranged from 30 to 62 dB HL in these listeners. The sensitivity thresholds for these hearing-impaired subjects are displayed in Table I. In the hearing-impaired subjects, the better ear was tested.

TABLE II. Formant frequencies for the synthetic CV tokens.

Consonant	Starting formant frequency (in Hz)		
	$F1$	$F2$	$F3$
/b/	200	1000	2200
/d/	200	1600	2600
/g/	200	1500	1700
Vowel	Steady-state formant frequency (in Hz)		
	$F1$	$F2$	$F3$
/a/	700	1220	2600

B. Stimuli and procedures

Subjects' performance scores were obtained for the discrimination of synthetic stop consonants plus vowel (CVs) as a function of formant transition duration (i.e., 5, 10, 20, 40, 80, and 160 ms). In each trial of the discrimination task, the subject was required to identify the synthetic token presented in a three-alternative, closed-set response task. The three consonant alternatives, /ba/, /da/, and /ga/, correspond to the three places of articulation, bilabial, alveolar, and velar, respectively. Sets containing three synthetic CV tokens were created digitally via a Macintosh version of the Klatt Speech Synthesis program (Sensyn; Sensimetrics Corp.). The three tokens within a set all had the same duration formant transitions. All tokens were 300 ms in duration, consisting of an initial formant transition portion for the specified time (5, 10, 20, 40, 80, or 160 ms) followed by a steady-state vowel portion for the remainder of the 300 ms. Each token was constructed as a “voiced” stop consonant syllable, with the voicing source turned on at stimulus onset and continuing throughout the entire token. No burst or aspiration was included in the tokens, in an attempt to restrict the listeners to using the formant transitions as the primary cues for place of articulation. The common vowel for all tokens was /a/, with steady-state formant values of 700, 1220, and 2600 for $F1$, $F2$, and $F3$, respectively. The three tokens within each set differed according to the place of articulation of the initial stop consonant, which was accomplished by setting the starting frequencies of the formant transitions appropriately. These values are shown in Table II. Formant bandwidths were set to 90 Hz for $F1$, 110 Hz for $F2$, and 170 Hz for $F3$. All three formants changed their center frequency as a linear function of time from stimulus onset to the end of the transition duration. Informal listening tests performed during the development of the stimuli showed that the tokens with formant transition durations of 10–40 ms were the easiest to identify as common stop consonants (/b/, /d/, and /g/), whereas the shortest (5 ms) and longest (80 and 160 ms) tokens were less easily heard as common stop consonants.

A single experimental run consisted of 180 trials, in which each of the three tokens was presented 60 times each in a randomized (without replacement) order. On a single trial, the subject heard the presented token, then responded by pressing one of three buttons, labeled /b/, /d/, or /g/, to indicate the identity of the token just heard. If subjects were not certain of a response, they were required to guess. A

feedback light was then illuminated above the button corresponding to the correct response. In view of the fact that most subjects found the shortest and longest transition duration stimuli somewhat unnatural, the subjects first practiced on the 40-ms transition stimulus set until asymptotic performance was attained. This typically required 1000 or more total trials. Then a final data set consisting of 360 trials for each data point was collected, also with individual trial feedback presented to the listener. Data for the remaining stimulus sets, corresponding to the other transition durations, were then collected in a similar manner, with practice sessions administered until asymptotic performance was reached preceding the final data set. The order of presentation of the remaining transition duration conditions was randomized for each subject.

Stimuli were generated and presented using a Macintosh IIX computer with a DigiDesign sound accelerator 16-bit digital-to-analog converter at a sampling rate of 44.1 kHz. An anti-aliasing low-pass filter of 20 kHz was used. Sennheiser HD 25 SP headphones were used for monaural presentation. The presentation levels of the stimuli are expressed as that level developed in an NBS 9-A coupler by this headphone when the steady-state vowel portion of the stimuli was played repeatedly. Presentation levels were 70 or 75 dB sound-pressure level (SPL) for the normal-hearing subjects. Pilot data indicated that normal-hearing subjects could attain maximum performance at these levels. Stimuli were initially presented to hearing-impaired subjects at 90 dB SPL. After a hearing-impaired subject reached asymptotic performance for the 40-ms transition duration stimuli at 90 dB SPL, the presentation level was increased in 5 dB steps until maximum performance was attained, unless the level was deemed too uncomfortable by the subject. Thus the data for the hearing-impaired listeners were collected at the highest comfortable level yielding maximum performance of the task for the 40-ms transition stimuli. The data for all other transition durations were collected at this same presentation level. For the hearing-impaired listeners, this final presentation level was 95, 100, or 105 dB SPL. In a further determination of stimulus audibility, a one-third octave band analysis of the concatenated, initial, consonant-only portion of the three stimuli was performed and revealed that the presentation levels of the average consonant spectra for all the hearing-impaired listeners (including the two with the most severe hearing loss (subjects WS and BB) were at least 5 dB suprathreshold for all frequencies up to and including 3000 Hz.

II. RESULTS

For all normal-hearing listeners, performance improved as transition duration was increased, with the nearly perfect performance achieved for the 20-ms transitions and longer. There was considerable intersubject variability among the eight hearing-impaired subjects. Figure 1 displays the recognition scores of the individual normal-hearing and hearing-impaired subjects as a function of the formant-transition duration. To facilitate fitting these functions with straight lines, as described below, the data are expressed as d' values on the ordinate as determined from the tables published by

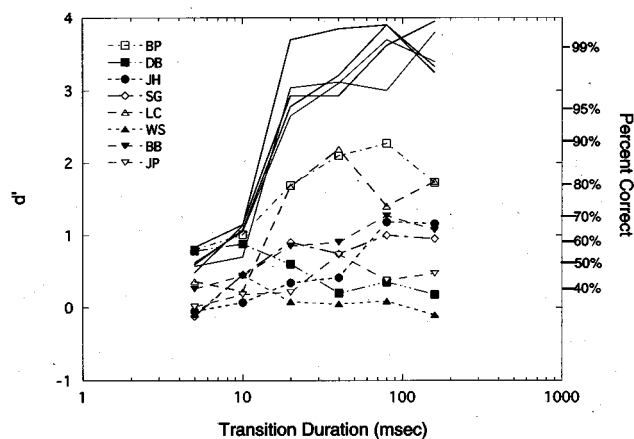


FIG. 1. Performance on the three-alternative, closed-set identification task for the normal-hearing and hearing-impaired subjects plotted as a function of the duration of the formant transitions. The normal hearing subjects' data are shown by solid lines without symbols and the hearing-impaired subjects' data are shown by broken lines with individual symbols (see legend). Performance is expressed as prime d' on the left axis and as percent correct on the right axis. 33% correct (or prime $d' = 0$) represents chance performance.

Hacker and Ratcliff (1979). The corresponding percent-correct values are also indicated on the right-hand axis. It is clear from the figure that none of the hearing-impaired listeners perform as well as the normal-hearing subjects for the longer duration transitions. Several of the hearing-impaired listeners performed similarly to the normal subjects for the 5-ms transition, where normal performance ranged from 47% to 58% and the scores for the hearing-impaired listeners ranged from about 33% (chance performance) to 57%. A closer examination of the results of individual hearing-impaired subjects in Fig. 1 reveals that there is a particularly wide range of performance for the longer transition durations, with several subjects showing an increase in performance as formant transition duration increased (the same trend as normal-hearing subjects), and others either no improvement or even decreasing performance as formant transition duration increases. However, the data do indicate that for formant transitions of 20 ms and longer, no hearing-impaired listener's performance was equivalent to that of the normal group.

In order to determine if performance as a function of formant transition duration was related to the degree of hearing loss, the data of each subject in Fig. 1 were fit with linear functions. Improvement in score as a function of transition duration was expressed by the slope of a linear function relating performance (in d') to the logarithm of the transition duration in ms. In Fig. 2, the slopes of the listeners' functions are plotted against the corresponding degree of hearing loss at 2 kHz, which roughly corresponds to the 2nd formant frequency region of these stimuli. There was a strong inverse relation ($r=0.98$) between the slope values and the degree of hearing loss, indicating that the more severe the hearing loss, the less performance improved as transition duration was increased. In other words, those listeners with more severe hearing losses, who presumably would suffer the greatest deficits in speech recognition and might most benefit from a signal-processing scheme to increase the salience of

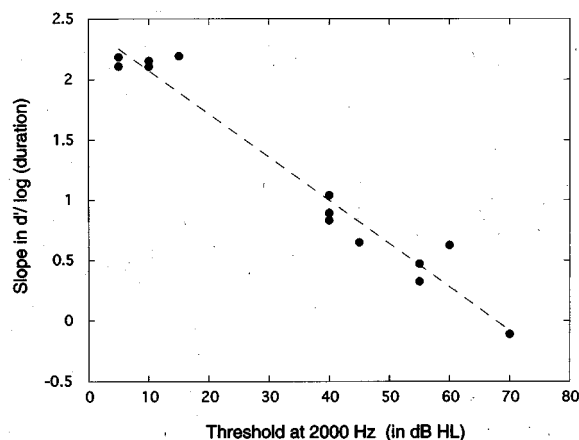


FIG. 2. The slope of the fitted lines for each individual subjects' data shown in Fig. 1 plotted as a function of the degree of hearing loss (in dB HL) at 2000 Hz. The linear correlation coefficient for these data is $r = 0.98$.

stop-consonant cues, would benefit the least from "slowing down" the formant transitions.

Although care was taken to present the second-formant frequency region of the stimuli at suprathreshold levels to all subjects, audibility may have still been a concern for some individuals. In particular, the subject with the most severe hearing loss at 2 kHz (WS) showed near-chance performance for all transition durations except the 10 ms condition where his score was 47%. Why audibility would be a factor for all but one transition duration for this subject is not clear, but since no totally reliable method exists for expressing the audibility of small temporal and spectral portions of a CV stimulus, audibility problems cannot be ruled out. Even without this subject's data, the trends described above are still evident. Since the degree of hearing loss does tend to be related to the sensation level of stimulus presentation, it is difficult to separate those two factors, especially when presentation level is limited by loudness discomfort levels.

An examination of the individual confusion matrices revealed that for many subjects, both normal-hearing and hearing-impaired, the largest number of errors was made in /d/ vs /g/ confusions. This is not surprising, as the second-formant starting frequencies of these two consonants are very similar (see Table II). There were exceptions, however, most notably hearing-impaired subjects BP and LC, who achieved scores of 80%–90% for some conditions. For these subjects, errors were much more equally distributed across incorrect response cells of the confusion matrices.

III. DISCUSSION

The findings of the present experiment, along with those of Summers and Leek's (1995) difference limen experiments using toneglides, point to the same conclusion with regard to the abilities of hearing-impaired listeners' use of formant transition information. In particular, we found that hearing-impaired subjects do not necessarily perform more closely to normal-hearing levels of performance for longer, slower transitions. Our work extends this generalization to a task more akin to actual speech recognition. This suggests that

speech processing schemes designed to improve recognition for hearing-impaired patients that are based upon "slowing down" rapid formant transitions might not be effective.

If one takes 20–40 ms as the formant transition duration of naturally spoken stop consonants, the present data show that increasing transition duration to 80 or 160 ms for any of our hearing-impaired listeners would not produce recognition performance equivalent to that of a normal-hearing person listening to speech with natural formant transition durations. It was true that for some of our hearing-impaired listeners, "slowing down" the formant transitions did tend to improve their recognition scores. However, the degree of hearing loss in our subjects was a strong predictor of a subject's improvement from using longer formant transitions. This might argue against increasing transition duration as a strategy to improve the speech recognition abilities of hearing-impaired patients since it is precisely those listeners with more severe hearing losses who would presumably need assistance in perceiving speech cues. While we can not rule out the possibility that low sensation levels of the formant transitions for the more severely impaired listeners did contribute to their problems, attempts to equate sensation levels of speech cues for such listeners are usually not practical solutions in the real world, due to loudness discomfort limitations.

ACKNOWLEDGMENTS

Helpful comments on an earlier version of this manuscript were provided by Karen Doherty and Pam Souza. Andrew Faulkner and a second, anonymous reviewer provided many helpful suggestions on an earlier version of the paper. This work was supported by NIDCD Grant No. DC 00377.

- Dorman, M. F., Marton, K., Hannley, M. T., and Lindblom, J. M. (1985). "Phonetic identification by elderly normal and hearing-impaired listeners," *J. Acoust. Soc. Am.* **77**, 664–670.
- Hacker, M. J., and Ratcliff, R. (1979). "A revised table of d' for M -alternative forced choice," *Percept. Psychophys.* **26**, 168–170.
- Ochs, M. T., Humes, L. E., Ohde, R. N., and Grantham, D. W. (1989). "Frequency discrimination ability and stop-consonant identification in normally hearing and hearing-impaired subjects," *J. Speech Hear. Res.* **32**, 133–142.
- Owens, E., Benedict, M., and Schubert, E. (1972). "Consonant phonemic errors associated with pure tone configurations and certain kinds of hearing impairment," *J. Speech Hear. Res.* **15**, 308–322.
- Summerfield, Q., Foster, J., Tyler, R., and Bailey, P. J. (1985). "Influences of formant bandwidth and auditory frequency selectivity on identification of place of articulation in stop consonants," *Speech Commun.* **4**, 213–229.
- Summers, V., and Leek, M. R. (1995). "Frequency glide discrimination in the $F2$ region by normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **97**, 3825–3832.
- Turner, C. W., Fabry, D. A., Barrett, S., and Horwitz, A. R. (1992). "Detection and recognition of stop consonants by normal-hearing and hearing-impaired listeners," *J. Speech Hear. Res.* **35**, 942–949.
- Turner, C. W., and Robb, M. P. (1987). "Audibility and recognition of stop consonants in normal and hearing-impaired subjects," *J. Acoust. Soc. Am.* **81**, 1566–1573.
- Zeng, F. G., and Turner, C. W. (1990). "Recognition of voiceless fricatives by normal and hearing-impaired subjects," *J. Speech Hear. Res.* **33**, 440–449.

An investigation of stop place of articulation as a function of syllable position: A locus equation perspective

Harvey M. Sussman,^{a)} Nicola Bessell, Eileen Dalston, and Tivoli Majors
Department of Linguistics, University of Texas, Austin, Texas 78712

(Received 29 August 1996; accepted for publication 16 January 1997)

Locus equations were employed to phonetically describe stop place categories as a function of syllable-initial, -medial, and -final position. Ten speakers, five male and five female, produced a total of 2700 CVC and 4500 VCV utterances that were acoustically analyzed to obtain $F2$ onset, $F2$ vowel, and $F2$ offset frequencies for locus equation regression analyses. In general, degree of coarticulation, as indexed by locus equation slope, was reduced for post-vocalic (VC) stops relative to pre-vocalic stops (pooled data from initial and medial positions), but significant differences were observed as a function of stop consonant. All stops showed significantly reduced R^2 values and increased standard errors of estimate for VC relative to CV productions. Separability of stop place categories in a higher-order slope \times y -intercept acoustic space also diminished for VC vs CV stop productions. The degradation of classic locus equation form (high correlation and linearity) for VC relative to CV productions was attributed to greater articulatory precision in the production of pre-vocalic compared to post-vocalic stops. This greater articulatory precision was interpreted as reflecting a greater need to normalize vowel context-induced variability of the $F2$ transition for syllable onset relative to final stops. The decline in acoustic lawfulness of syllable-final stops is discussed in terms of coarticulatory interactions and expected perceptual correlates. © 1997 Acoustical Society of America. [S0001-4966(97)05205-3]

PACS numbers: 43.70.Fq [AL]

INTRODUCTION

Locus equations are linear regressions of the frequency of the second formant transition sampled at its onset ($F2$ onset) on the frequency of the second formant sampled in the middle of the following vowel ($F2$ vowel). These coordinates are obtained for a single consonant coarticulated with a range of vowels. $F2$ onset is plotted on the y axis and $F2$ vowel on the x axis. Scatterplots are consistently characterized by highly correlated and linearly related $F2$ coordinates that tightly cluster around the regression line (Lindblom, 1963; Nearey and Shammass, 1987; Sussman *et al.*, 1991, 1993). The empirical phenomenon of locus equations shows that when sets of utterances forming an equivalence class (e.g., a stop place category) are plotted in $F2$ onset– $F2$ vowel acoustic space, vowel context-induced variability becomes lawful for the first time. We interpret this emergence of acoustic orderliness as reflecting a higher-order systematicity of the place of articulation feature in syllable-initial stop+vowel utterances. Derived locus equation parameters, slope, and y intercept, when used as predictor variables in discriminant analyses, have shown 100% accuracy in classification of stop consonants into place categories compared to an overall rate of 77% when token level $F2$ onset and $F2$ vowel frequencies are used (English: Sussman *et al.*, 1991; Spanish: Celdán and Villalba, 1995).

Place of articulation coding for syllable-initial obstruents has been illustrated across various manner classes, including aspirated voiceless stops, nasals, and fricatives (Sussman, 1994; Sussman and Shore, 1996). The paradigm

has been validated across various languages (Sussman *et al.*, 1993), and even the presence of articulatory perturbations, such as speaking with bite blocks, fails to alter the linearly correlated relationship between $F2$ onset and $F2$ vowel (Sussman *et al.*, 1995).

When two acoustic parameters covary as predictably as $\langle F2$ onset, $F2$ vowel \rangle , it is highly probable that feature extraction could occur by normalizing one acoustic parameter with respect to the other. The robust consistency of the form of locus equations has led to the interpretation that the linear functions reflect (in a noncomputational sense) a vowel normalization of the variable $F2$ transitions for CV utterances (Sussman *et al.*, 1995). Each locus equation scatterplot characterizes an equivalence class of CV transitions; each equivalence class possesses an acoustically coded commonality that phonetically describes a place of articulation category. The classes intersect to some extent, so they are only partially distinctive as to consonant (Nearey and Shammass, 1987), but higher-order parameters, slope and y intercept do not overlap when plotted as coordinates in locus equation acoustic space.

To summarize, locus equation studies to date have shown that two phonetic objectives are inherently achieved when CVs contrasting in place are produced: (i) a form of “normalization” of the vowel-context induced variability of the $F2$ transition; and (ii) speaker-controlled differential adjustments of the degree of CV coarticulation as a function of stop place.¹ The primary focus of this research is the second, or coarticulatory aspect of stop consonantal place coding by the $F2$ transition. The notion of coarticulatory “settings” varying with place of articulation provides a new perspec-

^{a)}Electronic mail: sussman@mail.utexas.edu

tive, wherein coarticulation is viewed as helping to establish, rather than blur, stop place contrasts.

Locus equation studies have unequivocally shown that labials are always characterized by steeper slopes than alveolars. This implies greater coarticulatory influence of the following vowel on the onset *F*₂ frequency of /b/ compared to /d/. Alveolars (and coronals, in general) maximally resist coarticulatory effects of the following vowel. Velars, in all languages thus far examined, are best characterized by two linear regression functions, one with a flat slope with front vowels, showing very little coarticulatory effects, and a steep function with back vowels showing maximal coarticulatory effects.²

If locus equations are to be regarded as a maximally general phonetic index for place of articulation, then their acoustic characterization of the place feature across additional syllable positions must be examined. The purpose of the present study was to investigate the capability of locus equations to phonetically distinguish stop place categories in syllable-initial, syllable-medial, and syllable-final positions. Production of CV(C), (V)CV, and (C)VC utterances also allowed for a systematic assessment of both consonantal and vowel context effects on CV and VC coarticulation as reflected by changes in locus equation coefficients and overall linearity and correlation.

A. Syllable position-dependent distributional differences

It is commonly remarked that linguistic data identify the open, CV syllable as the preferred or basic syllable structure. This understanding is based on distributional facts pointing to an asymmetry between onset and coda position: Codas are often constrained in ways that onsets are not. Coda constraints may prohibit codas altogether or limit the kinds of segments that occur in coda position. Although no truly exhaustive cross-linguistic survey of syllable types exists, the onset/coda asymmetry is supported by the surveys in Bell (1971) and Bell and Hooper (1978). These studies report that, while languages may require their syllables to have onsets, and every language has CV syllables, almost no language requires its syllables to begin with a vowel or end with a consonant. Thus, the CV syllable is never prohibited, but VC or CVC syllables may be. This basic asymmetry may be extended to constraints on coda content. Thus many languages with CVC syllables limit the segments that can occupy coda position to a subset of the phonemic inventory. Japanese is a classic example: While virtually any segment may occur in onset position, only nasals and the first half of geminate consonants may occupy coda position (Ito, 1989).

B. Syllable position-dependent phonetic differences

There is a long phonetic history of uncovering instances of segmental variation due to position effects within a syllable (e.g., Kahn, 1976). Krakow (1989) documented differences in anticipatory velic lowering such that syllable-final nasals were characterized by greater and earlier lowering of the velum compared to syllable-initial nasals. Thus medial vowels have greater nasal assimilation due to a final nasal than perseverative nasalization from a preceding nasal. In ad-

dition, labial-velic timing relationships for /m/ were found to differ as a direct function of syllable-initial vs syllable-final position. Fujimura (1990) reported more extensive anticipatory velic lowering in syllable-final than syllable-initial nasals. Component gestures for laterals have also been observed to differ as a function of syllable position. Sproat and Fujimura (1993), in a microbeam study tracking tongue tip and dorsum movement patterns for lateral /l/, found different phasing relations for syllable-final /l/ vs syllable-initial /l/. As with nasalized vowels, these different inter-gestural phasing relations between tongue tip and tongue dorsum resulted in lawful acoustic consequences reflected in light versus dark /l/. EPG investigations show that stops in coda position have reduced contact compared to onset segments (Manuel, 1991; Byrd, 1994; Keating, 1995).

Studies by Dorman *et al.* (1975) and Fujimura *et al.* (1978) point to a perceptual analogue of the asymmetry. Working with VCV stimuli that contained conflicting formant transitions for consonantal place of articulation, these studies found that formant transitions into a vowel (in the CV/onset portion) outweigh the perceptual contribution of transitions out of the vowel (in the VC/coda structure). Perceptual adaptation studies likewise identify an onset/coda asymmetry in perception: Onsets are subject to adaptation but codas systematically are not (e.g., Samuel, 1989).

C. Syllable position-dependent coarticulation differences

While we use the terms onset and coda, these are really shorthand terms for where there is a vowel relative to the C. Phonetically, it is the availability of a following as opposed to a preceding vowel that defines what we are calling “onset” as opposed to “coda.” This definition places initial and medial C together as a class as both interact with a following vowel, and only final C’s interact with preceding vowels. The position of the vowel relative to the consonant appears to be a key aspect of coarticulatory interaction. Lindblom (1983) has defined the degree of coarticulation as “...the extent to which the vowel environment is allowed to color the formant frequency pattern of the consonant (p. 225).” The extent to which the vowel influences *F*₂ of immediately adjacent consonants as a function of place of articulation can be ascertained by comparing CV to VC production within a CVC monosyllable.

In the original locus equation study by Lindblom (1963), performed on a single Swedish speaker, slopes for bV, dV, gV vs Vb, Vd, Vg were 0.69, 0.28, 0.95, and 0.57, 0.55, and 1.02, respectively. Using slope magnitude as a direct index of degree of coarticulation (Krull, 1988; Sussman *et al.*, 1993), there was greater coarticulation for initial /b/ vs final /b/, less coarticulation for initial /d/ compared to final /d/, and slightly less coarticulation for initial versus final /g/. No phonetic explanations were offered for these CV vs VC differences. Krull (1988) also compared locus equation slopes and y intercepts for initial versus final stops in a single Swedish speaker. She reported contrastive locus equation coefficients for initial stops, but very small slope differences across final stops, with all values approximating a slope of 1.0. Thus, in agreement with Lindblom (1963), final C loci

were much more vowel dependent (greater extents of coarticulation) than initial C loci. Krull did report that *y*-intercept values best differentiated among velar (the highest), dental (next highest), and labial (lowest) stops in final position.

Since production, perception, and phonological factors point to coda C as distinct from onset C, there should be clear correlates in locus equation parameters reflecting this syllabic asymmetry. The purpose of this study was to compare stop consonant production in pre- and post-vocalic positions and thereby assess the integrity of locus equations as a general phonetic descriptor of consonantal place.

I. METHODS

A. Subjects and stimulus materials

Ten speakers, five male and five female, were recruited for this study. Nine were native speakers of American-English and one (F#3) was a native speaker of Canadian-English. Dialectal heterogeneity was sought (midwest, east-coast, southern, westcoast, and Texas). The age range was 26–53 with a mean age of 37.8. Subjects were asked to produce both CVC and VCV syllables embedded in a carrier phrase format “Say _____ again.” CVCs consisted of /bVb/, /bVd/, /bVg/, /dVb/, /dVd/, /dVg/, /gVb/, /gVd/, and /gVg/ tokens with 10 medial vowels /i ɪ e ε æ ɑ ɒ ʊ u/. Each CVC token was repeated three times. Subjects were instructed to produce the CVCs with a released final stop. Locus equation scatterplots were first derived from all 90 tokens for comparing syllable-initial (CV) and syllable-final (VC) productions (e.g., for /b/ there were 10 vowels×3 repetitions=30 /bVb/ +30 /bVd/ +30 /bVg/ for syllable-initial /b/, and 30 /bVb/ +30 /dVb/ +30 /gVb/ for syllable-final /b/). In addition, the effect of C₂ on syllable onset CVs and the effect of C₁ on syllable coda VCs was also tested by comparing locus equation coefficients obtained from each subset of 30 CVCs. Medial stops in VCV format were also /b/, /d/, and /g/ with the same ten vowels as listed above produced in postconsonantal V₂ position. Locus equation coordinates were only derived from the CV₂ portion of the V₁CV₂ token as only three initial vowel contexts were used: /V1=i æ u/. Speakers were instructed to produce the VCVs with equal stress on each vowel. Each VCV token was repeated five times yielding a total of 150 tokens per locus equation plot (50 /i/ CV tokens +50 /æ/ CV tokens +50 /u/ CV tokens for each stop place). Thus, each speaker produced 270 CVCs and 450 VCVs yielding a total corpus of 7200 tokens that were analyzed for this study. All subjects first produced the CVCs followed by a later recording of the VCVs. Each speaker’s productions were recorded in either a soundproof room (Industrial Acoustics Corp.) or a sound attenuated recording room using a high-quality microphone (Shure, SM58) and a Marantz cassette tape recorder (model PMD 201).

B. Measurements

All CVC productions were analyzed with MacSpeech Lab II by HMS and ED (version 1.4, G. W. Instruments, Inc.). The recorded signal was sampled at 10 kHz and digi-

tized using an Apple Macintosh IIfx computer with MacAdios II hardware. For syllable-initial stops (CV) *F*₂ onset frequencies were plotted on the *y* axis as usual, but for syllable-final stops (VC) the *F*₂ offset frequency was plotted on the *y* axis. Both *F*₂ onsets and *F*₂ offsets were plotted with the same *F*₂ mid-vowel frequency.

When the same ten speakers were brought back to produce VCV tokens, CSL and SoundScope were used for some speakers.³ *F*₂ vowel values were mean frequencies obtained from three separate sources: (i) directly from cursor frequency readouts via mouse position placed on the wide-band spectrographic display; (ii) linear predictive coding (LPC) spectra; and (ii) fast Fourier transform (FFT) narrow-band displays. *F*₂ onsets (and *F*₂ offsets for VCs) were solely determined by direct cursor positioning on the wide-band spectrographic display. To better view CV and VC interfaces, portions of the speech waveform were expanded to show the initial and final glottal pulses. The criterion for locating *F*₂ offsets was to place the cursor on the last visible glottal pulse of the *F*₂ resonance, preceding occlusion and final release of the C₂ stop. The great majority, if not all, of the final stops in the CVCs were produced with a clearly discernible release. Figure 1 shows a typical wide-band spectrograph ([geg]) and below, two expanded portions to better view *F*₂ onset and *F*₂ offset glottal pulses (shown by the arrows). The following are the specifications used for the three analysis methods.

1. MacSpeech lab (MSL)

Tokens were digitized at 16 bits and sampled at 10 000 Hz. LPC spectra were derived from a fixed 24.6-ms window surrounding a time marker placed on the speech waveform. The default parameter of 13 coefficients was used for the autocorrelation function. Wide-band FFTs were computed for an 8-ms time segment centered around the time marker simulating a 300-Hz filter while narrow-band FFTs used a 25-ms time window simulating a 45-Hz filter. A 6-dB pre-emphasis option was used to enhance higher formant intensities.

2. SoundScope (SS)

Tokens were digitized on a Macintosh 7100 at 16 bits with a 22 050-Hz sampling rate. LPC spectra were calculated with 25 coefficients and a 20-ms frame length using 512 FFT points. The FFT spectra were calculated with 512 points using various filter settings varying from 184 Hz to 300 Hz. Spectrograms were generated with a 300-Hz filter setting, a frame advance of 0.1 ms, and 256 FFT points. A 6-dB pre-emphasis of the higher frequencies was used.

3. Computerized speech lab (CSL)

Tokens were digitized at 16 bits and 10 000 Hz. LPC spectra were calculated with 12 coefficients and a 20-ms frame. FFT spectra were calculated with a 512 point frame, emulating a 29-Hz filter. Wide-band spectrograms were generated using a 50 point (293 Hz) or 75 point (195 Hz) frame. Whichever setting gave the best resolution was used. A default setting of 0.9 was used for preemphasis.

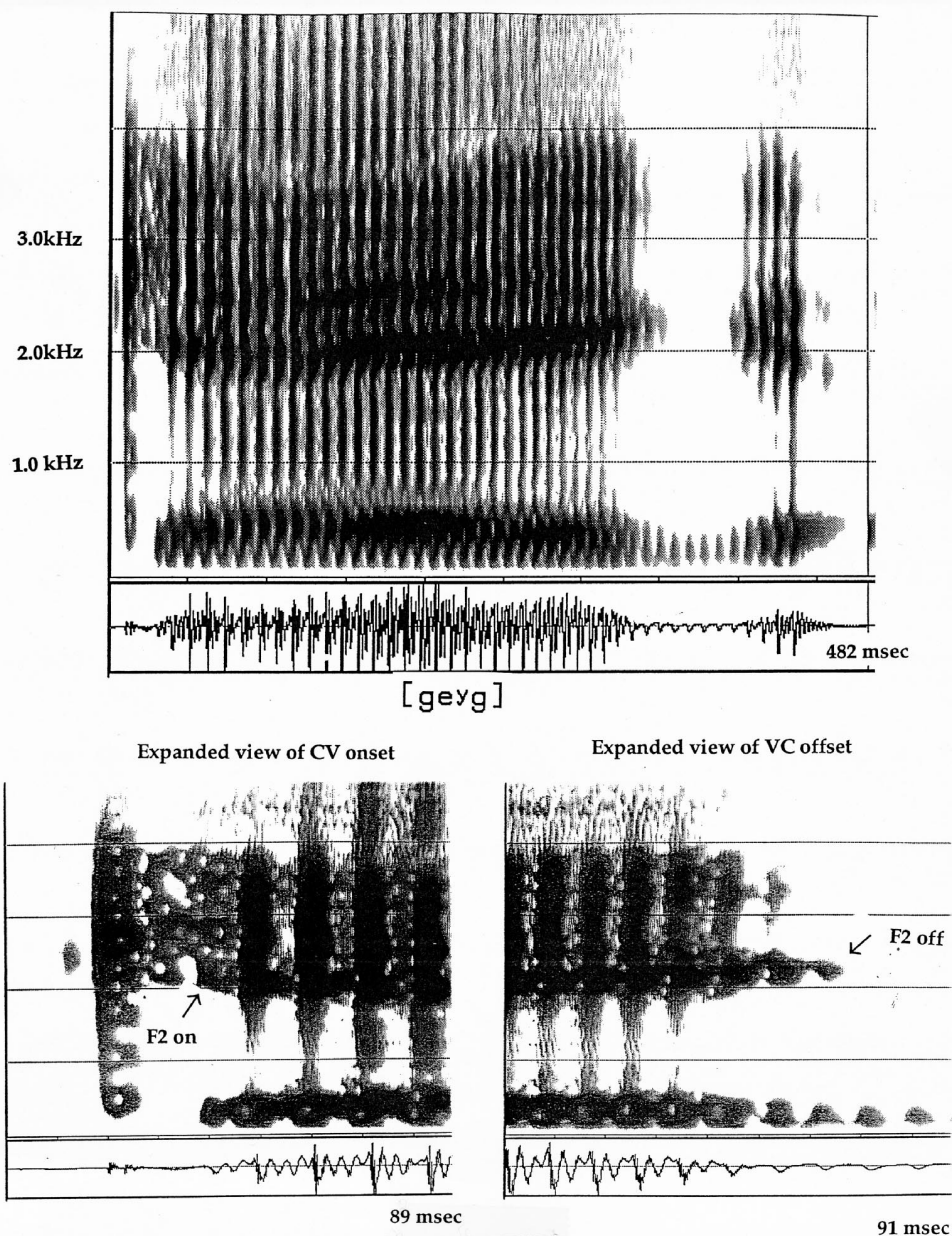


FIG. 1. A representative wide-band spectrogram of a CVC utterance ([gevg]) showing expanded portions for visually determining sampling points for F_2 onsets and F_2 offsets.

II. RESULTS

Four dependent variables were used to assess the phonetic distinctiveness and stability of locus equations as descriptors of stop place of articulation across syllable positions—slope, y intercept, R^2 , and standard error of estimate (SE). Following presentation of the main consonantal effects across syllable positions, specific context effects will be presented—initial and final C's on CV and VC locus equations, respectively, and V_1 effects on CV locus equations in VCV contexts.

A. Stop place characterization across syllable positions

Tables I–IV show slope (k), y intercept (c), R^2 , and SE values from locus equations obtained from each speaker for /b/, /d/, /g/ palatal, and /g/ velar as a function of syllable position—initial, medial, and final.

1. Slope

Figure 2 shows mean slope values for stops plotted across syllable positions. Slope values for /b/, /d/, and the two subsets of /g/ are quite similar for syllable-initial and syllable-medial contexts. Syllable-final slopes, except for /d/, can be seen to be different, as /b/ and /g/ velar values markedly decrease, while /g/ palatal slopes increase compared to syllable-initial and medial values. Only alveolar /d/ slope maintains consistency across the three syllable positions. The overall difference in slope values between /d/-initial vs /d/-final was only 0.05, which indicates the high stability of alveolar productions compared to labial (Δ slope = 0.265) and velar (Δ slope /g/ palatal = 0.17; Δ slope /g/ velar = 0.36) productions when onset and coda positions are compared.

Figure 3 shows distributional aspects of slope variability across the ten speakers as a function of stop consonant and

TABLE I. Locus equation statistics for labial /b/ across syllable positions for all speakers. k =slope; c = y intercept (Hz); SE=standard error of estimate (Hz).

Speaker:	Syllable-Initial				Syllable-Medial				Syllable-Final			
	k	c	R^2	SE	k	c	R^2	SE	k	c	R^2	SE
M#1	0.75	275	0.94	109	0.79	160	0.96	95	0.68	274	0.78	189
M#2	0.74	280	0.97	75	0.81	203	0.98	61	0.74	104	0.77	222
M#3	0.82	185	0.95	82	0.79	235	0.96	63	0.62	254	0.64	184
M#4	0.75	263	0.95	99	0.68	377	0.95	77	0.52	373	0.88	104
M#5	0.61	445	0.90	115	0.69	366	0.95	94	0.45	477	0.77	142
F#1	0.71	378	0.93	130	0.74	367	0.96	96	0.26	714	0.48	174
F#2	0.88	98	0.94	112	0.81	190	0.94	105	0.34	597	0.50	169
F#3	0.78	269	0.95	110	0.77	298	0.97	82	0.49	453	0.79	151
F#4	0.84	294	0.96	114	0.84	202	0.95	112	0.58	448	0.75	204
F#5	0.77	236	0.95	111	0.73	306	0.93	106	0.34	584	0.54	192
Mean:	0.765	272	0.94	106	0.765	270	0.96	89	0.50	428	0.69	173
s.d.:	0.075	94.8	0.019	16.3	0.053	82.3	0.014	17.7	0.157	183	0.14	33.8

syllable position. The upper left quadrant shows the syllable position effect for /b/. Syllable-initial and medial contexts have similar median values and interquartile ranges, while syllable-final /b/ is characterized by a lower median value and greater overall variability. The coefficient of variation (s.d./mean) for syllable-final /b/ was 0.313 compared to 0.098 and 0.07 for initial and medial positions, respectively. The upper right quadrant shows /d/ slopes for the three syllable positions. Median slope values are similar across all three positions with a greater range of variability for syllable-final positions (coefficient of variation=0.279 compared to 0.19 and 0.198 for initial and medial positions, respectively). The lower left quadrant shows /g/ palatal slopes across syllable positions. Once again median slope values and interquartile ranges for initial and medial position are quite similar compared to syllable-final values that show higher overall slopes and a greater range of variation. The lower right quadrant shows /g/ velar slope comparisons across syllable position. Syllable-initial and -medial slopes are again more similar compared to syllable-final values.

Slope values were entered into a 4 (Stop: /b/ vs /d/ vs /g/ pal vs /g/ vel) \times 3 (Position: initial versus medial versus final) repeated measures multivariate analysis. This MANOVA resulted in a significant effect for stop [$F(3, 27) = 56.11$,

$p < 0.01$]. Bonferroni tests explored the source of this effect. The results of these post hoc tests are given in Table V. Significant differences (at the 0.01 level) were obtained for all mean comparisons (collapsed across syllable positions) except /d/ (0.362) vs /g/ pal (0.316). A significant interaction was found between stop consonant and syllable position [$F(6, 54) = 8.72$, $p < 0.01$]. The sources for this effect were explored with Bonferroni tests and shown in Table VI. Final /b/ (0.502) was significantly different from initial (0.765) and medial (0.765) /b/ at the 0.01 level. No significant differences were found for /d/ slope means across the three syllable positions (0.371, 0.394, 0.321). For /g/ palatal, syllable-final slope (0.429) was found to be different from syllable-initial (at 0.05 level) and syllable-medial (at 0.01 level) means. Only the pairwise comparison of /g/ velar in final position (0.758) vs medial position (1.245) was found to be significant (at 0.01 level). Final /g/ velar versus initial /g/ velar means narrowly missed significance at the 0.05 level.

2. y-intercepts

Figure 4 shows distributional aspects of y-intercept variability across the ten speakers as a function of syllable position. For /b/ it can be seen that syllable-initial and medial

TABLE II. Locus equation statistics for alveolar /d/ across syllable positions for all speakers. k =slope; c = y intercept (Hz); SE=standard error of estimate (Hz).

Speaker:	Syllable-Initial				Syllable-Medial				Syllable-Final			
	k	c	R^2	SE	k	c	R^2	SE	k	c	R^2	SE
M#1	0.33	1377	0.84	77	0.29	1368	0.81	73	0.33	1257	0.57	163
M#2	0.31	1318	0.89	56	0.36	1204	0.87	61	0.36	1052	0.62	156
M#3	0.28	1304	0.69	40	0.33	1227	0.82	55	0.20	1262	0.61	58
M#4	0.32	1264	0.88	61	0.36	1214	0.84	67	0.40	1059	0.78	110
M#5	0.41	1140	0.82	109	0.45	1100	0.85	98	0.21	1403	0.47	131
F#1	0.42	1249	0.88	97	0.45	1251	0.90	87	0.20	1403	0.35	171
F#2	0.46	1157	0.86	89	0.43	1289	0.79	100	0.43	1017	0.68	142
F#3	0.41	1321	0.85	98	0.47	1272	0.84	111	0.31	1307	0.70	123
F#4	0.47	1296	0.87	100	0.51	1134	0.86	101	0.42	1263	0.64	185
F#5	0.30	1512	0.82	78	0.29	1629	0.69	95	0.35	1301	0.58	180
Mean:	0.37	1294	0.84	81	0.39	1269	0.83	85	0.32	1232	0.60	142
s.d.:	0.07	106	0.058	22.3	0.078	148	0.057	19.4	0.089	141	0.12	38.5

stops had similar median values with syllable-final *y* intercepts considerably higher. Variability (interquartile range) increased dramatically from initial-to-medial-to final position for labial productions. *Y* intercepts for /d/ were almost identically distributed for initial and medial positions, while considerably more variable for syllable final /d/. Both /g/ distributions show a greater range of values along the ordinate compared to labial and alveolar plots. For /g/ palatal utterances the syllable-medial position resulted in the smallest interquartile range, with the syllable-final position showing the largest spread. Interquartile ranges were basically similar for /g/ velar *y* intercepts with syllable-final position having the largest range between the 10th and 90th percentiles.

A second series of tests identical to those performed on the slope values were conducted on the *y*-intercept values. *Y*-intercept values were entered into a 4 (Stop: /b/ vs /d/ vs /g/ pal vs /g/ vel) × 3 (Position: initial versus medial versus final) repeated measures multivariate analysis. This MANOVA resulted in a significant effect for stop [$F(3, 27) = 103.6, p < 0.01$]. Bonferroni tests explored the sources of this effect and the results are shown in Table VII. All pairwise comparisons were significant except /b/ (323.5 Hz) vs /g/ velar (206.3 Hz). Once again, an interaction between stop and position was found [$F(6, 54) = 3.98, p < 0.01$] and Table VIII shows the results of the Bonferroni tests. For labial /b/ only the comparison between final position (427.8 Hz) vs medial position (270.4 Hz) reached significance at the 0.05 level. Once again no significant differences occurred for alveolar /d/ across syllable positions. For /g/ palatal the only significant comparison was between medial vs final position (at the 0.05 level), and for /g/ velar no pairwise comparisons reached significance at the 0.05 level.

Figure 5 shows the dispersion of “place clustering” in locus equation coefficients space (slope × intercept). Syllable-initial and -medial stops were combined for CV forms and compared to syllable-final stops (VC). It can be seen that there is considerably greater dispersion of place clusters for stops in CV form throughout locus equation space (unfilled shapes) compared to VC final stops. While syllable-final stops are kept apart in this space, they exhibit a more compact distribution with /d/ and /g/ palatal relatively close in both slope and intercept parameters.

3. R^2 and standard error of estimate analyses

A series of repeated measures ANOVAs were separately performed for each stop consonant using R^2 and SE values as dependent measures. Significant main effects for syllable position were found for /b/—(R^2 : [$F(2, 27) = 33.55, p < 0.0001$], SE: [$F(2, 27) = 30.02, p < 0.001$]); /d/—(R^2 : [$F(2, 27) = 25.67, p < 0.0001$], SE: [$F(2, 27) = 14.97, p < 0.0001$]); and /g/ velar—(R^2 : [$F(2, 27) = 19.66, p < 0.0001$], SE: [$F(2, 27) = 14.17, p < 0.0001$]). All subsequent comparisons between treatment means (using Fisher’s PLSD, the Scheffe, and the Bonferroni/Dunn) showed the same pattern of results, viz., no difference between means across syllable-initial and -medial positions, but highly significant differences between each of these positions versus syllable-final means. There was no significant main effect due to syllable position using R^2 for /g/ palatal productions,

but SE values showed a significant main effect [$F(2, 27) = 8.65, p = 0.002$] with subsequent comparisons all showing a significant difference between syllable-final (105 Hz) vs syllable-initial (61 Hz) and syllable-medial (62 Hz) means. In general, only syllable-initial and medial stops revealed the characteristic locus equation form of low SEs, indicating tight clustering of *F2* coordinates around a linear regression fit. VC locus equations showed a poorly correlated and relatively unlawful relationship between midvowel frequencies and offset loci for all the stop consonants.

B. Consonantal context effects on CV and VC coarticulation in CVCs

In general, the effect of C_2 on C_1V coarticulation and the effect of C_1 on VC_2 coarticulation, as documented by changes in locus equation slopes, was extremely minimal for all but one context— C_1 effects on VC_2 coarticulation for /g/ velars. Table IX presents mean slope values and standard deviations for each CV/(C) and (C)VC context. Looking first at C_2 consonantal effects on C_1V transitions, it can be seen that slope values remained fairly constant for all stop place categories regardless of final consonant identity. Initial labial and alveolar stops had the smallest net slope change across the three C_2 contexts with an overall standard deviation (averaged across the two categories) of only 0.077. Initial /g/ palatal slopes showed a slightly greater variation due to C_2 contextual effects, but nevertheless only varied by 0.06 between the lowest mean value (for final /d/ contexts) vs the highest (for final /g/ contexts). Initial velar /g/’s showed the highest standard deviations as a function of final consonant, but mean slope values only ranged from 1.20 to 1.24. Surprisingly, only initial /b/ VC contexts showed a significant context effect [$F(2, 18) = 7.598, p = 0.004$], with [b]Vb (slope = 0.744) being significantly different from [b]Vd (slope = 0.787). All other main effects were nonsignificant in repeated measures ANOVAs.

The effect of initial consonant on VC_2 slope values was slightly greater than the reverse contextual effect described above. Across stop categories there was a mean increase in standard deviations of 0.124, with the majority of the increase contributed by the extensive variability seen in CV[*g_v*] productions. Syllable-final alveolars changed very little from syllable-initial alveolars and also resisted C_1 contextual effects. A series of repeated measures ANOVAs showed nonsignificant context effects for all stops except final /g/ velars where all three slopes were found to be significantly different in post hoc tests. The extreme interspeaker variability in /g/ offset slopes in back vowel contexts precludes any definitive statements regarding the “significant” contextual effects. One subject in particular (M#3) evidenced very unusual slopes for CV /g/ velar productions (0.13, −0.89, −0.59 for C_1 = /b/, /d/, /g/, respectively) as three back vowel contexts (/ɑ/ɑ/) exhibited an unusual horizontally oriented distribution, similar to front vowel contexts, and hence a highly irregular regression fit resulted with negative slopes. Eliminating M#3’s CVCs with final /g/ velars, resulted in the following grand means (standard deviations) for bV[*g_v*], dV[*g_v*], gV[*g_v*] tokens, respectively: 1.18 (0.478), 0.78 (0.288), 1.01 (0.445).

TABLE III. Locus equation statistics for /g/ palatal across syllable positions for all speakers. k =slope; c = y intercept (Hz); SE=standard error of estimate (Hz).

Speaker:	Syllable-Initial				Syllable-Medial				Syllable-Final			
	k	c	R^2	SE	k	c	R^2	SE	k	c	R^2	SE
M#1	0.16	1886	0.24	44	0.27	1643	0.38	62	0.28	1763	0.33	50
M#2	0.37	1370	0.61	51	0.25	1689	0.49	50	0.71	764	0.63	132
M#3	0.19	1603	0.33	41	0.26	1511	0.50	46	0.33	1436	0.55	42
M#4	0.32	1264	0.60	69	0.16	1893	0.30	58	0.21	1823	0.14	131
M#5	0.41	1140	0.48	59	0.19	1940	0.54	52	0.48	1250	0.31	178
F#1	0.02	2623	0.01	65	0.35	1816	0.64	71	0.61	1138	0.51	135
F#2	0.23	1786	0.37	72	0.11	2132	0.14	71	0.16	2075	0.30	60
F#3	0.39	1602	0.54	70	0.31	1861	0.64	65	0.60	1134	0.66	123
F#4	0.29	1870	0.53	77	0.36	1711	0.67	69	0.39	1808	0.55	93
F#5	0.25	1963	0.55	64	0.30	1932	0.64	75	0.52	1370	0.67	103
Mean:	0.263	1711	0.43	61	0.256	1813	0.49	62	0.429	1456	0.465	105
s.d.:	0.12	425	0.19	12.2	0.08	178	0.174	10	0.184	404	0.182	43.6

C. Initial vowel context effects on CV coarticulation in VCVs

Mean slope and standard deviations averaged across the ten speakers for each of the initial vowel context conditions are shown in Table X. It can be seen that CV locus equation slopes remain constant despite initial vowel contexts ranging from a high front spread vowel, to a low front vowel, to a high back rounded vowel. No consistent pattern of slope changes emerged from these varying initial vowel contexts. Table XI shows absolute slope differences between each V_1 context (derived from an $N=50$) compared to the slope of the joint locus equation function when all three V_1 contexts were combined (derived from an $N = 150$). For labial, alveolar, and /g/ palatal productions slope differences from the mean ranged from 0.02 to 0.05. Only /g/ velar productions with initial /æ/ and /u/ caused a sizeable difference in slope (0.17 and 0.10, respectively), and this was found to be primarily attributable to one speaker (F#5) who had decreased slope values in these two contexts (1.36 and 1.32) compared to the mean /g/ velar slope of 1.78 across all three initial vowels. Despite the relatively small differences in slopes there were significant main effects due to V_1 context in a series of repeated measures ANOVAs. For VbV utterances [$F(2, 18)=7.67, p=0.004$] /i/ bV (0.787) was sig-

nificantly different from /u/ bV (0.721), for VdV utterances [$F(2, 18)=9.19, p=0.0018$] /i/ dV (0.36) was significantly different from /u/ dV (0.419), for VgV palatal utterances [$F(2, 18)=0.037, p=0.037$] /i/ gV (0.215) was significantly different from /æ/ gV (0.284). There was no significant effect of V_1 on /g/ velar productions.

III. DISCUSSION

A locus equation analysis of CV vs VC syllable structure provides three measures of comparison: (1) slope: an index of degree of coarticulation; (2) R^2 and SE: linearity and tightness of clustering of scatterplots; and (3) slope/ y -intercept coordinates plotted in locus equation-defined acoustic space. Significant differences were found between onset versus coda and medial versus coda position for all dependent variables, but, as expected, no differences were found between onset versus medial position, as both are produced with prevocalic stop consonants.

The mean slope of all CV locus equations (averaged across initial and medial stops) was 0.65 compared to a mean of 0.50 for VC forms. Similar to Lindblom (1963), labial slopes decreased from CV (slope=0.77) to VC (slope=0.50) form, implying greater coarticulatory influence of the medial

TABLE IV. Locus equation statistics for /g/ velar across syllable positions for all speakers. k =slope; c = y intercept (Hz); SE=standard error of estimate (Hz).

Speaker:	Syllable-Initial				Syllable-Medial				Syllable-Final			
	k	c	R^2	SE	k	c	R^2	SE	k	c	R^2	SE
M#1	1.01	398	0.64	176	1.57	-220	0.68	173	1.48	-398	0.42	299
M#2	1.02	344	0.40	163	1.14	152	0.39	154	1.04	112	0.27	252
M#3	0.92	600	0.68	94	0.80	671	0.44	120	-0.39	1835	0.03	329
M#4	1.22	130	0.64	159	1.06	329	0.43	197	0.64	490	0.24	197
M#5	1.18	110	0.84	101	1.03	231	0.59	131	1.11	-123	0.51	175
F#1	1.29	-139	0.72	164	1.27	-56	0.72	153	0.75	388	0.18	302
F#2	0.83	445	0.59	126	1.27	4	0.68	167	0.72	372	0.26	217
F#3	1.05	244	0.73	141	1.23	-34	0.92	104	0.41	712	0.41	125
F#4	1.37	-168	0.89	132	1.3	-176	0.82	151	1.08	226	0.37	369
F#5	1.33	23	0.65	187	1.78	-638	0.56	217	0.743	325	0.13	296
Mean:	1.12	199	0.68	144	1.24	26	0.62	157	0.76	394	0.28	252
s.d.:	0.18	253	0.135	31	0.28	353	0.174	34	0.50	596	0.15	80

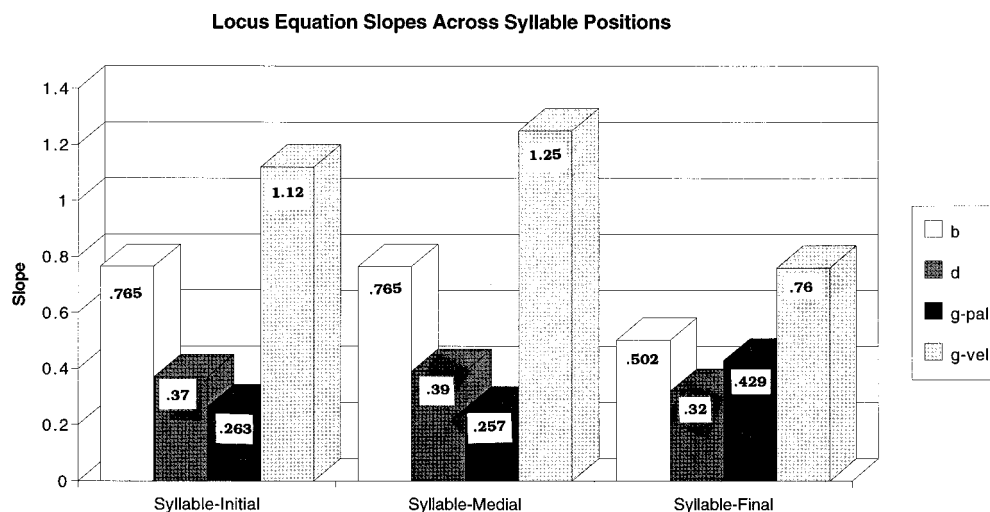


FIG. 2. Bar graph showing mean slope values for /b/, /d/, /g palatal/, and /g velar/ productions across syllable-initial, -medial, and -final positions.

vowel on initial /b/ compared to final /b/. Unlike Lindblom's single speaker, however, our results for /d/ showed highly stable coarticulatory effects with only a small decline in mean slope from CV to VC form (CV=.38; VC=.32). The velar data, when fit with a *single* regression function,

showed an increase in slope from CV (0.76) to VC (1.02) form as did Lindblom's /g/ data. However, this effect was not uniform across vowel place. When analyzed by allophones, opposite effects emerged. In front vowel contexts there was an *increase* in slope from 0.26 to 0.43, and in back

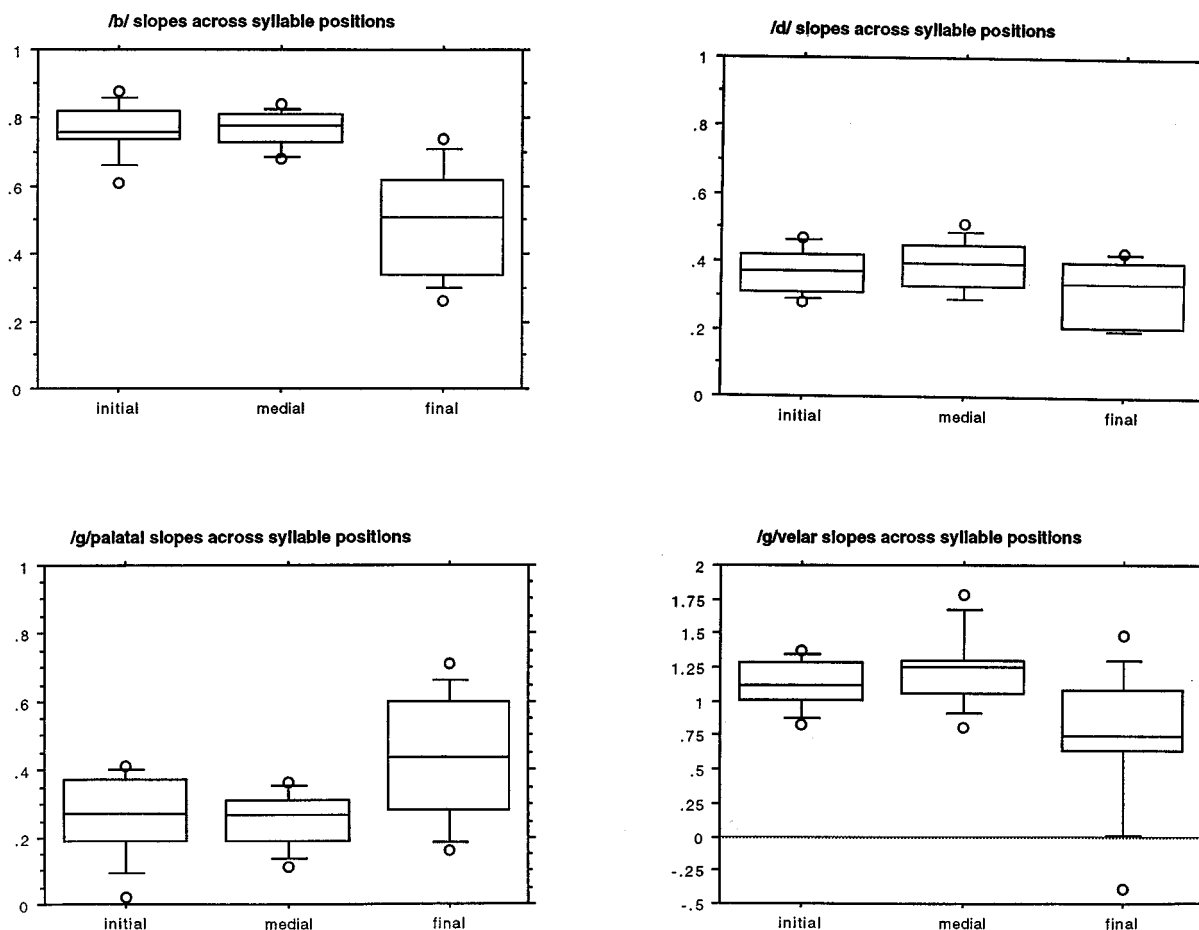


FIG. 3. Distributional ranges of slope values for each stop place category across syllable positions. Each box represents the interquartile range (bottom =25th percentile, top=75th percentile); the horizontal line within each box shows median slope; dashes above and below each box represent the 90th and 10th percentiles, respectively, with outliers shown as circles above and below these limits.

TABLE V. Results of Bonferroni post hoc tests on pairwise comparisons of slope means for each consonantal category collapsed across syllable positions.

	/b/	/d/	/g/ palatal	/g/ velar
Mean (s.d.)	0.677 (0.072)	0.362 (0.059)	0.316 (0.09)	1.04 (0.259)
/b/	...	*	*	*
/d/	ns	*
/g/ palatal	*
/g/ velar

*=*t* is significant at the 0.01 level.

vowel contexts there was a *decrease* in slope from 1.18 to 0.76. Overall, degree of coarticulation, as reflected by locus equation slope, showed a declination from syllable-initial to syllable-final position, with the sole exception of /g/ with front vowels.

R^2 and SE provide a clearer picture of the relative disintegration of the characteristic form of locus equation scatterplots for VCs compared to CVs. Pooling syllable-initial and syllable-medial stops, mean R^2 was 0.72 compared to a mean of 0.44 in coda position. The decrease in R^2 was uniformly seen across all stops (except for /g/ in front vowel contexts where R^2 was constant): /bV/=0.95, /Vb/=0.69; /dV/=0.83, /Vd/=0.60; /gV_{velar}=0.65, /Vg_{velar}=0.28. The mean SE across all onset and medial stops was 98 Hz compared to 168 Hz for stops in final position. The degradation of classic locus equation form in VC plots reflects a decline in acoustic orderliness for stops in coda position relative to pre-vocalic stops. CV stops were also maximally dispersed

TABLE VI. Results of Bonferroni post hoc tests on pairwise comparisons of slope means for the interaction between stop consonant and syllable position.

	Initial /b/	Medial /b/	Final /b/
Mean (s.d.)	0.765(0.075)	0.765(0.053)	0.502(0.157)
Initial /b/	...	ns	**
Medial /b/	**
Final /b/

**=*t* is significant at the 0.01 level.

	Initial /d/	Medial /d/	Final /d/
Mean (s.d.)	0.371(0.07)	0.394(0.078)	0.321(0.089)
Initial /d/	...	ns	ns
Medial /d/	ns
Final /d/

*=*t* is significant at the 0.05 level.
**=*t* is significant at the 0.01 level.

	Initial /g/ pal	Medial /g/ pal	Final /g/ pal
Mean (s.d.)	0.263(0.12)	0.256(0.081)	0.429(0.184)
Initial /g/ pal	...	ns	*
Medial /g/ pal	**
Final /g/ pal

*=*t* is significant at the 0.05 level.
**=*t* is significant at the 0.01 level.

	Initial /g/ vel	Medial /g/ vel	Final /g/ vel
Mean (s.d.)	1.122(0.183)	1.245(0.276)	0.758(0.504)
Initial /g/ vel	...	ns	ns
Medial /g/ vel	**
Final /g/ vel

**=*t* is significant at the 0.01 level.

within locus equation-defined space relative to a tighter clustering for VCs, although it appears that the latter still maintain a sufficient degree of separability (see Fig. 5).

A. The nature of coarticulatory interactions in CVCs

Coarticulation remains controversial as the term evokes a variety of theoretical conceptualizations for speech motor control and its neural instantiation (e.g., Henke, 1966; Daniloff and Hammarberg, 1973; Hammarberg, 1976, 1982; Fowler, 1980; Whalen, 1990). Applying locus equation results to better understand coarticulation has theoretical limitations as the metric merely quantifies *degree* of coarticulation. However, if the underlying basis for this quantification is accepted (see Krull, 1988; Sussman *et al.*, 1993), then some insights related to the relative contextual influence of consonants and vowels on each other can be logically arrived at. The premise for locus equation slope indexing degree of coarticulation is based on the view that the vowel exerts primary contextual influence on the F_2 pattern of preceding and following stops. A vowel-dominant view supports a traditional rather than a co-production account of coarticulation (see Fowler, 1980). Rather than a shared overlap of C and V gestural realizations, a directional account arises—a R-L anticipatory effect on C_1 onset frequency and concurrently a “carryover” L-R effect on C_2 offset frequency. C-to-V effects (carryover C_1 effects and anticipatory C_2 effects on the vowel) are not considered as “phonetic players” shaping locus equation plots. The “vowel-dominant” view, if taken to a logical conclusion, would also claim that the V-to- C_1 anticipatory interaction is more likely a planned articulatory sequence both demanding and realizing greater articulatory precision compared to VC₂ articulation which is conceived more as a defaultlike, mechanical contextual effect. Another way of stating this is: If greater articulatory precision is required for onsets relative to offsets, then let the brain handle onsets. A carefully orchestrated and systematic pattern of “absorbing” F_2 variability would more likely derive from neural programming control than simple bio-mechanical effects.⁴

The type of coarticulatory interaction usually attributed to C_2 effects on the medial vowel is the inertia-related anticipatory lowering of the velum for a final nasal stop. Krakow (1989) has demonstrated greater nasal assimilation on the vowel due to a final nasal compared to an initial nasal. C_2 anticipatory effects on medial vowels have not been evidenced when place cues, F_2 transitions, are scrutinized. In CVCs with initial and final *oral* stops, the speech motor control system appears to opt for greater articulatory primacy for the initial obstruent. If final oral stops exerted anticipatory (hence, planned) coarticulation on the vowel, then one would expect more systematic locus equation orderliness as consistently shown for syllable-onset stops. It is difficult to conceive of a medial vowel exerting R-L anticipatory effects on C_1 , while at the same time being influenced by R-L effects of C_2 . It is also difficult to imagine a purely coproduction account with uniformly shared overlapping of C and V elements since there is phonetic and acoustic asymmetry in onset relative to coda positions. At present it is not possible to dissociate independent consonantal effects from those prima-

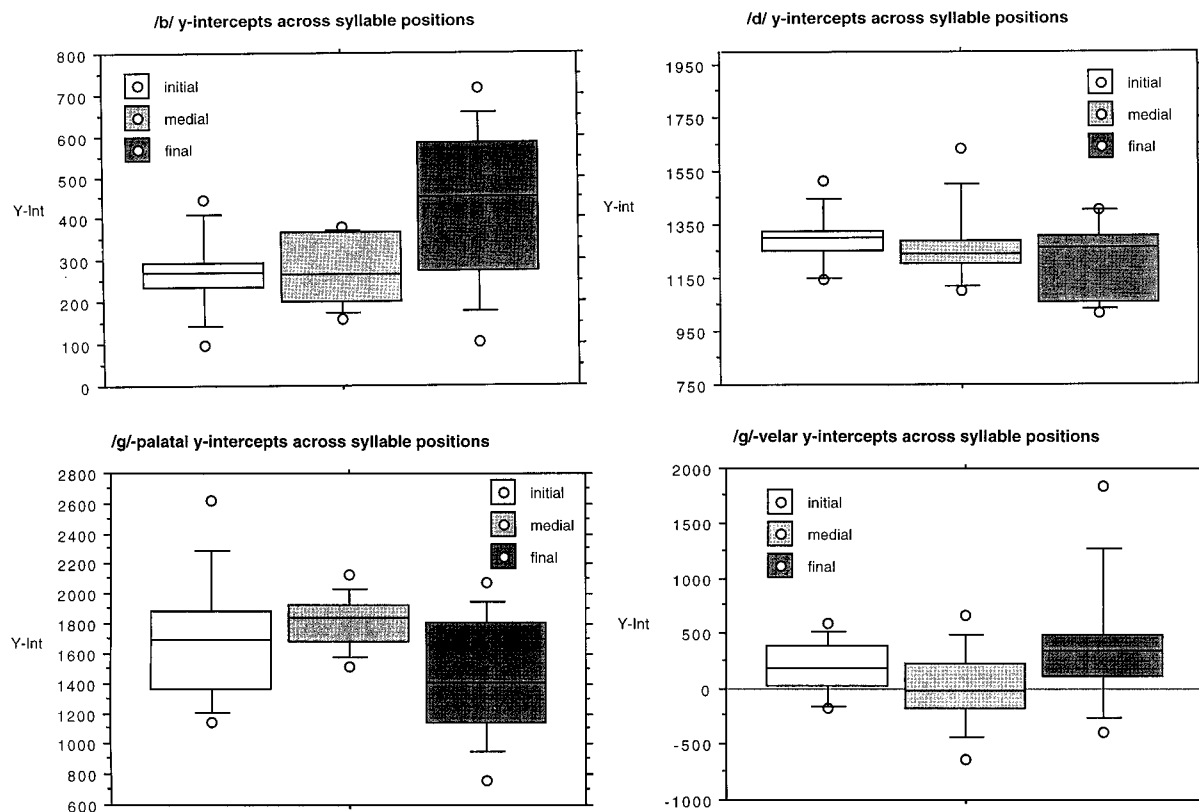


FIG. 4. Distributional ranges of y-intercept values for each stop place category across syllable positions.

rily contributed by the vowel. One can remove the coarticulated F_2 vowel as the independent variable in the regression analysis and substitute a coarticulation-neutral vowel (e.g., /hV/), but F_2 onset frequencies would still reflect the following vowel context and thus preclude assessment of the consonant's contribution to the coarticulated CV unit.

With respect to the acoustic coding of stop place of articulation, the role of the F_2 transition for VCs is clearly diminished relative to CVs. The acoustic breakdown of locus equation orderliness for coda relative to initial stops is interpreted in terms of articulatory declination. An aerodynamic correlate of articulatory declination has been documented by Stathopoulos and Weismer (1985) who showed systematic decreases in peak oral airflow for /t/ and /d/ from syllable-initial to syllable-final position. Perhaps the presence of a final release, in combination with top-down semantic context effects, can mitigate (for perceptual purposes) the articulatory and acoustic declination of syllable-final stops as cued

by the F_2 transition. In sum, a locus equation view of CV coarticulation posits an elegantly controlled articulatory process that helps to encode and contrast stop place of articula-

TABLE VIII. Results of Bonferroni post hoc tests on pairwise comparisons of y-intercept (Hz) means for the interaction between stop consonant and syllable position.

	Initial /b/	Medial/b/	Final /b/
Mean (s.d.)	272(94.8)	270(82.3)	428(182.6)
Initial /b/	...	ns	ns
Medial /b/	*
Final /b/
*=t is significant at the 0.05 level.			
	Initial /d/	Medial /d/	Final /d/
Mean (s.d.)	1294(106.1)	1269(147.6)	1232(141.4)
Initial /d/	...	ns	ns
Medial /d/	ns
Final /d/
	Initial /g/ palatal	Medial /g/ palatal	Final /g/ palatal
Mean (s.d.)	1711(424.6)	1813(178.5)	1456(404)
Initial /g/ pal	...	ns	ns
Medial /g/ pal	*
Final /g/ pal
*=t is significant at the 0.05 level.			
	Initial /g/ velar	Medial /g/ velar	Final /g/ velar
Mean (s.d.)	199(253.1)	26(352.6)	394(595.8)
Initial /g/ vel	...	ns	ns
Medial /g/ vel	ns
Final /g/ vel

TABLE VII. Results of Bonferroni post hoc tests on pairwise comparisons of y intercepts for each consonantal category collapsed across syllable positions.

	/b/	/d/	/g/ palatal	/g/ velar
Mean (s.d.)	324(91.6)	1265(95.8)	1660(217.2)	206(330.4)
/b/	...	**	**	ns
/d/	**	**
/g/ palatal	**
/g/ velar

**=t is significant at the 0.01 level.

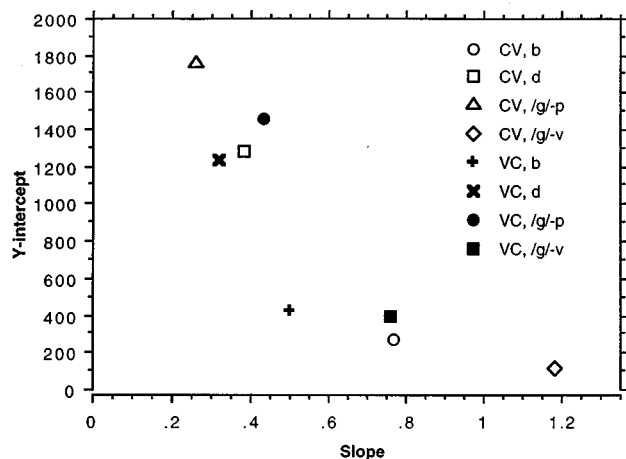


FIG. 5. Slope \times y-intercept coordinates for CV vs VC stops.

tion by (i) reducing variability due to vowel contexts, and (ii) differentially tweaking degree of coarticulation as a function of place of articulation. Such a view is in contrast to stereotyped impressions of coarticulation as an inelegant process analogous to the smashing of Easter eggs (Hockett, 1955), or accounts relying simply on temporal adjustments or phase shifts in overlapping gestures (Browman and Goldstein, 1992).

The data of this study can also address one of the classic findings in experimental phonetics. Ohman (1966) showed that initial and final vowel contexts in V_1CV_2 utterances greatly influenced the CV_2 and V_1C transition *across* the intervocalic consonant. Ohman's Fig. 5 (p. 158) shows stylized F_2 trajectories with extensive ranges of overlap across [b], [d], and [g] when produced in medial position with different V_1 and V_2 contexts. This context-dependent variability in the F_2 transitional pattern across the diphthongal (V_1 -to- V_2) gesture lent support to the view that the superimposed and independent stop consonant gesture could not be acoustically characterized with an invariant F_2 signal. Our data present a more encouraging picture of stop consonant acoustic contrastability as locus equation scatterplots were

TABLE X. Mean locus equation slopes and standard deviations for V_1 context effects on CV coarticulation in VCV tokens.

VCV	Slope	Standard deviation
[i]bV	0.787	0.059
[æ]bV	0.777	0.073
[u]bV	0.721	0.055
[i]dV	0.360	0.072
[æ]dV	0.398	0.080
[u]dV	0.419	0.088
[i]g _p V	0.215	0.059
[æ]g _p V	0.284	0.089
[u]g _p V	0.241	0.121
[i]g _v V	1.226	0.266
[æ]g _v V	1.126	0.262
[u]g _v V	1.199	0.199

shown to be only minimally influenced by V_1 contexts ([i æ u]). A crucial distinction between the Ohman (1966) results and those of the present study is that Ohman's plots always included a *fixed* vowel context in either the V_1 or V_2 position in relation to only four transconsonantal vowel contexts. When (V)CV transitions are plotted across a more complete vowel space, the linear and lawful vowel-normalization patterning of the F_2 transitions is readily apparent and not significantly altered by initial vowel contexts.

B. Perceptual implications

The decline of acoustic lawfulness of syllable-final stops relative to syllable-initial stops should have parallels in perceptual distinctiveness. Word-initial phonemes have long been considered to be especially salient in guiding both lexical access (Marslen-Wilson, 1987) and auditory word recognition (Pisoni *et al.*, 1985), but direct perceptual tests of consonant identification as a function of syllable-initial vs syllable-final position are not plentiful. Ahmed and Agrawal (1969) generated separate confusion error matrices for syllable-initial and syllable-final Hindi consonants ($N=29$) surrounding 10 vowels in CVC nonsense syllables. For stop plosives the number of correct identifications were 516, 518,

TABLE IX. Mean locus equation slopes and standard deviations for C_2 context effects on CV coarticulation and C_1 context effects on VC coarticulation.

CV(C)s	Slope	SD	(C)VCs	Slope	SD
[b]Vb	0.744	0.081	bV[b]	0.526	0.139
[b]Vd	0.787	0.073	dV[b]	0.497	0.191
[b]Vg	0.766	0.068	gV[b]	0.485	0.166
[d]Vb	0.364	0.091	bV[d]	0.311	0.079
[d]Vd	0.365	0.068	dV[d]	0.354	0.089
[d]Vg	0.388	0.080	gV[d]	0.307	0.129
[g _p]Vb	0.226	0.114	bV[g _p]	0.401	0.190
[g _p]Vd	0.203	0.080	dV[g _p]	0.407	0.148
[g _p]Vg	0.264	0.124	gV[g _p]	0.466	0.261
[g _v]Vb	1.24	0.302	bV[g _v]	1.079	0.561
[g _v]Vd	1.20	0.211	dV[g _v]	0.613	0.594
[g _v]Vg	1.22	0.390	gV[g _v]	0.849	0.657

TABLE XI. Mean absolute slope difference between each V_1 context and the slope derived from utterances summing across all three V_1 contexts.

$V_1 =$	/i/	/æ/	/u/
/b/	0.03	0.03	0.04
/d/	0.03	0.02	0.03
/g/ palatal	0.05	0.03	0.05
/g/ velar	0.07	0.17	0.10

533 for syllable-initial /bdg/ compared to 452, 485, 496 for syllable-final /bdg/, respectively. In analyzing the error substitutions they noted that “*in the initial position, the confusions are generally according to the manner of articulation rather than according to the place of articulation while in the final position the confusions are more frequently according to the place of articulation*” (p. 760). This finding nicely dovetails with our results showing locus equations for CV transitions encoding systematic differences across place contrasts relative to poorer place contrastiveness for the VC transitions of final stops. Slightly different results were obtained by Wang and Bilger (1973) who analyzed consonant confusions with and without noise for both CV and VC syllable sets using three vowel contexts (/i a u/). In the “no noise” condition there was very little difference between total number of correct identifications of /bdg/ from CV vs VC stimulus sets. For listening in noise initial /b/<final /b/, initial /d/=final /d/, and initial /g/>final /g/. Thus results from /d/ and /g/ agree with the locus equation results, while results for /b/ are opposite to what we would predict. A recent study (Redford and Diehl, 1996) using natural speech CVC nonsense syllables tested syllable-initial and syllable-final consonant perception for /p t k f θ s j/ with three medial vowels (/i a u/). For the three voiceless stops errors were significantly greater for syllable-final position relative to syllable-initial position (as was also found for all consonants examined). The authors concluded that “*in naturally produced CVC syllables, initial consonants are more perceptually distinctive than final consonants.*”

The phonetic differences shown by CV and VC phonological structures are also evident from the noneffect consistently reported in adaptation studies whereby initial stops fail to adapt the same stop when in final position (Ades, 1974; Miller and Eimas, 1976; Pisoni and Tash, 1975; Samuel *et al.*, 1984; Sawusch, 1977; Wolfe, 1978). Samuel (1989) using several paradigms to counteract the acoustic cancellation of phonetic effects during the adaptation procedure, concluded that “*the perceptual process treats syllable-initial consonants and syllable-final ones as inherently different* (p. 485).” Both the adaptation studies and the articulatory/acoustic results of this study add credence to Fujimura’s (1976) idea that the demisyllable might be a more realistic unit for perceptual processing than the syllable per se.

In sum, CV and VC entities appear to be phonetically and hence perhaps, representationally, fundamentally dissimilar. Locus equation data strongly suggest that the CV F_2 transition is more precisely controlled than the VC transition to bring about a lawful and emergentlike property—the vowel-normalization of the variable F_2 transitions encoding stop place contrasts.

ACKNOWLEDGMENTS

This research was supported by Grant No. R01 DC2014-01A1 from the National Institute on Deafness and Other Communication Disorders to the first author. The authors would like to acknowledge the constructive comments of two anonymous reviewers and the editorial assistance of Anders Löfqvist. Special thanks to Gary Weismer and Randy Diehl for providing some helpful suggestions during the final stages of revising the manuscript.

¹Traditional phonetic descriptions of the place feature deal with surface properties—where along the vocal tract is the occlusion/constriction formed? Such descriptions fail to consider derived, higher-order, and possibly more crucial, linguistic properties of the speech signal that only first emerge when relevant parameters are viewed at a higher, categorical level.

²Admittedly, /g/ is a difficult category to reconcile in classic locus equation terms. Much debate has been devoted to this issue (e.g., Fowler, 1994) without achieving satisfactory closure. Our rationale for depicting /g/ via two subgroups is based on the fact that each allophonic cluster is distinctively different in its linear distribution. Forcing one “Procrustean” fit (as Lindblom originally referred to the /g/ regression line) through two clearly distinct populations of coordinates does not seem intuitively reasonable at this stage of our understanding of this empirical phenomenon. Since it is felt that a significant aspect of the locus equation signature is their highly linear nature, it seems reasonable to treat the allophonic split as a “dual-mapping” phenomenon that does not prejudice a higher level categorical identity for /g/.

³Extremely high agreement was found among the three analysis methods and across measurers. The mean standard deviation across slope and intercept values obtained for [bdg] utterances ($N=150$ coordinates per locus equation plot for each stop) using Computerized Speech Lab, MacSpeech Lab, and Sound Scope was 0.013 and 32 Hz, respectively. The absolute frequency differences for F_2 onset and F_2 vowel among the three analysis systems ranged from a low of 25 Hz to a high of 41 Hz.

⁴The picture that emerges for CV relative to VC forms is analogous to driving a car with a standard transmission. Getting started without stalling is crucial and demands precise motor coordination (=CVs), while stopping (=VCs) is less precisely controlled as it matters little if the car stops two feet or one foot from the stop sign.

- Ades, A. E. (1974). “How phonetic is selective adaptation? Experiments on syllable position and vowel environment,” *Percept. Psychophys.* **16**, 61–66.
- Ahmed, R., and Agrawal, S. S. (1969). “Significant features in the perception of (Hindi) Consonants,” *J. Acoust. Soc. Am.* **45**, 758–763.
- Bell, A. (1971). “Some patterns of occurrence and formation of syllable structure,” *Working Papers in Language Universals* **6**, 23–137.
- Bell, A., and Hooper, J. B. (1978). “Issues and evidence in syllabic phonology,” in *Syllables and Segments*, edited by A. Bell and J. B. Hooper (North-Holland, Amsterdam), pp. 3–22.
- Browman, C. P., and Goldstein, L. (1992). “Articulatory phonology: An overview,” *Phonetica* **49**, 155–180.
- Byrd, D. (1994). “Articulatory timing in English consonant sequences,” *UCLA Working Papers in Phonetics* **86**, 1–196.
- Celdran, E. M., and Villalba, X. (1995). “Locus equations as a metric for place of articulation in automatic speech recognition,” *Proceedings of the XIIIth International Congress of Phonetic Sciences*, **1**, 3033, Stockholm, Sweden.
- Daniiloff, R. G., and Hammarberg, R. (1973). “On defining coarticulation,” *J. Phonetics* **1**, 239–248.
- Dorman, M. F., Raphael, L. J., Liberman, A. M., and Repp, B. (1975). “Some maskinglike phenomena in speech perception,” *Haskins Laboratories Status Report on Speech Research* **SR-42/43**, 265–276.
- Fowler, C. A. (1980). “Coarticulation and theories of extrinsic timing,” *J. Phonetics* **8**, 113–133.
- Fowler, C. A. (1994). “Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation,” *Percept. Psychophys.* **55**, 597–610.
- Fujimura, O. (1976). “Syllables as concatenated demisyllables and affixes,” *J. Acoust. Soc. Am. Suppl.* **1** **59**, S55.

- Fujimura, O. (1990). "Methods and goals of speech production research," *Language Speech* **33**, 195–258.
- Fujimura, O., Macchi, M. J., and Streeter, L. A. (1978). "Perception of stop consonants with conflicting transitional cues: A cross-linguistic study," *Language Speech* **21**, 337–347.
- Hammarberg, R. (1976). "The metaphysics of coarticulation," *J. Phonetics* **4**, 353–363.
- Hammarberg, R. (1982). "On redefining coarticulation," *J. Phonetics* **10**, 123–137.
- Henke, W. L. (1966). "Dynamic articulatory model of speech production using computer simulation," doctoral dissertation, Massachusetts Institute of Technology.
- Hockett, C. (1955). *Manual of phonology*, Publications in Anthropology and Linguistics, No. 11 (Indiana U.P., Bloomington, IN).
- Ito, J. (1989). "A prosodic theory of epenthesis," *Natural Language and Linguistic Theory* **7**, 217–259.
- Kahn, D. (1976). "Syllable-based generalizations in English phonology," Bloomington, IN: Indiana University Linguistics Club.
- Keating, P. (1995). "Effects of prosodic position on /t,d/ tongue/palate contact," *Proceedings of the XIIIth International Congress of Phonetic Sciences* **3**, 432–435.
- Krakow, R. A. (1989). "The articulatory organization of syllables: A kinematic analysis of labial and velar gestures," Ph.D. dissertation, Yale University, New Haven, CT.
- Krull, D. (1988). "Acoustic properties as predictors of perceptual responses: A study of Swedish voiced stops," *Phonetic Experimental Research at the Institute of Linguistics, University of Stockholm (PERILUS)* **VII**, 66–70.
- Lindblom, B. (1963). "On vowel reduction," Report #29, The Royal Institute of Technology, Speech Transmission Laboratory, Stockholm, Sweden.
- Lindblom, B. (1983). "Economy of speech gestures," in *The Production of Speech*, edited by P. MacNeilage (Springer-Verlag, New York), pp. 217–245.
- Manuel, S. (1991). "Some phonetic bases for the relative malleability of syllable-final versus syllable-initial consonants," *Proceedings of the XIIth International Congress of Phonetic Sciences* **5**, 118–121.
- Marslen-Wilson, W. (1987). "Functional parallelism in spoken word recognition," *Cognition* **25**, 71–102.
- Miller, J. L., and Eimas, P. D. (1976). "Studies on the selective tuning of feature detectors for speech," *J. Phonetics* **4**, 119–127.
- Nearey, T. M., and Shammass, S. E. (1987). "Formant transitions as partly distinctive invariant properties in the identification of voiced stops," *Can. Acoust.* **15**, 17–24.
- Ohman, S. E. G. (1966). "Coarticulation in VCV utterances: Spectrographic measurements," *J. Acoust. Soc. Am.* **39**, 151–168.
- Pisoni, D. B., and Tash, J. (1975). "Auditory property detectors and processing place features in stop consonants," *Percept. Psychophys.* **18**, 401–408.
- Pisoni, D. B., Nusbaum, H. C., Luce, P. A., and Slowiaczek, L. M. (1985). "Speech perception, word recognition and the structure of the lexicon," *Speech Commun.* **4**, 75–95.
- Redford, M. A., and Diehl, R. L. (1996). "A study on the relative perceptibility of syllable-initial and syllable-final consonants," *J. Acoust. Soc. Am.* **100**, 2693(A).
- Samuel, A. (1989). "Insights from a failure of selective adaptation: Syllable-initial and syllable-final consonants are different," *Percept. Psychophys.* **45**, 485–493.
- Samuel, A. G., Kat, D., and Tartter, V. C. (1984). "Which syllable does an intervocalic stop belong to? A selective adaptation study," *J. Acoust. Soc. Am.* **76**, 1652–1663.
- Sawusch, J. R. (1977). "Processing of place information in stop consonants," *Percept. Psychophys.* **22**, 417–426.
- Sproat, R., and Fujimura, O. (1993). "Allophonic variation in English /l/ and its implications for phonetic implementation," *J. Phonetics* **21**, 291–311.
- Stathopoulos, E. T., and Weismer, G. (1985). "An aerodynamic study of stress in children and adults," *J. Phonetics* **13**, 343–355.
- Sussman, H. M. (1994). "The phonological reality of locus equations across manner class distinctions: Preliminary observations," *Phonetica* **51**, 119–131.
- Sussman, H. M., and Shore, J. (1996). "Locus equations as phonetic descriptors of consonantal place of articulation," *Percept. Psychophys.* **58**, 936–946.
- Sussman, H. M., Fruchter, D., and Cable, A. (1995). "Locus equations derived from compensatory articulation," *J. Acoust. Soc. Am.* **97**, 3112–3124.
- Sussman, H. M., Hoemeke, K., and Ahmed, F. (1993). "A cross-linguistic investigation of locus equations as a relationally invariant descriptor for place of articulation," *J. Acoust. Soc. Am.* **94**, 1256–1268.
- Sussman, H. M., McCaffrey, H. A., and Matthews, S. A. (1991). "An investigation of locus equations as a source of relational invariance for stop place categorization," *J. Acoust. Soc. Am.* **90**, 1309–1325.
- Whalen, D. H. (1990). "Coarticulation is largely planned," *J. Phonetics* **18**, 3–35.
- Wang, M. D., and Bilger, R. C. (1973). "Consonant confusions in noise: A study of perceptual features," *J. Acoust. Soc. Am.* **5**, 1248–1266.
- Wolfe, C. G. (1978). "Perceptual invariance for stop consonants in different positions," *Percept. Psychophys.* **24**, 315–326.

Concurrent vowel identification. I. Effects of relative amplitude and F_0 difference^{a)}

Alain de Cheveigné

Centre National de la Recherche Scientifique/Université Paris 7, 2 place Jussieu, case 7003, F-75251 Paris
Cédex 05, France and ATR Human Information Processing Research Laboratories, 2-2 Hikaridai,
Seika-cho Soraku-gun, Kyoto 619-02, Japan

Hideki Kawahara, Minoru Tsuzaki, and Kiyooki Aikawa

ATR Human Information Processing Research Laboratories, 2-2 Hikaridai, Seika-cho Soraku-gun,
Kyoto 619-02, Japan

(Received 19 December 1995; revised 12 August 1996; accepted 22 November 1996)

Subjects identified concurrent synthetic vowel pairs that differed in relative amplitude and fundamental frequency (F_0). Subjects were allowed to report one or two vowels for each stimulus, rather than forced to report two vowels as was the case in previously reported experiments of the same type. At all relative amplitudes, identification was better at a fundamental frequency difference (ΔF_0) of 6% than at 0%, but the effect was larger when the target vowel amplitude was below that of the competing vowel (-10 or -20 dB). The existence of a ΔF_0 effect when the target is weak relative to the competing vowel is interpreted as evidence that segregation occurs according to a mechanism of cancellation based on the harmonic structure of the competing vowel. Enhancement of the target based on its own harmonic structure is unlikely, given the difficulty of estimating the fundamental frequency of a weak target. Details of the pattern of identification as a function of amplitude and vowel pair were found to be incompatible with a current model of vowel segregation.

© 1997 Acoustical Society of America. [S0001-4966(97)03904-0]

PACS numbers: 43.71.An, 43.71.Es, 43.66.Ba, 43.66.Lj [WS]

INTRODUCTION

Various experiments have shown that identification of two synthetic vowels that are mixed together improves when the vowels differ in fundamental frequency (F_0) (Scheffers, 1983; Summerfield and Assmann, 1991; Culling and Darwin, 1993, 1994; Assmann and Summerfield, 1994). The results of several of these studies are shown in Fig. 1. Despite differences in task, vowel set, subjects, etc., a common trend is that identification improves as the F_0 difference (ΔF_0) increases between vowels. However, it has been noted that identification is far above chance at $\Delta F_0=0\%$, and yet remains less than perfect when F_0 's are different.

The limited effect of F_0 differences may be due partly to ceiling effects. If identification is perfect at $\Delta F_0=0\%$ for certain subjects or conditions, there is no room left for improvement with ΔF_0 . De Cheveigné *et al.* (1995a) reasoned that such might be the case for one vowel within a pair if it dominated its companion. They tried to determine corrective amplitude factors to balance relative dominance and mutual masking. However, that reasoning was flawed: There is nothing to guarantee that, once the balance is attained, identification of *both* vowels will not be at ceiling. In the present study, we followed the opposite course and introduced a systematic amplitude imbalance to reduce ceiling effects for the weaker vowel. The first aim of the study was to test that idea, and determine good amplitude imbalance factors for future

experiments. For that purpose, we measured identification rate as a function of both ΔF_0 and relative amplitude between vowels. This experiment allows a link to be made between the classic paradigm in which identification rates are measured with constant stimuli (Fig. 1), and more recent paradigms in which thresholds of identification are determined adaptively (Demany and Semal, 1990; Assmann and Summerfield, 1990; Summerfield, 1992; Summerfield and Culling, 1992; Culling *et al.*, 1994; Culling and Summerfield, 1995). Finally, the dependency of identification on relative amplitude and ΔF_0 may be used to test theories of vowel perception or segregation.

McKeown (1992) performed a similar experiment using double-vowel stimuli with ΔF_0 's of 0%, 25%, and 100% (relative to the lower F_0). Relative amplitude between vowels was varied between -14 and 14 dB in 2-dB steps, by reducing the amplitude of either vowel from the equal-amplitude condition (thereby reducing the overall stimulus level). The experiment we report here used a larger range (-20 – 20 dB), a larger step (10 dB), and overall rms amplitude was held constant for all stimuli. McKeown required subjects to identify a "dominant vowel" and a second vowel, and analyzed results separately for each. We also scored the constituents of vowel pairs separately, but according to stimulus properties rather than a subjective judgement.

In the classic double vowel identification paradigm, subjects are presented with two vowels and requested to identify both of them, whether they are both audible or not. This has several drawbacks: (a) The task is uncomfortable when one vowel is inaudible and the subject must guess; (b) the subject may use one particular vowel as a default response, and thus

^{a)}Portions of this work were presented at a meeting of the Acoustical Society of Japan and as an ATR Human Information Processing Research Laboratories technical report (de Cheveigné, 1995; de Cheveigné *et al.*, 1995b).

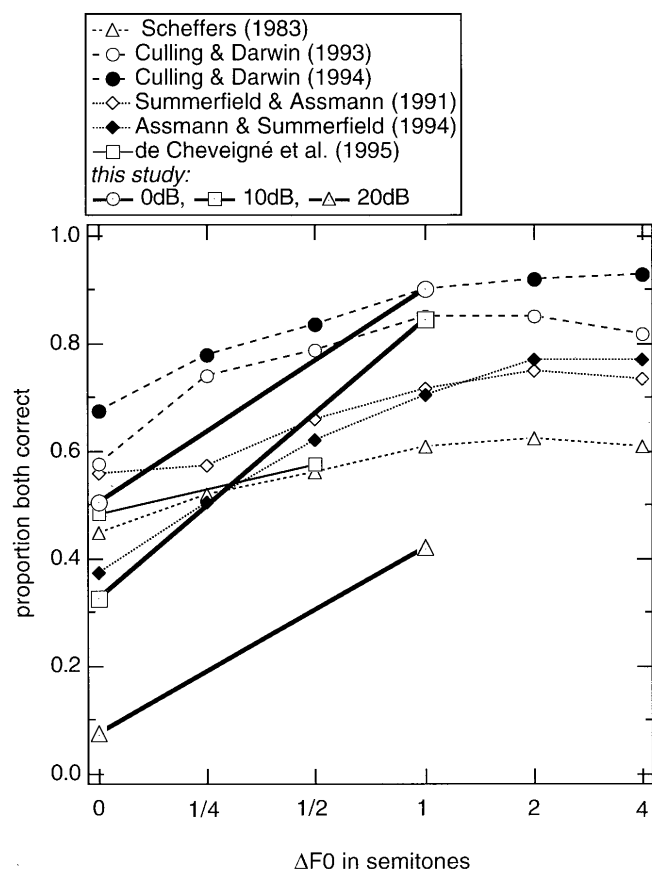


FIG. 1. Proportion of "both correct" identification as a function of ΔF_0 for previous studies (thin lines), and for this report (thick lines).

unwittingly score perfect identification on that vowel; (c) segregation cues that signal the multiplicity of sources are ignored; and (d) subjects are under pressure to improve their identification. This might enhance training effects that may contribute to reduce the size of ΔF_0 effects (Assmann and Summerfield, 1994). Instead of requiring two vowel responses, we told our subjects that the stimuli contained either single or double vowels, and we requested them to give either one or two answers. This task is typical of natural situations where segregation occurs, and the number of vowels reported is an interesting measure.

I. METHODS

Steady-state Japanese vowels /a/, /i/, /u/, /e/, and /o/ were synthesized at two fundamental frequencies: 125 and 132.5 Hz, with equal rms signal amplitude. Details of the vowels and synthesis technique are given in the Appendix. Double vowels were created by scaling one vowel by a factor (−20, −10, 0, 10, or 20 dB), adding the two vowels, and setting the rms amplitude of the sum to a standard value. Stimuli were 200 ms in duration, with 20-ms raised-cosine onset and offset ramps. They were presented to subjects via headphones (Stax SR-Λ), at a sound-pressure level between 63 and 70 dBA. The sound system was calibrated using a Bruel&Kjaer artificial ear (sound level meter type 2231, half-inch microphone type 4134).

For a given vowel pair, all four combinations of the two F_0 's were used to produce ΔF_0 's of 0% and 6%. Vowels within a pair were always different. There were three repetitions of each condition, for a total of 600 double-vowel stimuli ($2 \Delta F_0$'s) \times (5 amplitude differences) \times (10 unordered vowel pairs) \times (2 F_0 orders) \times (3 repetitions). To these were added 240 single vowel stimuli ($2 F_0$'s) \times (5 vowels) \times (24 repetitions). A relatively large number of single vowels were included to ensure that the stimulus set was consistent with the description given to the subjects. It also allowed us to verify that vowel quality was acceptable and unaffected by synthesis parameters such as F_0 .

Subjects were six native speakers of Japanese, aged 18 to 26 years. Two were male (K, S) and four female (N, M, T, U). Two were members of the ATR staff (N, M), and the other four were students paid for their services. Each subject performed five sessions on different days. Subjects were seated in a sound-treated booth or room, in front of a computer terminal that was used to give prompts and gather results. Each stimulus was presented once, and the subjects were requested to give one or two responses, according to the number of vowels they perceived to be present. The subjects were informed that the stimulus contained one or two vowels belonging to the set /a/, /i/, /u/, /e/, /o/, and that, in the case of double vowels, vowels within a pair were different. They were not allowed to answer twice the same vowel, but they had the option to answer "x" instead of a vowel that they could not identify.¹ They could pause at will, in which case the last stimulus before the pause was presented again after the pause. A session typically lasted between one and two hours. There was no feedback.

The answers were scored to obtain three measures: (a) Average number of vowels reported per stimulus. This is the proportion of stimuli that elicited two responses, regardless of whether they were correct or not, and regardless of whether the stimulus contained one vowel or two (b) *Combination-correct identification rate*. This is the proportion of double-vowel stimuli for which both constituents were correctly identified. It is the measure most commonly reported for double-vowel experiments (Fig. 1) (c) *Constituent-correct identification rate* (Lea, 1992; de Cheveigné *et al.*, 1995a). In the case of double vowels, this rate was obtained by scoring each stimulus twice, once for each vowel. A constituent was scored as correctly identified if it was matched by the response vowel or, if two responses were given, by either of the two response vowels. In the case of single vowels, each stimulus was scored as correct if the stimulus vowel was matched by the response vowel (or either of the response vowels if two responses were given).

For the first two measures, the order of vowels within a pair is ignored. In counting conditions we consider three levels of intervowel amplitude difference (0, 10, 20 dB), and ten conditions of unordered vowel pair, in addition to two levels of ΔF_0 (0%, 6%). The third measure distinguishes the order of vowels within a pair (target/background). There are 20 levels of ordered vowel pair, and 5-levels of target-to-background rms level (−20, −10, 0, 10, 20 dB) according to whether the vowel being considered (the "target") is the weaker or the stronger within the pair. In addition, there are

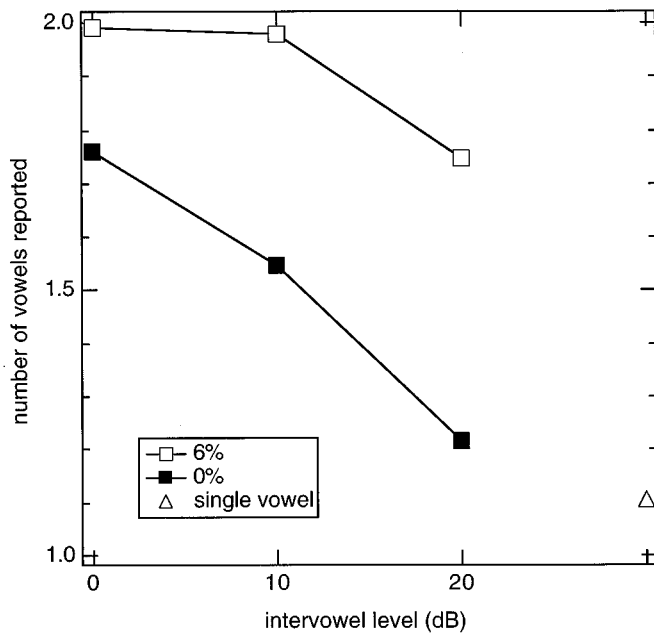


FIG. 2. Number of vowels reported as a function of rms intervowel amplitude difference, for $\Delta F_0=0\%$ (filled symbols) and $\Delta F_0=6\%$ (open symbols). The triangle represents single vowels.

two levels of ΔF_0 (0%,6%), and also two levels of absolute F_0 (low/low versus high/high, and low/high versus high/low).

II. RESULTS

A. ΔF_0 and relative amplitude

The number of vowels reported per stimulus is plotted in Fig. 2 as a function of amplitude difference for both values of ΔF_0 . The triangle is for single vowels. For double vowels at $\Delta F_0=0\%$, subjects tended to report two vowels when both stimulus vowels had the same rms amplitude (0 dB). When either vowel was stronger (10 and 20 dB), they more often reported a single vowel, but they still reported two vowels for a certain proportion of trials, even when the stimulus contained only one vowel (triangle). At $\Delta F_0=6\%$ they almost always reported two vowels. Thus ΔF_0 seems to function as a cue indicating the *multiplicity* of sources within the stimulus.

Combination-correct scores are plotted in Fig. 1 as a function of ΔF_0 , for all three values of amplitude difference. At 0 dB the ΔF_0 effect is similar, if somewhat larger, to that reported in previous studies. At 10 and 20 dB the overall rates are lower but the ΔF_0 effect remains large.

Constituent-correct identification rates for double vowels were analyzed in more detail. Rates averaged over session (5) and repetition (3) were subjected to a repeated-measures analysis of variance (ANOVA). Probabilities reflect, where necessary, an adjustment of the degrees of freedom by a factor that corrects for the inherent correlation of repeated measurements (Geisser and Greenhouse, 1958). Only the three lowest (-20, -10, 0 dB) of the five levels of target-to-background AMPLITUDE were retained, as responses were essentially perfect at 10 and 20 dB, with rela-

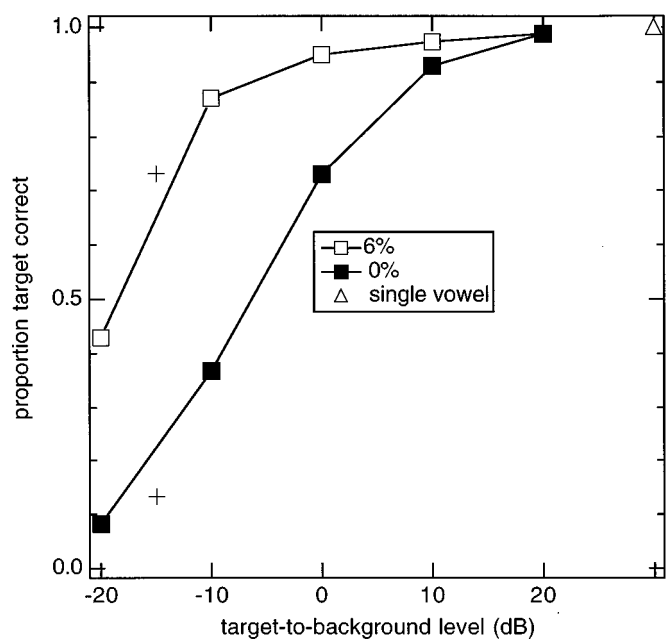


FIG. 3. Target-correct identification rate as a function of rms target-to-background level, for $\Delta F_0=0\%$ (filled symbols) and $\Delta F_0=6\%$ (open symbols). Crosses represent rates obtained with the same subjects at -15 dB in a similar experiment (de Cheveigné *et al.*, 1997). The triangle represents single vowels.

tively little variability. The other factors were $\Delta F_0(2)$, ordered vowel PAIR (20) and absolute $F_0(2)$. Significant effects were noted for factors ΔF_0 [$F(1,5)=51.80$, $p=0.0008$], AMPLITUDE [$F(2,10)=200.31$, $p<0.0001$, $GG=0.64$] and PAIR [$F(19,95)=5.91$, $p=0.0045$, $GG=0.18$], as well as for interactions between ΔF_0 and AMPLITUDE [$F(2,10)=11.45$, $p=0.008$, $GG=0.7$], AMPLITUDE and PAIR [$F(38,190)=4.00$, $p=0.013$, $GG=0.11$], and ΔF_0 and PAIR [$F(19,95)=4.65$, $p=0.02$, $GG=0.15$]. The main effect of F_0 order was not significant, nor were any of the interactions involving it. In other words, at $\Delta F_0=0\%$ it made no difference whether both vowels were at 125 or 132.5 Hz, and, at $\Delta F_0=6\%$, it made no difference whether the target was the lower or the higher of the two frequencies, whatever the vowel pair or amplitude difference. Given the various ways in which F_0 can interact with formant structure, this result is perhaps surprising.

Constituent-correct identification rates are plotted in Fig. 3 as a function of amplitude difference for both values of ΔF_0 . As one might expect, identification of a vowel was better when its relative amplitude was greater. It was also better at $\Delta F_0=6\%$ than at $\Delta F_0=0\%$, particularly when the target amplitude was low (-20 or -10 dB).

B. Vowel pairs

Results for two particular pairs are described in detail. The first pair, /o/+u/, serves later on in Sec. III E to test a model of double-vowel segregation (Meddis and Hewitt, 1992). The second pair, /e/+o/, was chosen to illustrate the variety of response patterns for different pairs. Figure 4 shows the number of vowels reported for both pairs. The pattern for /o/+u/ is typical of the average (Fig. 2), but that

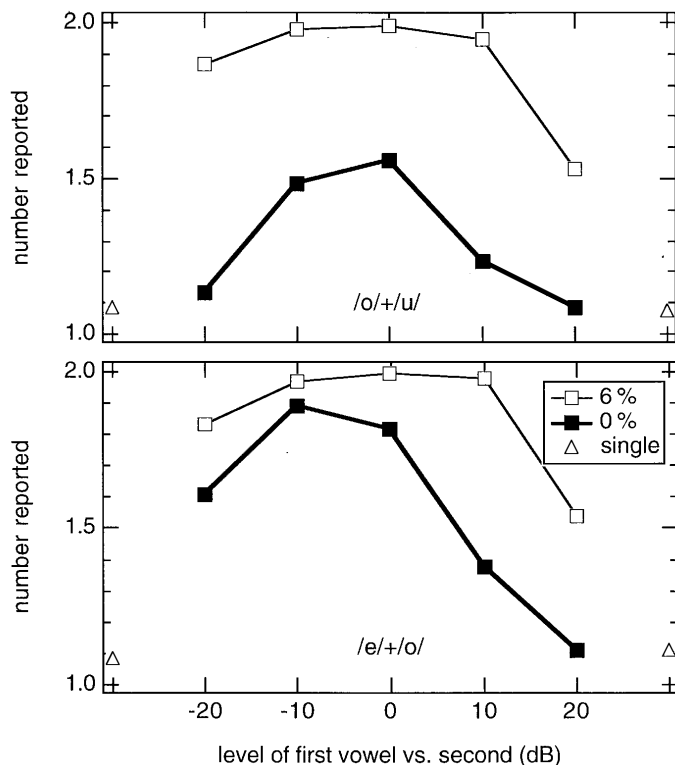


FIG. 4. Number of vowels reported for vowel pairs /o/+/u/ (top) and /e/+/o/ (bottom) as a function of the amplitude of the first vowel relative to the second, at $\Delta F_0=0\%$ (filled symbols) and $\Delta F_0=6\%$ (open symbols). Triangles represent single vowels.

for /e/+/o/ is asymmetric: ΔF_0 has less effect on the response count when /o/ is dominant and /e/ is weak than vice versa. Figure 5 shows the identification rate for the same pairs. The pattern of identification for components of /o/+/u/ is typical of the average (Fig. 3), and the same is true for the /o/ vowel of /e/+/o/ (Fig. 5, lower panel, descending, dotted lines). However for /e/ the effect of ΔF_0 is very small, mainly because the identification rate is high at $\Delta F_0=0\%$. The patterns for the vowels within this pair are quite asymmetric. Vowel-pair specificities are discussed in terms of possible models of vowel segregation in Secs. III D and E.

C. Subject differences

Results of three subjects (K, T, and M) are presented in Figs. 6–8 to illustrate differences between subjects. The effect of ΔF_0 on the number of vowels reported was overall larger for T and M than for K (Fig. 6). The same was true for the identification rate (Fig. 7). This was mainly due to the fact that K had higher identification rates at $\Delta F_0=0\%$ than T and M, and also a lower rate at -20 dB and $\Delta F_0=6\%$. The identification rate *conditional* on a two-vowel response (Fig. 8) reveals marked intersubject differences, mainly at $\Delta F_0=0\%$. Conditional rates for subject M are high, perhaps because she was more conservative than other subjects in reporting two vowels, and thus more often correct when she did so. However, they are also high for subject K, despite the fact that he was the least conservative in reporting two vowels. The one-or-two response task probably exaggerates in-

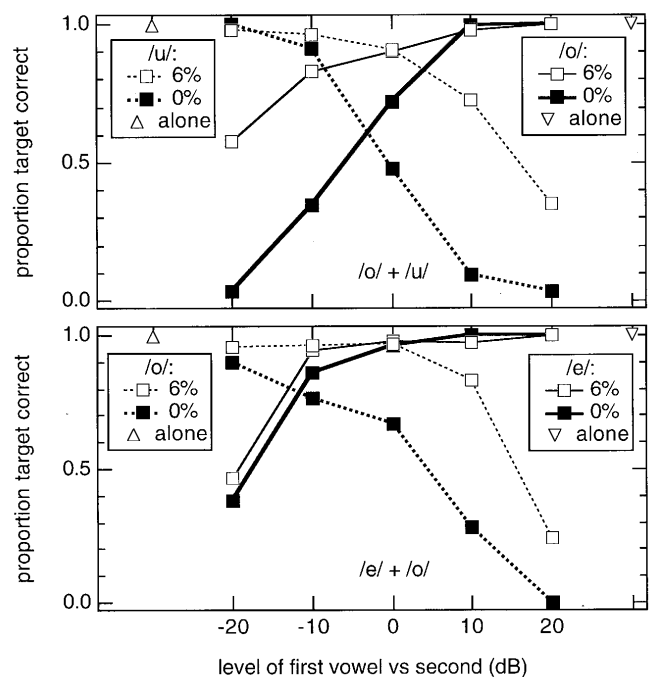


FIG. 5. Target-correct identification rate for vowel pairs /o/+/u/ (top) and /e/+/o/ (bottom) as a function of the amplitude of the first vowel relative to the second, at $\Delta F_0=0\%$ (filled symbols) and $\Delta F_0=6\%$ (open symbols). Ascending lines (continuous) are for the first vowel, descending lines (dotted) are for the second vowel in the pair. Triangles represent single vowels.

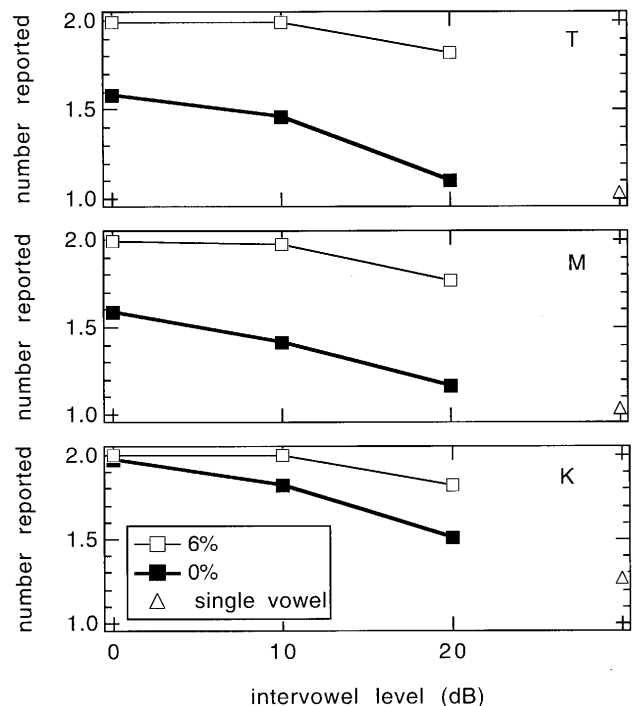


FIG. 6. Number of vowels reported for three subjects as a function of rms intervowel amplitude difference, at $\Delta F_0=0\%$ (filled symbols) and $\Delta F_0=6\%$ (open symbols). Triangles represent single vowels.

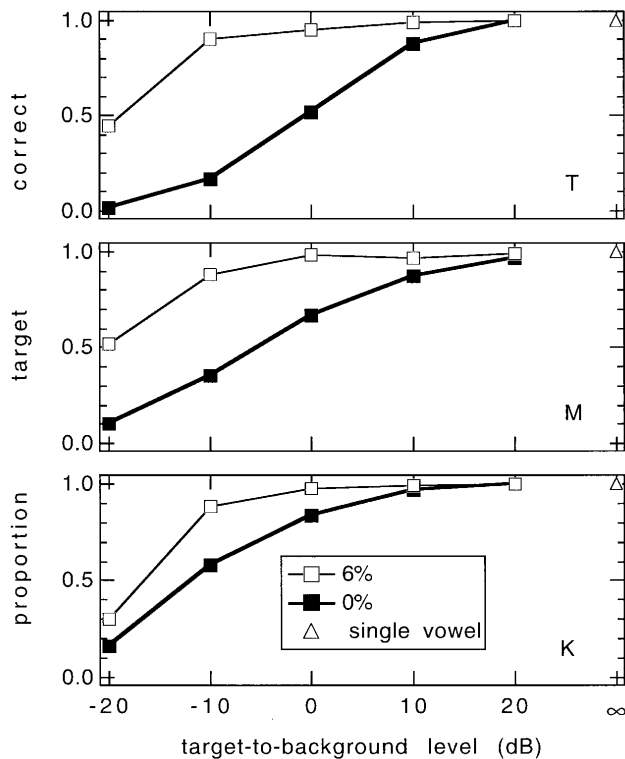


FIG. 7. Identification rate as a function of rms target-to-background level for three subjects, at $\Delta F_0=0\%$ (filled symbols) and at $\Delta F_0=6\%$ (open symbols). Triangles represent single vowels.

tersubject differences, as subjects may use their freedom differently when permitted to report one or two vowels.

D. Single vowels

The stimulus set contained 240 single vowels in addition to 600 double vowels. For single vowels, the overall identification rate was 99.75%. The lowest rate for a subject (subject N) was 99.3%, and the lowest rate for a vowel (/i/) was 99.2%. Evidently subjects had no difficulty identifying any of the vowels. About 10% of all single vowels evoked two-vowel responses, with considerable differences between subjects (27% for subject K, 2% for subject U), but only small differences between vowels. No effect of F_0 was evident.

III. DISCUSSION

A. ΔF_0 effect

To allow comparison with previous results, combination-correct scores were plotted in Fig. 1. The ΔF_0 effect at 0 dB appears to be larger than in previous studies. Our one-or-two response task probably enhanced effect size, as discussed by Cheveigné *et al.* (1997). Assmann and Summerfield (1994) also found relatively large effects, possibly because their subjects were allowed to make “both-same” responses. These may have played a role similar to “single-vowel” responses in our task, with the result of relatively low identification rates at $\Delta F_0=0\%$, as in our experiment.

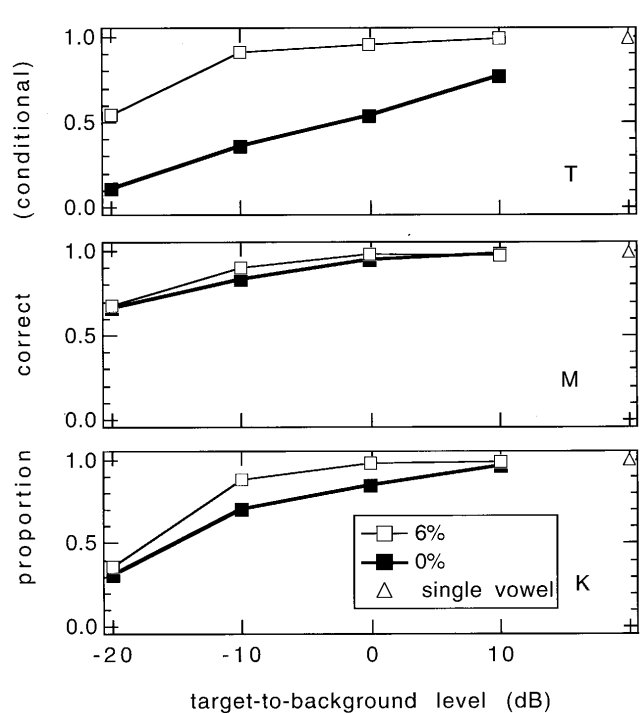


FIG. 8. Identification rate conditional on a two-vowel response (proportion of correct responses for stimuli that evoked two responses) for three subjects as a function of rms target-to-background level at $\Delta F_0=0\%$ (filled symbols) and at $\Delta F_0=6\%$ (open symbols). Triangles represent single vowels.

B. Interaction with relative amplitude

The effects of ΔF_0 at low target amplitudes are relevant to voice segregation in everyday life, since naturally competing voices rarely stay at the same amplitude. When a target voice is weaker it can benefit from F_0 -guided segregation, whereas when it is stronger, segregation is less necessary. Drawing a horizontal line in Fig. 3 at a performance level of 70%, the ΔF_0 effect appears to be equivalent to an amplitude difference of about 13 dB. This may be compared with the 17-dB shift in masked threshold (70% correct) found by Culling *et al.* (1994) using an adaptive technique. There are at least two possible explanations for the increase in effect size as targets get weaker. One is that lower identification rates reduce ceiling effects, as we aimed for when we designed the experiment. The other is that estimation of the competing vowel's F_0 (required by segregation models that invoke harmonic cancellation) is more accurate when targets are weak. If the first explanation were correct, then other manipulations that make the task more difficult should also produce large effects. Instead, Assmann and Summerfield (1990) found that reducing stimulus duration from 200 to 52 ms practically abolished ΔF_0 effects. De Cheveigné *et al.* (1995a) used a stimulus set with high intravowel variability to reduce overall performance, yet they obtained relatively small ΔF_0 effects. Shackleton *et al.* (1994) manipulated formant frequencies to increase the chance of confusions and reduce ceiling effects, but the ΔF_0 effect sizes they obtained were not particularly large. Within their data, identification rate and effect size both varied considerably across subjects and binaural conditions, but there was no systematic relationship between the two quantities.

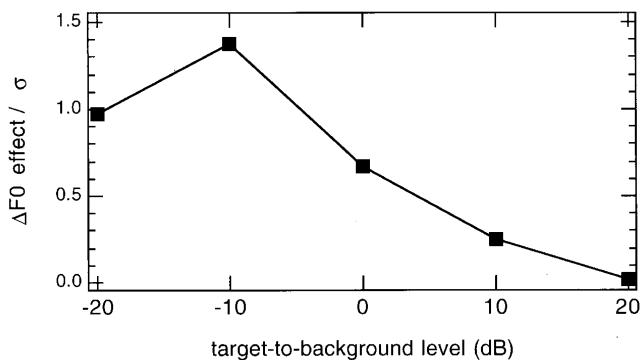


FIG. 9. Ratio between the ΔF_0 effect (difference between identification rate at $\Delta F_0=6\%$ and 0%) and its standard deviation (calculated over levels of SUBJECT, SESSION, and PAIR), as a function of rms target-to-background level.

Whatever the explanation, an amplitude mismatch increases ΔF_0 effect size for the weaker vowel. Large effects are of practical interest because they are relatively easy to demonstrate with statistical confidence. However, that benefit would be lost if large effects were accompanied by an equally large variability. Figure 9 shows that this is not the case: the ratio of effect size to standard deviation (calculated over levels of SUBJECT, PAIR, and SESSION) is greater at -20 and -10 dB than at 0 , 10 , and 20 dB. In general, if an experiment measures small effects relative to a baseline, the identification rate for that baseline should not be too high (to avoid ceiling effects) nor too low (to avoid floor effects and subject frustration). Taking 70% as a compromise, and the $\Delta F_0=6\%$ condition of our experiment as a baseline, a target amplitude of about -15 dB should be appropriate to avoid ceiling effects. This value is used by de Cheveigné *et al.* (1997) to design a sensitive test of phase and harmonicity effects. If the baseline were $\Delta F_0=0\%$, a target amplitude of 0 or -5 dB might be better. The most effective amplitude bias depends on the experiment.

C. Enhancement versus cancellation

The ΔF_0 effect was strong when the amplitude of the target vowel was -10 or -20 dB relative to the competing vowel. In contrast, the ΔF_0 effect measured by McKeown (1992) vanished beyond a 12 -dB amplitude mismatch, but that may have been the result of floor effects (identification rates were overall lower). A ΔF_0 effect at low amplitudes can hardly be explained by a segregation strategy using the target's F_0 (harmonic enhancement), because that parameter is difficult to estimate when the target-to-background ratio is small. Harmonic enhancement might work when the target is strong, but it is difficult to demonstrate segregation in that case because of ceiling effects.

On the other hand the result is compatible with the hypothesis of harmonic cancellation, already supported by other experimental data (Lea, 1992; Summerfield and Culling, 1992; de Cheveigné, 1994; de Cheveigné *et al.*, 1995a, 1997). Cancellation requires estimation of the competing vowel's F_0 , and this is relatively easy when the target's amplitude is low.

D. Segregation based on beats

It has been suggested that, for small ΔF_0 's, segregation mechanisms might exploit beat patterns that occur when there is a difference in F_0 (Assmann and Summerfield, 1994; Culling and Darwin, 1994; Culling and Summerfield, 1995). For example, beats near the formants of a weaker vowel might reveal the presence of those formants (de Cheveigné *et al.*, 1997, Fig. 1). Beats at a vowel formant will be strong if the spectral envelopes of both vowels have the same amplitude near that formant.² If formant F_1 (respectively, F_2) determines a target vowel's identity, ΔF_0 effects should be large at a relative amplitude for which spectral envelopes of both vowels coincide in the F_1 (respectively, F_2) region of the target. In other words, for each vowel pair we should be able to predict the size of the ΔF_0 effect given the difference in envelope amplitude of the two vowels in the region of the formants of the target.

To test whether beats might have played a role in the current experiment, we chose two parameters representing the absolute amplitude difference between vowel envelopes at formant F_1 (respectively, F_2) of the target vowel. (This corresponds to the vertical distance between curves in Fig. A1 at each of the first two formants of the target.) A linear regression model was then formed, based on these two parameters, to predict the difference between identification rate at $\Delta F_0=0\%$ and at $\Delta F_0=6\%$. This model was compared to a second regression model predicting simply that the ΔF_0 effect is uniformly greatest at a target amplitude of -10 dB. The two-parameter model fits *less* well than the one-parameter model ($r^2=0.20$ vs 0.26), despite its larger number of parameters. This model is clearly inadequate, possibly because it is too crude, possibly because beats did not determine segregation in this experiment. The question of beats is examined in further detail in a companion paper (de Cheveigné *et al.*, 1997).

E. Meddis and Hewitt's channel selection model

Meddis and Hewitt (1992) proposed a model for the identification of concurrent vowels with same or different F_0 's. The model is based on an array of autocorrelation functions (ACF) calculated within individual channels output by a model of peripheral filtering and hair-cell transduction. The pattern is summed across channels to obtain a summary autocorrelation function (SACF). The largest peak in the summary autocorrelation serves to estimate the period of the "dominant" vowel. In the same authors' pitch perception model (Meddis and Hewitt, 1991), a similar peak served as a cue to the pitch. To account for better identification of double vowels when there is a difference in F_0 , the model supposes that individual ACFs that show a peak at the period of the dominant vowel are selected and attributed to that vowel. The remaining channels are attributed to the other vowel. ACFs within each group are summed, and the sums are matched to templates to identify both vowels. When there is an amplitude mismatch between vowels, the partition between channels is likely to be determined on the basis of the periodicity of the stronger vowel, and thus the weaker vowel can be said to have been segregated according to the

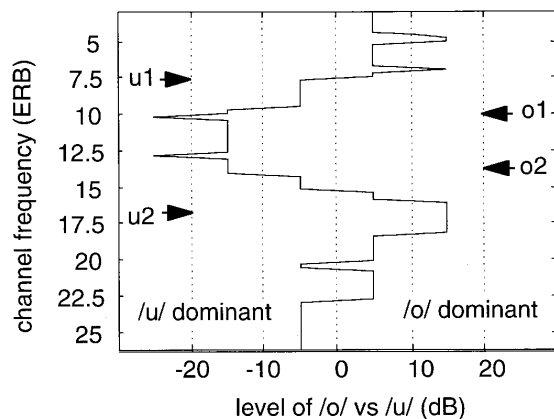


FIG. 10. Dominance of auditory-model channels by periods corresponding to the fundamental of /u/ (125 Hz), left of crooked line, and the fundamental of /o/ (132.5 Hz), right of crooked line, as a function of channel frequency in ERB (vertical scale) and amplitude of /o/ relative to /u/ (horizontal scale). Arrows indicate formant frequencies of each vowel.

principle of harmonic cancellation. In this respect the model is compatible with our findings, as discussed in Sec. III C.

We implemented Meddis and Hewitt's model using a software model of peripheral filtering and hair-cell transduction (Holdsworth *et al.*, 1988; Meddis, 1988; Culling, 1996) that produced a nerve fiber discharge firing probability within channels spaced along an ERB scale (equivalent rectangular bandwidth; Moore and Glasberg, 1983) at a density of four channels per ERB, between 80 Hz and 4 kHz. ACFs were calculated and summed to obtain a SACF from which the dominant period was determined. Each channel was assigned to one or the other vowel according to whether or not its ACF attained its highest value within 3% of the period of the dominant vowel.³ The partition between channels at different amplitudes is illustrated in Fig. 10 for the vowel pair /o/+ /u/ at $\Delta F_0 = 6\%$. To the left of the crooked line channels are dominated by /u/, to the right they are dominated by /o/. As either vowel is made stronger, the partition changes to its advantage, as more channels respond with that vowel's periodicity. When /o/ is 20 dB below /u/ (left-hand side) the partition isolates channels near F_1 and F_2 of /o/, from which that vowel might be identified. However when /o/ is 20 dB above /u/, its periodicity dominates *all* channels and leaves no channels to represent /u/. As ΔF_0 does not produce a partition, the model cannot predict the improvement with ΔF_0 that we observed for the vowel /u/ when it was 20 dB weaker than /o/ (Fig. 5, top panel, dotted lines, abscissa = 20 dB). This problem occurs for 5 pairs out of 20.

Our implementation of Meddis and Hewitt's model failed to predict ΔF_0 effects for weak targets. However, it would be unwise to reject the model on that account, for at least two reasons. The first is that it might perform better with filters narrower than those used in the gammatone simulation. The second is that the auditory system might use a more sensitive criterion to detect that a channel is not completely dominated by the period of the stronger vowel. A scheme involving such a criterion is investigated in a companion paper (de Cheveigné, 1997), together with a model that allows for segregation *within* channels instead of be-

tween channels. To summarize, Meddis and Hewitt's model can explain the main aspects of our data, but not the ΔF_0 effects observed at low target-to-background levels for certain vowel pairs.

IV. CONCLUSIONS

Identification of concurrent synthetic vowels was measured as a function of amplitude difference and ΔF_0 , using a task in which subjects could report one or two vowels. Main results were as follows.

(1) The number of vowels reported and the identification rate were greater at $\Delta F_0 = 6\%$ than at $\Delta F_0 = 0\%$. For equal-amplitude vowel pairs, the effect of ΔF_0 on identification was larger than previously reported, presumably because effects were enhanced by the one-or-two response task.

(2) The ΔF_0 effect size was greater when the target vowel was weaker by 10 to 20 dB relative to the competing vowel. This may be explained as due either to reduced ceiling effects, or to more effective harmonic cancellation, since the competing vowel's F_0 is easier to estimate when the target is weak.

(3) The existence of a ΔF_0 effect when the target is weak relative to the competing vowel is difficult to reconcile with the hypothesis of harmonic enhancement of the target.

(4) Patterns of response for specific vowel pairs were not compatible with a simple model that assumed that identification should be better when large beats occur near formant frequencies of the target vowel.

(5) ΔF_0 effects observed at low target-to-background levels (−20 dB) for certain vowel pairs were not compatible with Meddis and Hewitt's (1992) channel selection model.

ACKNOWLEDGMENTS

The experiments were carried out at ATR Human Information Processing Laboratories, within a research collaboration agreement with the Centre National de la Recherche Scientifique. The first author thanks ATR for its kind hospitality and CNRS for leave of absence. Cécile Marin, Jean Laroche, and Steve McAdams participated in the preparation of these experiments. Hiroaki Kato and Ikuyo Masuda contributed useful ideas and advice and Rieko Kubo supervised the experiments. Thanks to John Culling of the MRC Institute of Hearing Research for providing the software for stimulus synthesis, and to him and Peter Assmann for detailed comments on previous versions of the manuscript.

APPENDIX: VOWEL SYNTHESIS

Vowels were /a/, /i/, /u/, /e/, and /o/ of Japanese. The first four formants had frequencies suggested by Hirahara and Kato (1992), the fifth was set to 4200 Hz for all vowels. The same bandwidths were used for all vowels, as in Culling and Summerfield (1995). Formant frequencies and bandwidths are shown in Table AI. Vowels were synthesized using an implementation of Klatt's synthesizer (Klatt, 1980; Culling, 1996). Table AI indicates the rms levels produced by the synthesizer (so-called "equal effort" levels). Waveforms were then scaled so the rms amplitudes of all vowels were the same. Table AI shows the resulting sound-pressure

TABLE A1. Frequencies and bandwidths of vowels, together with the rms amplitude produced by the synthesizer for each vowel, and the sound-pressure level at the earphone measured with an artificial ear (after rms amplitudes had been made equal).

	/a/	/i/	/u/	/e/	/o/	BW
<i>F</i> 1	750	281	312	469	468	90
<i>F</i> 2	1187	2281	1219	2031	781	110
<i>F</i> 3	2595	3187	2469	2687	2656	170
<i>F</i> 4	3781	3781	3406	3375	3281	250
<i>F</i> 5	4200	4200	4200	4200	4200	300
dB rms after synthesis	46.9	40.6	40.4	41.8	44.5	
dB (A) SPL	70.0	63.0	63.6	67.4	66.2	

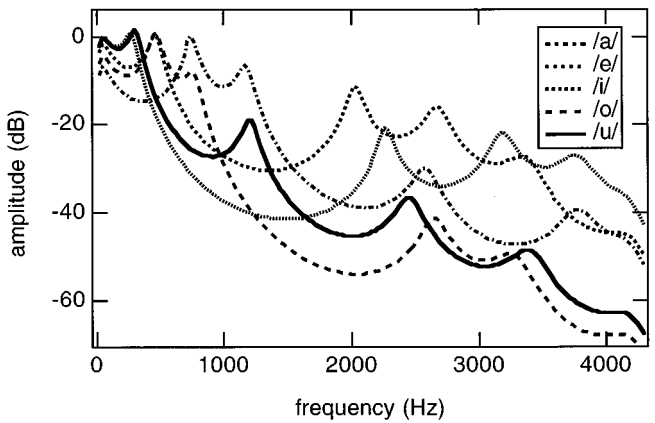


FIG. A1. Spectral envelopes of all five vowels.

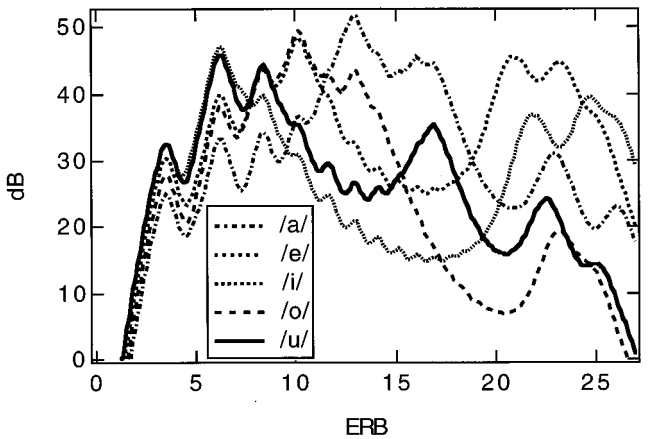


FIG. A2. Estimates of basilar-membrane excitation as a function of channel frequency (on an ERB scale) for all five vowels (fundamental is 125 Hz).

levels delivered by the earphones, as measured by the artificial ear. Spectral envelopes (scaled by the same amount as the waveforms) are plotted in Fig. A1. Estimates of basilar membrane excitation patterns for each vowel are plotted in Fig. A2. Excitation patterns were calculated by taking the FFT of a 16-ms Hanning-shaped window of a 125-Hz vowel (two periods), and applying spectral smoothing according to formulas of Moore and Glasberg (1983).

¹The “x” answer was counted as if the subject had reported an incorrect vowel. The option was rarely used by the subjects, and in subsequent experiments it was removed.

²In the low-frequency region where partials are resolved, the amplitude of beats of adjacent partials depends upon the levels of their excitation patterns within each channel. Excitation may locally be equal even if spectral envelopes do not coincide. However, given the small ΔF_0 , excitation patterns of the lowest partials tend to overlap, and the range of intervowel levels over which they may coincide is therefore narrow and near the intervowel level for which spectral envelopes coincide. At higher frequencies where partials are not resolved, excitation patterns are wide and best beats also occur when the spectral envelopes coincide.

³This selection criterion is a departure from Meddis and Hewitt’s model, in the interest of clarity. In his model, the dominant period was derived from the largest peak in the summary ACF (within a certain range), and individual channels were classified according to whether their ACF had a “peak” at the dominant period. A peak was defined as any point higher than its left and right neighbors. This definition can lead to somewhat erratic results when ACF’s are “noisy.” Our modification produced patterns that are better suited for illustration purposes. It does not betray the spirit of the model or reduce its chance of success.

Assmann, P. F., and Summerfield, Q. (1990). “Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies,” *J. Acoust. Soc. Am.* **88**, 680–697.

Assmann, P. F., and Summerfield, Q. (1994). “The contribution of waveform interactions to the perception of concurrent vowels,” *J. Acoust. Soc. Am.* **95**, 471–484.

Culling, J. (1996). “Signal processing software for teaching and research in psycholinguistics under UNIX and x-windows,” *Behav. Res. Methods Instrum. Comput.* **28**, 376–382.

Culling, J. F., and Darwin, C. J. (1993). “Perceptual separation of simultaneous vowels: Within and across-formant grouping by F_0 ,” *J. Acoust. Soc. Am.* **93**, 3454–3467.

Culling, J. F., and Darwin, C. J. (1994). “Perceptual and computational separation of simultaneous vowels: Cues arising from low frequency beating,” *J. Acoust. Soc. Am.* **95**, 1559–1569.

Culling, J., and Summerfield, Q. (1995). “The role of frequency modulation in the perceptual segregation of concurrent vowels,” *J. Acoust. Soc. Am.* **98**, 837–846.

Culling, J. F., Summerfield, Q., and Marshall, D.-H. (1994). “Effects of simulated reverberation on the use of binaural cues and fundamental frequency differences for separating concurrent vowels,” *Speech Commun.* **14**, 71–95.

de Cheveigné, A. (1994). “Strategies for voice separation based on harmonicity,” *Proceedings of the International Conference on Speech and Language Processing, Yokohama* (Acoustical Society of Japan), pp. 1071–1074.

de Cheveigné, A. (1995). “Experiments in vowel segregation,” *ATR Human Information Processing Research Labs Tech. Report No. TR-H-154*.

de Cheveigné, A. (1997). “Concurrent vowel segregation. III. A neural model of harmonic interference cancellation,” *J. Acoust. Soc. Am.* **101**, 2857–2865.

de Cheveigné, A., McAdams, S., Laroche, J., and Rosenberg, M. (1995a). “Identification of concurrent harmonic and inharmonic vowels: A test of the theory of harmonic cancellation and enhancement,” *J. Acoust. Soc. Am.* **97**, 3736–3748.

de Cheveigné, A., Kawahara, H., Tsuzaki, M., and Aikawa, K. (1995b). “Sensitive experimental techniques for the study of sound segregation,” *ASJ autumn meeting* (Acoustical Society of Japan), pp. 373–374.

de Cheveigné, A., McAdams, S., and Marin, M. (1997). “Concurrent vowel segregation. II. Effects of phase, harmonicity and task,” *J. Acoust. Soc. Am.* **101**, 2848–2856.

- Demany, L., and Semal, C. (1990). "The effect of vibrato on the recognition of masked vowels," *Percept. Psychophys.* **48**, 436–444.
- Hirahara, T., and Kato, H. (1992). "The effect of F_0 on vowel identification," in *Speech Perception, Production and Linguistic Structure*, edited by Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka (Ohmsha, Tokyo), pp. 89–112.
- Holdsworth, J., Nimmo-Smith, I., Patterson, R. D., and Rice, P. (1988). "Implementing a GammaTone filter bank (SVOS final report, annex C)," MRC Applied Psychology Unit Tech. Report.
- Geisser, S., and Greenhouse, S. W. (1958). "An extension of Box's results on the use of the F distribution in multivariate analysis," *Ann. Math. Stat.* **29**, 885–889.
- Klatt, D. H. (1980). "Software for a cascade/parallel formant synthesizer," *J. Acoust. Soc. Am.* **67**, 838–844.
- Lea, A. (1992). "Auditory models of vowel perception," Ph.D. dissertation, Nottingham (unpublished).
- McKeown, J. D. (1992). "Perception of concurrent vowels: The effect of varying their relative level," *Speech Commun.* **11**, 1–13.
- Meddis, R. (1988). "Simulation of auditory-neural transduction: further studies," *J. Acoust. Soc. Am.* **83**, 1056–1063.
- 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.
- Meddis, R., and Hewitt, M. J. (1992). "Modeling the identification of concurrent vowels with different fundamental frequencies," *J. Acoust. Soc. Am.* **91**, 233–245.
- Moore, B. C. J., and Glasberg, B. R. (1983). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," *J. Acoust. Soc. Am.* **74**, 750–753.
- Scheffers, M. T. M. (1983). "Sifting vowels," Ph.D. thesis, Gröningen (unpublished).
- Shackleton, T. M., Meddis, R., and Hewitt, M. J. (1994). "The role of binaural and fundamental frequency difference cues in the identification of concurrently presented vowels," *Q. J. Exp. Psychol. A* **47**, 545–563.
- Summerfield, Q. (1992). "Roles of harmonicity and coherent frequency modulation in auditory grouping," in *The auditory processing of speech: from sounds to words*, edited by M. E. H. Schouten (Mouton de Gruyter, Berlin), pp. 157–166.
- Summerfield, Q., and Assmann, P. F. (1991). "Perception of concurrent vowels: Effects of harmonic misalignment and pitch-period asynchrony," *J. Acoust. Soc. Am.* **89**, 1364–1377.
- Summerfield, Q., and Culling, J. F. (1992). "Periodicity of maskers not targets determines ease of perceptual segregation using differences in fundamental frequency," *J. Acoust. Soc. Am.* **92**, 2317(A).

Concurrent vowel identification. II. Effects of phase, harmonicity, and task^{a)}

Alain de Cheveigné

Centre National de la Recherche Scientifique/Université Paris 7, 2 place Jussieu, case 7003, F-75251 Paris
Cédex 05, France and ATR Human Information Processing Research Laboratories, 2-2 Hikaridai,
Seika-cho Soraku-gun, Kyoto 619-02, Japan

Stephen McAdams

Laboratoire de Psychologie Expérimentale (Centre National de la Recherche Scientifique, URA316),
Université René Descartes, EPHE, 28 rue Serpente, F-75006 Paris, France and Institut de Recherche et de
Coordination Acoustique/Musique (IRCAM), 1 place Stravinsky, F-75004 Paris, France

Cécile M. H. Marin

IRCAM, 1 place Stravinsky, F-75004 Paris, France

(Received 19 December 1995; revised 12 August 1996; accepted 22 November 1996)

Subjects identified concurrent synthetic vowel pairs in four experiments. The first experiment found that improvements in vowel identification with a difference in fundamental frequency (ΔF_0) do not depend on component phase. The second investigated more precisely whether phase patterns resulting from ongoing phase shifts in inharmonic stimuli can by themselves produce effects similar to those attributed to differences in harmonic state of component vowels. No such effects were found. The third experiment found that identification was better for harmonic than for inharmonic backgrounds, and that it did not depend on target harmonicity. The first three experiments employed a task in which subjects were free to report one or two vowels for each stimulus. The fourth experiment reproduced several conditions with a more classic task in which subjects had to report two vowels. Compared to the classic task, the new task gave larger effects and provided an additional measure of segregation: the number of vowels reported per stimulus. Overall, results were consistent with the hypothesis that the auditory system segregates targets by a mechanism of harmonic cancellation of competing vowels. They did not support the hypothesis of harmonic enhancement of targets. The lack of a phase effect places strong constraints on models that exploit pitch period asynchrony (PPA) or beats. © 1997 Acoustical Society of America.
[S0001-4966(97)04004-6]

PACS numbers: 43.71.An, 43.71.Es, 43.66.Ba, 43.66.Nm [WS]

INTRODUCTION

Speech is easier to understand when there is a difference in fundamental frequency (ΔF_0) between a target voice and an interfering voice (Brokx and Nooteboom, 1982). When two steady-state synthetic vowels are presented simultaneously, identification is better when their fundamental frequencies (F_0) are different than when they are the same (Scheffers, 1983; Darwin, 1981; Zwicker, 1984; Assmann and Summerfield, 1990; McKeown, 1992; Culling and Darwin, 1993, 1994). A variety of models and methods of “ F_0 -guided segregation” have been proposed to explain or emulate this effect [see de Cheveigné (1993) for a review]. They may be classified according to whether they exploit target harmonicity (harmonic *enhancement*) or background harmonicity (harmonic *cancellation*). Some evidence has been found in favor of harmonic cancellation (Lea, 1992; Summerfield and Culling, 1992; de Cheveigné, 1994; de Cheveigné *et al.*, 1995a), but so far there is little to support the harmonic enhancement hypothesis. Recently, other

mechanisms have been proposed that do not depend directly on harmonicity or F_0 , but rather on waveform interactions that co-occur with F_0 differences.

A. Pitch period asynchrony (PPA)

An F_0 difference is equivalent to a gradually increasing time shift of one waveform relative to another. A natural vowel's short-term energy is pulsatile, so the masking it causes or receives may vary with time alignment relative to the other vowel. A ΔF_0 might in this way cause either vowel or both to be better perceived. This is known as the pitch period asynchrony (PPA) mechanism (Assmann and Summerfield, 1988, 1994; Summerfield and Assmann, 1991; Carlyon and Shackleton, 1994). Summerfield and Assmann (1991) investigated whether a time shift *per se* is sufficient to produce segregation in the absence of mistuning. They presented subjects with synthetic vowels at the same F_0 (50 or 100 Hz), with and without a time shift of half a period (both vowels were ramped on and off simultaneously, so the shift did not affect onset times). The time shift produced a significant improvement at 50 Hz, but not at 100 Hz. Although in a later experiment Assmann and Summerfield (1994) did find a significant effect of intervowel alignment at 100 Hz, as well

^{a)}Portions of this work were presented at a meeting of the Acoustical Society of Japan and as an ATR Human Information Processing Research Laboratories technical report (de Cheveigné, 1995; de Cheveigné *et al.*, 1995b).

as other indirect evidence that PPA contributes to segregation, they failed to replicate this effect in a further experiment with inexperienced subjects.

Estimates of the equivalent rectangular duration (ERD) of the auditory temporal window (Plack and Moore, 1990) are of the same order (6–13 ms) as the fundamental periods used in double-vowel experiments. One might therefore expect features of a 10-ms period to be smoothed out too much for PPA to be effective. However, Kohlrausch and Sander (1995) found that masking of a short pure-tone target varied by as much as 17 dB within the period of a 100-Hz masker. The variation was smaller but still appreciable (about 6 dB) at a fundamental of 220 Hz. The effect was dependent on the component phase of the masker and largest for a phase relationship designed to produce highly modulated patterns of activity within auditory channels.

Several experiments suggest that vowel identification might depend on uneven masking within a masker's fundamental period. Moore and Alcántara (1995, 1996) synthesized harmonic "vowels" with a fundamental of 100 Hz and a spectral envelope that was flat on average. "Formants" were defined by amplitude modulation of groups of two consecutive harmonics at a rate of 10 Hz. For cosine phase, the stimuli could be identified as vowels, despite their flat average spectrum. For random phase, identification was at chance level. Stimuli with cosine phase and a flat spectrum have a peaked waveform that produces strongly modulated activity within peripheral channels, provided the F_0 is low enough and the channel CF high enough (Horst *et al.*, 1986). Within the dips of this modulation, masking may be relatively weak. Raising or lowering the level of a group of components is equivalent to adding those components (in same or opposing phase) so that they stand out during the modulation dips.

Palmer *et al.* (1987) observed a change with phase of the position of the F_1 phoneme boundary along a /e/-/I/ continuum (Darwin and Gardner, 1986). The manipulated partial was the fourth harmonic (500 Hz) of a 125-Hz fundamental. The boundary moved down from 450 to 430 Hz when the phase shifted by 90° relative to the phase produced by a Klatt synthesizer. In other words, this shift is equivalent to a 20-Hz rise in the perceived F_1 of the stimuli. The authors also performed a physiological experiment in which similar stimuli (with a fundamental of 100 Hz) were presented to guinea pigs and the response was recorded from a population of auditory-nerve fibers. Without the phase shift fibers below 1 kHz were equally dominated by frequencies of 400 or 500 Hz. With the 90° phase shift they were dominated mainly by the higher component, which is congruent with a rise in perceived F_1 in the human subjects. Stimuli in Klatt phase (phase produced by the Klatt synthesizer) produce highly modulated patterns of activity in auditory channels. Shifting the phase of a component is equivalent to adding a component with the same frequency and suitable phase and amplitude. The added component may be perceptible within the valleys of modulation.

Traunmüller (1987) modulated the amplitude spectrum of a glottal source with the phase spectrum of a glottal tract (simulated as a cascade synthesizer) to produce nine Swedish

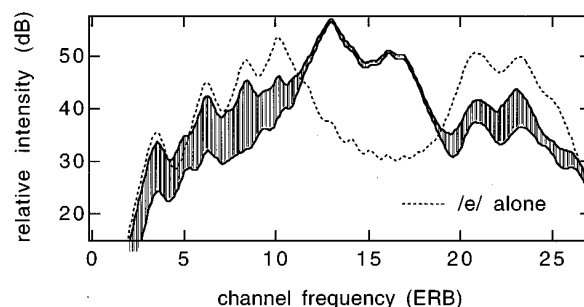


FIG. 1. Striped zone: Range of variation of the excitation pattern due to beats within a vowel pair /a/+ /e/ in which /e/ is 12 dB weaker than /a/. Thin dotted line: Excitation pattern produced by /e/ alone. Excitation patterns were calculated by taking the FFT of a 16-ms Hanning-shaped window and applying spectral smoothing according to formulas of Moore and Glasberg (1983). The origin of the ordinate (dB scale) is arbitrary.

"vowels." There were no spectral amplitude peaks present to signal the formants, but several subjects could label the stimuli consistently if the F_0 was low enough (71 or 100 Hz). Labeling was less consistent at higher frequencies (141 and 200 Hz) and at 283 Hz it fell to chance level. The "phase vowels" were intelligible via earphones, but not when presented through a loudspeaker in an ordinary room. "Flat-spectrum" diphthongs produced by Schroeder and Strube (1986) were also unintelligible if presented via loudspeakers in a reverberant room, because of phase randomization.

These experiments all show that phase may in some circumstances affect vowel identification. The PPA hypothesis depends on the particular phase patterns that produce peaked waveforms. One might therefore suspect that the ΔF_0 effects observed in "double vowel" experiments are specific to the particular phase employed. If so, they should be reduced for random phase stimuli that lack the temporal cues upon which PPA depends.

B. Waveform interaction (beats)

The PPA explanation involves intravowel phase patterns that produce peaked waveforms, together with the particular intervowel phase relationship that is equivalent to a time shift. It also supposes that the temporal resolution of the auditory system is fine enough to resolve patterns on the time scale of a period. However, waveform interaction may also produce patterns that are static (at $\Delta F_0 = 0$) or that vary on a slower time scale. When a ΔF_0 is introduced between vowels, beats occur between corresponding partials at a rate equal to their frequency difference, and with a depth that depends on their relative amplitudes. The short-term spectrum thus varies with time, and it may assume a shape that favors the identification of one vowel or the other, or possibly both together. Alternatively, the pulsation itself might reveal spectral cues too weak to stand out in the average spectrum. Figure 1 illustrates this idea. The vowel /e/ at 132 Hz was added to the vowel /a/ at 124 Hz, with a 12-dB mismatch in favor of /a/. The two vowels have the same spectrum level at the formants F_1 and F_2 of /e/, causing the spectrum of their sum to undergo relatively deep beats near those formants. The excitation pattern for the sum varies over the range shown in Fig. 1. The pulsation might reveal

the formants, despite the fact that the average spectrum does not show clear evidence of their presence at any instant.

Culling and Darwin (1994) suggested that beats in the low-frequency (F_1) region might explain improvements in identification performance with ΔF_0 's smaller than 1 semitone. In agreement with this hypothesis, Assmann and Summerfield (1994) found that successive 50-ms intervals excised from a 200-ms double vowel were not equally identifiable. Identification rates for the best interval were consistent with the idea that the auditory system takes advantage of beats to choose, within the 200-ms stimulus, a favorable interval on which to base identification.

If the ΔF_0 is small or zero, the overall spectrum of a double-vowel stimulus depends on the pattern of intervowel starting phase. In particular, the spectrum of the $\Delta F_0=0$ condition of double-vowel experiments is phase-dependent. It is conceivable that the commonly used Klatt or sine phase patterns might produce at $\Delta F_0=0$ a spectrum that is particularly unfavorable for identification, contributing artificially to the size of the ΔF_0 effects observed. Like PPA, the beats hypothesis leads us to suspect that the classic ΔF_0 effect might be phase dependent.

In a previous experiment (de Cheveigné *et al.*, 1995a) we presented subjects with double-vowel stimuli in which each vowel was either harmonic or inharmonic (with partials randomly mistuned). Our aim was to determine whether F_0 -guided segregation mechanisms used the harmonic structure of the vowel being identified, that of the competing vowel, or both. We found that harmonicity of the competing vowel improved identification of the target, but that harmonicity of the target itself did not. However, all our stimuli were synthesized with a sine starting phase. Each partial of an inharmonic vowel can be interpreted as gradually shifting in phase, due to its mistuning. Consequently, inharmonic vowels shifted towards a random phase pattern, whereas harmonic vowels kept their original phase throughout the stimulus. If phase affected identification, it might conceivably have been responsible for the pattern of results that we attributed to harmonicity.

C. The present investigation

The experiments described in this paper were designed to reveal phase effects and to test the generality of our results on harmonicity. Experiment 1 examined whether the classic ΔF_0 effect depends on intra- or intervowel phase relationships. Experiment 2 investigated more particularly whether phase effects could have constituted an artifact in our previous harmonicity experiment. Experiment 3 reproduced three crucial conditions of that experiment with stimuli designed to minimize the usefulness of beat or PPA cues. Experiment 4 compared the particular task we used (one-or-two response task) to the task classically used in double-vowel experiments (two vowel forced-choice task).

I. GENERAL METHODS

A. Stimuli

The subjects (six), vowel set (five), synthesis method, and presentation conditions were described in a companion

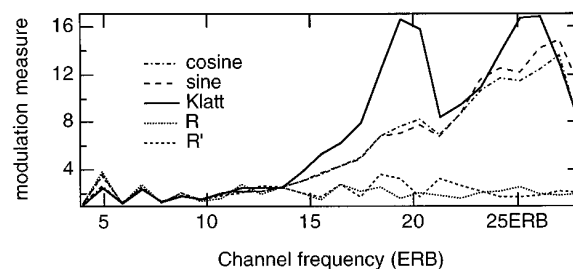


FIG. 2. Modulation of output channels of an auditory model as a function of CF for several phase patterns, averaged over vowels. F_0 was 100 Hz.

paper (de Cheveigné *et al.*, 1997). The present experiments used a different choice of F_0 , phase, duration, intervowel amplitude, harmonic state, and task, as described here. Single vowels were synthesized at F_0 's of 124 and 132 Hz, with a duration of 270 ms including 20-ms raised-cosine onset and offset ramps. Both F_0 's were chosen to be multiples of 4 Hz, the reciprocal of the effective stimulus duration (250 ms between -6 -dB points), so that all beat patterns between partials would have an integer number of periods, and the overall spectrum would be the same whatever the starting phases of beating partials. The spectrum did, however, depend on the relative phase of partials that had the same frequency. This was the case of all partials at $\Delta F_0=0$, but of only one partial (4092 Hz) when the F_0 's differed (this frequency is beyond the range that largely determines vowel identification). Partial started either in sine phase or in one of two "random" phase patterns (R, R'). The amount of modulation produced by different phase patterns within output channels of a peripheral filter model (Holdsworth *et al.*, 1988) is illustrated in Fig. 2 for a fundamental of 100 Hz. Modulation was estimated by taking the largest ratio between rms output calculated over two consecutive frames, each one-half period in length (5 ms). A large value of this measure indicates that the energy is localized within the period. Modulation is small up to about 14 equivalent rectangular bandwidth (ERB) (845 Hz) for all phases. It increases rapidly for Klatt, sine and cosine phase patterns, but remains small for both "random" phase patterns. Single vowels were added to obtain double vowels, with an amplitude mismatch of 15 dB to reduce ceiling effects for the weaker vowel (de Cheveigné *et al.*, 1997).

Stimuli were harmonic in experiments 1 and 2, and either harmonic or inharmonic in experiments 3 and 4. Inharmonic stimuli were produced by shifting partial frequencies of a harmonic vowel by random amounts less than 6.45%, or less than half the spacing between partials, whichever was smaller. The shifts obeyed further constraints that were designed to reduce the usefulness of phase or beat cues: (1) All partial frequencies had to be multiples of 4 Hz to ensure that the effective length of the stimulus (250 ms) was a superperiod of all beat patterns; (2) any given partial deviated by at most $F_0/2$, or $8 \times n$ Hz (where n was the partial's rank) whichever was smaller, from the harmonic series, to ensure that the spectral density was not too different from that of a harmonic stimulus; (3) Each partial was at least 16 Hz from any other component in the stimulus to ensure that all beats

between adjacent partials were faster than 16 Hz; and (4) within these constraints, the partial was chosen at random. Care was taken to ensure that the constraints did not introduce a systematic shift towards either higher or lower frequencies.

In order to satisfy constraint 3, different patterns had to be used at each of the nominal frequencies employed. Constraint 3 was relaxed for the second harmonic because it was incompatible with constraint 2. Since a random choice of frequencies may produce, by chance, patterns that are locally harmonic, a measure of inharmonicity was used to screen out such patterns. The measure was defined as the sum of absolute differences between consecutive partial frequencies divided by their rank. It is sensitive to local rather than cross-spectrum harmonicity patterns, and puts relatively less weight on higher partials. In this respect it differs from the measure used by de Cheveigné *et al.* (1995a).

The “ F_0 ” of an inharmonic vowel is defined as the F_0 of the harmonic vowel from which it is derived. Inharmonic vowels sounded odd but were unambiguously identifiable as vowels. They had a relatively clear pitch that depended on the particular vowel.

B. Task and experiment design

Experiments 1–3 had several stimulus conditions in common. In the interest of economy, their stimuli were pooled and presented together (in other words, they formed a single experiment that we describe as three in the interest of clarity). The stimulus set consisted of 200 single and 400 double vowels in random order. The subject’s task was to report one or two vowels for each stimulus, as in de Cheveigné *et al.* (1997).

The stimuli of experiment 4 were pooled with stimuli of another experiment not reported here. The stimulus set consisted of 400 double vowels in random order. Each stimulus was presented once and the subjects had to report *two* vowels. Subjects were warned that identification of both vowels might sometimes be impossible and were asked to make their “best guess” in that case. Again there was no feedback. All six subjects participated in experiments 1–3. Of the six, five also participated in experiment 4. Each subject performed five sessions on different days.

C. Scoring methods

Scoring methods are the same as used by de Cheveigné *et al.* (1997). Each double-vowel stimulus was scored twice, once for each vowel. A stimulus vowel was scored as correctly identified if it was matched by the response vowel (or either response vowel if two were given). Each single-vowel stimulus was scored once, in a similar fashion. Responses were classified according to the vowel’s nature (phoneme, F_0 , phase, harmonicity), the nature of the competing vowel, and their relationship (ΔF_0 , relative amplitude) to obtain *target-correct* identification rates for each of these conditions. Results for the more intense vowel (15 dB) were essentially perfect and are not reported. We report only rates for the weaker (–15 dB) vowel. From previous results (de Cheveigné *et al.*, 1997) no effect of absolute F_0 (low/low

versus high/high or low/high versus high/low) was expected, so scores were averaged over that factor. For all stimuli, the number of vowels reported was noted.

II. EXPERIMENT 1: PHASE DEPENDENCY OF THE ΔF_0 EFFECT

Experiment 1 was designed to check whether the classic ΔF_0 effect depends on the component phases of constituent vowels. All vowels were harmonic. There were two ΔF_0 conditions: 0% and 6.45%. There were also three phase conditions: S/S (sine/sine), R/R (random/random, same pattern) and R/R’ (random/random, different patterns). In the S/S condition both vowels have peaked waveforms that could support a PPA mechanism. The R/R condition lacks peaked waveforms, but might conceivably support a weak form of the PPA hypothesis, based on the alignment of whatever temporal features are present. These features would be aligned at $\Delta F_0=0\%$ and misaligned otherwise. The R/R’ condition should defeat PPA altogether: the waveforms lack large peaks and have no features in common. As far as waveform interaction is concerned, S/S and R/R are equivalent: Both have the same intervowel phase pattern (0) and produce the same particular spectrum at $\Delta F_0=0$ (sum of the spectra of the constituents). In the R/R’ condition, intervowel phase is random and remains so with ongoing phase shifts due to ΔF_0 . The spectrum produced at $\Delta F_0=0$ is the result of random vector summation.

There were $(2\Delta F_0\text{'s}) \times (3 \text{ phase conditions}) \times (10 \text{ unordered vowel pairs}) \times (2 F_0\text{'s}) \times (2 \text{ amplitudes}) = 240$ different stimuli, repeated within each of five sessions. The scoring process described in Sec. I C retained responses for only one amplitude (–15 dB), but distinguished 20 ordered vowel pairs. Identification rates, averaged over F_0 ’s (2) and sessions (5), were subjected to a repeated-measures analysis of variance (ANOVA) with factors ΔF_0 (2), PHASE (3), and ordered vowel PAIR (20). Probabilities reflect, where necessary, an adjustment of the degrees of freedom by a factor that corrects for the inherent correlation of repeated measurements (Geisser and Greenhouse, 1958). The main effects of ΔF_0 [$F(1,5)=28.07$, $p=0.003$] and PAIR [$F(19,95)=5.03$, $p=0.006$, GG=0.2] were significant, indicating that same- F_0 pairs were identified with more difficulty than different- F_0 pairs (15% vs 66% overall) and that identification rate varied across vowel pairs (from 27% to 54% overall). Their interaction was not significant, nor were the main effect of PHASE and its interactions with the other factors. Identification rates averaged over pairs and subjects are plotted in Fig. 3(a). The ΔF_0 effect is large, and phase effects are negligible. These data do not support the hypothesis that the ΔF_0 effect observed in classic “double-vowel” experiments is specific to the phase patterns (Klatt or sine) that were employed. The average number of vowels reported per stimulus [Fig. 3(b)] is also strongly dependent on ΔF_0 but not on phase.

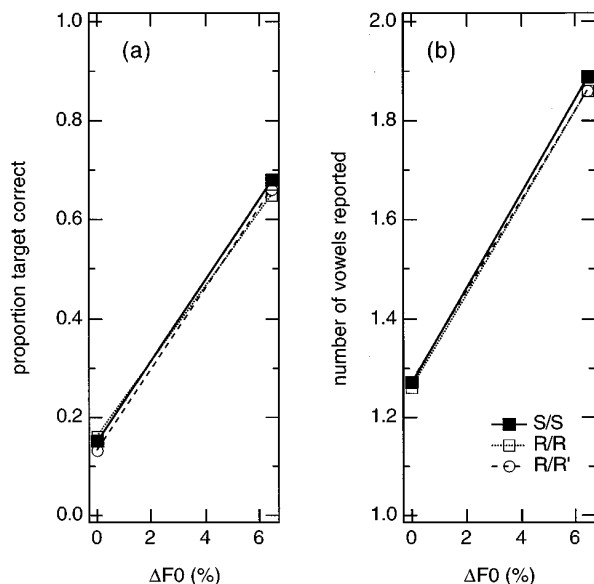


FIG. 3. (a) Target vowel identification rate as a function of ΔF_0 , for several phase patterns. (b) Number of vowels reported.

III. EXPERIMENT 2: TESTING FOR A PHASE ARTIFACT

In a previous experiment (de Cheveigné *et al.*, 1995a), we presented subjects with vowel pairs in which each vowel was either harmonic or inharmonic. When the F_0 's differed by 2.9%, we found that target identification was better when targets were inharmonic rather than harmonic. It was also better when the competing vowel was harmonic rather than inharmonic. We attributed that pattern of results to a particular segregation strategy (harmonic cancellation) that is sensitive to harmonicity. However, as pointed out in the Introduction, harmonic vowels used in that experiment were in sine phase, whereas inharmonic vowels shifted to random phase. If phase *per se* were sufficient to explain the results, then we should expect similar results for harmonic vowels with the same phase patterns. If such an outcome were observed, it would cast doubt on the generality of the conclusions of the harmonicity experiment.

Experiment 2 tested four phase relations that arose in the harmonicity experiment (the notation x/y means target phase x with background phase y): S/S (sine/sine), S/R (sine/random), R/S (random/sine), and R/R' (random/random, different random pattern), plus a fifth one: R/R (random/random, same random pattern), at a ΔF_0 of 6.45%. Identification rates were averaged over F_0 's (2) and sessions (5), and were subjected to a repeated-measures ANOVA with factors PHASE (5) and ordered vowel PAIR (20). The main effect of PAIR barely missed the 5% significance level [$F(19,95)=2.89$, $p=0.06$, $GG=0.18$]. Neither PHASE nor its interaction with PAIR were significant. Identification rates averaged over subjects, pairs, and sessions are plotted in Fig. 4. Phase does not appear to affect identification at this ΔF_0 , and there is no evidence of the hypothesized artifact.

IV. EXPERIMENT 3: HARMONICITY

Experiment 2 argued against the role of a phase artifact in the experiment reported by de Cheveigné *et al.* (1995a).

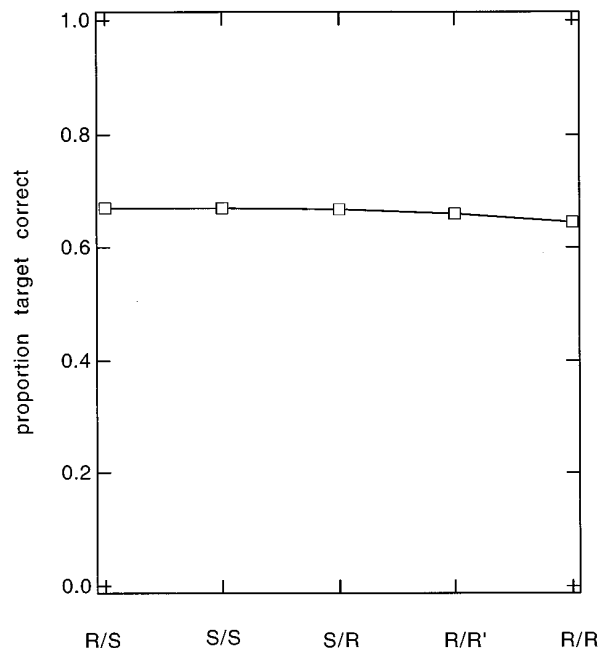


FIG. 4. Identification rate as a function of target and background vowel phase, at $\Delta F_0=6.45\%$.

Experiment 3 confirmed this conclusion by reproducing three crucial conditions of that experiment, using stimuli designed to reduce the usefulness of PPA or beat cues as follows. (1) Intravowel starting phase was "random," to reduce the salience of temporal cues, and each vowel had a different random phase, so residual temporal features, if any, were not common to both vowels. (2) Intervowel phase was also "random" and remained random with ongoing phase shifts due to ΔF_0 or inharmonicity. Beats were not eliminated, but as they occurred with random phases within different channels, there is no reason why the pattern arising in any particular condition should favor that condition over others. (3) Pairs containing inharmonic vowels had no partials closer than 16 Hz. To use spectral cues caused by beats, the auditory system would therefore have had to sample the beat pattern with a resolution better than about 30 ms. This cannot be excluded, but we expect it to be more difficult than with slower beats (note that making all beats faster than 16 Hz would have required a larger minimum spacing between partials, which is hard to reconcile with other constraints described in Sec. I A). (4) As explained in Sec. I A, all components of all vowels were multiples of 4 Hz, so all beats admitted the effective length of the stimulus (250 ms) as a period or subperiod. The long-term spectrum of the stimulus was therefore independent of starting phase.

The stimulus conditions were I/H, H/H, and H/I, with a ΔF_0 of 6.45% and an R/R' phase pattern. Following the scoring process described in Sec. I C, identification rates for the weaker vowel were averaged over F_0 's (2) and sessions (5), and subjected to a repeated-measures ANOVA with factors HARMONICITY (3) and ordered PAIR (20). The main factors of HARMONICITY [$F(2,10)=47.89$, $p=0.0004$, $GG=0.46$] and PAIR [$F(19,95)=3.10$, $p=0.04$, $GG=0.21$] were significant, as was their interaction [$F(31,190)=2.89$,

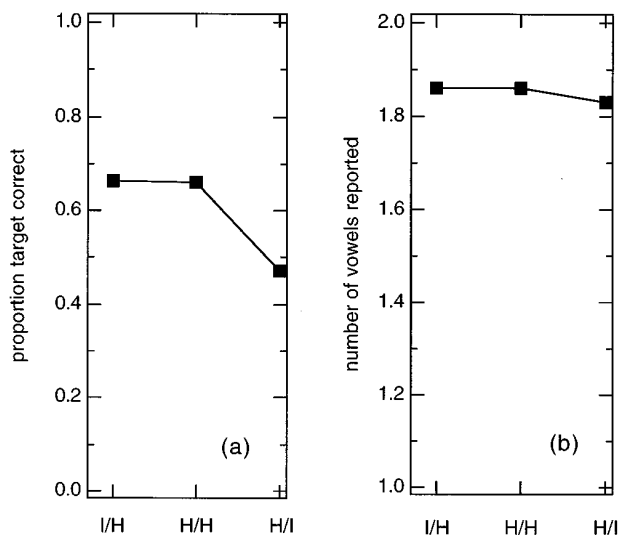


FIG. 5. (a) Identification rate as a function of target/ground harmonic state. (b) Number of vowels reported.

$p=0.04$, $GG=0.12$]. Identification rates averaged over pairs, subjects, and sessions are plotted in Fig. 5(a). Identification was better by about 21% when the background was harmonic rather than inharmonic. This effect is in the same direction as found by de Cheveigné *et al.* (1995a), but seven times larger. However, whereas they found that identification was better for inharmonic targets, here we observed no effect of target harmonicity.

In all conditions the compound stimulus was inharmonic. Inharmonicity of the *stimulus* seems to function as a strong multiplicity cue: The proportion of two-vowel responses was greater (86%) than when the stimulus was harmonic (27% in experiments 1 and 2 at $\Delta F_0=0$). This is evident also in the tendency of inharmonic single vowels to evoke more two-vowel responses than harmonic vowels (next paragraph). On the other hand for the inharmonic stimuli of experiment 3 it made no difference whether the *component* vowels were harmonic or not: the number of vowels reported did not differ significantly between conditions [Fig. 5(c)].

V. SINGLE VOWELS

The stimulus set used in experiments 1–3 comprised 200 single vowels in addition to 400 double vowels. All single vowels were identified correctly more than 99% of the time; there is nothing to suggest that the phonetic quality of the constituents of the double vowels used in experiments 1–3 was affected by their component phases or harmonicity (despite the fact that inharmonic vowels sounded unnatural). On the other hand, fewer than 9% of harmonic vowels but more than 63% of inharmonic single vowels evoked two-vowel responses.

VI. EXPERIMENT 4: TASK

The one-or-two response task we used differs from the classic two-response task used in double-vowel experiments. It is sensitive to “multiplicity” cues that influence the num-

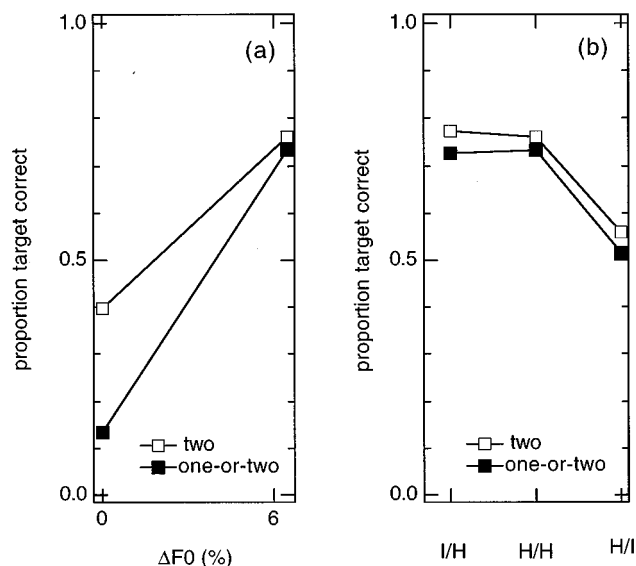


FIG. 6. (a) Identification rate as a function of ΔF_0 in the H/H condition for the two response task (open symbols) and the one-or-two response task (filled symbols). (b) Identification rate as a function of target/ground harmonic state at $\Delta F_0=6\%$, for both tasks.

ber of vowels reported, and also the identification rate. To see how the task affected identification rate, and permit comparison with previous reports, we reproduced several conditions of experiments 1–3 with the same subjects, but using the classic two-response task.

Subjects were five of the six that participated in experiments 1–3. Conditions were H/H at $\Delta F_0=0$, and I/H, H/H, and H/I at $\Delta F_0=6.45\%$. Phase was R/R'. Within the stimulus set there were (10 unordered pairs) \times (4 conditions) \times (2 F_0 orders) \times (2 amplitude orders). After scoring as explained in Sec. I C, identification rates were paired with those obtained in experiments 1–3 by the same subjects for the same conditions, and submitted to a repeated-measures ANOVA with factors TASK (2), CONDITION (4), and ordered PAIR (20). The main effect of CONDITION [$F(3,12)=56.47$, $p=0.0008$, $GG=0.39$] was significant as was the TASK by CONDITION interaction [$F(3,12)=21.37$, $p=0.008$, $GG=0.36$]. Results are plotted in Fig. 6. Identification rates at $\Delta F_0=0$ in the H/H condition were higher when the subjects were forced to report two vowels [$F(1,12)=115.20$, $p=0.0007$, $GG=0.36$]. The result is a smaller ΔF_0 effect size for the classic task [Fig. 6(a)]. Task had no significant influence on the pattern of results for harmonicity [Fig. 6(b)].

VII. DISCUSSION

A. ΔF_0

The ΔF_0 effect plotted in Fig. 3(a) is large compared to ΔF_0 effects usually observed. This results from the combined benefit of the 15-dB amplitude mismatch (de Cheveigné *et al.*, 1997) and the one-or-two-vowel task (Sec. I E). The effect was not reduced with intravowel phase patterns that eliminated waveform cues required by the PPA hypothesis, nor was it affected by the intervowel phase pattern that

determines both the relative phase of beats within different channels, and the spectrum of the double-vowel stimulus at $\Delta F_0=0\%$.

B. Phase

We found no measurable effect of phase at either ΔF_0 . Effects for factors other than phase were large, and our experiments did not lack statistical power. This result is surprising and hard to reconcile with the PPA hypothesis that presumably requires peaked waveforms. A possible explanation is that the 15-dB mismatch was too great for a PPA-type unmasking effect to occur, even with peaked waveforms. Another explanation is that harmonic cancellation was highly effective because the background F_0 was easy to estimate, and all other effects were dwarfed. If so, the amplitude mismatch that we introduced to enhance sensitivity actually had the opposite result.

The beat hypothesis was introduced to explain effects of ΔF_0 's smaller than one semitone (Culling and Darwin, 1994), but that mechanism might be expected to still have some effect at one semitone as in our experiment. In its simplest form, the beat hypothesis supposes that ongoing waveform interaction due to F_0 differences produces spectra that temporarily favor the identification of one vowel or the other (or perhaps the two together). If such were the case, one might expect identification to be affected by static, phase-dependent differences in waveform interactions at $\Delta F_0=0\%$, at least for individual vowel pairs. Instead, we found neither a main effect of phase, nor an interaction between phase and vowel pair. Once again, a possible explanation is that the 15-dB amplitude mismatch reduced spectral differences between phase conditions. An alternative form of the beat hypothesis is that identification depends on *dynamic* features of the beat pattern not present in our fixed-phase stimuli. Dynamic features also imply frequency cues exploitable by F_0 -guided mechanisms, so it is difficult to design an experiment that triggers one type of mechanism and not the other. Further evidence against the beat hypothesis may be found in the results of experiment 3. The I/H and H/I conditions are symmetrical and should produce similar beats, so a segregation mechanism based on beats cannot explain the asymmetry observed between these two conditions.

C. Background harmonicity

We found a strong effect of background harmonicity for both tasks [Figs. 5(a) and 6(b)]. The effect is the same as found previously (de Cheveigné *et al.*, 1995a), but about seven times larger. Several factors may explain the difference in effect size: (1) the larger ΔF_0 (6.45% rather than 2.9%); (2) the different inharmonic patterns, with larger mistunings; (3) the 15-dB amplitude mismatch that may have made harmonic cancellation more effective. There were also differences in vowel set, stimulus generation, and subjects. As previously, the results support the hypothesis that the auditory system uses a strategy of *harmonic cancellation* to separate vowels. Vowels that we called “inharmonic” were only mildly so (they retained a relatively clear pitch), which

may explain why the background harmonicity effect [Fig. 6(b), open symbols] was only about half the size of the ΔF_0 effect [Fig. 6(a), open symbols].

D. Target harmonicity

We found no effect of target harmonicity. This result contradicts our previous finding that a target was easier to identify when it was inharmonic rather than harmonic (de Cheveigné *et al.*, 1995a). That effect was paradoxical in that it was the opposite of the effect predicted by the hypothesis of harmonic enhancement. A tentative explanation that we offered was that cancellation is employed indiscriminately by the auditory system whenever segregation is called for. Harmonic targets are more likely to be victims of cancellation than inharmonic targets, so they are less well identified, hence the paradoxical effect. In the present experiment, targets were weak so the cancellation system would have found it more difficult to lock onto their harmonic structure, which may account for the lack of effect. In any case, neither experiment supported the hypothesis of *harmonic enhancement*.

E. Task

Subjects found the one-or-two response task in experiments 1–3 natural and easy to perform, and complained when they were forced to report two vowels in experiment 4. The one-or-two response task is sensitive to segregation cues that signal the *multiplicity* of sources. The classic task ignores these cues, since the subject must report two vowels whether they are heard or not. The one-or-two task produced larger ΔF_0 effects [Fig. 6(a)], mainly because identification was less good at $\Delta F_0=0$ where subjects tended to report only one vowel. Conditions that elicited double responses were less affected by the change of task [Fig. 6(b)].

One can object to the one-or-two task on the grounds that it taps into two different processes that both affect identification (one which senses the “multiplicity” of sources, the other which performs “unmasking”). Different subjects may give different weights to each, so one is not sure exactly what is being measured. Indeed, de Cheveigné *et al.* (1997) found that the pattern of identification conditional on two-vowel responses varied between subjects, suggesting differences in strategy. The classic two-response task is easier to interpret because subjects are encouraged to ignore “multiplicity” cues, so identification rates depend only on the “unmasking” process. Similar remarks might be made for identification thresholds measured by an adaptive technique used by Summerfield and Assmann (1991), Summerfield (1992), Summerfield and Culling (1992), Culling and Darwin (1994), and Culling and Summerfield (1995). In those experiments, subjects had to decide which interval contained the target and to identify the target. The background was a random vowel-like sound, different for each trial and each interval. It is possible that identification of the correct interval was aided by the presence of a “multiplicity” cue similar to those discussed here. However, according to J. Culling

(personal communication), listeners rarely made errors with regard to target interval in this paradigm, even at identification threshold.

The two measures (identification rate and number of vowels reported) are neither independent nor equivalent. In some cases they covaried [Fig. 3(a) and (b)]. In others, the response count was constant while identification rate varied [Fig. 5(a) and (b)]. In others the opposite was true: harmonic and inharmonic single vowels were recognized with the same accuracy, but the latter evoked double responses more often than the former (63% vs 9%). The number of vowels reported may be a useful measure in future studies of segregation.

VIII. CONCLUSIONS

(1) The ΔF_0 effect measured in a double-vowel experiment was not affected by the particular phase patterns chosen. This suggests that segregation was not the result of mechanisms sensitive to phase-dependent waveform interactions due to pitch-period asynchrony or beats.

(2) Phase effects did not constitute an artifact in a previous experiment on harmonicity. We reproduced our previous finding that identification is better when background vowels are harmonic. The result is consistent with the hypothesis of *harmonic cancellation*.

(3) We failed to reproduce our previous paradoxical finding that identification was better when targets were inharmonic rather than harmonic. Here we found no effect of target harmonicity. In either case the conclusion is the same; we found no evidence of *harmonic enhancement*.

(4) A task in which subjects may report one or two vowels is easier to perform and tends to produce larger ΔF_0 effects than the two-response task used in classic experiments. The *number of vowels reported* is an interesting measure of segregation, sensitive to cues that signal the multiplicity of sources.

ACKNOWLEDGMENTS

The experiments were carried out at ATR Human Information Processing Laboratories, within a research collaboration agreement with the Centre National de la Recherche Scientifique. The first author thanks ATR for its kind hospitality and the CNRS for leave of absence. Hideki Kawahara, Minory Tsuzaki, Kiyoaki Aikawa, Hiroaki Kato, and Ikuyo Masuda contributed ideas and advice and Rieko Kubo supervised the experiments. Thanks to John Culling of the MRC Institute of Hearing Research for providing the software for stimulus synthesis, and to him and Peter Assmann for detailed comments on previous versions of the manuscript.

- Assmann, P. F., and Summerfield, Q. (1988). "Pitch-pulse asynchrony and the perceptual segregation of competing voices," Speech 88 Conference (7th FASE), Institute of Acoustics, Edinburgh, pp. 531–538.
- Assmann, P. F., and Summerfield, Q. (1990). "Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies," J. Acoust. Soc. Am. **88**, 680–697.
- Assmann, P. F., and Summerfield, Q. (1994). "The contribution of waveform interactions to the perception of concurrent vowels," J. Acoust. Soc. Am. **95**, 471–484.

- Brox, J. P. L., and Nooteboom, S. G. (1982). "Intonation and the perceptual separation of simultaneous voices," J. Phon. **10**, 23–36.
- 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.
- Culling, J. F., and Darwin, C. J. (1993). "Perceptual separation of simultaneous vowels: Within and across-formant grouping by F_0 ," J. Acoust. Soc. Am. **93**, 3454–3467.
- Culling, J. F., and Darwin, C. J. (1994). "Perceptual and computational separation of simultaneous vowels: Cues arising from low frequency beating," J. Acoust. Soc. Am. **95**, 1559–1569.
- Culling, J. (1996). Personal communication.
- Culling, J. F., and Summerfield, Q. (1995). "The role of frequency and modulation in the perceptual segregation of concurrent vowels," J. Acoust. Soc. Am. **98**, 837–846.
- Darwin, C. J. (1981). "Perceptual grouping of speech components differing in fundamental frequency and onset-time," Q. J. Exp. Psychol. **A 33**, 185–207.
- Darwin, C. J., and Gardner, R. B. (1986). "Mistuning of a harmonic of a vowel: Grouping and phase effects on vowel quality," J. Acoust. Soc. Am. **79**, 838–845.
- de Cheveigné, A. (1993). "Separation of concurrent harmonic sounds: Fundamental frequency estimation and a time-domain cancellation model of auditory processing," J. Acoust. Soc. Am. **93**, 3271–3290.
- de Cheveigné, A. (1994). "Strategies for voice separation based on harmonicity," Proceedings of the International Conference Speech and Language Processing, Yokohama (Acoustical Society of Japan), pp. 1071–1074.
- de Cheveigné, A. (1995). "Experiments in vowel segregation" ATR Human Information Processing Research Labs Tech. Report TR-H-154 (unpublished).
- de Cheveigné, A., McAdams, S., Laroche, J., and Rosenberg, M. (1995a). "Identification of concurrent harmonic and inharmonic vowels: A test of the theory of harmonic cancellation and enhancement," J. Acoust. Soc. Am. **97**, 3736–3748.
- de Cheveigné, A., Kawahara, H., Tsuzaki, M., and Aikawa, K. (1995b). "Sensitive experimental techniques for the study of sound segregation," ASJ Autumn meeting (Acoustical Society of Japan), pp. 373–374.
- de Cheveigné, A., Kawahara, H., Tsuzaki, M., and Aikawa, K. (1997). "Concurrent vowel segregation I. Effects of relative amplitude and F_0 difference," J. Acoust. Soc. Am. **101**, 2839–2847.
- Geisser, S., and Greenhouse, S. W. (1958). "An extension of Box's results on the use of the F distribution in multivariate analysis," Ann. Math. Stat. **29**, 885–889.
- Holdsworth, J., Nimmo-Smith, I., Patterson, R. D., and Rice, P. (1988). "Implementing a GammaTone filter bank (SVOS final report, annex C)," MRC Applied Psychology Unit Tech. Report (unpublished).
- 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.
- 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.
- Lea, A. (1992). "Auditory models of vowel perception," Ph.D. thesis, Nottingham (unpublished).
- McKeown, J. D. (1992). "Perception of concurrent vowels: The effect of varying their relative level," Speech Commun. **11**, 1–13.
- Moore, B. C. J., and Alcántara, J. I. (1995). "Identification of flat-spectrum vowels on the basis of amplitude modulation," J. Acoust. Soc. Am. **97**, 3274.
- Moore, B. C. J., and Alcántara, J. I. (1996). "Vowel identification based on amplitude modulation," J. Acoust. Soc. Am. **99**, 2332–2343.
- Moore, B. C. J., and Glasberg, B. R. (1983). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," J. Acoust. Soc. Am. **74**, 750–753.
- Palmer, A. R., Winter, I. M., Gardner, R. B., and Darwin, C. J. (1987). "Changes in the phonemic quality and neural representation of a vowel by alteration of the relative phase of harmonics near F_1 ," in *The Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Martinus Nijhoff, Dordrecht), pp. 371–376.
- Plack, C. J., and Moore, B. C. J. (1990). "Temporal window shape as a function of frequency and level," J. Acoust. Soc. Am. **87**, 2178–2187.
- Scheffers, M. T. M. (1983). "Sifting vowels," Ph.D. thesis, Gröningen (unpublished).

- Schroeder, M. R., and Strube, H. W. (1986). "Flat-spectrum speech," J. Acoust. Soc. Am. **79**, 1580–1583.
- Summerfield, Q. (1992). "Roles of harmonicity and coherent frequency modulation in auditory grouping," in *The Auditory Processing of Speech: from Sounds to Words*, edited by M. E. H. Schouten (Mouton de Gruyter, Berlin), pp. 157–166.
- Summerfield, Q., and Assmann, P. F. (1991). "Perception of concurrent vowels: Effects of harmonic misalignment and pitch-period asynchrony," J. Acoust. Soc. Am. **89**, 1364–1377.
- Summerfield, Q., and Culling, J. F. (1992). "Periodicity of maskers not targets determines ease of perceptual segregation using differences in fundamental frequency," J. Acoust. Soc. Am. **92**, 2317(A).
- Traunmüller, H. (1987). "Phase vowels," in *The Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Martinus Nijhoff, Dordrecht), pp. 377–384.
- Zwicker, U. T. (1984). "Auditory recognition of diotic and dichotic vowel pairs," Speech Commun. **3**, 256–277.

Concurrent vowel identification. III. A neural model of harmonic interference cancellation

Alain de Cheveigné

Centre National de la Recherche Scientifique/Université Paris 7, 2 place Jussieu, case 7003, F-75251 Paris
Cédex 05, France and ATR Human Information Processing Research Laboratories, 2-2 Hikaridai,
Seika-cho Soraku-gun, Kyoto 619-02, Japan

(Received 19 December 1995; revised 12 August 1996; accepted 22 November 1996)

This paper presents a “neural cancellation filter” capable of segregating weak targets from competing harmonic backgrounds, and a model of concurrent vowel segregation based upon it. The elementary cancellation filter comprises a delay line and an inhibitory synapse. Filters within each peripheral channel are tuned to the period of the competing sound to suppress its correlates within the neural discharge pattern. In combination with a pattern matching model based on autocorrelation functions summed over channels, the cancellation filter forms a model of concurrent vowel identification. The model predicts the number of vowels reported for each stimulus (when subjects are allowed to report one or two) and identification rates. It belongs to the class of “harmonic cancellation” models that are supported by experimental evidence that vowel identification is better when competing sounds are harmonic than inharmonic. Two alternative schemes using the same filter are also considered. One derives a “place” representation from the magnitude of the filter output. The other uses the ratio of filter input/output to select channels. © 1997 Acoustical Society of America. [S0001-4966(97)04104-0]

PACS numbers: 43.71.An, 43.71.Cq, 43.64.Bt, 43.66.Hg [WS]

INTRODUCTION

In “double-vowel” experiments, subjects identify concurrent steady-state synthetic vowels better when they differ in fundamental frequency (F_0) than when F_0 's are the same. Identification is also improved when vowels are modulated in frequency in a fashion that introduces instantaneous F_0 differences (ΔF_0) (Summerfield, 1992; Summerfield and Culling, 1992a; Culling and Summerfield, 1995b), or when concurrent synthetic or natural voices have different intonation patterns (Brokx and Nootboom, 1982). A number of “ F_0 -guided” segregation models have been proposed to explain this effect [see de Cheveigné (1993a) for a review], that can be classified into three categories: spectral, spectro-temporal, and temporal. Spectral models, derived from Parsons' (1976) harmonic selection method, require a spectral representation that can resolve individual harmonics. Assmann and Summerfield (1990) implemented such a model based on a computer model of peripheral filtering, and found that it lacked the necessary resolution. Spectro-temporal models (Assmann and Summerfield, 1990; Lea, 1992; Meddis and Hewitt, 1992; Weintraub, 1985) are less demanding in terms of resolution: Whereas a spectral model needs to resolve individual partials, a spectro-temporal model needs only to ensure that some channels are dominated by one vowel, and others by the other vowel.

Among spectro-temporal models, the model of Meddis and Hewitt (1992) is perhaps the most effective in terms of predictive power and physiological plausibility (Palmer, 1992). It works by partitioning auditory channels between those that respond with the same period as the “dominant” vowel, and those that do not. The dominant period is determined from the largest peak in a summary autocorrelation function (Licklider, 1959, 1962; Meddis and Hewitt, 1991).

Each group of channels is used to identify a vowel, and the model works as long as all channels are not dominated by the same vowel. If all channels are dominated by the same vowel, then there is no partition, and thus no advantage when vowels have different F_0 's. The model was applied to the stimulus set of de Cheveigné *et al.* (1997a) that included conditions for which one vowel was weaker than the other by 10 or 20 dB. For certain vowel pairs all channels were dominated by the stronger vowel, and the model therefore predicted no improvement with ΔF_0 . Nevertheless, strong ΔF_0 effects were observed for these vowel pairs.

F_0 -guided segregation models of the third category (temporal) do not explicitly require peripheral filtering to split information between vowels. De Cheveigné (1993a) proposed a time-domain filtering model that partitioned neural information *within* channels, rather than between channels, using a “neural cancellation filter.” That model was tested with physiological data recorded from the guinea pig auditory nerve, but it was not tested in detail with stimuli of double-vowel identification experiments. The purpose of this paper is to do so. The “neural cancellation filter” is first described in a simplified form that operates on discharge probability. The behavior of the filter is then illustrated with stimuli employed in double-vowel identification experiments, and three possible segregation schemes are outlined. Finally, one of these schemes is developed into a quantitative model of concurrent vowel identification, and its predictions are compared to the response counts and identification rates that were measured experimentally using the same stimuli.

I. THE NEURAL CANCELLATION FILTER

The time-domain “neural comb filter” involves delay and inhibition, and operates within auditory channels

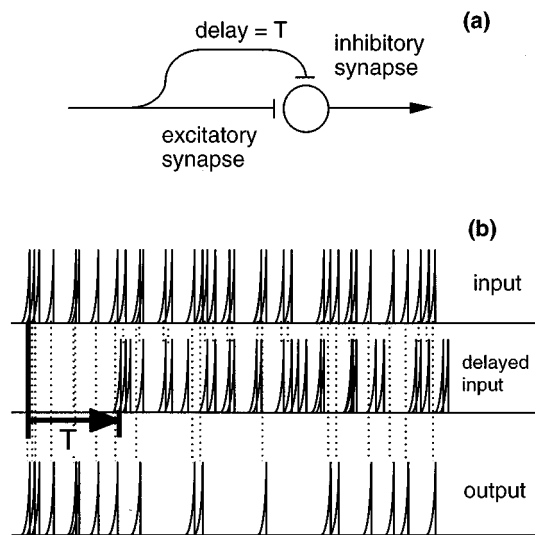


FIG. 1. (a) Schema of neural filter. (b) Input and output spike trains. The delay is equal to delay parameter T of the filter. Input spikes that coincided with delayed spikes are no longer present in the output.

(auditory-nerve fibers or groups of similar fibers). It removes those spikes that form, with a preceding spike, an interval equal to the delay. The result is that the spike train contains fewer such intervals. Supposing that the filter is tuned to remove intervals equal to the period of some harmonic interference, correlates of that interference are *cancelled*. In this respect the filter is compatible with recent data showing that vowel identification is better when interference is harmonic than inharmonic (Summerfield and Culling, 1992b; Lea, 1992; de Cheveigné *et al.*, 1995, 1997a, b). As in Meddis and Hewitt's (1992) model, the neural filter partitions neural data into correlates of the dominant vowel, that are removed, and a remainder that is used to identify the weaker vowel. In contrast to their model, however, the partition occurs within individual channels rather than between channels.

A "neural cancellation filter" is implemented by a neuron with two synapses [Fig. 1(a)]. One is excitatory and fed by the input spike train, the other is inhibitory and fed by a delayed version of the input. Synapse and neuron properties are such that the neuron fires each time an input spike arrives, *unless* a spike arrives simultaneously along the delayed path. The result is that some spikes are weeded out, reducing the number of intervals equal to the delay [Fig. 1(b)]. de Cheveigné (1993a) tested the filter with auditory-nerve fiber data recorded in the guinea pig in response to double-vowel stimuli (Palmer, 1990), and found that it partitioned neural data into patterns that clearly reflected the individual constituent vowels.

The behavior of a physiological mechanism following this principle would depend on many factors (number of afferents, pattern of innervation, spike generation process, spike rate, integration time constants, etc.). These are ignored in a simpler formulation, appropriate to a large ensemble of neurons, that captures the "vowel segregation" properties of the filter:

$$o(t) = \text{MAX}(0, i(t) - \Phi[i](t - T)),$$

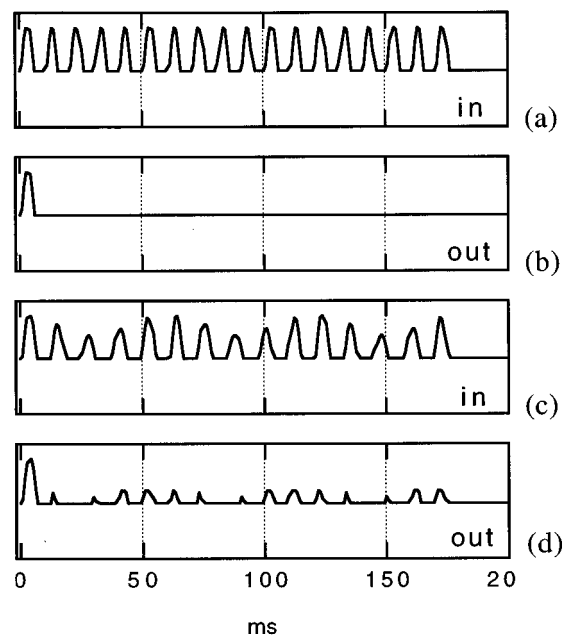


FIG. 2. (a) 100-Hz half-wave rectified sine wave input and (b) output of a neural cancellation filter. (c) Half-wave rectified sum of two sines of frequencies 80 and 100 Hz, differing in amplitude by 10 dB; (d) output of neural cancellation filter. The neural filter was tuned to cancel the stronger 80-Hz component.

where $o(t)$ is the output neuron's firing probability, $i(t)$ is input spike probability, and T is a lag parameter. The "MAX" operation reflects the fact that firing probability cannot be negative. $\Phi[\]$ is an operator designed to represent temporal smearing due to temporal integration, etc. In the following we further simplify the filter by supposing $\Phi = \text{identity}$:

$$o(t) = \text{MAX}(0, i(t) - i(t - T)). \quad (1)$$

The behavior of the filter is illustrated in Fig. 2.

Figure 2 represents discharge probability within an auditory-nerve fiber (or group of fibers) in response to a 100-Hz pure-tone pulse. The representation is idealized—ringing and adaptation effects are ignored, and transduction is represented as a simple half-wave rectification. The first peak of the cancellation filter input (a) finds its way to the output (b), but the following peaks are suppressed. In Fig. 2(c), the filter input is a half-wave rectified sum of sines, as might occur for example within an auditory-nerve channel in response to concurrent vowels with different F_0 's. One of the sines, the "target," is the same 100-Hz sine as in (a), the other competing sine is lower in frequency (80 Hz) and higher in level by 10 dB. The filter is tuned to cancel the stronger competing sine. The output (d) consists of a peak followed by a modulated series of small peaks with the same spacing as the weaker target sine. This filter can be exploited to segregate two vowels on the basis of ΔF_0 according to at least three schemes that all suppose that the neural cancellation filter is tuned to cancel the period of the dominant vowel. The segregation schemes are illustrated in Sec. II with double-vowel stimuli used in experiments described in companion papers (de Cheveigné *et al.*, 1997a,b). In Sec. III

TABLE I. Formant frequencies in units of Hz and ERB (Moore and Glasberg, 1983), and equivalent formant periods in ms of the two vowels used to illustrate the segregation schemes.

	/u/			/o/		
	Hz	ERB	ms	Hz	ERB	ms
F_0	125	3.7	8.00	132.5	3.8	7.55
F_1	312	7.5	3.21	468	9.9	2.14
F_2	1219	16.9	0.820	781	13.4	1.28
F_3	2469	22.7	0.405	2656	23.3	0.377
F_4	3406	25.3	0.294	3281	25.0	0.304
F_5	4200	27.0	0.238	4200	27.0	0.238

one scheme is further developed into a model of concurrent vowel segregation, and compared with the experimental results.

The stimuli were based on Japanese vowels /a/, /i/, /u/, /e/, and /o/, synthesized at F_0 's of 125 and 132.5 Hz at a sampling rate of 16 kHz. Each stimulus consisted of either a single vowel or the sum of two vowels at $\Delta F_0=0\%$ or $\Delta F_0=6\%$. The rms level of all stimuli was the same, but the contribution of each vowel within a pair was varied so that a vowel could be at -20 , -10 , 0 , 10 , or 20 dB relative to its companion. Stimuli were 200 ms in duration, with 20-ms raised-cosine onset and offset ramps. In Sec. II the segregation schemes are illustrated using a subset of the stimuli: /u/ at 125 Hz, and /o/ at 125 and 132.5 Hz. The F_0 's and formant frequencies of these two vowels are listed in Table I in units of Hz and ERB (equivalent rectangular bandwidth) (Moore and Glasberg, 1983) together with the corresponding periods in ms. Stimuli were fed to a model of peripheral filtering (Holdsworth *et al.*, 1988; Patterson *et al.*, 1988; Culling, 1996). The center frequencies (CF) of peripheral filter channels were spaced along an ERB scale with a density of 4 channels per ERB. Each channel was processed by a model of hair-cell transduction (Meddis, 1986, 1988). To illustrate better the effects of the cancellation filter in the "place" segregation scheme described below, the input level of the hair-cell model was chosen to be relatively high to drive the hair cells to saturation over most of the spectrum, and reduce place cues (Sachs and Young, 1979). The output of the hair-cell model (instantaneous density of auditory-nerve fiber firing probability) was fed to the cancellation filter defined by Eq. (1). If the delay parameter was not a multiple of the sampling period, delay was implemented by linear interpolation between adjacent samples.

II. THREE SEGREGATION SCHEMES

A. Place scheme

In channels that respond only to the dominant vowel, the output of the cancellation filter should show an initial peak followed by zero output [Fig. 2(a)]. If the channel is influenced by the other vowel, even only weakly, the output following the initial pulse should not be zero [Fig. 2(b)]. Supposing that the amplitude of this residue is proportional to the relative amplitude of the weaker vowel at the frequency of the channel, the profile of cancellation filter outputs over channels should produce a place representation from which

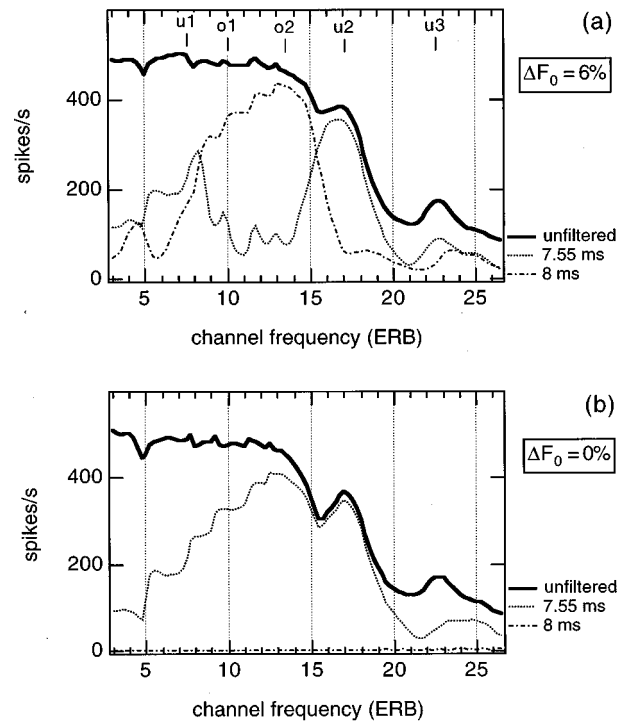


FIG. 3. (a) Input (continuous line) and output (dotted lines) of a neural cancellation filter as a function of channel frequency on an ERB scale, for the vowel pair /u/ (125 Hz, 8 ms) + /o/ (132.5 Hz, 7.55 ms) at a relative amplitude of 0 dB. The filter was tuned to cancel each of the constituent vowels in turn. Symbols indicate the frequencies of the first formants of each vowel. (b) Same, for vowel pair /u/ + /o/ (125 Hz, 8 ms) at a relative amplitude of 0 dB. The filter is tuned to the same two periods as in (a), one of which is the common period of both vowels.

that vowel might be identified. This is illustrated in Fig. 3(a) for the vowel pair /u/ + /o/ at $\Delta F_0=6\%$. The thick, continuous line corresponds to the profile of unfiltered activity over channels of the gammatone/hair-cell model. The dotted lines correspond to the profile of cancellation filter outputs when the filter delay parameter is set to each of the vowel periods in turn. When the filter is tuned to the period of /o/ (7.55 ms) the profile has peaks at the first two formants of /u/ (indicated by symbols). When it is tuned to the period of /u/ (8 ms) the profile may be interpreted as reflecting the first two formants of /o/. Filtered profiles might thus serve as a substrate for vowel identification. When the vowel F_0 's are the same, almost all activity is suppressed if the filter is tuned to the common period, 8 ms [Fig. 3(b), dot-dash line near axis], whereas if it is tuned to a different period (7.55 ms) the profile is not specific to either vowel (dotted line). The fact that filtered profiles reflect constituent vowels better at $\Delta F_0=6\%$ than at $\Delta F_0=0\%$ could explain why identification is better in the former case than in the latter.

The filter is still effective when /u/ is at 0, -10 , or even -20 dB (lower thin lines in Fig. 4) relative to the competing vowel /o/. At 10 dB the profile is not so clear, and at 20 dB it is the same as that obtained for /u/ in isolation (thick continuous line). Both the filtered and the unfiltered profiles for /u/ in isolation (uppermost in Fig. 4) lack clear peaks at formants F_1 and F_2 of /u/, due to the compressive saturation properties of the hair-cell model. The driving level was cho-

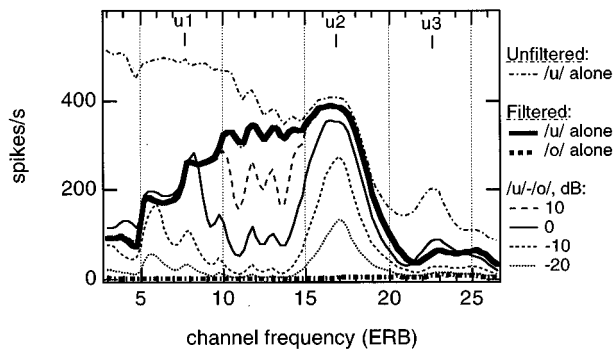


FIG. 4. Input (uppermost dot-dash line) and output (other lines) of a neural cancellation filter as a function of channel frequency on an ERB scale. Thin lines are for vowel pair /u/ (125 Hz, 8 ms)+/o/ (132.5 Hz, 7.55 ms) at several relative amplitudes. Thick lines are for single vowels. The filter was tuned to cancel the period of /o/.

sen relatively high so as to illustrate this point.¹ Paradoxically, formant-related features are clearer for weak segregated vowels (-20 to 0 dB) than for the same vowels in isolation. This segregation scheme might be developed into a model of concurrent vowel identification, but we shall not attempt to do so here.

B. Time-domain scheme

Instead of place information, time-domain information at cancellation filter outputs may be exploited to identify vowels. One way to summarize this information is to calculate autocorrelation functions (ACFs) within each channel and add them to obtain a summary ACF (SACF), as in the model of Meddis and Hewitt (1992). The SACF tends to vary with the square of input level, which makes it difficult to plot SACFs for different levels on the same graph. To compensate for this dependency and make the figures easier to read, we plot the *square root* of the SACF. Figure 5(a) (continuous line) shows the square-root SACF for the vowel pair /u/+/o/ at 0 dB relative amplitude and $\Delta F_0=6\%$. There is a peak at the period of /o/ (7.55 ms), suggesting that the unfiltered response is overall dominated by that vowel. When the filter is tuned to the period of /o/ (7.55 ms), there is a peak at the period of /u/ (8 ms). When it is tuned to the period of /u/, the peak is instead at the period of /o/. The two filtered SACFs differ from each other, and from the unfiltered SACF. When both vowels have the same F_0 (125 Hz), filtering is not effective [Fig. 5(b)]. If the filter is tuned to the common period of the vowels (8 ms), it wipes out all activity. If it is tuned to a different period (7.55 ms), it produces a pattern similar to the unfiltered pattern. This could explain why subjects identify vowels less well at $\Delta F_0=0\%$ than when F_0 's are different.

The filter tuned to cancel /o/ remains effective even when /u/ is weak. Figure 6 displays square-root SACFs restricted to lags shorter than 4.5 ms (the "timbre region" that Meddis and Hewitt's model used to identify vowels). The unfiltered square-root SACF for the isolated vowel /u/ is represented by the topmost line. Other lines represent SACFs obtained after filtering with a cancellation filter tuned to suppress 8 ms (period of /o/). Each line represents a different

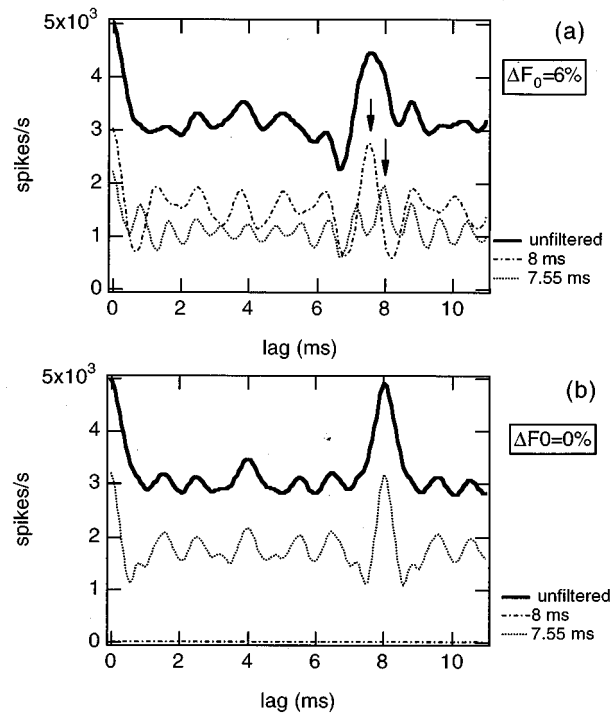


FIG. 5. (a) Square root SACF calculated from input (continuous line) and output (dotted lines) of neural cancellation filter for vowel pair /u/ (125 Hz, 8 ms)+/o/ (132.5 Hz, 7.55 ms) at a relative amplitude of 0 dB. The filter is tuned to cancel each of the periods in turn. Arrows indicate the periods of each vowel. (b) Same, for vowel pair /u/+/o/ (125 Hz, 8 ms) at a relative amplitude of 0 dB. The filter is tuned to the same two periods as in (a), one of which is the period common to both vowels.

proportion of /u/ in the input stimulus. The filtered response for /u/ alone is represented by the thick continuous line, and thin lines represent responses for the vowel pair /u/+/o/. All filtered patterns are similar to the filtered pattern for /u/ alone, which itself is similar (with the exception of the dip near 0.5 ms) to the unfiltered pattern for /u/ alone. None of these patterns reflects the presence of /o/, even at -20 dB where the /u/+/o/ stimulus contains mostly /o/. A concurrent vowel identification model based on this scheme is presented in Sec. III.

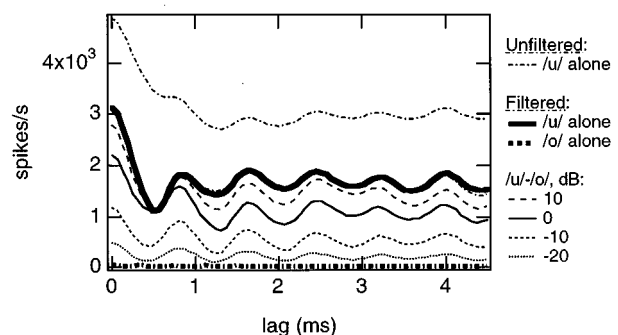


FIG. 6. Uppermost line: Square-root SACF in response to /u/, without filtering. Other lines: Square-root SACF's obtained after filtering to suppress the period of /o/ (8 ms). Thin lines are in response to vowel pair /u/ (125 Hz, 8 ms)+/o/ (132.5 Hz, 7.55 ms) at several relative amplitudes. Thick lines are for single vowels. Lags are restricted to the "timbre" region (Meddis and Hewitt, 1992).

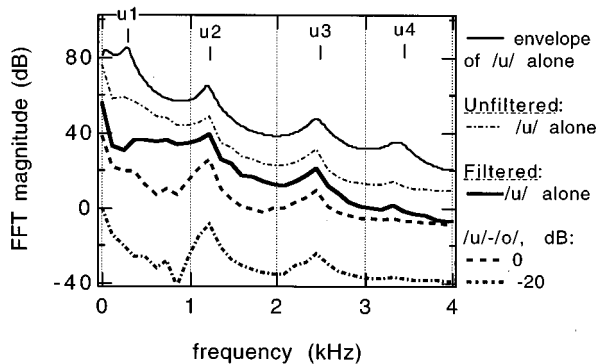


FIG. 7. Uppermost line: spectral envelope of /u/. Uppermost thick line: Fourier transform (FT) of SACF of unfiltered response to /u/. Full thick line: FT of SACF in response to /u/ (125 Hz), filtered at 7.55 ms. Lowest two lines: FT of SACF in response to /u/ (125 Hz)+/o/ (132.5 Hz) at 0 and -20 dB relative amplitude of target /u/. The origin of the ordinate (dB) is arbitrary. The lowest four plots were given an extra 10-dB spacing for visibility.

The SACFs of Fig. 6 are not easy to interpret in terms of spectral features, apart from the fact that peaks at 0.8 ms and multiples seem to reflect the F_2 of /u/ (1219 Hz). Figure 7 shows Fourier transforms derived from the first 8-ms portion of the SACFs, together with the spectral envelope used to synthesize /u/ (topmost line). For the vowel pair /u/+/o/ at 0 and -20 dB, the filtered SACF Fourier transform shows peaks at F_2 and F_3 of /u/ (lowest thin lines). Both patterns resemble the filtered and unfiltered patterns for /u/ alone (thick lines). All show similarities with the spectral envelope used to synthesize the vowel /u/ (uppermost line), apart from a complete lack of evidence of F_1 . The Fourier transform reinforces the conclusion that the filtered SACFs carry evidence of /u/ (F_2 and F_3), and no evidence of /o/. It is, however, debatable whether the auditory nervous system could carry out a Fourier transform to exploit this evidence.

C. Channel selection scheme

Instead of calculating ACFs for all cancellation-filtered channels and pooling them, we may use the filter output-to-input ratio as a criterion to select channels, as in Meddis and Hewitt's (1992) model. The main difference with their scheme is that the ratio threshold may be tuned to be more or less stringent in the selection of channels. Channel selection is also analogous to the scheme of Sec. II B, in that channels are given nonuniform weights before summation. The weighting is graded in one case, all-or-none in the other. A more important difference is that channels were filtered by the cancellation filter in the previous scheme, whereas here they are not. Filtering causes distortion that might interfere with template matching (compare filtered and unfiltered SACFs for /u/ in Fig. 6), so the lack of filtering in the present channel selection scheme might be an advantage.

Channel selection was implemented by filtering all channels with the cancellation filter tuned to the period of the competing vowel, and calculating the ratio between rms filter output and input (integrated over a 100-ms square window starting 50 ms from stimulus onset). Channels for which the output/input ratio fell below a threshold (arbitrarily set to 0.25) were considered dominated by the competing vowel,

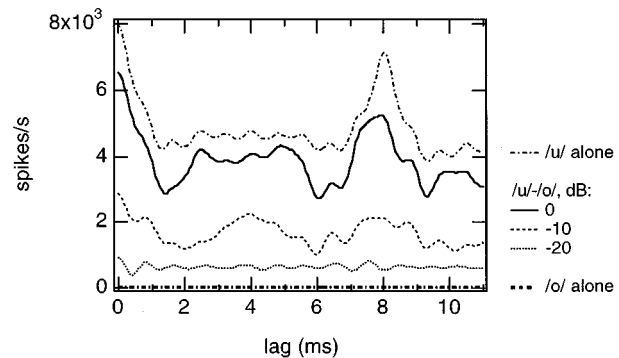


FIG. 8. SACF's calculated from selected channels not dominated by the period of the competing vowel (/o/, 132.5 Hz, 7.55 ms). A channel was "not dominated" if the rms output of the filter tuned to that period was greater than 0.25 times the rms input. Dotted uppermost line is the SACF calculated from all channels in response to /u/ alone. Other lines are SACF's calculated from selected channels, for /u/ alone (continuous line) and /u/ (125 Hz)+/o/ (132.5 Hz) at several values of the relative amplitude.

and discarded. ACFs for remaining *unfiltered* channels were calculated and summed to form a SACF representative of the target vowel. Figure 8 illustrates the result of this process for target /u/ mixed with competing vowel /o/. The uppermost line is the SACF for /u/ alone, calculated from all channels. The lowermost line is for /o/ alone, for which no channels were selected. The remaining lines are SACFs in response to the vowel pair /u/+/o/, calculated by summing channels that survived the selection process. When /u/ is at -20 dB, a few channels are selected in the vicinity of 17 ERB, close to the F_2 of /u/ (see Fig. 4), as suggested by the period of the ripple in the SACF. However, the peak near 7.5 ms in the SACF shows that those channels were predominantly responding with the period of /o/. The same is true when /u/ is at -10 dB, and to a certain degree, 0 dB. At all three amplitudes the SACFs are rather different in shape from the SACF obtained in response to /u/ alone. In the previous scheme of Sec. II B, a cancellation filter processed each channel and eradicated all evidence of the competing vowel. In the present scheme, unfiltered selected channels are more subject to "talk through" from the competing vowel. The channel selection scheme could form the basis of a concurrent vowel identification model similar to that of Meddis and Hewitt (1992), but we do not attempt to do so in this paper.

III. A MODEL OF CONCURRENT VOWEL IDENTIFICATION

This section takes the segregation scheme of Sec. II B and develops it into a model of concurrent vowel identification. For each stimulus, the model predicts (a) the number of vowels reported and (b) the name(s) of the vowel(s) reported. The aim is to prove that such a model can be built on the basis of the neural cancellation filter, and to work out enough detail to make a comparison with experimental data. The important question is whether the neural cancellation filter is consistent with major qualitative features of experimental results (dependency on ΔF_0 , amplitude, harmonicity, etc.). The quantitative match and specific details of the identification model are considered secondary.

A. Implementation details

After peripheral filtering and hair-cell transduction, the running autocorrelation function (ACF) of the auditory-nerve discharge probability within each channel was calculated with a time constant of integration of 12.5 ms. ACFs were sampled at 150 ms from stimulus onset, summed, and an estimate of the “dominant period” was obtained from the largest peak within the SACF in the range 5–10 ms. That range was chosen to include both periods of the stimulus vowels, and exclude their subperiods (harmonics) and superperiods. Next, the auditory-nerve discharge probability within each channel was fed through a cancellation filter tuned to the dominant period, and ACFs were calculated and summed to obtain a second, filtered SACF pattern. The average rms output-to-input ratio of the cancellation filter was also calculated. Following Meddis and Hewitt (1992), vowel identification was performed by matching SACFs to templates obtained for single vowels. SACFs were restricted to the portion between 0 and 2.5 ms to reduce the influence of F_0 , and normalized to have a mean of zero and a variance of one. The uniform Euclidean distance was calculated between each token and all templates, and the closest template was chosen as the match.

Based on these elements (output-to-input ratio, filtered and unfiltered SACF patterns, templates) and two threshold parameters T1 and T2, the model predicts the number of vowels reported for each stimulus and their names. The algorithm is as follows. The output-to-input ratio is compared to threshold T1. If it is larger, the model decides that two vowels are present. The first is determined from the best match to the unfiltered SACF. The second is determined from the best match to the filtered SACF, unless this produces the same match as the first vowel. In that case the second best match to the unfiltered SACF is taken instead (this rule enforces the restriction that both vowels must be different). If the output-to-input ratio is smaller than T1, the model discards the filtered SACF and attempts two matches of the unfiltered SACF. If the distance ratio between the two matches is above threshold T2, the model predicts two response: the two matches. Otherwise it predicts one response: the best match. The model thus predicts a single response only if the cancellation filter residual is small *and* the nonfiltered SACF can unambiguously be matched by a single template.

B. Predicted effects of ΔF_0 and amplitude

The symbols in Fig. 9 represent experimental results obtained by de Cheveigné *et al.* (1997a), and the lines are predictions of the model for the same stimuli. Consider first $\Delta F_0=0\%$ (filled symbols and thick lines). Threshold T2 was adjusted to obtain a reasonable match for the number of vowels reported [Fig. 9(a)]. The value chosen (0.2) also produced identification rates that are close to experimental values [Fig. 9(b)].

Consider now results at $\Delta F_0=6\%$ (open symbols and thin lines). Threshold T2 was maintained at its previous value and threshold T1 was set to 0.01. The number of vowels reported is well predicted [Fig. 9(a)], but the prediction

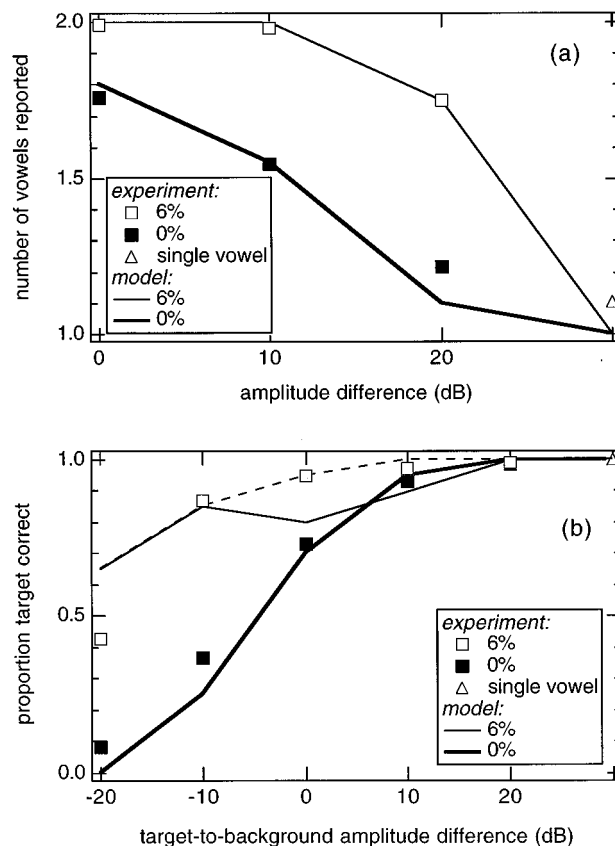


FIG. 9. (a) Symbols: Number of vowels reported for each double-vowel stimulus as a function of relative amplitude, at $\Delta F_0=0\%$ (filled) and $\Delta F_0=6\%$ (open), measured experimentally. Lines: Prediction of the model. (b) Symbols: Identification rate as a function of amplitude of target relative to background, at $\Delta F_0=0\%$ (filled) and $\Delta F_0=6\%$ (open), measured experimentally. Lines: Prediction of the model. The dotted line is the rate that would be predicted at $\Delta F_0=6\%$ if the model's period estimates were accurate.

for identification rate is somewhat less good [Fig. 9(b)]. The discrepancy at 0 and 10 dB may be attributed to inaccurate period estimation in the first stage of the model: If estimates are corrected, the predictions follow the dotted line in Fig. 9(b). The discrepancy at -20 dB may be attributed to the fact that the cancellation filter is implemented linearly and works well over a large dynamic range (Fig. 6), whereas a physiological implementation would certainly have a more limited dynamic range. The model successfully explains the one aspect of the experimental data that Meddis and Hewitt's (1992) model had difficulty accounting for the improvement in identification with ΔF_0 when the target is very weak. It also nicely predicts the number of vowels reported for each stimulus.

One should not give excessive weight to the quantitative match between predictions and experimental data: Each condition was tested using only 20 vowel pairs, so predicted rates have a granularity of 5%. Small changes in parameters may cause large jumps in predicted rate. Assmann and Summerfield (1989) and Assmann (1995) used techniques to transform distances (or neural net activations) to better-behaved “probability” estimates. We did not follow that course, because it introduces extra assumptions and param-

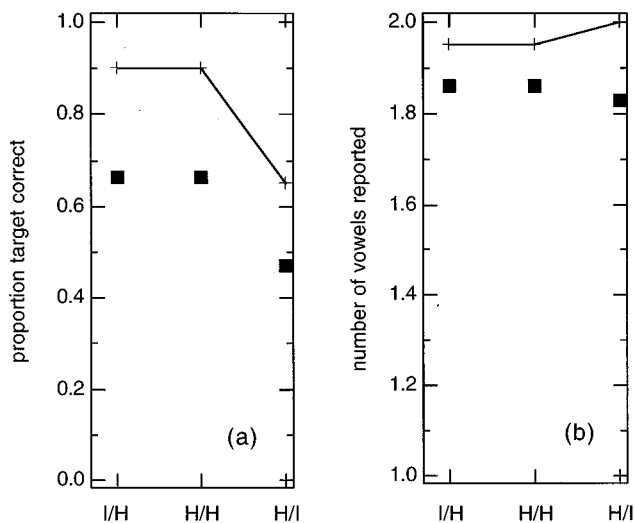


FIG. 10. (a) Symbols: Identification rate as a function of the harmonicity states of target and ground (notation of abscissa: target/ground). Line: Prediction of the model. (b) Symbols: Number of vowels reported. Line: Predictions of the model.

eters, and because our simpler method was sufficient for the points we wished to make.

C. Predicted effects of harmonicity

The model was applied to the stimuli of de Cheveigné *et al.* (1997b), using the same parameters as before. The symbols in Fig. 10 are experimental results, and the lines are predictions of the model. The model correctly predicts that target harmonicity has no effect on identification rate, and that identification is better when the competing vowel is harmonic rather than inharmonic [Fig. 10(a)]. Overall, predicted identification rates are too high, as was observed in Fig. 9(b) at $\Delta F_0 = 6\%$ and -20 dB. The number of responses predicted and observed is plotted in Fig. 10(b). The predictions are again a bit too high, and neither the predictions nor the experimental data vary much between conditions.

IV. DISCUSSION

Section I presented the neural cancellation filter. Section II showed that it could be exploited in several ways to segregate concurrent vowels that differ in fundamental frequency. Finally Sec. III elaborated one of those schemes into a model of concurrent vowel identification.

A. Neural cancellation

Removing certain spikes from the spike train (Fig. 1) suppresses evidence of a competing vowel, but we would be wrong in concluding that that evidence was “contained” in those particular spikes. Rather, the filter adjusts the interval statistics of the spike train so that they no longer reflect the competing vowel. Equation (1) is a crude account of the statistics of the spike selection process, but it produces patterns similar to simulations with recordings from the auditory nerve (de Cheveigné, 1993a). A linear version of Eq. (1) (time-domain comb filter) was used for speech interference rejection by de Cheveigné (1993b). Subtraction in Eq. (1) is

analogous to the cancellation step in the equalization and cancellation model of binaural interaction (Durlach, 1963; Culling and Summerfield, 1995a), which could be implemented by a spike selection process similar to the one illustrated in Fig. 1.

So far there is little physiological evidence for delay lines of the length required to cancel vowels with fundamental periods of the order of 10 ms (100 Hz), although similar delays would be needed to extract their pitch with an autocorrelation model (Licklider, 1959, 1962; Meddis and Hewitt, 1992; Cariani and Delgutte, 1996). Shorter delay lines have been shown to exist for binaural processing of interaural time delay cues (Yin and Chan, 1990).

B. Peripheral selectivity

Whereas most segregation models rely critically on peripheral filtering (including the schemes of Sec. II A, and C), the scheme of Sec. II B developed in Sec. III does not give it a major role. Cancellation is performed within channels, to be sure, but channel bandwidth is not a critical parameter, and channels separated by the cochlea are merged after the cancellation filter. Peripheral filtering can, however, be assigned two roles as follows. (a) It isolates channels in which the target-to-background ratio is more favorable than within the stimulus itself, which is useful if the linearity or dynamic range are limited. Thus peripheral selectivity might play a secondary role in the principle of the model but be essential for its implementation. (b) A signal that is split into band-limited channels can be reconstructed despite within-channel nonlinearities (Slaney *et al.*, 1994). Peripheral filtering may thus protect acoustic information from being degraded by the cascaded nonlinearities of peripheral transduction and cancellation filter.² In other words, the cochlea splits the acoustic signal redundantly into channels that differ by their phase and spectral content so that information that is degraded within each channel may survive in the ensemble. The neural cancellation filter should be seen as a segregation mechanism that operates together with, rather than instead of, peripheral filtering.

C. The place alternative

The scheme of Sec. II B was chosen as a basis for the identification model of Sec. III on criteria of ease of description and similarity, in its pattern-matching stage, with the model of Meddis and Hewitt (1992). The alternative schemes are worth considering. The place profile (Sec. II A) has the advantage of displaying information in a familiar spectrum-like form. A difficulty is that the profile for isolated vowels shows little vowel-specific evidence (Fig. 4), as observed by Sachs and Young (1979) in auditory-nerve fiber populations. Profiles might be better preserved in a population of low spontaneous rate fibers. One of our examples showed that the place profile can carry information, such as the formant F_1 of /u/ (Fig. 4) that is apparently not coded in the SACF (Fig. 7). The auditory system does not share the need for simplicity that constrains our models, so it might exploit both the SACF and the place profile, or even the entire ACF array

from which both derive. This is reminiscent of a many-to-many neural network proposed by Licklider (1959) for pitch perception.

D. The channel selection alternative

The channel selection scheme of Sec. II C eliminated channels that left little residual after cancellation, whereas Meddis and Hewitt (1992) eliminated channels that had a peak in the ACF at the competing vowel's period. The scheme of Sec. II C has the advantage that it depends upon a threshold that may be tuned to avoid rejecting all channels when a target vowel is weak. A version of Meddis and Hewitt's model based on this scheme might thus account for the improvement of identification with ΔF_0 at low target-to-background levels (de Cheveigné *et al.*, 1997a). That idea is not without problems, however, as a less stringent selection criterion retains channels that are more strongly affected by the competing vowel.

E. Recent experimental results

McKeown and Patterson (1996) found that one vowel of a concurrent pair was always identifiable whatever the ΔF_0 , even if the stimulus was very short (one fundamental period). Identification of the other vowel improved with ΔF_0 , but only if the stimulus was at least three of four periods in duration. The model of Sec. III is consistent with this behavior; the first vowel is determined independently from any segregation process, and the second after application of a cancellation filter tuned to an estimate of the period of the first. An F_0 estimate of sufficient precision for effective cancellation probably requires some time to develop. Indeed, Robinson and Patterson (1995) found that perception of the pitch of a vowel required several fundamental periods, whereas perception of the timbre needed only a single period.

Culling and Darwin (1993) synthesized concurrent vowels with ΔF_0 's interchanged between vowels in the region of the second formant (the F_2 of one vowel belonged to the same harmonic series as the F_1 of the other). A harmonic cancellation model tuned to either period should produce chimerical patterns comprising the F_1 of one vowel and the F_2 of the other. Nevertheless, Culling and Darwin found that identification improved with ΔF_0 , although less than for normal concurrent vowels. Several tentative explanations can be offered: (a) The chimerical segregated patterns may support identification better than the unsegregated pattern available at $\Delta F_0=0\%$ (for example, the pattern-matcher might make do with one formant while ignoring the other); (b) the ΔF_0 -induced inharmonicity of the stimulus may make it easier to interpret as containing two vowels, and thus indirectly improve identification; (c) harmonic cancellation may use different delays in different frequency regions, in a fashion analogous to Culling and Summerfield's (1995a) modification of the equalization and cancellation model of binaural interaction, and (d) Segregation at small ΔF_0 's may be due to beats in the F_1 region rather than F_0 -guided segregation. This last explanation was favored by Culling and Darwin (1993).

Culling and Darwin (1994) found further that identification improves with ΔF_0 (up to one semitone) for concurrent vowels represented by alternating partials of two harmonic series (successive partials of each vowel belonged to different series). This result is also difficult to explain on the basis of harmonic cancellation only. Darwin (1996) presented subjects with pairs of concurrent sentences synthesized each with a fixed F_0 . Identification increased greatly with ΔF_0 , but only beyond one semitone. At one semitone, identification was no better than at $\Delta F_0=0\%$. This result casts some doubt on the relevance of effects demonstrated in double-vowel experiments to the segregation of real speech.

V. SUMMARY AND CONCLUSION

The first two articles in this series confirmed earlier data indicating that the auditory system may segregate sounds by exploiting the harmonic structure of interference to suppress it. Alternative hypotheses such as enhancement of harmonic targets, beats, or pitch-period-asynchrony are unlikely to account for segregation of weak target vowels (-10 to -20 dB) at relatively large ΔF_0 's (6%), although we cannot exclude that they might play a role at higher target levels or smaller ΔF_0 's. Together, the results support a rather wide class of models capable of *harmonic cancellation*. Among those models, that of Meddis and Hewitt (1992) stands out because of its plausibility and predictive power. However, it could not explain one aspect of our data: the improvement with ΔF_0 of the identification of weak targets (-20 dB). For some vowel pairs, the model predicted no effect of ΔF_0 for the weaker vowel, yet for the same pairs we observed a clear effect. The problem appeared to lie in the model's all-or-nothing selection principle that assigns each peripheral channel as a whole to one or the other vowel.

The "neural cancellation filter" described in the present paper processes auditory-nerve fiber discharge patterns within channels to isolate correlates of concurrent vowels. Certain spikes are removed from the spike train, so that its statistics no longer reflect the competing vowel. The competing vowel must be harmonic for the filter to work, which is consistent with the principle of harmonic cancellation supported by our data. The filter allows individual channels to be searched for residual information reflecting a vowel too weak to dominate any one channel. It can in principle be exploited in a variety of ways, either to implement the channel selection stage of a revised version of Meddis and Hewitt's model, or else to produce segregated "temporal" or "place" representations of each vowel. Of the two physiological elements required by the filter, one (inhibitory gating) is plausible but the other (delay lines of up to 10 or 20 ms) has yet to be found in the auditory system.

The concurrent vowel identification model of Sec. III added to the neural cancellation filter a set of assumptions that allowed it to make quantitative predictions. The details of those assumptions are secondary; their purpose was to produce an incarnation of the model for which reasonable predictions could be formulated. The focus of this paper is the neural cancellation filter that underlies the model, and beyond it, the concept of harmonic cancellation that has

emerged from various experimental results as an important principle underlying F_0 -guided segregation.

ACKNOWLEDGMENTS

This paper is based on experimental data obtained at ATR Human Information Processing Research Laboratories under a research agreement between ATR and the Centre National de la Recherche Scientifique. John Culling of the MRC Institute of Hearing Research provided part of the software used to build the models. Thanks to him, Ray Meddis, Alan Palmer, and one anonymous reviewer for detailed comments on previous versions of the draft.

¹The illustration is valid for a population of high-spontaneous rate fibers. If low-spontaneous rate fibers are considered, place features may be preserved at high driving levels.

²On average, 75% of the samples of the signal within a channel are set to 0: Half-wave rectification sets all negative samples to zero, and of the remainder, all those such as $s(t) < s(t-T)$ are set to 0 in Eq. (1). Only 25% of the samples retain their original value.

- Assmann, P. F. (1995). "The role of formant transitions in the perception of concurrent vowels," *J. Acoust. Soc. Am.* **97**, 575–584.
- Assmann, P. F., and Summerfield, Q. (1989). "Modeling the perception of concurrent vowels: Vowels with the same fundamental frequency," *J. Acoust. Soc. Am.* **85**, 327–338.
- Assmann, P. F., and Summerfield, Q. (1990). "Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies," *J. Acoust. Soc. Am.* **88**, 680–697.
- Brox, J. P. L., and Nooteboom, S. G. (1982). "Intonation and the perceptual separation of simultaneous voices," *J. Phonetics* **10**, 23–36.
- Cariani, P. A., and Delgutte, B. (1996). "Neural correlates of the pitch of complex tones. I. Pitch and pitch salience," *J. Neurophysiol.* **76**, 1698–1716.
- Culling, J. F. (1996). "Signal processing software for teaching and research in psycholinguistics under UNIX and X-Windows," *Behav. Res. Methods Instrum. Comput.* **28**, 376–382.
- Culling, J. F., and Darwin, C. J. (1993). "Perceptual separation of simultaneous vowels: Within and across-formant grouping by F_0 ," *J. Acoust. Soc. Am.* **93**, 3454–3467.
- Culling, J. F., and Darwin, C. J. (1994). "Perceptual and computational separation of simultaneous vowels: Cues arising from low frequency beating," *J. Acoust. Soc. Am.* **95**, 1559–1569.
- Culling, J. F., and Summerfield, Q. (1995a). "Perceptual segregation of concurrent speech sounds: absence of across-frequency grouping by common interaural delay," *J. Acoust. Soc. Am.* **98**, 785–797.
- Culling, J. F., and Summerfield, Q. (1995b). "The role of frequency modulation in the perceptual segregation of concurrent vowels," *J. Acoust. Soc. Am.* **98**, 837–846.
- Darwin, C. J. (1996). (Personal communication.)
- de Cheveigné, A. (1993a). "Separation of concurrent harmonic sounds: Fundamental frequency estimation and a time-domain cancellation model of auditory processing," *J. Acoust. Soc. Am.* **93**, 3271–3290.
- de Cheveigné, A. (1993b). "Time-domain comb filtering for speech separation" (TR-H-016), ATR Human Information Processing Laboratories Tech. Report (unpublished).
- de Cheveigné, A., McAdams, S., Laroche, J., and Rosenberg, M. (1995). "Identification of concurrent harmonic and inharmonic vowels: A test of the theory of harmonic cancellation and enhancement," *J. Acoust. Soc. Am.* **97**, 3736–3748.
- de Cheveigné, A., Kawahara, H., Tsuzaki, M., and Aikawa, K. (1997a). "Concurrent vowel segregation. I. Effects of relative amplitude and F_0 difference," *J. Acoust. Soc. Am.* **101**, 2839–2847.
- de Cheveigné, A., McAdams, S., and Marin, C. (1997b). "Concurrent vowel segregation. II. Effects of phase, harmonicity and task," *J. Acoust. Soc. Am.* **101**, 2848–2856.
- Durlach, N. I. (1963). "Equalization and cancellation theory of binaural masking-level differences," *J. Acoust. Soc. Am.* **35**, 1206–1218.
- Holdsworth, J., Nimmo-Smith, I., Patterson, R. D., and Rice, P. (1988). "Implementing a GammaTone filter bank" (SVOS final report, annex C), MRC Applied Psychology Unit Tech. Rep. (unpublished).
- Lea, A. (1992). "Auditory models of vowel perception," Ph.D. thesis, Nottingham (unpublished).
- Licklider, J. C. R. (1959). "Three auditory theories," in *Psychology, a Study of a Science*, edited by S. Koch (McGraw-Hill, New York), Vol. I, pp. 41–144.
- Licklider, J. C. R. (1962). "Periodicity pitch and related auditory process models," *Int. Audiol.* **1**, 11–36.
- McKeown, J. D., and Patterson, R. D. (1996). "The time course of auditory segregation: Concurrent vowels that vary in duration," *J. Acoust. Soc. Am.* **98**, 1866–1877.
- Meddis, R. (1986). "Simulation of mechanical to neural transduction in the auditory receptor," *J. Acoust. Soc. Am.* **79**, 702–711.
- Meddis, R. (1988). "Simulation of auditory-neural transduction: further studies," *J. Acoust. Soc. Am.* **83**, 1056–1063.
- 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.
- Meddis, R., and Hewitt, M. J. (1992). "Modeling the identification of concurrent vowels with different fundamental frequencies," *J. Acoust. Soc. Am.* **91**, 233–245.
- Moore, B. C. J., and Glasberg, B. R. (1983). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," *J. Acoust. Soc. Am.* **74**, 750–753.
- Palmer, A. R. (1990). "The representation of the spectra and fundamental frequencies of steady-state single- and double-vowel sounds in the temporal discharge patterns of guinea pig cochlear-nerve fibers," *J. Acoust. Soc. Am.* **88**, 1412–1426.
- Palmer, A. R. (1992). "Segregation of the responses to paired vowels in the auditory nerve of the guinea-pig using autocorrelation," in *Audition Speech and Language*, edited by M. E. H. Schouten (Mouton-de Gruyter, Berlin), pp. 115–124.
- Parsons, T. W. (1976). "Separation of speech from interfering speech by means of harmonic selection," *J. Acoust. Soc. Am.* **60**, 911–918.
- Patterson, R. D., Holdsworth, J., Nimmo-Smith, I., and Rice, P. (1988). "An efficient auditory filterbank based on the gammatone function" (SVOS final report, annex B), MRC Applied Psychology Unit Tech. Rep. (unpublished).
- Robinson, K., and Patterson, R. D. (1995). "The stimulus duration required to identify vowels, their octave, and their pitch chroma," *J. Acoust. Soc. Am.* **98**, 1858–1865.
- Sachs, M. B., and Young, E. D. (1979). "Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate," *J. Acoust. Soc. Am.* **66**, 470–479.
- Slaney, M., Naar, D., and Lyon, R. F. (1994). "Auditory model inversion for sound separation," *Proc. ICASSP* **5**, 213–216.
- Summerfield, Q. (1992). "Roles of harmonicity and coherent frequency modulation in auditory grouping," in *The Auditory Processing of Speech: from Sounds to Words*, edited by M. E. H. Schouten (Mouton-de Gruyter, Berlin), pp. 157–166.
- Summerfield, Q., and Culling, J. F. (1992a). "Auditory segregation of competing voices: absence of effects of FM or AM coherence," *Phil. Trans. R. Soc. London Ser. B* **336**, 357–366.
- Summerfield, Q., and Culling, J. F. (1992b). "Periodicity of maskers not targets determines ease of perceptual segregation using differences in fundamental frequency," *J. Acoust. Soc. Am.* **92**, 2317(A).
- Weintraub, M. (1985). "A theory and computational model of auditory monaural sound separation," Ph.D. thesis, Stanford (unpublished).
- Yin, T. C. T., and Chan, J. C. K. (1990). "Interaural time sensitivity in medial superior olive of cat," *J. Neurophysiol.* **64**, 465–488.

Intraspeech spread of masking in normal-hearing and hearing-impaired listeners

Van Summers^{a)} and Marjorie R. Leek

Army Audiology & Speech Center, Walter Reed Army Medical Center, Washington, DC 20307-5001

(Received 30 June 1996; revised 5 November 1996; accepted 27 November 1996)

Hearing-impaired and normal-hearing listeners labeled synthetic consonant-vowel stimuli (/ba/, /da/, /ga/, /be/, /de/, /ge/) presented at moderate and high signal levels. First formant ($F1$) regions were synthesized at normal and at attenuated levels to test whether $F1$ attenuation might reduce upward spread of masking, making information contained in higher formant regions more available. Performance was tested in quiet and in broadband noise sufficient to mask initial release bursts. Although complete removal of $F1$ consistently reduced performance, $F1$ attenuation of up to 18 dB led to increased labeling accuracy, particularly in the /a/ vowel context. Benefit associated with $F1$ attenuation was more consistently seen for hearing-impaired than for normal-hearing listeners and, in particular, for listeners with steep increases in audiometric thresholds between the first and second formant regions of the test stimuli. The availability of initial bursts as a source of place cues during testing in quiet did not reduce the benefit associated with $F1$ attenuation. [S0001-4966(97)06204-8]

PACS numbers: 43.71.Ky, 43.71.Es [WS]

INTRODUCTION

Upward spread of masking (USM) refers to the process by which the audibility of an acoustic stimulus is reduced in the presence of acoustic energy at lower frequencies. Listeners with sensorineural hearing impairments may be especially susceptible to USM for several reasons. First, frequency resolution in hearing-impaired (HI) listeners is typically poorer than normal, resulting in increased effects of low-frequency maskers on high-frequency signals (Glasberg and Moore, 1986, 1989; Tyler, 1986). Further, both HI and normal-hearing (NH) listeners experience greater USM at high presentation levels (Leek and Summers, 1993; Zwicker and Jaroszewski, 1982; Egan and Hake, 1950). These presentation levels may be necessary for HI listeners in order to assure signal audibility. As a consequence, amplification provided by a hearing aid to accommodate a listener's sensitivity loss may increase the likelihood of USM. A number of commercially available hearing aids employ noise-reduction techniques aimed at reducing USM. These aids reduce low-frequency gain in response to increases in input level. Under controlled experimental conditions, this processing strategy can provide significant release from USM and result in improved speech recognition (Rankovic *et al.*, 1992; Fabry *et al.*, 1993).

There are several possible sound sources that may contribute to USM of cues contained in the primary speech signal. Most obvious is low-frequency energy from competing environmental noise and from the speech of other talkers in the listening environment. Reverberation of low-frequency components of an acoustic signal may also be a source of USM. This may include reverberation of energy present in the primary signal itself or from other sound sources. Finally,

low-frequency portions of the primary signal may mask higher frequency regions. USM from competing sound sources or from signal reverberation is specific to a given listening environment and may be reduced by changes in that environment. Intraspeech USM in the absence of reverberation, on the other hand, is a more general phenomenon that HI listeners must deal with continuously. This form of USM, which may influence performance regardless of the noise or reverberant characteristics of the listening environment, is the focus of the present study. Specifically, the masking of second-formant ($F2$) frequency transitions by energy in the first-formant ($F1$) region will be examined.

Figure 1 shows acoustic energy as a function of frequency averaged over a lengthy sample of normal adult speech produced at a typical conversational level of 63 dB SPL (Pavlovic, 1982). Acoustic energy is greatest in the region below 800 Hz and steadily decreases at higher frequencies. The top of the figure indicates frequency ranges associated with $F1$ and $F2$ based on adult productions of ten English vowels (Peterson and Barney, 1952). In general, acoustic energy associated with $F1$ is present below 800 Hz, whereas energy associated with $F2$ is seen at higher frequencies. As the figure suggests, greater acoustic energy is associated with $F1$ than with $F2$ or higher formants. The greater intensity of $F1$ relative to $F2$ and the fact that low-frequency components of a signal tend to be effective maskers of higher-frequency components, indicate that $F1$ energy could, in some instances, mask $F2$.

A series of studies by Danaher, Pickett, and colleagues at Gallaudet College in the mid-1970s tested for masking of $F2$ by $F1$ in NH and HI listeners (Danaher *et al.*, 1973; Danaher and Pickett, 1975; Pickett and Danaher, 1975). The difference in initial onset frequency required to discriminate an $F2$ transition from a steady-state $F2$ was determined for listeners with moderate-to-severe sensorineural impairments. The effect of $F1$ on $F2$ discrimination was examined by

^{a)}Address correspondence to: Van Summers, Ph.D., Army Audiology and Speech Center, Walter Reed Army Medical Center, Washington, DC 20307-5001.

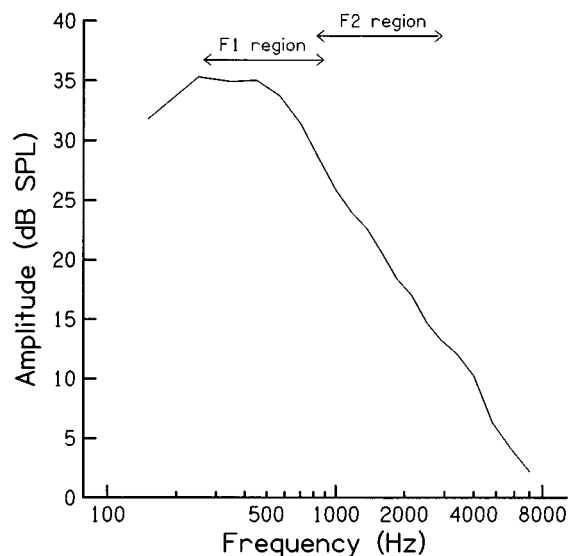


FIG. 1. Average long-term spectrum of speech produced with normal vocal effort (Pavlovic, 1982). Overall level=63 dB SPL.

testing with $F2$ presented in isolation and in combination with $F1$. For most HI subjects, larger onset frequency differences were required in the presence of $F1$ than in the $F2$ -alone condition. The authors suggested that USM from $F1$ to $F2$ caused the poorer performance in the $F1+F2$ condition. This interpretation was supported by their additional findings that the $F1$ effect was reduced by presenting $F1$ and $F2$ to separate ears of a subject or by reducing $F1$ intensity relative to $F2$ (Danaher and Pickett, 1975; Pickett and Danaher, 1975).

The Gallaudet group also examined the effect of presentation level on USM from $F1$ to $F2$. When tested at comfortable listening levels, NH listeners did not show the $F1$ on $F2$ masking effects reported for HI listeners. However, when tested at equal, high-presentation levels (95 and 105 dB SPL), both groups showed susceptibility to $F1$ -on- $F2$ masking (Danaher *et al.*, 1973). These findings suggest that high presentation levels (which may be necessary to assure audibility for some HI listeners) increase the likelihood of USM from $F1$ to $F2$, even in listeners with normal hearing.

The results reported by the Gallaudet group pertained to listeners with moderate to severe sensorineural losses. Using similar methods, Van Tasell (1980) examined $F2$ transition discrimination for listeners with mild sensorineural losses. Stimuli were modeled after those used by Danaher *et al.* (1973) and were presented at 85 dB SPL. Van Tasell observed significant variability in susceptibility to USM across four HI listeners. However, all HI subjects required larger $F2$ onset frequency differences in the $F1+F2$ condition than in the $F2$ -alone condition in order to reliably discriminate a rising $F2$ transition from an 1100-Hz steady state. The results suggested that $F1$ masking of $F2$ information may be present for less severely impaired listeners and at lower presentation levels than those examined in the original studies at Gallaudet.

These $F2$ discrimination results led to studies examining the influence of $F1$ on *identification* of two-formant

consonant-vowel (CV) stimuli where initial $F2$ transitions contain important place-of-articulation cues. Hannley and Dorman (1983) examined recognition of a synthetic /ba/-/da/-/ga/ continuum in which initial $F2$ transitions provided the only place information. They tested whether attenuation of $F1$ relative to $F2$ improved identification performance. Presentation levels of 80 and 100 dB SPL were used, with $F1$ amplitude set to 0, -6, -12, and -18 dB relative to $F2$ amplitude. At 80 dB SPL, HI subjects showed poor labeling performance when $F1$ and $F2$ amplitudes were equal. Performance improved with $F1$ attenuation and was within the normal range for the -12 and -18 dB conditions. At 100 dB SPL, both HI and NH listeners benefitted from $F1$ attenuation. HI listeners showed their most reliable labeling with $F1$ amplitude set -12 or -18 dB below $F2$. For NH listeners, labeling was most consistent with less $F1$ attenuation (-6 dB condition). These results suggest that $F1$ attenuation may improve labeling performance, particularly for HI subjects. In addition, $F1$ attenuation provides more benefit at high presentation levels than at moderate levels.

The Hannley and Dorman (1983) results were obtained with synthetic stimuli that did not contain initial release bursts as found in natural speech. Thus their results do not address the importance of $F1$ -on- $F2$ masking when burst cues are also available as a source of place information. This is an important consideration since these additional place cues may be available to listeners in many instances, such as quiet listening environments. Dorman *et al.* (1985) examined this issue using a synthetic /be/-/de/-/ge/ continuum containing initial bursts. HI subjects labeled two versions of the continuum, one containing $F1-F5$ at appropriate relative amplitudes and one containing only $F2-F5$. Subjects were tested at a range of presentation levels. There was little evidence that the complete removal of $F1$ benefitted labeling performance at any presentation level. Dorman *et al.* (1985) suggested that the inclusion of release bursts in their stimuli was the likely explanation for the absence of $F1$ -on- $F2$ masking effects on labeling performance. That is, $F1$ masking of $F2$ place cues may be of less importance when release bursts are present as an additional source of place information.

Several unanswered questions remain regarding the role of $F1$ -on- $F2$ masking in the speech-understanding difficulties of HI listeners. The discrimination data of the Gallaudet group (Danaher *et al.*, 1973; Pickett and Danaher, 1975) and Van Tasell (1980) along with the identification data of Hannley and Dorman (1983) suggest that USM of $F2$ by $F1$ may frequently occur for HI listeners. Conversely, using stimuli that included initial bursts, Dorman *et al.* (1985) failed to demonstrate any consistent negative effect of $F1$ -on- $F2$ masking for their HI listeners. Their results suggest that USM from $F1$ to $F2$ may not have an important influence on place identification in quiet listening conditions in which burst cues may supplement formant transition cues. However, given that consonantal bursts are generally produced at lower intensities than vowel formants (Turner *et al.*, 1992), release bursts may be masked in certain noisy environments while formant transitions cues remain audible. Therefore, $F1$ -on- $F2$ masking may have an important influence on

speech understanding under certain noisy conditions. In addition to availability of burst cues, the unnaturalness of the stimulus continuum used by Dorman *et al.* (1985), in which *F1* was entirely removed, may have limited the benefit of reduced *F1-on-F2* masking in their study. Finally, vowel context may also have an important impact on the amount of *F1-on-F2* masking (Danaher *et al.*, 1973). The Dorman *et al.* findings were based on stimuli containing the /e/ vowel which has a fairly large frequency difference between *F1* and *F2*, relative to some other vowels. Therefore, the lack of any benefit from *F1* attenuation in the Dorman *et al.* study could also be due to choice of vowel.

The present study was a replication and extension of Dorman *et al.* (1985). We tested for *F1* masking of *F2* place cues in two vocalic contexts, using a range of *F1* to *F2* amplitude ratios. Testing was carried out at two signal presentation levels both in quiet and in noise sufficient to effectively mask burst cues. Successful replication of the Dorman *et al.* (1985) findings under quiet listening conditions would support their suggestion that place cues present within initial bursts minimize the significance of possible *F1* masking of *F2*. However, even given this result, susceptibility to *F1-on-F2* masking may still have an important influence on performance in noisy environments in which burst cues are less available. Finally, amount of *F1-on-F2* masking may vary with signal level. As noted above, both NH and HI listeners experience greater USM as level increases. Recent data from Leek and Summers (1993) suggest that this tendency might be especially true of HI listeners.

I. METHOD

A. Subjects

Seven normal-hearing and seven hearing-impaired listeners were tested. Normal-hearing listeners had thresholds no greater than 15 dB HL (*re*: ANSI, 1989) at audiometric test frequencies from 250 to 4000 Hz. They ranged in age from 19 to 74 years with a mean age of 46.4 years. Hearing-impaired listeners had sensorineural losses representing a range of severities and audiometric configurations (see Table

TABLE I. Age and test ear audiometric thresholds for HI subjects (dB HL *re*: ANSI, 1989).

Subj	Age	Frequency (Hz)							
		250	500	1000	1500	2000	3000	4000	6000
JM	53	15	20	15	55	65	80	75	70
JS	49	10	15	70	75	65	65	65	75
JL	55	10	15	15	15	35	65	75	75
RL	70	20	30	40	50	65	70	75	65
NW	39	25	35	45	45	40	45	45	70
PN	80	20	25	35	40	40	50	70	60
SK	66	30	30	30	30	35	55	60	80

I). Results from bone conduction and/or immittance audiometry were consistent with cochlear site of lesion for these subjects. Hearing-impaired listeners ranged from 39 to 80 years of age with a mean of 58.8.

B. Instrumentation

All stimuli were played monaurally to subjects over TDH-49 earphones. Responses were collected using a touch screen terminal. The terminal display provided visual feedback when appropriate.

Speech stimuli were synthesized in software and converted to analog form by a D/A converter (TDT DD1) at a sample rate of 20 kHz/s. Waveforms were low-pass filtered at 8.5 kHz (TDT FLT3, >70 dB/octave rolloff above 8.5 kHz), attenuated as required (TDT PA4), mixed with noise when appropriate, and led via a headphone buffer to the earphones in a sound-isolated booth. Gaussian noise was produced by a waveform generator (TDT WG1), and was low-pass filtered at 8.5 kHz (TDT FLT3). Noise was gated on and off using a cosine switch (TDT SW3, 10-ms cosine-squared rise and fall), attenuated, and mixed with the speech signal. Levels of all noises and signals were measured using a sound-level meter and spectral characteristics were verified with a signal analyzer.

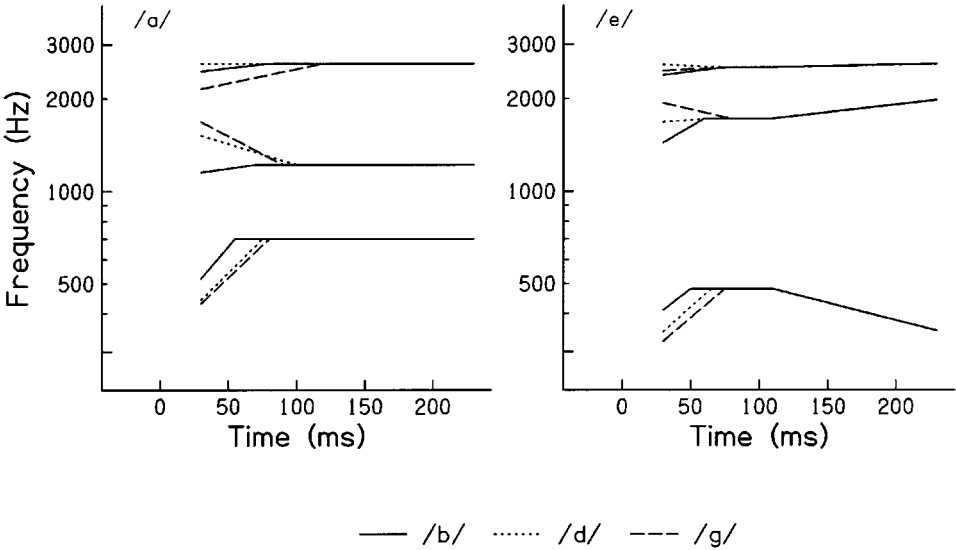


FIG. 2. Formant frequency trajectories for CV stimuli, vowel=/a/ in left-hand panel, /e/ in right-hand panel. Initial release bursts which preceded vowel onset are not represented.

C. Base stimuli

Six synthetic CV stimuli, with /b/, /d/, and /g/ as initial consonants and with /a/ (as in “dot”) and /e/ (as in “day”) as following vowels, were generated using the parallel configuration of a cascade/parallel software speech synthesizer (Klatt, 1980).¹ Base stimuli contained initial release bursts and formant transitions appropriate to the consonantal place of articulation. Stimulus vowels were 200 ms in duration with F_0 held constant at 120 Hz. Schematic representations of the formant frequency values for each base stimulus appear in Fig. 2. For a given initial consonant, burst characteristics were fairly similar across vowel contexts. For /ba/ and /be/ stimuli, initial bursts were 5 ms in duration with diffuse, flat amplitude spectra. For /da/ and /de/, bursts were 10 ms in duration with rising amplitude spectra with a peak around 3600 Hz. For /ga/ and /ge/, bursts were 20 ms in duration with compact spectra with energy concentrated between F_2 and F_3 of the following vowel (approximately 2200 Hz for /a/ and 2800 Hz for /e/).

These base stimuli were the starting point for stimulus generation for three experimental tasks. The first task was a screening test to verify that each subject could detect the initial consonantal burst in each of the base stimuli at the lowest presentation level to be used in the experiment. The second task determined the level of noise required to mask the initial bursts for each individual subject. The third task measured the percent correct identification of CV stimuli under various conditions of F_1 attenuation and noise masking. For each of these tasks, modifications to the base stimuli were carried out as described in the sections below.

D. Initial screening task: Burst detection in quiet

Initial screening was carried out to verify that each subject could detect initial consonantal bursts in quiet at the presentation levels to be tested.

1. Stimuli

The base stimuli were equated in terms of long-term rms level and edited to extract the initial consonantal burst region of each stimulus. Twenty repetitions of each burst stimulus were randomized and tested for detection in quiet in a single block of 120 trials (6 stimuli \times 20 repetitions/stimulus).

2. Procedure

Burst detection was tested using a cued two-alternative forced-choice paradigm. On each trial, lights on a touch screen panel indicated three successive presentation intervals. A single consonantal burst was presented in the first interval and again in either the second or third interval, with the remaining interval silent. Subjects indicated which of the final two intervals contained a stimulus matching the one presented initially by touching the appropriate box on the touch screen. Correct answer feedback was provided following each trial. Performance was tested at a nominal signal level of 80 dB SPL, corresponding to the burst levels observed when overall level for a given base stimulus was 80 dB SPL. Burst amplitudes differed by less than 1 dB across

stimuli and were generally about 11 dB lower than overall base stimulus levels. Thus the nominal 80 dB SPL level corresponded to an actual presentation level of about 69 dB SPL.

E. Burst detection in noise task: Signal-to-noise ratio threshold

This task was carried out to determine the levels of broadband noise necessary to mask the initial consonantal burst portion of each stimulus for each subject. The levels of noise determined here were used to eliminate burst cues in a subsequent CV labeling task, forcing listeners to rely on formant transition cues.

1. Stimuli

The individual consonantal bursts were concatenated to create a stimulus containing all six bursts appended end-to-end with 50-ms regions of silence between bursts. This “bursts-only” stimulus was approximately 400 ms in duration and was the target stimulus for the detection-in-noise task. Masking was provided by low-pass filtered Gaussian noise with a bandwidth of 8.5 kHz. Noise was gated on 50 ms prior to target stimulus onset and continued for 50 ms following target offset.

2. Procedure

Four separate blocks of trials were presented. Two blocks tested performance at each of two nominal signal levels (80 and 100 dB SPL). Within a block of trials, signal level was held constant while noise level was varied using a simple up-down adaptive procedure that estimates the 50% correct point on the psychometric function (Levitt, 1971). On each trial, subjects heard three presentation intervals, with each containing broadband masking noise. The “bursts-only” target stimulus was presented within the noise during the first trial interval. The target stimulus was again presented in noise in either the second or third presentation interval with each interval having a 50% probability of selection on a given trial. The remaining (second or third) interval contained the masking noise alone. Subjects indicated whether interval 2 or interval 3 contained the burst stimulus by touching the appropriate box on a touch screen panel. At the start of a trial block, noise level was set to 50 dB SPL. Noise level was increased by 4 dB after each correct response and decreased by 4 dB after each incorrect response until four reversals in the track had occurred. The 4-dB step-size was then reduced to 1 dB and the run continued for an additional ten reversals. The mean noise level at these final 10 reversals was taken as an estimate of the threshold (lowest) noise level capable of masking the set of consonantal bursts to chance levels of detection. The mean of the threshold estimates from the two blocks of trials carried out at a given presentation level was taken as the threshold noise level for that subject at that signal presentation level. The threshold signal-to-noise ratio (S/N) levels reported are defined as the dB difference between the nominal signal presentation level and the threshold noise level.

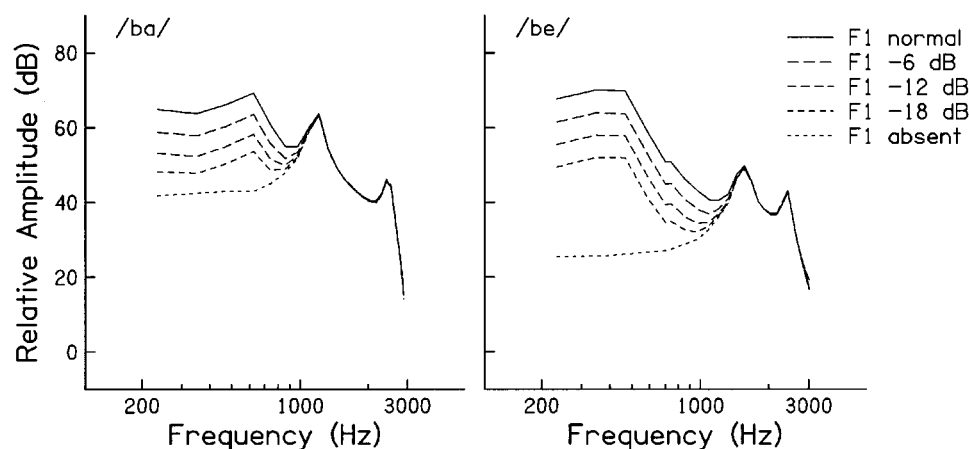


FIG. 3. Short-term spectra at voicing onset of /b/-vowel stimuli for each level of $F1$ attenuation. Values are based on the initial 25.6 ms of each vowel (512-point FFT).

F. Consonant identification task

The main experimental task involved identification of CV stimuli that differed in the level of $F1$ relative to higher formant regions. Identification data were collected in quiet and in noise sufficient to mask initial burst cues (as determined in the burst detection task). The original base stimuli were tested along with four additional sets of stimuli in which the level of $F1$ was attenuated to reduce possible $F1$ -on- $F2$ masking. Stimuli were presented at overall levels corresponding to 80 and 100 dB SPL for base stimuli ($F1$ at full strength). These overall presentation levels were reduced for all other stimuli as a consequence of $F1$ attenuation without any adjustment of amplitudes at higher frequencies. The 80- and 100-dB SPL presentation levels referred to below are therefore nominal levels for all but base stimuli.

1. Stimuli

Relative formant amplitudes for base stimuli were determined based on the rules for parallel formant synthesis described in Klatt (1980). This set of stimuli was tested along with four additional sets in which $F1$ levels were attenuated. In three of these stimulus sets, $F1$ level was decreased by 6, 12, and 18 dB, respectively, relative to levels present in the base stimulus set. In the final set, the $F1$ spectral prominence was entirely removed. A total of 30 stimuli were tested for identification (five stimulus sets \times six stimuli per set). Vowel onset spectra for /ba/ and /be/ tokens from each of the $F1$ -attenuation conditions are shown in Fig. 3. Values are based on the initial 25.6 ms of each vowel (512-point FFT).

2. Procedure

This task involved eight separate blocks of trials, with two blocks testing each possible combination of presentation level (80 and 100 dB SPL) and listening condition (in quiet and in noise). Within each block, 10 repetitions of the 30 stimuli were randomized to produce a block of 300 trials. On each trial a single stimulus was presented in a six-alternative forced-choice labeling task. Subjects responded by touching

the appropriately labeled box on a touch screen panel. Response alternatives were: “ba,” “da,” “ga,” “bay,” “day,” and “gay.”

One block of trials representing each of the four experimental conditions was presented before any were tested a second time. In addition, within the first set of four trial blocks, the “in-quiet” conditions were tested prior to the “in-noise” conditions. Block order was randomized for the second presentation of each block (trial blocks 5–8).

For blocks testing identification in noise, the S/N was set to the subject’s burst-detection threshold level as previously determined. This was intended to preclude the use of burst cues in identification since, at this S/N, consonantal bursts were not detected at greater than chance levels.

II. RESULTS

A. Initial screening—burst detection in quiet

Accurate detection of consonantal bursts in quiet was required in order to continue in the study. Burst detection accuracy below 80% for any of the six initial bursts was taken as evidence that a subject could not reliably detect a given burst under quiet listening conditions. All subjects detected each burst stimulus with at least 90% accuracy (at least 18 of 20 correct detections).

B. Burst detection in noise: S/N threshold

This task provided estimates of the broadband noise levels necessary for effective masking of initial bursts when the full CV stimuli (base stimuli) were presented at 80 and 100 dB SPL. Threshold S/N levels are broken down by group and signal presentation level in Table II. Threshold values were similar across groups and level conditions with all values listed within 1 dB of the overall mean of 12.29 dB. A two-way analysis of variance (ANOVA) confirmed the lack of any significant effect of group membership or presentation level on threshold S/N values ($p > 0.5$ for the group and level main effects and for the group by level interaction).

TABLE II. Mean S/N thresholds for masking of initial bursts (standard deviations in parentheses). S/N is defined as nominal burst level (80 or 100 dB SPL) relative to broadband noise level at burst threshold.

	80 dB	100 dB	Mean
Normal hearing	11.49 (4.32)	12.51 (2.86)	12.00 (3.56)
Hearing impaired	12.70 (2.85)	12.44 (2.65)	12.57 (2.65)
Mean	12.09 (3.57)	12.48 (2.65)	12.29 (3.09)

C. CV identification

The identification data were analyzed in terms of percent-correct CV identification (%CV), i.e., both consonant and vowel had to be accurately identified for responses to be scored as correct. An alternative means of scoring would be in terms of consonant correct (%C) regardless of vowel response. Since few vowel errors occurred in conditions where $F1$ was present at either normal or attenuated levels (mean vowel correct score across groups=98.3% for these conditions), either method of scoring led to the same pattern of results for these stimuli. Vowel errors did occur when $F1$ was completely removed (mean vowel correct=79.5% for stimuli containing /e/ with $F1$ absent). Percent-correct scores in $F1$ -absent conditions would, as a result, improve using %C rather than %CV scoring. However, the overall pattern of results we report below is true for either %CV or %C scoring. The %CV measure is reported to assure that any benefits to consonant identification provided by $F1$ attenuation are not offset by possible negative effects of the change in $F1$ on cues to the following vowel.

Mean percent-correct CV identification scores are plotted in Fig. 4 as a function of $F1$ attenuation. The panels of the figure separate the data according to stimulus vowel and noise condition. Within each panel, the results are broken down according to subject group and presentation level. Analysis of variance was carried out on these data using vowel, noise condition, presentation level, and $F1$ attenuation as within-subjects factors and using group membership as a between-subjects factor. Follow-up simple main-effects tests were also conducted as appropriate. Each of these analyses involve an assumption of homogeneity of variance across experimental conditions. As the error bars in Fig. 4 indicate, this equality of variance was not strictly present across all conditions. In particular, little variability was seen for the NH listeners tested in quiet relative to other conditions. To increase homogeneity of variance, arcsine transformations (Cohen and Cohen, 1983) of the percent-correct CV scores were used as the dependent variable in the statistical analyses. The arcsine transform was successful in reducing, but not entirely eliminating, heterogeneity of variance in the present data. As a further precaution against spuriously significant results, the conservative Greenhouse–Geisser (1959) adjustment to the degrees of freedom of the F statistic was used in evaluating statistical significance.

The focus of this work is on the possible benefit of $F1$ attenuation for initial consonant identification and on the stimulus conditions where this benefit occurs. However, prior to examining how $F1$ attenuation influenced identifica-

tion accuracy, we will briefly describe how the other factors influenced performance in terms of main effects.

First, there was a clear main effect of hearing status on identification accuracy [$F(1,12)=17.85$, $p=0.0012$]. In every instance, performance by NH listeners for a given stimulus condition (solid symbols) exceeded performance by HI listeners in the same condition (open symbols). Across stimulus conditions, NH listeners showed a mean accuracy of 78.8% compared to a mean of 57.4% for HI listeners.

Noise condition also had a clear main effect on performance [$F(1,12)=230.44$, $p<0.0001$]. Across stimuli and subject groups, performance was always better in quiet than in broadband noise. Collapsing across other factors (including subject group) mean accuracy in quiet was 81.7% compared to 54.6% in noise.

Vowel context also had a significant main effect [$F(1,12)=67.86$, $p<0.0001$]. Holding other stimulus and subject variables constant, performance was nearly always better in the /a/ vowel context than in the /e/ context. Collapsing across subject groups, mean accuracy in the /a/ context was 74.6% while accuracy for stimuli containing /e/ was 61.7%.

Presentation level did not have a significant main effect on identification performance [$F(1,12)=0.89$, $p=0.37$]. Although performance was generally better at the higher presentation levels for stimuli tested in quiet, this benefit was not observed for testing in noise.

The influence of $F1$ attenuation on labeling accuracy was of primary interest. Specifically, we were interested in any benefit to labeling associated with $F1$ attenuation. However, within each panel of Fig. 4, the most obvious and consistent effect of $F1$ attenuation is a decrease in labeling accuracy in “ $F1$ absent” conditions (mean performance =58.1%) relative to other $F1$ conditions (mean=70.6%). This decrease in performance with $F1$ removal contributed to a significant main effect for the $F1$ attenuation factor, in the opposite direction to the hypothesized $F1$ -attenuation benefit [$F(2.39, 28.67)=47.36$, $p<0.0001$].

Both consonants and vowels were more poorly identified in $F1$ -absent conditions than in conditions with $F1$ present at normal or attenuated levels. Clearly, complete attenuation of $F1$ did not improve CV labeling for either group in any stimulus condition. However, more limited $F1$ attenuation did improve performance in some instances. Focusing on HI listeners’ data for stimuli containing /a/ (left-hand panels of Fig. 4), performance was nearly always poorer with $F1$ at full strength than in any of the conditions with $F1$ attenuated but still present. The most common pattern was a gradual improvement in labeling as $F1$ attenuation increased from 0 up to 18 dB with a 6%–7% increase in accuracy for the 18-dB condition relative to the 0-dB condition. This pattern can also be seen in some of the NH listeners’ data and in some of the results for /e/ stimuli: both groups show this pattern for /e/ stimuli presented in quiet at 100 dB (Fig. 4, upper right-hand panel).

$F1$ attenuation improved labeling performance under certain conditions but complete $F1$ removal always lowered performance. In order to focus on the four levels of the $F1$ factor that appeared to show some benefit of $F1$ attenuation,

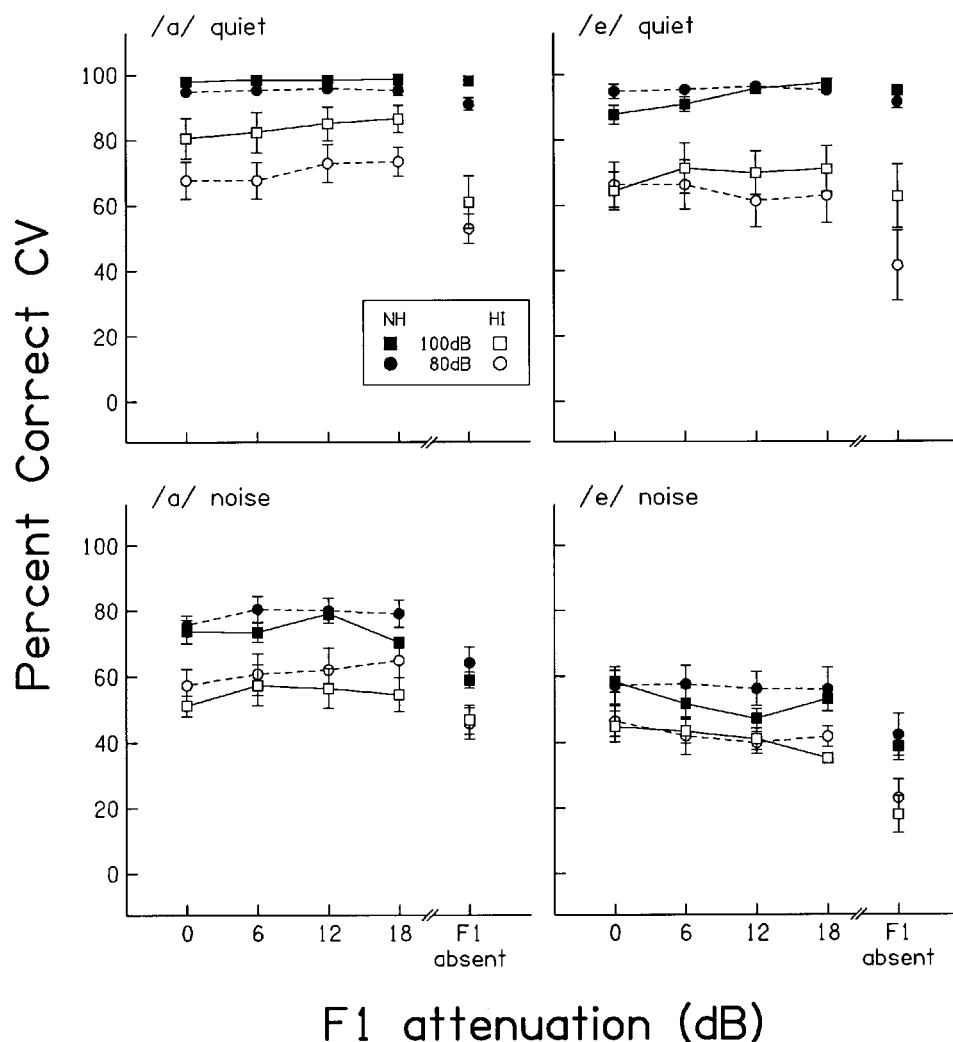


FIG. 4. Percent-correct CV identification scores. Upper panels show performance in quiet, lower panels in noise. Left-hand panels show performance with vowel=/a/, right-hand panels show vowel=/e/.

the original ANOVA was repeated, excluding the $F1$ -absent conditions. That is, the $F1$ attenuation factor was treated as having only four levels (0-, 6-, 12-, and 18-dB $F1$ attenuation) in the new ANOVA (henceforth, ANOVA2). With the $F1$ -absent conditions no longer included in the analysis, $F1$ attenuation did not have a significant main effect on performance [$F(2.47, 29.63)=1.80, p=0.1762$]. However, three of the interaction terms involving the $F1$ attenuation factor were statistically significant and were explored further.

First, a significant vowel by group by $F1$ -attenuation interaction [$F(2.67, 32.03)=3.14, p=0.0437$] suggested differences in the benefit associated with $F1$ attenuation for various combinations of group and stimulus vowel. Simple effects tests, holding group and stimulus vowel constant, showed a significant effect of $F1$ attenuation on HI listeners' responses to consonant-/a/ (C/a/) stimuli [$F(2.73, 32.76)=4.17, p=0.0154$]. This result is consistent with the observation made earlier that a pattern of benefit associated with $F1$ attenuation can be seen in Fig. 4 for HI listeners' responses to C/a/ stimuli. The $F1$ attenuation factor was not significant for HI listeners' responses to consonant-/e/ (C/e/) stimuli or for NH listeners' responses to stimuli containing either vowel.

A significant noise by level by $F1$ -attenuation interaction was also present in ANOVA2 [$F(2.64, 31.62)=4.94, p=0.0082$]. Simple effects tests, holding noise and level conditions constant, showed significant benefit associated with $F1$ attenuation only for stimuli presented in quiet at the high (100 dB) presentation level [$F(1.64, 19.66)=6.38, p=0.0102$]. The data in Fig. 4 reflect this result: at the 100-dB presentation level, performance by each group improved with $F1$ attenuation for each stimulus vowel, although performance by NH listeners was near ceiling for /a/ stimuli and improved only slightly with $F1$ attenuation. $F1$ attenuation did not have a significant effect on performance for the other combinations of noise and presentation level.

Finally, the noise by vowel by $F1$ -attenuation interaction in ANOVA2 was marginally significant [$F(2.01, 24.15)=3.09, p=0.0638$]. Simple effects tests showed a significant benefit of $F1$ attenuation for C/a/ stimuli presented in noise [$F(2.67, 32.02)=3.23, p=0.0401$] and a marginally significant $F1$ -attenuation benefit for C/a/ stimuli in quiet [$F(2.81, 33.72)=2.75, p=0.0615$]. No significant benefit of $F1$ attenuation was present for C/e/ stimuli in either quiet or noise.

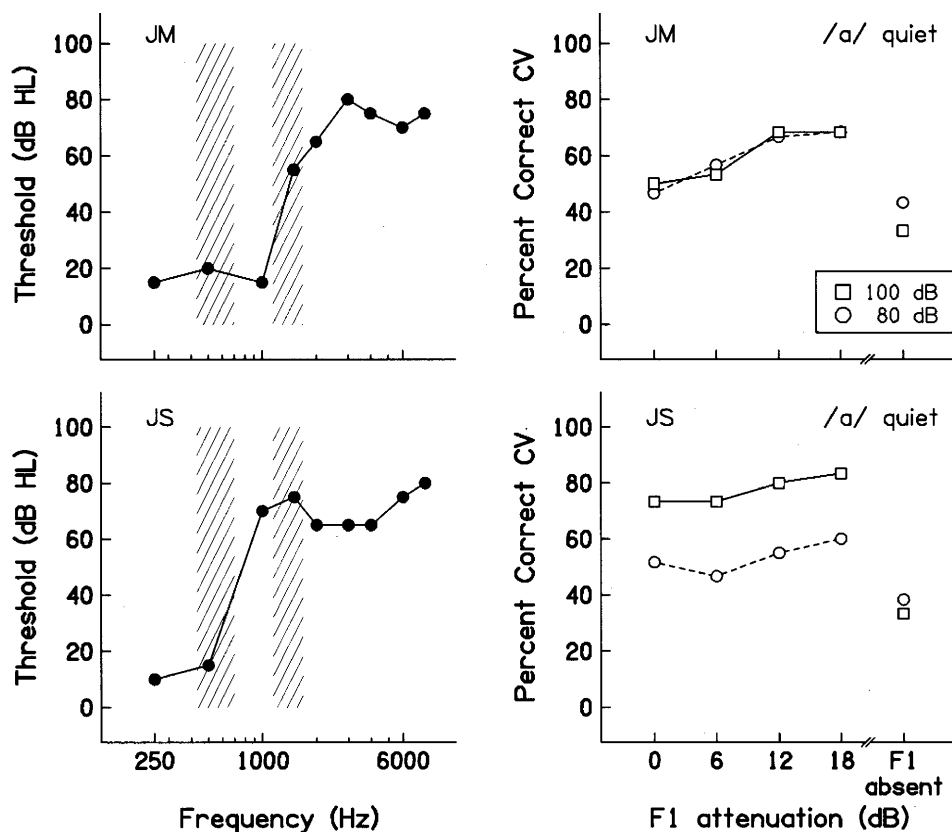


FIG. 5. Audiometric thresholds and CV identification performance on C/a/ stimuli presented in quiet for two HI subjects with steep increases in threshold in $F1$ – $F2$ region for /a/.

D. Relating $F1$ attenuation benefit to individual age and hearing loss

Overall, for C/a/ stimuli, CV labeling tended to improve with $F1$ attenuation of up to 18 dB. In terms of group performance, this benefit of $F1$ attenuation was most consistently observed for HI listeners. However, not all HI listeners showed the pattern of benefit reflected in the group means. Performance was examined for individual subjects to determine if the benefit associated with $F1$ attenuation was related to a given listener's age or to the configuration or severity of the hearing loss. Age and overall severity of the impairment in the $F1$, $F2$ region for /a/ (pure tone average at 500, 1000, and 1500 Hz) were poor predictors of which HI subjects showed benefit from $F1$ attenuation. For C/a/ stimuli presented in quiet, the influence of $F1$ attenuation on labeling accuracy appeared to relate to the slope of the audiogram between the $F1$ and $F2$ regions.

The left-hand panels of Fig. 5 show audiometric thresholds for HI subjects JM and JS. The hatched areas in these panels represent the frequency regions containing $F1$ and $F2$ energy for C/a/ stimuli. Both subjects show a steep increase in threshold between $F1$ and $F2$ for these stimuli (threshold increase of at least 35 dB between 500 and 1500 Hz). Labeling performance by these subjects for C/a/ stimuli in quiet (right-hand panels) shows smooth increases in labeling accuracy as $F1$ attenuation is increased from 0 up to 18 dB. The benefit of $F1$ attenuation is apparent at each presentation level for each subject and the percent improvement between the 0-dB and 18-dB conditions (approximately 20%

for JM and 10% for JS) is greater in every case than the mean percent improvement for the HI group. Three HI listeners (PN, NW, and RL) had more moderate increases in threshold between the $F1$ and $F2$ regions of the /a/ stimuli (threshold increases of 10–20 dB between 500 and 1500 Hz). The data from two of these listeners (PN and NW) also contributed to the significant benefit of $F1$ attenuation on labeling C/a/ stimuli seen in ANOVA2, however, the benefit was not as consistent for these listeners and generally was not as large as seen for JM and JS (subject RL did not show benefit from $F1$ attenuation). Finally, two HI listeners (JL and SK) had flat audiograms in the low-frequency region with no increase in threshold between 500 and 1500 Hz. $F1$ attenuation had no clear positive (or negative) effect on performance for these subjects at either presentation level.

The pattern just described, relating slope of the hearing loss to $F1$ -attenuation benefit, was observed for C/a/ stimuli presented in quiet. There was also a benefit of $F1$ attenuation when these stimuli were presented in noise. However, in noise, the slope of hearing loss between $F1$ and $F2$ did not predict which subjects would show the clearest benefit of $F1$ attenuation. Subjects PN, NW, and SK showed the most consistent benefit from $F1$ attenuation in noise, with subjects JL and JS also contributing to the group effect.

III. DISCUSSION

A. Importance of release bursts

Hearing-impaired listeners, as a group, showed improvements in labeling accuracy with up to 18 dB of $F1$ attenuation for C/a/ stimuli. This benefit of $F1$ attenuation was seen

for testing in quiet and in noise sufficient to mask initial release bursts. The results in quiet indicate improved labeling associated with *F1* attenuation even under favorable listening conditions where release bursts were available as an additional source of place cues. These results may appear inconsistent with the earlier findings of Dorman *et al.* (1985), who did not observe benefit from *F1* attenuation when release bursts were present. Actually, the present results are in agreement with Dorman *et al.* for matching stimuli across the two studies. The difference in overall conclusion is based on stimulus conditions tested here that were not examined by Dorman *et al.* In testing for possible benefit of *F1* attenuation, those investigators compared performance with *F1* at full strength to performance with *F1* entirely removed. Based only on these two conditions, the present results also show no benefit associated with *F1* attenuation. It is in the intermediate levels of *F1* attenuation, not tested by Dorman *et al.*, where benefit was observed in the present data.

Although *F1* removal had no consistent positive or negative effect on labeling in Dorman *et al.* (1985), there was a significant decrease in labeling accuracy for *F1*-absent conditions in the present data. Stimulus differences between the two studies may have contributed to this difference in results. The present stimuli contained small differences in *F1* associated with the place of the initial consonant (see Fig. 2). This contrasts with Dorman *et al.* where *F1* characteristics were held constant regardless of place. As a result, removal of *F1* in the present stimuli eliminated some possible place cues. No place information was lost by *F1* removal in the earlier study since none was present in *F1*. This stimulus difference may partially explain why *F1* removal had a more obvious negative effect in the present data than in Dorman *et al.*

Task differences between the studies may have also contributed to the differing results with regard to *F1* removal. In Dorman *et al.* (1985), subjects identified only the initial consonant with vowel context held constant. In the present study, subjects performed a more difficult labeling task, identifying both consonant and vowel. Here, complete *F1* removal had a negative effect on vowel identification (see Sec. II C). Any effect of *F1* removal on vowel identification could not be observed in the Dorman *et al.* data since only one vowel was used.

Although neither study supports complete *F1* removal as beneficial, less extreme *F1* attenuation resulted in improved labeling performance for several stimulus conditions in the present data. This benefit was seen in quiet and in noise sufficient to mask initial release bursts. It was originally expected that masking of *F2* place cues by *F1* (and masking release through *F1* attenuation) might be particularly significant in noise since ambiguity in processing *F2* cues could not be offset by reliance on burst cues. The results provide little support for any increased benefit from *F1* attenuation in noise. For C/a/ stimuli, benefit of *F1* attenuation was seen both in quiet and in noise with no clear change in the amount of benefit between these conditions. For C/e/ stimuli, the only benefit of *F1* attenuation was seen for testing in quiet at the high presentation level. With this one exception, the presence of broadband noise did not enhance

or reduce the benefit to labeling provided by *F1* attenuation. These findings suggest that appropriate attenuation of *F1* may benefit labeling performance independent of the availability of release bursts whereas *F1* removal will consistently reduce labeling accuracy.

B. Vowel context and presentation levels

The /a/ and /e/ vowels provided very different contexts in which to test *F1*-on-*F2* masking in that they vary both in formant-frequency spacing and in relative formant amplitudes (see Fig. 3). For /a/, *F1* and *F2* are closely spaced in frequency with *F2* amplitude about 5 dB below *F1* amplitude. For /e/, *F1* and *F2* are more widely spaced, with *F2* level much less intense relative to *F1* (−20 dB). The closer formant spacing in the /a/ context would be expected to produce greater USM for /a/ than /e/. On the other hand, the greater amplitude difference seen in the /e/ context would make masking more likely for /e/. In the present results the more consistent benefit associated with *F1* attenuation for C/a/ stimuli suggests that *F1*-on-*F2* masking had a greater impact on performance for these stimuli than on C/e/ stimuli. Apparently, in this case, the relative intensities of *F1* and *F2* had less influence on USM than formant spacing.

An early discrimination study also examined how vowel context affects *F1*-on-*F2* masking (Danaher *et al.*, 1973). Formant frequencies and relative amplitudes of *F1* and *F2* approximated /a/ and /I/. The /a/–/I/ contrast involves an even greater difference in formant spacing and in relative formant amplitudes than the present /a/–/e/ case in that the frequency separation and intensity difference between *F1* and *F2* is greater for /I/ than /e/. Testing listeners with more severe hearing impairments than those tested here, Danaher *et al.* reported a vowel effect as well, but in the opposite direction: greater *F1*-on-*F2* masking in the /I/ context than for /a/. In that study, HI subjects were tested at most comfortable listening levels that were generally greater than 105 dB SPL. When tested at 105 dB SPL, NH listeners also showed greater *F1*-on-*F2* masking for /I/ and /a/. The high presentation levels apparently contributed to USM in the /I/ context where broad formant spacing should have allowed separate resolution of the formant peaks. The Danaher *et al.* (1973) results are in agreement with the present data in demonstrating differences in the amount of *F1*-on-*F2* masking across vowel contexts. The specific vowels most likely to demonstrate this masking may depend on the characteristics of the hearing loss and the presentation levels involved.

With regard to presentation level, it is well-known that USM increases with level for both NH and HI listeners (Leek and Summers, 1993; Zwicker and Jaroszewski, 1982; Egan and Hake, 1950). It would therefore be expected that *F1*-on-*F2* masking might be particularly evident at high presentation levels. The Danaher *et al.* (1973) discrimination data provide support: at high presentation levels (95 and 105 dB SPL), both NH and HI subjects showed susceptibility to *F1*-on-*F2* masking. At lower presentation levels NH listeners did not show this pattern. In the present data, the most consistent benefit associated with *F1* attenuation was seen at the high (100 dB) presentation level. At this level, for presentations in quiet, *F1* attenuation of up to 18 dB improved

performance of both groups across vowel contexts. For C/e/ stimuli, this was the only condition where $F1$ attenuation improved labeling. These results for /e/ parallel the discrimination results for /I/ reported by Danaher *et al.* (1973). For these vowels, characterized by greater frequency spacing between $F1$ and $F2$ than is present for /a/, evidence of $F1$ -on- $F2$ masking may only be seen at high presentation levels. Overall, the present results suggest that high stimulus levels, which may be necessary to assure audibility for some HI listeners, increase the likelihood of USM from $F1$ to $F2$.

C. Effect of hearing loss and configuration

Some of the clearest cases of benefit from $F1$ attenuation occurred for HI listeners with steeply sloping audiograms across the $F1$ – $F2$ region of the test stimuli. These results parallel the Danaher *et al.* (1973) discrimination data where greater $F1$ -on- $F2$ masking was observed for subjects with sloping losses than for subjects with flat losses. Studies using other materials and methods (e.g., tone-on-tone masking, noise masking) have reached similar conclusions. In a review of a number of studies examining spread of masking in listeners with high-frequency hearing loss, Humes (1982) concluded that the “transition” or sloping portion of the loss was the region most likely to show USM not accounted for by quiet thresholds. More recent work by Trees and Turner (1986) reported similar findings, with HI subjects frequently showing abnormal USM in the transition region of their loss.

The present results attest to the potential danger of over-amplifying low-frequency portions of the acoustic signal for listeners with sloping hearing loss. These listeners may be particularly susceptible to USM when low-frequency portions of the signal are presented at high levels. In a number of studies examining aided speech recognition by listeners with sloping losses, it has been shown that reductions in low-frequency amplification can lead to improved performance (Gordon-Salant, 1984; Kamm *et al.*, 1982; Pascoe, 1975).

IV. CONCLUSIONS

Up to 18 dB of $F1$ attenuation led to improved CV identification for both HI and NH listeners under a number of stimulus conditions. This benefit to labeling associated with $F1$ attenuation was more consistently present for HI than for NH subjects. The magnitude of this benefit was similar for testing in quiet and in noise, suggesting that cues available from release bursts during presentations in quiet did not negate the benefit of reduced $F1$ -on- $F2$ masking.

Improvements in CV labeling associated with $F1$ attenuation were more reliably present for C/a/ stimuli than C/e/ stimuli. The proximity of $F1$ and $F2$ in /a/ probably underlies this result, outweighing the greater $F1$ – $F2$ level differences present in /e/. For C/a/ stimuli presented in quiet, the labeling benefit associated with $F1$ attenuation related to the steepness of the audiogram in the $F1$, $F2$ region for /a/. Listeners with steep increases in hearing thresholds across

this region showed the largest and most consistent improvements in labeling accuracy with $F1$ attenuation.

There was a consistent labeling benefit associated with $F1$ attenuation for stimuli presented in quiet at the high signal levels. The benefit was observed across groups and in both vowel contexts. This is consistent with earlier studies suggesting that at high presentation levels, $F1$ -on- $F2$ masking may occur for vowels with considerable frequency spacing between $F1$ and $F2$ and may be seen independent of cochlear pathology.

Although partial attenuation of $F1$ often improved performance, complete removal of $F1$ consistently reduced labeling accuracy. The removal of $F1$ led to poorer performance on both the consonant and vowel portion of the CV stimulus.

Overall, the results emphasize the balance that must be achieved in meeting the amplification needs of HI listeners. Underamplification of low-frequency regions may leave important speech cues inaudible. Obviously this may include segmental and suprasegmental cues unrelated to consonantal place of articulation, (e.g., vowel cues). Conversely, overamplification of low frequencies may lead to masking of speech cues present at higher frequencies. This speech-on-speech masking may also interfere with other important speech cues in addition to the consonantal place cues examined here.

ACKNOWLEDGMENTS

This research was supported by Grant No. DC 00626 from the National Institute on Deafness and Other Communication Disorders to Walter Reed Army Medical Center (PI: Marjorie Leek), and by the Clinical Investigation Service, Walter Reed Army Medical Center, under Work Unit #2538. 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.

¹Synthesis parameters are available from the first author upon request.

- ANSI (1989). ANSI S3.6-1989, “Specifications for audiometers” (American National Standards Institute, New York).
- Cohen, J., and Cohen, P. (1983). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 2nd ed. (Erlbaum, Hillsdale, NJ).
- Danaher, E. M., Osberger, M. J., and Pickett, J. M. (1973). “Discrimination of formant frequency transitions in synthetic vowels,” *J. Speech Hear. Res.* **16**, 439–451.
- Danaher, E. M., and Pickett, J. M. (1975). “Some masking effects produced by low-frequency vowel formants in persons with sensorineural hearing loss,” *J. Speech Hear. Res.* **18**, 261–271.
- Dorman, M. F., Lindholm, J. M., and Hannley, M. T. (1985). “Influence of the first formant on the recognition of voiced stop consonants by hearing-impaired listeners,” *J. Speech Hear. Res.* **28**, 377–380.
- Egan, J. P., and Hake, H. W. (1950). “On the masking pattern of a simple auditory stimulus,” *J. Acoust. Soc. Am.* **22**, 622–630.
- Fabry, D. A., Leek, M. R., Walden, B. E., and Cord, M. (1993). “Do adaptive frequency (AFR) hearing aids reduce ‘upward spread’ of masking?” *J. Rehab. Res. Devel.* **30**, 318–325.
- Glasberg, B. R., and Moore, B. C. J. (1986). “Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments,” *J. Acoust. Soc. Am.* **79**, 1020–1033.

- Glasberg, B. R., and Moore, B. C. J. (1989). "Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech," *Scand. Audiol. Suppl.* **32**, 1–25.
- Gordon-Salant, S. (1984). "Effects of reducing low-frequency amplification on consonant perception in quiet and noise," *J. Speech Hear. Res.* **27**, 483–493.
- Greenhouse, S. W., and Geisser, S. (1959). "On methods in the analysis of profile data," *Psychometrika* **24**, 95–112.
- Hannley, M., and Dorman, M. F. (1983). "Susceptibility to intraspeech spread of masking in listeners with sensorineural hearing loss," *J. Acoust. Soc. Am.* **74**, 40–51.
- Humes, L. E. (1982). "Spectral and temporal resolution by the hearing impaired," in *The Vanderbilt Hearing Aid Report*, edited by G. A. Studebaker and F. H. Bess (Monographs in Contemporary Audiology, Upper Darby, PA), pp. 16–31.
- Kamm, C. A., Dirks, D. D., and Carterette, E. C. (1982). "Some effects of spectral shaping on recognition of speech by hearing-impaired listeners," *J. Acoust. Soc. Am.* **71**, 1211–1224.
- Klatt, D. (1980). "Software for a cascade/parallel formant synthesizer," *J. Acoust. Soc. Am.* **67**, 971–995.
- Leek, M. R., and Summers, V. (1993). "Auditory filter shapes of normal-hearing and hearing-impaired listeners in continuous broadband noise," *J. Acoust. Soc. Am.* **94**, 3127–3137.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Pascoe, D. P. (1975). "Frequency responses of hearing aids and their effects on the speech perception of hearing-impaired subjects," *Ann. Otol. Rhinol. Laryngol.* **84** (Suppl 23), 5–40.
- Pavlovic, C. V. (1982). "Derivation of primary parameters and procedures for use in speech intelligibility predictions," *J. Acoust. Soc. Am.* **82**, 413–422.
- Peterson, G. E., and Barney, H. L. (1952). "Control methods in the study of vowels," *J. Acoust. Soc. Am.* **24**, 175–184.
- Pickett, J. M., and Danaher E. M. (1975). "On discrimination of formant transitions by persons with severe sensorineural hearing loss," in *Auditory analysis and perception of speech*, edited by G. Fant and M. A. A. Tatham (Academic, New York), pp. 275–292.
- Rankovic, C. M., Freyman, R. L., and Zurek, P. M. (1992). "Potential benefits of adaptive frequency-gain characteristics for speech reception in noise," *J. Acoust. Soc. Am.* **91**, 354–362.
- Trees, D. E., and Turner, C. W. (1986). "Spread of masking in normal subjects and in subjects with high-frequency hearing loss," *Audiology* **25**, 70–83.
- Turner, C. W., Fabry, D. A., Barrett, S., and Horowitz, A. R. (1992). "Detection and recognition of stop consonants by normal-hearing and hearing-impaired listeners," *J. Speech Hear. Res.* **35**, 942–949.
- Tyler, R. S. (1986). "Frequency resolution in hearing-impaired listeners," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London) pp. 309–371.
- Van Tasell, D. J. (1980). "Perception of Second-Formant Transitions by Hearing-Impaired Persons," *Ear Hear.* **1**, 130–136.
- Zwicker, E., and Jaroszewski, A. (1982). "Inverse frequency dependence of simultaneous tone-on-tone masking patterns at low levels," *J. Acoust. Soc. Am.* **71**, 1508–1512.

A targeting-and-extracting technique to enhance hearing in the presence of competing speech

Chen Liu,^{a)} Judith Rosenhouse, and Samuel Sideman

The Julius Silver Institute, Department of Biomedical Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel

(Received 18 October 1995; revised 28 October 1996; accepted 20 December 1996)

A targeting-and-extracting procedure of speech enhancement for hearing aids in the presence of background noise, especially competing speech, is proposed. The procedure is composed of two steps: targeting by a fixed (or deterministic) beamforming array, followed by a post-targeting extracting step. Emphasis is placed on the extracting step, which performs noise cancellation based on the acoustic difference between the desired speech and interfering speech. Either comb filtering or attenuation is applied to the signal in accordance with the current voiced/unvoiced/silence state of the desired signal. The comb filter design is based on the fundamental pitch frequency of the desired speech. Algorithms for deciding the voiced/unvoiced/silence state and determining the fundamental frequency are developed. The performance of the system is evaluated through computer simulation. The simulation results indicate significant noise cancellation and intelligibility improvement. © 1997 Acoustical Society of America. [S0001-4966(97)06804-5]

PACS numbers: 43.72.Ew, 43.66.Ts, 43.38.Hz [JS]

INTRODUCTION

Fixed microphone arrays have recently been extensively studied as a very promising tool for reducing the background noise that affects the usefulness of hearing aids (Soede *et al.*, 1993; Stadler and Rabinowitz, 1993; Kates, 1993; Liu and Sideman, 1996). However, the noise reduction capability of the fixed microphone arrays for hearing aids is limited by their small size, since they are usually designed to be worn as eyeglasses, or otherwise attached to the body. For example, consider a conventional delay-and-sum five-omnidirectional-microphone array spanning 10.0 cm in an endfire configuration, which is suitable for being mounted on the legs of a normal frame of eyeglasses. The size of such an array corresponds to one-half a wavelength for an audio signal of 1.73 kHz, which is, however, too high relative to the whole frequency band of human speech.

Several enhancing methods have been considered to increase the directivity of the fixed arrays for hearing aid applications. They include using various directional microphones as the array sensors instead of the omnidirectional microphones (Soede *et al.*, 1993), optimizing the array weights to obtain a superdirectivity (Kates, 1993), and a combination of the two (Stadler and Rabinowitz, 1993; Liu and Sideman, 1996). However, the largest directivity that can be achieved is still insufficient to cope with adverse practical noise situations. To further improve speech enhancement, it is proposed here to exploit the acoustic difference between the desired speech and noise, specifically the differences in the voiced/unvoiced/silence state and the differences in the fundamental pitch frequency, so as to achieve further noise reduction after the beamforming.

Exploiting pitch differences between different speakers

has been studied from the auditory modeling point of view by many researchers aiming to separate *two* concurrent speech events (Scheffers, 1983; Weintraub, 1986; Stubbs and Summerfield, 1988; Assmann and Summerfield, 1990; Meddis and Hewitt, 1992; De Cheveigné, 1993), or for the purpose of speech enhancement from the signal processing point of view (Parsons, 1976; Frazier *et al.*, 1976; Hanson and Wong, 1984; Naylor and Boll, 1987; Lee and Childers, 1988). Unfortunately, those methods are not suitable for implementation in hearing aids in the environment with *many* competing speakers. Here we present a technique to determine the voiced/unvoiced/silence state and the pitch of the desired speech corrupted by other concurrent speech events. This determination is based on comparing the signals before and after the beamforming, as well as on analysis of each signal itself. Once the acoustic feature of the desired speech is determined, either a comb filter is designed and implemented in the frequency domain so as to remove the noise components filling the interharmonic valleys, or an attenuation is applied to make noise reduction.

Simpson *et al.* (1990), Stone and Moore (1992), and Wang *et al.* (1993) identified speech formants by different methods and removed the noise in the spectral valleys between the *formants*. We think that removing the noise that fills in the valleys between the *pitch harmonics* should yield a much cleaner result. Moreover, it is practically much easier and more robust to detect the pitch when speech constitutes the background noise, than to detect the formants.

First, we will outline the rationale of the proposed targeting-and-extracting system, followed by a detailed description of each part. The extraction part, which includes the voiced/unvoiced/silence decision, pitch determination, and comb filter design, constitutes the major part of this paper. Finally, computer simulations of characteristic sample situations are used to demonstrate the performance of the proposed system.

^{a)}Present address: Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801.

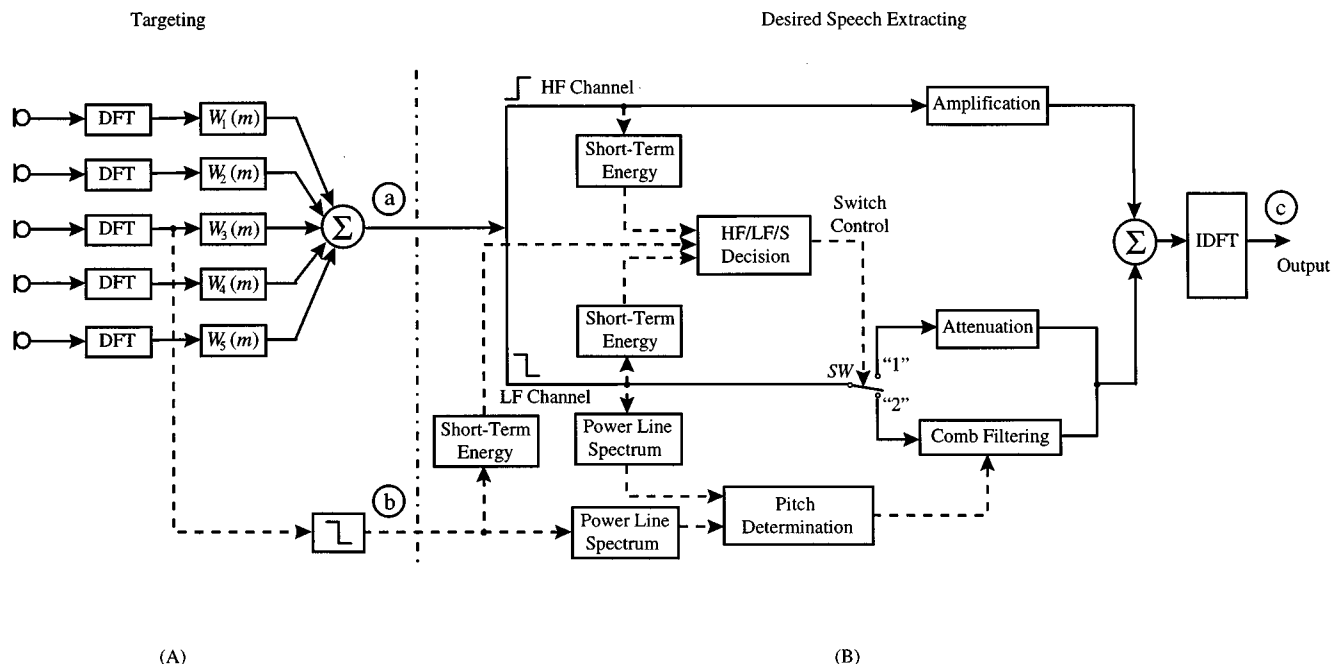


FIG. 1. Block diagram of the targeting-and-extracting system. (A) The broadband frequency-domain beamforming array, i.e., the targeting step. (B) The desired speech extracting step. The solid lines denote the channel of the signal to be processed, and the dashed lines denote the channel of the signal used in the determination algorithms.

I. METHODOLOGY

A. General outline

The targeting and extracting parts are schematically presented in Fig. 1. The *targeting part* is a microphone beamforming array that receives the sound of interest coming from a certain direction, generally directly in front of the listener, while suppressing any sounds coming off-axis. The *extracting part* includes a high-frequency (HF) channel, a low-frequency (LF) channel, a comb filter, and an attenuator, as well as other supporting elements which function to identify the current acoustic features of the desired speech. This part aims at further reduction of the competing background noises, which are generally different in acoustic features from the desired speech.

The method for selectively receiving the speech of interest, by taking advantage of both the spatial cues and acoustic features of the signal, represents an approximate imitation of acoustic perception by human beings. It is still unclear how these two mechanisms work in nature, either separately or in unison, toward the final acoustic perception. Here, these two parts are arranged in series: The microphone array serves as a spatially discriminative sensor submitting the preliminary noise-suppressed result to the extracting process for further acoustic-feature-based noise cancellation.

The present scheme thus attempts to enhance the inadequate noise cancellation provided by portable microphone arrays (Soede *et al.*, 1993; Stadler and Rabinowitz, 1993; Kates, 1993; Liu and Sideman, 1996) which, due to the size limitations, are rather limited in suppressing directional interference, especially at low frequencies. Even though the technique of superdirective arrays provides significant improvement in noise cancellation, there is still a considerable

degradation in array gain at low frequencies relative to that at high frequencies (Kates, 1993; Liu and Sideman, 1996).

The underlying rationale of the present approach is based on the assumption (see Sec. I B) that the head-size microphone array is able to provide sufficient discrimination for high frequencies, e.g., the frequencies above 3 kHz. Thus, we denote the signal components above 3 kHz as *high frequencies* (HFs), and the components below 3 kHz as *low frequencies* (LFs). The HFs are mostly unvoiced phonemes and the LFs are mostly voiced phonemes. Obviously, there is no definitive dividing point between the voiced and unvoiced phonemes. However, either the amount of unvoiced phonemes below 3 kHz, or the amount of voiced phonemes above 3 kHz, is generally small enough so as not to significantly affect the feasibility of the present method.

Table I summarizes the possible dynamic situations defined by the different phonetic characteristics of the desired speaker and the competing speakers, as well as the corresponding tactics devised for implementing the technique. A

TABLE I. Different phonetic events by the desired and the competing speakers, and the corresponding noise cancellation tactics by the targeting-and-extracting technique. HF—high-frequency above 3 kHz; LF—low-frequency below 3 kHz.

Desired speaker	Competing noise	Noise cancellation tactics	Position of switch SW in LF channel in Fig. 1(B)
HF	HF	Microphone array	"1"
HF	LF	Deep attenuation	"1"
LF	HF	Microphone array	"2"
LF	LF	Comb filter	"2"
Silent	Whatever	Microphone array + Deep attenuation	"1"

number of simplified situations are considered.

First case: The desired HFs coincide with the competing HF noise; we assume that the microphone array alone suffices for canceling the competing HFs.

Second case: The desired HFs coincide with the competing LFs; the microphone array not only serves as the sensor to pick up the signals but also suppresses the competing LFs to some extent. The residual LF noise can be removed by the attenuator in the LF channel.

Third case: The desired LFs coincide with the HF noise, and, as in the first case, the HF noise is to be removed by the microphone array.

Fourth case: The desired LFs coincide with the competing LFs. As is well known, this case is the most frequently encountered situation in conversations (Flanagan, 1972). Here, the competing LFs are first suppressed to some extent by the microphone array. Since generally different speakers feature different pitches in the production of voiced phonemes, the array output is spectrally an overlap of the spectral harmonics of different fundamental pitch frequencies. This is where the *comb filter* comes in. The periodicity of the comb filter's transfer function, i.e., the interval between the adjacent comb "teeth" is adjusted to match the fundamental frequency F_0 of the desired speaker, so as to allow the desired voiced phonemes to pass through. Simultaneously, the voiced phonemes of different fundamental frequencies, mostly those with spectral components which fall in the "valley area" of the transfer function of the comb filter, are rejected.

Fifth case: The desired speaker is silent. The system is not switched off; instead, a deep attenuation is used in order (1) to keep the listener alert to the environment; and (2) to retain a constant comfort noise level so as to eliminate unpleasant subjective effects of switching between silence and speech in high noise (Freeman *et al.*, 1989).

Obviously, the above cases represent five simplified situations, whereas real life is usually a combination of these simplified situations, especially when more than one competing speaker is participating in the conversation. In this case, however, there is not conflict in using suitably combined corresponding noise cancellation tactics.

B. Targeting

As illustrated in Fig. 1(B), the post-targeting extracting step involves a manipulation in the frequency domain. Consequently, we employed a broadband *frequency-domain* beamforming technique (for details, see Liu and Sideman, 1995) for the targeting step [Fig. 1(A)]. Therefore, the well-known short-term Fourier analysis and synthesis technique (Allen, 1977; Allen and Rabiner, 1977; Crochiere and Rabiner, 1983), was used.

We now consider an array with n omnidirectional microphones. Denote the digitized microphone signal in the i th channel as $x_i(k)$ ($i = 1, \dots, n$), the beamforming output as $y_b(k)$, and their spectra as $X_i(m)$ ($i = 1, \dots, n$) and $Y_b(m)$, respectively. Theoretically, the broadband frequency-domain beamforming operation is

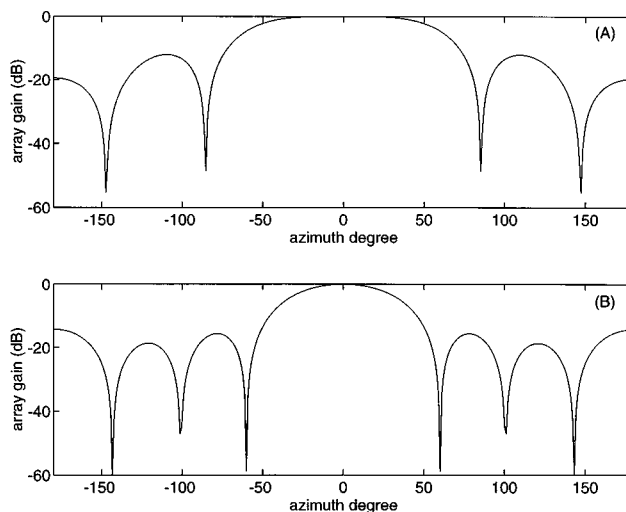


FIG. 2. (A) Beampattern of the five-omnidirectional-microphone delay-and-sum array in a 10-cm endfire configuration at 3 kHz. (B) Beampattern of the five-omnidirectional-microphone superdirective array in a 10-cm endfire configuration at 3 kHz.

$$Y_b(m) = \sum_{i=1}^n W_i(m) X_i(m), \quad (1)$$

where $W_i(m)$ ($i = 1, \dots, n$) are the optimum broadband beamforming weights. In this research, we optimized the weights so as to achieve a broadband superdirective array (Cox *et al.*, 1986). The detailed implementation of Eq. (1) was provided in Liu and Sideman (1996). For a variety of other optimization criteria for the beamformer, please refer to Monzingo and Miller (1980).

It is noted that the targeting step can also be implemented in the form of broadband *time-domain* beamforming array, which comprises transversal filters (or tapped-delay lines), thus demanding the use of phase shifters or delay lines. The difference between the two beamforming schemes is that the Fourier analysis is performed before beamforming for the frequency-domain array, while after beamforming for the time-domain array.

To illustrate the noise reduction provided by the array for frequencies above 3 kHz, we look into the beampattern of, for example, a five-omnidirectional-microphone array in the endfire configuration with a span of 10.0 cm, at 3 kHz. Figure 2(A) and (B) shows the beampatterns for the conventional uniformly weighted delay-and-sum beamforming weights and the superdirective beamforming weights, respectively. In both cases, no grating lobes occur, thus a maximum array gain is ensured exclusively for the target. Moreover, the directivity index is 11 dB for the superdirective beamformer, compared with 6.5 dB for the delay-and-sum beamformer, thus demonstrating a stronger ability to cancel the environmental noise.

C. Extracting

1. General outline

As shown in Fig. 1(B), the extracting step consists of the low- and high-frequency channels, comb filtering, and deep attenuation, which, with the support of voiced/unvoiced/

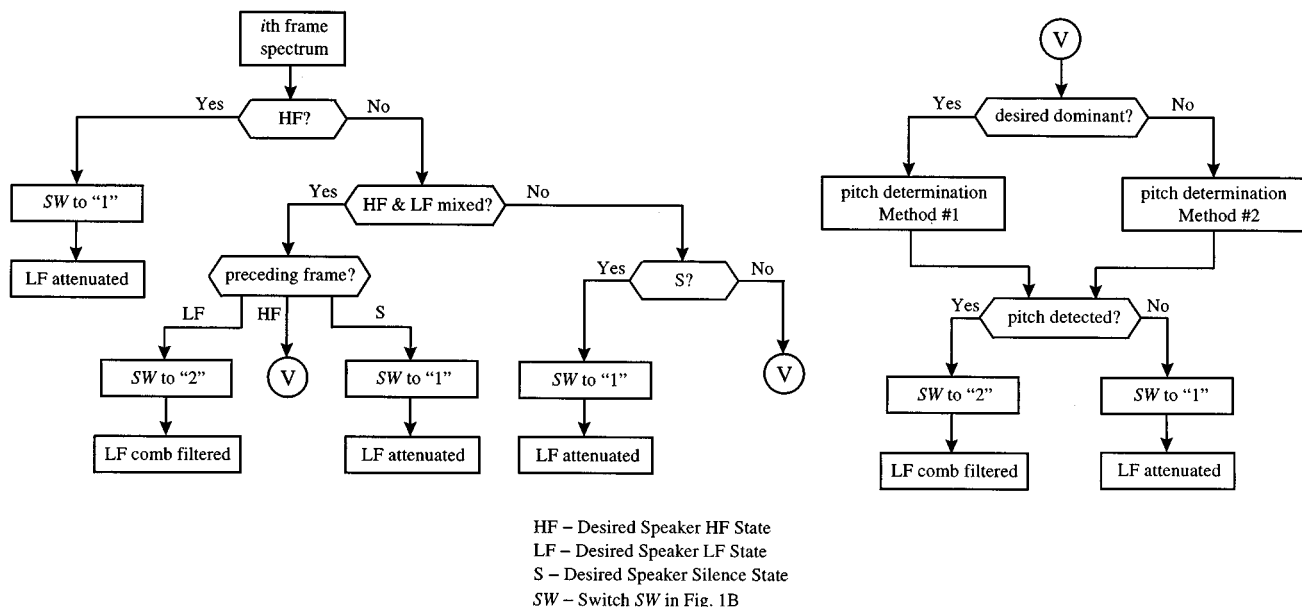


FIG. 3. Flowchart of the desired extracting algorithm.

silence decision and pitch determination, perform the extracting work for the various acoustic situations listed in Table I. The output of the DFT filterbank is split into the low- and high-frequency channels, with the cutoff frequency being 3 kHz. Due to the sufficient cancellation of noise components above 3 kHz by the targeting step, the frequencies above 3 kHz (i.e., HF) are transmitted through the HF channel to the system output, while the LFs are processed according to the different situations.

The algorithm for the extracting step is outlined in Fig. 3. For clarity, Fig. 3 shows all the major parts of the algorithm, while the details of each part will be elaborated in the following text. First, the spectrum of the present frame is analyzed to determine the high-frequency dominant/low-frequency dominant/desired silence (HF/LF/S) state, for the desired speech. Then, based on the detected HF/LF/S state, a choice is made [represented by the position of the switch SW in Fig. 1(B)] so as to apply either the comb filtering or attenuation on the LF signal. Some typical situations are given as follows.

As discussed in Sec. I A, the HF, if detected, originates from the desired speaker, while the LF may come from the desired speaker and/or the competing speakers. If the present frame is determined to be mainly HF, the switch SW in Fig. 1(B) is turned to position "1" so that any LF, if it exists, is further attenuated before output.

When the present frame contains a significant amount of both HF and LFs, then the present frame may be in one of several possible transition states, which can be determined by referring to the HF/LF/S state and the pitch information (if available) of the preceding frame. The following potential situations are considered:

(1) If the preceding frame is dominated by the desired LF, the present frame may be a transition from the desired LF to the desired HF. Consequently, the switch SW is turned to position "2" and the comb filter, the same as for the

preceding frame, is applied to the desired LF tail in the LF channel.

(2) If the preceding frame is HF dominant, the present frame may be either a transition from the desired HF to the desired LF, or the desired HF accompanied by the beginning of the competing LF. In any case, a pitch determination is made and the corresponding comb filter is applied for the present frame. However, an occurrence of more than two HF+LF mixed frames in succession will suggest the latter possibility, i.e., the desired HF accompanied by the undesired LF, hence the switch SW is turned to "1" and the components in the LF channel are attenuated. Otherwise, if the succeeding frames are LF dominant, it is likely the former possibility, i.e., a transition of the desired speech from HF to LF. Consequently, SW is turned to "2" and the comb filter is applied to the signal in the LF channel.

(3) If the desired speaker is silent in the preceding frame, the present frame may be a transition from the desired speaker's silence to the desired HF. The switch SW is turned to "1" and the components in the LF channel are attenuated.

If the present frame is initially determined to be the desired speaker's silence state while its preceding frame is the desired LF, it could either mean that the desired LF is absent or that the desired LF is becoming very weak with respect to the background noise. Note that the initial HF/LF/S decision is made based on the ratio of energies of the higher- and lower-frequency components as well as changes of the energy between successive frames; this point will be elaborated in Sec. I C 2. Therefore, a final determination is made by checking in the present frame for the existence of a pitch which may be attributed to the desired speech by comparison with the pitch detected in the preceding frame for the desired speech. If such a pitch exists, even though it may not be due to the desired speech, the switch SW is turned to position "2" and a corresponding comb filter is applied; this is a failsafe measure. Otherwise, the switch SW is turned to

“1” and the LF is attenuated. However, if the desired speaker is silent in the preceding frame and the total energy for the present frame is still below a certain threshold (see Sec. I C 2 b), the desired speaker is also regarded as silent in the present frame.

As illustrated in Fig. 3, the algorithm checks the present frame for HF *before* checking for the desired speaker's silence. Note that because the HF components are generally weak in energy, checking for HF *after* checking for the desired speaker silence would lead to most of the HF frames being determined as “silence” by the algorithm, which determines the desired speaker's silence on the energy basis (see Sec. I C 2 b).

If the present frame cannot be assigned to any of the above situations (HF, HF+LF mixed, or S) for the desired speaker, then it is determined to be mainly LF by the desired speaker. The switch SW is turned to position “2” and the pitch of the desired speech is then determined based on the LF components, both before and after the beamforming (see Sec. I C 3). A frequency-domain comb filter, designed so that its between-the-teeth interval matches the detected fundamental pitch frequency F_0 , is applied to the signal components in the LF channel so as to further remove the spectral components of the competing speech.

2. HF/LF/S decision

The HF/LF/S decision is actually a voiced/unvoiced/silence decision problem, which has been extensively studied in the past two decades (Atal and Rabiner, 1976; Siegel, 1979; Un and Lee, 1980; Siegel and Bessey, 1982; Kobatake, 1987; Krubsack and Niederjohn, 1991). The voice activity detection problem was sometimes investigated as an independent problem (Brady, 1965; Lee and Un, 1986; Freeman *et al.*, 1989; Tucker, 1992). Nevertheless, in the presence of noise, especially when the noise is the competing speech, the voiced/unvoiced/silence determination remains a very difficult and still open problem, which is beyond the scope of this work. Here, we use a simple but effective method which is made up of two tasks: The voiced/unvoiced/mixed determination and the desired voice activity detection.

a. Voiced/unvoiced/mixed determination. The voiced/unvoiced/mixed determination is performed based on monitoring the interframe changes of the spectral distribution of the energy in a frame. If we denote the spectrum of the beamforming output as $Y_b(m)$, the energy e_L below 2 kHz (without considering dc) and the energy e_H between 3 and 5 kHz are given by

$$e_L = \sum_{m=1}^{2000N/f_s} |Y_b(m)|^2 \quad (2)$$

and

$$e_H = \sum_{m=3000N/f_s}^{5000N/f_s} |Y_b(m)|^2, \quad (3)$$

respectively, where f_s is the sample rate.

The ratio (e_H/e_L) is a well-proved feature for voiced/unvoiced determination (Siegel and Bessey, 1982). However, as analyzed in Sec. I B, since the HF components of noise

signals are supposed to be largely removed by the targeting procedure, the signal components in the HF channel (≥ 3 kHz) are assumed to be due to the desired speech signal, while the components in the LF channel (< 3 kHz) are attributed to the desired as well as the noise signals. Thus, the ratio (e_H/e_L) varies with the power of the noise. A fixed threshold for the voiced/unvoiced/mixed determination is obviously not robust. By taking the ratio (e_H/e_L) for a vowel frame in a short time interval, say every 1 s, as a reference, we define a relative ratio for each frame within the interval,

$$(e_H/e_L)_{rv} = \frac{(e_H/e_L)}{(e_H/e_L)_v}, \quad (4)$$

where (e_H/e_L) is the ratio for each frame in the interval and $(e_H/e_L)_v$ is for an arbitrary vowel frame at the beginning of the interval. A threshold ϵ_v for making the determination is now defined:

$$\begin{aligned} \text{HF dominant} & \quad \text{if } (e_H/e_L)_{rv} > \epsilon_v, \\ \text{HF+LF mixed} & \quad \text{if } \epsilon_v/2 < (e_H/e_L)_{rv} \leq \epsilon_v, \\ \text{LF dominant} & \quad \text{if } (e_H/e_L)_{rv} \leq \epsilon_v/2. \end{aligned} \quad (5)$$

The reference $(e_H/e_L)_v$ is dynamic, being updated for each new interval. There must be an overlap between two successive intervals so that the vowel nature of the reference frame for a new interval may be determined in the preceding interval. Two ways for setting a value for the initial $(e_H/e_L)_v$ are suggested here. One way is to rely on the user to adjust its value until a best processing result is obtained. The other way is to define a second reference $(e_H/e_L)_s$, which represents the ratio (e_H/e_L) for a desired-speech-silent frame. Thus, the voiced/unvoiced/mixed determination for the first several intervals is made based on the (e_H/e_L) ratio for each frame with respect to this reference,

$$(e_H/e_L)_{rs} = \frac{(e_H/e_L)}{(e_H/e_L)_s}. \quad (6)$$

It is noted that for a desired-speech-silent frame, both e_H and e_L are from the interference speech only. Due to the beamforming, the residual HF components are rather small, so that the value of $(e_H/e_L)_s$ is quite small. If the first phoneme that the desired speaker produces is unvoiced, whether it lasts long enough to fill the whole frame or not, there should be an abrupt increase in (e_H/e_L) . In view of this situation, another threshold ϵ_s is defined for the voiced/unvoiced/mixed determination,

$$\begin{aligned} \text{HF dominant} & \quad \text{if } (e_H/e_L)_{rs} > \epsilon_s, \\ \text{HF+LF mixed} & \quad \text{if } \epsilon_s/2 < (e_H/e_L)_{rs} \leq \epsilon_s, \\ \text{LF dominant} & \quad \text{if } (e_H/e_L)_{rs} \leq \epsilon_s/2. \end{aligned} \quad (7)$$

In practice, the user will be required to turn on the system when the desired speaker is silent. The reference $(e_H/e_L)_s$, together with the determination criterion in Eq. (7), is in effect until the desired speaker begins speaking and the first vowel frame is detected. At that time the reference $(e_H/e_L)_s$ is replaced by $(e_H/e_L)_v$, and the determination criterion in Eq. (7) is replaced by the criterion in Eq. (5). The $(e_H/e_L)_s$, together with the criterion in Eq. (7), is not used

for the whole procedure because: (1) The interfering speakers and locations vary, affecting the reliability of $(e_H/e_L)_s$ as a reference, and (2) $(e_H/e_L)_s$ does not relate quantitatively to the ratio $(e_H/e_L)_v$ from the desired speech as closely as does $(e_H/e_L)_v$, which itself is from the desired speech.

The criteria in (5) and (7), which define the distinction between HF dominant/HF+LF mixed/LF dominant, are obviously coarse, but practical. By using the middle state, i.e., the HF+LF mixed state, the determination as well as the processing is more precise than when using only a two-state HF/LF determination. The state of a frame that is preliminarily determined as the HF+LF mixed state is finally determined by referring to the state of its preceding frame (Sec. I C 1).

b. Desired voice activity detection. A method of exploiting the energy relation between the signals at the beamforming input and output is now proposed. It is based on the fact that in the absence of the desired voice activity, the signal which is supposed to be off-axis will be suppressed to a significant extent by the beamforming. Since the beamforming operation does not cancel the desired signal, the value of the energy ratio between the signals at the beamforming output and input will be greater when the desired voice activity is present, than its value when the desired voice activity is absent. Thus, we define a relative measure

$$(e_A/e_B)_{rs} = \frac{(e_A/e_B)}{(e_A/e_B)_s}, \quad (8)$$

where (e_A/e_B) is the energy ratio between the signals at the beamforming output and input for each current frame; $(e_A/e_B)_s$ is the energy ratio for the initial desired-speech-silent frame. The user is thus required to turn the system on when the desired speaker is silent. The symbols e_A and e_B represent the frame energy of the signals after and before beamforming, respectively. In view of the symmetry property in the spectrum domain, e_A and e_B are defined as follows, neglecting the symmetrical last half of the spectral components at $m=N/2, \dots, N-1$,

$$e_A = \sum_{m=0}^{N/2-1} |Y_b(m)|^2 \quad (9)$$

and

$$e_B = \sum_{m=0}^{N/2-1} |X(m)|^2, \quad (10)$$

where $X(M)$ denotes the signal spectrum in any one of the n input channels (In our computer simulation with a five-microphone linear uniform array, we used the signal spectrum from the middle microphone for calculation of e_B).

The desired voice activity detection in the inspected frame depends partially on a threshold ϵ_d ,

$$\begin{aligned} \text{desired speech is present} & \quad \text{if } (e_A/e_B)_{rs} > \epsilon_d, \\ \text{desired speech may be absent} & \quad \text{if } (e_A/e_B)_{rs} \leq \epsilon_d. \end{aligned} \quad (11)$$

Thus, when $(e_A/e_B)_{rs} > \epsilon_d$, the present frame is regarded as containing the desired speech. Otherwise, the desired speaker may or may not be silent in the present frame. In other

words, the condition $(e_A/e_B)_{rs} \leq \epsilon_d$ is a necessary but not sufficient condition for determining the absence of the desired speech. This is because the relative measure $(e_A/e_B)_{rs}$ is insensitive to the presence of the desired signal when its energy relative to the noise signals is rather low. As discussed in Sec. I C 1, the final determination in this case is made based on an additional investigation into the HF/LF/S state for the preceding frame, or the existence of the pitch in the present frame, possibly belonging to the desired speaker.

It is easy to see that the effectiveness of this desired-voice-activity-detection method largely depends on the extent to which the beamforming suppresses the noise signals. The larger the amount of noise signals is suppressed by the beamforming, the more effective is the measure $(e_A/e_B)_{rs}$ in detecting the desired voice activity. However, the suppression of the noise signals by deterministic beamformers in not highly reverberant conditions depends, in turn, on the location of the noise signals with respect to the microphone array. Consequently, the proposed method is expected to work better when the location of the noise signals is not within the range of the mainlobe of the beampattern.

To accommodate for the location changes of the noise sources, the reference $(e_A/e_B)_s$ is also updated, similar to $(e_H/e_L)_v$.

3. Pitch determination

Pitch determination, an essential component in a variety of speech processing systems, has been extensively studied and a number of effective algorithms have been suggested (Rabiner *et al.*, 1976; Hess, 1983). The problem gets more complicated when background noise is involved, especially when the noise is due to competing speech from other speakers. Much effort has been directed towards separating the voices of two concurrent speakers by means of their different pitches (Parsons, 1976; Darwin, 1981; Gardner *et al.*, 1989; Assmann and Summerfield, 1990; Stubbs and Summerfield, 1991; Meddis and Hewitt, 1992; De Cheveigné, 1993). Nevertheless, pitch determination in a multiple-speech-interference environment remains elusive.

Due to the adverse interference, the pitch determination algorithm for the "cocktail party" environment is required to be rather robust. Following the classification of Hess (1983), two classes of methods are known to be more robust than others among the available algorithms. One class is the *spectral compression*, including the harmonic product spectrum (Schroeder, 1968) and the harmonic sum spectrum (Noll, 1970). The other class is the *harmonic matching* (Martin, 1982; Paliwal and Rao, 1983; Duifhuis *et al.*, 1982; Terhardt, 1979; Sreenivas and Rao, 1981). Both classes exploit the harmonic structure in the spectrum domain (either amplitude spectrum or power spectrum), and are, in effect, similar. Specifically, they coherently accumulate the harmonics in a series, thus magnifying the salience of the pitch information in the spectrum. Note that these two classes of methods are considered to be robust for a signal contaminated by noise or by one concurrent speech, as well as for a signal with some of its harmonics missing, which is not a rare occurrence in speech and music.

It is also noteworthy that these robust methods are advantageous for signals with contaminated parts in their spectrum, e.g., when speech signals are disturbed by background sound or the lower harmonics are masked. This is especially true in our system, since the noise components of higher frequencies are removed by the beamforming more than the lower frequencies, even in the LF channel. However, our computer simulation demonstrates that the above two classes of methods cannot be used directly for pitch determination of speech contaminated by multiple speech interference. To overcome this limitation, we propose here a modified procedure for pitch determination which is based on the above methods.

a. Modified harmonic sum spectrum. The pitch determination is made based on part of the spectral components in the LF channel, say frequencies up to $500Q$ (Hz), i.e., $|Y_b(m)|$ ($m = 0, \dots, 500QN/f_s$), where Q is a positive integer (described below). The resolution of the pitch determination is increased by using interpolation on $Y_b(m)$ and the interpolated $|Y_b(m)|$ is denoted as $|Y_b(m')|$ ($m' = 0, \dots, 500pQN/f_s$), where p is the interpolation factor. The corresponding periodogram¹ is calculated by

$$P_{yb}(m') = |Y_b(m')|^2, \quad m' = 0, \dots, 500pQN/f_s. \quad (12)$$

The formant influence on pitch determination is removed by detecting the envelope of $P_{yb}(m')$ by selecting a piecewise-linear envelope $EP_{yb}(m')$, i.e., a chain of linear segments connecting the local maxima in $P_{yb}(m')$. Specifically,

$EP_{yb}(m')$

$$= \begin{cases} P_{yb}(m'), & m' = 0 \text{ or } 500pQN/f_s, \\ \frac{P_{yb}(M_j) - P_{yb}(M_{j-1})}{M_j - M_{j-1}} \cdot (m' - M_{j-1}) \\ \quad + P_{yb}(M_{j-1}), & M_{j-1} \leq m' < M_j, \end{cases} \quad (13)$$

while the locations of the local maxima are detected by checking for the points M_j ($j = 1, \dots, J$) which satisfy

$$[P_{yb}(M_j - 1) < P_{yb}(M_j)] \cap [P_{yb}(M_j + 1) < P_{yb}(M_j)]. \quad (14)$$

Finally, the formants are removed by

$$P'_{yb}(m') = \frac{P_{yb}(m')}{EP_{yb}(m')}, \quad m' = 0, \dots, 500pQN/f_s. \quad (15)$$

This procedure can similarly be performed on the interpolated spectrum $X(m')$ ($m' = 0, \dots, 500pQN/f_s$) of the signal $X(m)$ ($m = 0, \dots, 500QN/f_s$) before beamforming, i.e., on one of the channel signals shown in Fig. 1(A); its periodogram, and the periodogram after formant removing, are respectively denoted as $P_x(m')$ and $P'_x(m')$. Thus,

$$P_x(m') = |X(m')|^2, \quad (16)$$

and

$$P'_x(m') = \frac{P_x(m')}{EP_x(m')}, \quad m' = 0, \dots, 500pQN/f_s. \quad (17)$$

The modified harmonic sum spectra $\sigma_{yb}(m'')$ and $\sigma_x(m'')$ are derived based on $P'_{yb}(m')$ and $P'_x(m')$, respectively,

$$\sigma_{yb}(m'') = \sum_{q=1}^Q P'_{yb}(qm''), \quad m'' = 0, \dots, 500pN/f_s, \quad (18)$$

and

$$\sigma_x(m'') = \sum_{q=1}^Q P'_x(qm''), \quad m'' = 0, \dots, 500pN/f_s, \quad (19)$$

where Q is the number of spectra involved in the coherent summation. Since the fundamental pitch frequencies of human speech are lower than 500 Hz, the harmonic sum spectra are defined in the range of 0–500 Hz. Since only the LF components, which are strong in representing pitch harmonics, are adopted for formulating the harmonic sum spectrum, the value of Q can be 6, at most.

b. Pitch determination. Depending on the SNR level in the present frame, two different pitch determination procedures can be developed. One, when the SNR is high and there is little difference between $\sigma_{yb}(m'')$ and $\sigma_x(m'')$; the pitch of the desired speech can be determined directly based on $\sigma_{yb}(m'')$. In the second procedure, the pitch of the desired speech is determined by comparison between $\sigma_{yb}(m'')$ and $\sigma_x(m'')$. To distinguish between these two situations so as to utilize the appropriate method, we define the ratio γ of the mean values $\overline{EP_{yb}}$ and $\overline{EP_x}$ of the two corresponding envelopes $EP_{yb}(m')$ and $EP_x(m')$. Thus,

$$\gamma = \frac{\overline{EP_{yb}}}{\overline{EP_x}}, \quad (20)$$

where

$$\overline{EP_{yb}} = \frac{1}{500pQN/f_s + 1} \sum_{m'=0}^{500pQN/f_s} EP_{yb}(m') \quad (21)$$

and

$$\overline{EP_x} = \frac{1}{500pQN/f_s + 1} \sum_{m'=0}^{500pQN/f_s} EP_x(m'). \quad (22)$$

The argument for choosing γ as the indicator for the SNR level in a frame is that when SNR is high, both $P_{yb}(m')$ and $P_x(m')$ are dominated by the desired signal, and there is no large difference between them. Otherwise, the difference between $P_{yb}(m')$ and $P_x(m')$ is large because of the noise reduction by the beamforming.

A threshold Γ is empirically determined so that when $\gamma \geq \Gamma$, the desired pitch can be detected on $\sigma_{yb}(m'')$ alone; this is denoted as “pitch determination method #1” in Fig. 3. Otherwise, the desired pitch has to be detected based on the comparison between $\sigma_{yb}(m'')$ and $\sigma_x(m'')$; this is labeled as “pitch determination method #2” in Fig. 3.

Pitch determination by method #1 is done by locating the peaks coherently added in the harmonic spectrum $\sigma_{yb}(m'')$. If the fundamental frequency $F_0 \leq 250$ Hz, there will be more than one harmonic peak present in $\sigma_{yb}(m'')$. Clearly, the first of these harmonic peaks is situated at the frequency corresponding to F_0 . However, care should be taken to avoid confusion with the low-frequency noise, especially the compressed spectral harmonic peaks incurred in

the harmonic sum spectrum. While this problem is not serious, more attention should be given to it when the SNR decreases.

A harmonic difference-sum spectrum $\delta(m'')$ is defined for the situation when the SNR is sufficiently low so that the harmonic sum spectrum comparison between the signals after and before beamforming satisfies $\gamma < \Gamma$. Thus,

$$\delta(m'') = \sigma_{yb}(m'') - \sigma_x(m''). \quad (23)$$

Depending on the number of concurrent voiced speech sources, several series of harmonic peaks may be present in $\delta(m'')$. Since the desired speech passes through the beamforming without attenuation, the highest harmonic series is ideally that of the desired speech. However, when the number of concurrent voiced speech sources is large, the harmonics of the desired speech in $P_x(m)$ are seriously overlapped by the spectral components of the other speakers and the beamforming may probably also result in an absolute decrease in the spectrum at the harmonic locations of the desired speech. Therefore, while it is uncertain that every harmonic peak of the desired speech in $\delta(m'')$ is indeed always higher than the peaks belonging to the competing speech, the desired speech in $\delta(m'')$ is *often* the highest in terms of the sum of peak values in a harmonic series.

To ensure that the correct harmonic peaks are chosen in $\sigma_{yb}(m'')$ or $\delta(m'')$ in the presence of many other harmonic peaks, we tentatively exploit a statistical fact in speech production: The fundamental frequency F_0 for a conversational speech waveform varies continuously and slowly in time, and the variation rate is less than 10 Hz (Saito and Nakata, 1985). Thus, the F_0 that is the closest to the F_0 of the desired speech in the preceding frame is chosen from the several F_0 candidates in $\sigma_{yb}(m'')$ or $\delta(m'')$. The risk involved in using this method is that once the initial F_0 is inaccurately determined, it would wrongly affect the pitch determination for the following frames. This problem is frequently encountered when using pitch determination method #2 on $\delta(m'')$, i.e., when SNR is low. Much less uncertainty exists in determining F_0 when the SNR is high, and pitch determination method #1 on $\sigma_{yb}(m'')$ is used.

4. Comb filter design

The comb filter is designed according to the fundamental frequency F_0 of the desired speech. A transfer function $C(m')$ ($m' = 0, \dots, 3000pN/f_s$) of the comb filter is designed to pass through the harmonic peaks and their mainlobes while making a suppression for the spectral components between the comb teeth. The sinewave-shaped peaks correspond to the harmonics of the desired speech; the peak value equals one, and the valley value equals zero. Thus,

$$C(m') = 0.5 \cos\left(\frac{2\pi m' f_s}{F_0 N}\right) + 0.5, \quad 0 \leq m' < 3000pN/f_s. \quad (24)$$

It is important to retain the dc value for each frame when doing the comb filtering, since the mean value of the signal in each frame does not necessarily equal zero; losing the dc value will result in distortion when doing speech synthesis.

The comb filtering is performed in the amplitude spectrum domain by

$$|Y(m')| = |Y_b(m')| \cdot C(m'), \quad m' = 0, \dots, 3000pN/f_s. \quad (25)$$

The filtered $|Y(m')|$ is then decimated back to the original frequency resolution $|Y(m)|$ ($m = 0, \dots, 3000N/f_s$) so as to be ready for the inverse Fourier transform.

5. Deep attenuation

When the present frame is determined to be either unvoiced or silence, the switch SW in Fig. 1(B) is turned to position ‘‘1’’ and a deep attenuation is applied in the LF channel

$$Y(m) = B \cdot Y_b(m), \quad m = 0, \dots, 3000N/f_s, \quad (26)$$

where B is the attenuation factor.

D. Speech synthesis

As illustrated in Fig. 1(B), the HF part of the beamforming output $Y_b(m)$ may be amplified, i.e.,

$$Y(m) = A \cdot Y_b(m), \quad m = 3000N/f_s + 1, \dots, N/2. \quad (27)$$

This optional amplification, similar to a prewhitening operation, actually increases the consonant/vowel intensity ratio which is known to help improve the intelligibility (Hecker, 1974; Gordan-Salant, 1986; Montgomery and Edge, 1988; Preves *et al.*, 1991).

It is interesting to note that the underlying assumption used here for comb filtering in the spectrum domain is identical to the one used by Boll (1979) for spectral subtraction, i.e., it is principally the short-time spectral amplitude, rather than phase, that is important for speech intelligibility and quality. Therefore, the LF output $Y(m)$ ($m = 0, \dots, 3000N/f_s$) can be estimated with the amplitude spectrum $|Y(m)|$ after comb filtering and the phase spectrum $\angle Y_b(m)$ before comb filtering,

$$Y(m) = |Y(m)| \cdot \exp[j\angle Y_b(m)], \quad m = 0, \dots, 3000N/f_s. \quad (28)$$

Combining Eqs. (27) and (28) and using the symmetry property of DFT yields a complete spectrum for the present frame, $Y(m)$ ($m = 0, \dots, N-1$),

$$Y(m) = \begin{cases} |Y(m)| \cdot \exp[j\angle Y_b(m)], & m = 0, \dots, 3000N/f_s, \\ A Y_b(m), & m = 3000N/f_s + 1, \dots, N/2, \\ [Y(N-m)]^*, & m = N/2 + 1, \dots, N-1, \end{cases} \quad (29)$$

where the asterisk (*) denotes complex conjugation. The spectrum is then converted to the output in the time domain $y(k)$ by using IFFT:

$$y(k) = F^{-1}\{Y(m)\}. \quad (30)$$

E. Implementation of the targeting-and-extracting system

In order to achieve a high precision of phase delay for the broadband superdirective array beamforming, the beamforming procedure in Eq. (1) and the IFFT in Eq. (30) may

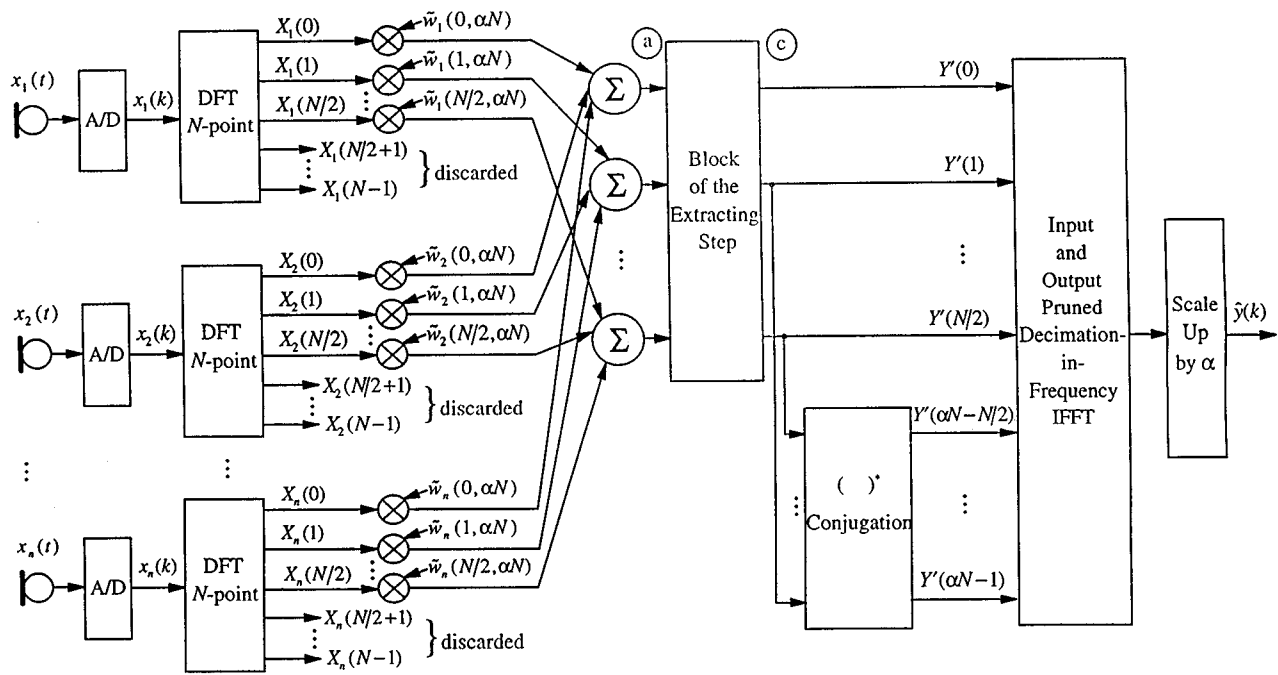


FIG. 4. Block diagram of the targeting-and-extracting system realized by adopting the digital frequency-domain optimum broadband beamforming technique.

be realized by the method of digital frequency-domain implementation of optimum broadband beamformer (Liu and Sideman, 1995, 1996). The extracting stage is inserted between the channel summation in the frequency domain and the IFFT operation. In this case, all the calculations described in Secs. I C and I D can be adopted, except that Eq. (29) should be replaced by

$$Y'(m) = \begin{cases} |Y'(m)| \cdot \exp[j\angle Y_b(m)], & m=0, \dots, 3000N/f_s, \\ AY_b(m), & m=3000N/f_s+1, \dots, N/2, \\ 0, & m=N/2+1, \dots, \alpha N - N/2 - 1, \\ [Y'(\alpha N - m)]^*, & m=\alpha N - N/2, \dots, \alpha N - 1, \end{cases} \quad (31)$$

where $Y'(m)$ ($m = 0, \dots, 3000N/f_s$) denote either the decimated version of the comb filtering output [Eq. (25)] or the output of the deep attenuation [Eq. (26)], and α is the interpolation factor defined in the above references. The block diagram of the targeting-and-extracting system realized by the digital frequency-domain optimum broadband beamforming technique is shown in Fig. 4; the block of the extracting step is detailed in Fig. 1(B). The symbols (a) and (c) in Fig. 4 represent the input and output ports of the extracting step, respectively (cf. Fig. 1). For simplicity, the LF channel before the beamforming, which is marked by the symbol (b) in Fig. 1, is omitted in Fig. 4.

II. EXPERIMENT

A. Implementation method

The effectiveness of the proposed targeting-and-extracting system is demonstrated by sample simulations for five different input SNR conditions: -10 dB, -5 dB, 0 dB,

+5 dB, and +10 dB. The targeting part was a 10-cm-long linear uniform endfire five-omnidirectional-microphone array. The simulation was made in an anechoic free-field environment and the effect of head shadowing was not considered.

The target source material contained two sentences which included various kinds of voiced and unvoiced phonemes, "Mr. Smith might discuss the bill," and "We are considering the cheers," spoken with natural intonation by a native American male. The interference source material was a 12-talker SPIN babble (Bilger *et al.*, 1984). Different sections of the babble recording were used as noise for each of the above signals. In the following experiments, we always used the same section of the babble recording for the corresponding sentence. The target source was assumed to be located directly in front of the listener; the interference source was 45° in azimuth.

The signals were low-pass filtered at 5 kHz and sampled at a 12.5-kHz rate with 12-bit amplitude quantization. In the short-term spectral analysis, a 30-ms segment of signal was weighted by a 30-ms Hamming window and then padded with zeros to 1024 points for FFT. The 30-ms-long segment generally corresponds to several pitch periods, but is much shorter than $\frac{1}{8}$ s, the approximate average duration of individual speech sounds or phonemes. The 1024-point FFT resulted in a 12-Hz frequency resolution.

The application of the superdirective beamforming is described elsewhere (Liu and Sideman, 1996). The beamforming output was fed into the extracting part for further noise cancellation before it was converted into the signal in the time domain.

In the extracting part, the HF/LF/S decision was first performed. The two thresholds for voiced/unvoiced/mixed determination, ϵ_v in Eq. (5) and ϵ_s in Eq. (7), were empirically determined to be 5 and 18, respectively. The threshold

ϵ_d used in the desired voice activity detection was empirically determined to be 1.01. The pitch determination was applied when an LF frame, or a transition frame changing to LF, was detected. The parameter Q for formulating the harmonic sum spectra in Eqs. (18) and (19) was 5. The threshold Γ for choosing between the two pitch determination methods was empirically set to be 0.8. In the attenuation element, the signal amplitude spectrum in the LF channel was attenuated by about 14 dB [see parameter B in Eq. (26)]. For the results to be comparable, the amplification [parameter A in Eq. (27)] in the HF channel was set to be unity.

The evaluation was performed by simulation in the intelligibility-weighted measure as well as by informal listening. We adopted the intelligibility-weighted measure that was first defined in Peterson (1989) based on the articulation index, and given in details in Liu and Sideman (1996). It includes $\Gamma(T_i)$, $\Gamma(T_o)$, $\Gamma(I_i)$, and $\Gamma(I_o)$, where T and I denote target and interference, while the subscripts i and o denote input and output, respectively. The system improvement for the target desired signal and the interfering noise is respectively given by

$$\Delta\Gamma(T) = \Gamma(T_o) - \Gamma(T_i) \quad (32)$$

and

$$\Delta\Gamma(I) = \Gamma(I_i) - \Gamma(I_o). \quad (33)$$

The overall intelligibility-weighted gain, G_I , that the system provides to suppress the interference noise against the desired speech is the sum of these two measures, i.e.,

$$G_I = \Delta\Gamma(T) + \Delta\Gamma(I). \quad (34)$$

The separate effects of the targeting-and-extracting operation on the desired signal and the environmental noise are determined by adding two processing attachments to the simulation system. The basic system then processes the combined desired signal and noise, while the two added attachments do the same processing separately on the desired signal and noise. The measures $\Gamma(T_i)$, $\Gamma(T_o)$, $\Gamma(I_i)$, and $\Gamma(I_o)$ can now be calculated, since all the processing operations in this system are linear.

B. Results and discussion

The performance of the targeting part was evaluated and reported earlier (Liu and Sideman, 1996). When the well-known sensitivity-constrained superdirective beamforming weights are applied to the array, and when the weights are designed to provide a white noise gain of -11.6 dB, a noise reduction of approximately 7.0 dB and 7.1 dB in the intelligibility-weighted measure can be obtained for the two sample babble sentences used here, respectively. Since beamforming is a linear operation, both the improvement in intelligibility and the improvement in quality are constant for the same signal, regardless of its SNR value. In fact, since the desired speech, which is the actual target of the array, is received without any change, the overall improvement in intelligibility and quality at the targeting stage depends solely on the ability of the array to reduce the noise. These beam-

forming results were fed into the extracting stage, allowing us to evaluate the noise reduction provided by the extracting process.

1. Evaluation of the determination algorithms

In order to evaluate the performance of each determination algorithm, including the HF/LF/S decision (i.e., both the voiced/unvoiced determination and desired voice activity detection) and the pitch determination, the HF/LF/S state and the F_0 corresponding to the dominant voiced phoneme in each frame were determined manually, in advance, for the two sample sentences. Since a small number of frames were located at the transition regions between the HF/LF/S states, or between two LFs of different F_0 values, it was difficult to decide which state they belonged to. These irregular frames were not counted in the following evaluation of the performance of each determination algorithm. The statistical results of the manual determination of the HF/LF/S states are shown in Tables II and III.

a. HF/LF/S determination. The results of the HF/LF/S determination, including the voiced/unvoiced determination and the desired voice activity detection, for the two sample sentences are presented in Tables IV and V, and compared with the results of the manual determination given in Tables II and III. Tables IV and V show the results of the HF/LF/S determination for the five different system input SNR conditions. The results are presented in the form of the number of frames whose HF/LF/S state, as determined by the determination algorithms, is identical with, or different from, that of the manual determination.

As shown in Tables IV and V, unvoiced frames are mistaken for “voiced” when the system input SNR is low. We ascribe this error to the effect of the HFs of the noise at the beamforming output: Although suppressed by the beamforming, the residual HF noise is still comparable in energy to the HFs of the desired signal for very low SNR situations; hence, the value of the reference $(e_H/e_L)_v$ increases. On the other hand, when the system input SNR is high, the energy of HFs of the noise decreases. Hence, $(e_H/e_L)_v$ decreases as SNR is high, leading to more voiced frames being mistaken for “unvoiced.”

The results in Tables IV and V also show that when the system input SNR is low, more nonsilence (including both the voiced and unvoiced) frames are mistaken for “silence.” It turned out that in these frames the energy of the desired speech was too weak relative to the noise, i.e., it was buried in the noise and failed to be detected. On the other hand, most of the silence frames are correctly determined by the algorithm, regardless of the system input SNR, except that a few manually determined silence frames are determined as “voiced” when the system input SNR is very high. A visual examination of these frames revealed that they contained a very small amount of starting or ending parts of phonemes, which were too weak to be perceived by hearing. However, when input SNR is very high, the energy of the noise at the beamforming output is lower than the energy of these weak parts, thus leading to a nonsilence detection. Nevertheless, the silence frames are rarely determined as “unvoiced,” since the HFs of the noise signals are suppressed by the

TABLE II. Statistics of manual determination of the HF/LF/S state for the sentence “*Mr. Smith might discuss the bill.*”

Frame nature		Number of frames
Nonsilence	Voiced	130
	Unvoiced	55
Silence		41
Irregular		36
Total		262

beamforming to a much greater extent than the LFs, which result in a very low value of (e_H/e_L) during the silence of the desired speaker.

Obviously, no harm is incurred to the unvoiced phoneme (HFs) when unvoiced frames are mistaken for “voiced,” because the components in the HF channel are not subject to any filtering or suppression; rather, they are always transferred intact to the system output. On the other hand, the LFs in these frames, which are from the noise sources and hence could have been largely suppressed by the attenuation element, can partially leak to the system output by way of the comb filtering (Fig. 1). In this case, a relatively dominant pitch among the noise signals will often be detected by the pitch determination algorithm. Consequently, the LFs of the noise signal with this pitch will be admitted by the comb filter.

Since the extracting process carries the HFs to the system output and makes a deep attenuation of the LFs for the unvoiced and silence frames, there is no penalty for either mistaking the unvoiced frames for “silence” or vice versa.

When the silence frames are determined as voiced, the adverse effect is that the noise signals are not removed by the attenuator as they should have been. Rather, a relatively dominant pitch among the noise signals will most often be detected by the pitch determination algorithm and consequently, the comb filtering will admit the noise components with this pitch (relative to the other noise components).

The worst mistake occurs when the voiced frames are mistaken for either unvoiced or silence. Generally, this happens to frames with weak voiced phonemes. When this mistake occurs, the desired signal in those frames is canceled. If we assume that the desired speaker is not speaking in a rather low voice, then those voiced frames containing weak vowels are generally located at the beginning or end of a strong vowel. This mistake sometimes also occurs to the frames containing weak voiced consonants which are quite short in duration with respect to the frame length. There seems to be no good way to compensate for the LF loss incurred by this mistake. For a listener with normal hearing, this LF loss may

TABLE III. Statistics of manual determination of the HF/LF/S state for the sentence “*We are considering the cheers.*”

Frame nature		Number of frames
Nonsilence	Voiced	121
	Unvoiced	33
Silence		10
Irregular		35
Total		199

TABLE IV. Results of the HF/LF/S determination algorithm, including voiced/unvoiced determination and the desired voice activity detection, under five different input SNR conditions, for the sentence “*Mr. Smith might discuss the bill.*” Comparison with results of manual determination. Results are presented in the number of frames.

System input SNR (dB)	Manual determination	HF/LF/S determination algorithm			
		Voiced	Unvoiced	Silence	Total
-10	Voiced	98	0	32	130
	Unvoiced	15	31	9	55
	Silence	4	0	37	41
-5	Voiced	119	0	11	130
	Unvoiced	6	43	6	55
	Silence	5	0	36	41
0	Voiced	117	4	9	130
	Unvoiced	3	51	1	55
	Silence	4	0	37	41
+5	Voiced	120	8	2	130
	Unvoiced	1	53	1	55
	Silence	6	1	34	41
+10	Voiced	126	4	0	130
	Unvoiced	2	53	0	55
	Silence	11	0	30	41

be compensated for semantically, by the speech context. Informal listening to the simulation results suggests that this is quite feasible. Moreover, the listeners with normal hearing do not find it difficult, or uncomfortable, when listening to it. However, the influence of this loss on the intelligibility of the hearing impaired may be quite serious and needs to be further studied. Clearly, one way to reduce this kind of mistake now is to raise the thresholds ϵ_v and ϵ_s and lower the threshold ϵ_d , at the price of increasing the errors of mistaking the unvoiced or silence frames for “voiced.”

TABLE V. Results of the HF/LF/S determination algorithm, including voiced/unvoiced determination and the desired voice activity detection, under five different input SNR conditions, for the sentence “*We are considering the cheers.*” Comparison with results of manual determination. Results are presented in the number of frames.

System input SNR (dB)	Manual determination	HF/LF/S determination algorithm			
		Voiced	Unvoiced	Silence	Total
-10	Voiced	97	1	23	121
	Unvoiced	8	2	23	33
	Silence	0	0	10	10
-5	Voiced	106	5	10	121
	Unvoiced	4	14	15	33
	Silence	0	0	10	10
0	Voiced	114	7	0	121
	Unvoiced	0	30	3	33
	Silence	0	0	10	10
+5	Voiced	109	12	0	121
	Unvoiced	0	33	0	33
	Silence	0	0	10	10
+10	Voiced	101	20	0	121
	Unvoiced	0	33	0	33
	Silence	0	1	9	10

TABLE VI. Results of the two pitch determination algorithms under five different system input SNR conditions. The results are presented as the fraction of the number of frames correctly determined out of the number of frames input to the algorithm. As the standard of judgment, we use the results of manual pitch determination for each frame.

Pitch determination method	System input SNR (dB)				
	-10	-5	0	+5	+10
<i>“Mr. Smith might discuss the bill.”</i>					
Method #1	3/8	18/27	53/58	71/79	86/104
(%)	(37.5)	(66.67)	(91.38)	(89.87)	(82.69)
Method #2	34/90	32/92	24/59	10/41	9/22
(%)	(37.78)	(34.78)	(40.68)	(24.39)	(40.91)
Total, correctly determined	37/98	50/119	77/117	81/120	95/126
(%)	(37.76)	(42.02)	(65.81)	(67.5)	(75.40)
<i>“We are considering the cheers.”</i>					
Method #1	1/2	24/29	61/70	81/97	90/99
(%)	(50)	(82.76)	(87.14)	(83.51)	(90.91)
Method #1	24/95	22/77	17/44	5/12	1/2
(%)	(25.26)	(28.57)	(38.64)	(41.67)	(50)
Total, correctly determined	25/97	46/106	78/114	86/109	91/101
(%)	(25.77)	(43.40)	(68.42)	(78.90)	(90.10)

b. Pitch determination. Table VI shows the results of the two pitch determination algorithms under five different system input SNR conditions, for the two sample sentences. Choosing the proper pitch determination algorithm for a specific frame depends on the value of γ in Eq. (20), with respect to the threshold Γ (Sec. I C 3 b). The results are listed in the form of the number of frames correctly determined by each pitch determination algorithm, out of the total number of the frames which invoked that algorithm. The number of all the frames correctly determined by the two algorithms is also given with respect to the total number of voiced frames detected in the sentence. Judging the determination to be either correct or incorrect is based on the results of our manual pitch determination. However, the pitch period is variant even within the same frame since the frame is rather long (30 ms in our simulation) and can contain more than one pitch period. Therefore, the result of the pitch determination for a frame is regarded to be correct if it is within ± 5 Hz of the manual determination result.

As demonstrated by the results presented in Table VI, higher system input SNR leads to more frames to be processed by the pitch determination method #1. Conversely, a lower system input SNR leads to more frames to be processed by the pitch determination method #2. This result proves the effectiveness of the parameter γ in distinguishing between low and high SNR frames. Generally, pitch determination method #1 performs with a high correction rate, whereas pitch determination method #2 does not work so satisfactorily. As can be seen, the higher the system input SNR, the better the performance of the two pitch determination algorithms as a whole. In our simulation, up to 70% correction rate could be attained for 0-dB system input SNR. Clearly, the performance of each pitch determination algorithm also varies for different desired and noise signals.

A careful inspection of the results of the pitch determination algorithms against the manual pitch determination reveals that most of the incorrectly determined frames hap-

pened at the beginning or the end of the voiced phonemes, or at the transitions between two successive phonemes, i.e., in the irregular voiced frames. Most often, the signals in these frames are relatively weak with respect to the major part of the voiced phonemes. A simulation test showed that losing all the LF components in these frames results in only a 0.1–0.2 dB loss of the desired speech in the intelligibility-weighted measure for each sample sentence. Some other incorrect determinations happened to low F_0 voiced phonemes, e.g., the semivowel /l/ in the first sample sentence. The frame length of 30 ms is capable of containing two, or less than two, complete pitch periods for $F_0 < 70$ Hz, while it largely limits the expression of the harmonics of a lower pitch in the Fourier spectrum. One solution to this problem is to increase the frame length, but this is limited by the requirement on the quasistability of the signal within a frame. Another possible solution is to have the frame length increased adaptively when the detected F_0 decreases.

2. Intelligibility improvement in the extracting stage

In order to study the potential of the extracting procedure for speech enhancement, and establish a reference for comparison, we utilize the results of the manual determination to control the signal processing for each frame. However, there are several options for treating the unvoiced frames, the silence frames, and the irregular frames. Option #1 is to zero, or infinitely attenuate, all the LFs for the above three kinds of frames. This results in the largest cancellation of noise and the largest overall intelligibility-weighted gain. Thus, option #1 can approximately provide maximal potential of the extracting procedure for noise cancellation and intelligibility improvement. Option #2 is to attenuate all the LFs amplitude by 14 dB for the above three categories of frames. This is similar to our simulation with the two sample sentences, except that our algorithm had no “irregular” state; instead, those manually determined irregular frames may normally be determined as “voiced.” Therefore, option #2 may be used as a standard for judging our simulation results.

Informal listening demonstrates that there is almost no perceived loss of the desired signal when applying option #1 to the signal with 0-dB system input SNR; there is almost no residual noise being heard, and there is almost no perceived difference between the processing output and the original clean signal. In comparison, option #2 gives almost no perceived loss of the desired signal and the residual noise can only be heard with special attention and effort.

Table VII shows the intelligibility-weighted improvement in the desired signal and the noise, and the overall intelligibility-weighted improvement, attained by the extracting stage. These results are obtained either through the signal processing controlled by the results of manual determination (with treatments defined by options #1 and #2), or from the determination algorithms under the five different input SNR conditions. Since all the processing operations are linear, the intelligibility-weighted improvement results obtained for the given two options are the same for different SNR conditions. In other words, the SNR can only affect the performances of the determination algorithms.

TABLE VII. Intelligibility-weighted improvement attained at the extracting stage in the desired signal and noise, and the overall intelligibility-weighted improvement at the extracting stage, respectively, obtained via signal processing controlled by the results either from the manual HF/LF/S state and F_0 determination, or from the determination algorithms, under five different system input SNR conditions. Results for the sentences, “*Mr. Smith might discuss the bill,*” and “*We are considering the cheers.*” Option #1—The undesired LFs are infinitely attenuated; Option #2—The undesired LFs are attenuated by 14 dB.

Intelligibility-weighted improvement (dB)	System input SNR (dB)						
	Option #1	Option #2	-10	-5	0	+5	+10
“Mr. Smith might discuss the bill.”							
Desired signal	-2.7	-2.7	-4.0	-3.6	-2.7	-2.6	-2.5
Noise	7.5	6.7	7.3	6.6	6.5	6.2	5.9
Overall	4.8	4.0	3.3	3.0	3.8	3.6	3.4
“We are considering the cheers.”							
Desired signal	-2.0	-2.0	-4.0	-3.0	-2.2	-2.0	-1.8
Noise	6.0	5.6	6.5	6.0	5.2	5.0	4.8
Overall	4.0	3.6	2.5	3.0	3.0	3.0	3.0

It can be seen that even when all the HF/LF/S states and F_0 determinations are sufficiently accurate as in options #1 and #2, there is still a small amount of loss in the desired speech. This is due to several reasons: (1) Since the signal in each frame is time variant, there is a deviation in its spectrum from the exact harmonic structure, and the deviated parts will be largely reduced by the comb filtering; (2) The non-deviating harmonic peaks will be sharpened by the comb filtering, with a consequent loss of part of the components; (3) The window sidelobes will also suffer a reduction from comb filtering. Fortunately, the amount of loss is very small, e.g., 2.7 dB for the first sample sentence and 2.0 dB for the second sample sentence, and informal listening demonstrated that the loss in the desired speech was imperceptible.

It should be noted that the amount of noise that is removed by comb filtering depends on its spectral distribution. Specifically, the noise components in the valley areas between the harmonic peaks of the desired speech are reduced, whereas the noise components located close to the harmonic peaks are subject to only little cancellation. Consequently, there may be situations where more signal than noise is removed by comb filtering. However, both the losses in the signal and noise in those cases are quite small and could hardly be perceived.

In contrast, LF attenuation for the unvoiced or silence frames results in a significant amount of noise cancellation. For example, the amount of noise cancellation for both the unvoiced and silence frames due to the LF attenuation is up to 5 dB for the first sample sentence and up to 3.5 dB for the second sample sentence, accounting for 60%–70% of the total noise cancellation.

Table VII shows that approximately 6–7 dB of noise cancellation is achieved at the cost of 2–3 dB of the desired speech cancellation. The lower the SNR, the larger the cancellation of the desired signal. This is because the correction rate of the determination algorithms decreases with SNR. Tables IV and V reveal that more voiced frames were mistaken for “silence,” and more unvoiced frames were mis-

taken for “voiced.” Table VI shows that under these conditions, the correction rate of the pitch determination, either by the method #1 or by method #2, also decreases. Moreover, more frames are to be determined by method #2, which generally has a much lower correction rate than method #1. On the other hand, when more voiced frames are mistaken for “silence” as the SNR lowers, larger amount of noise is also removed. Consequently, both the signal loss and noise cancellation increase for lower SNR and, in the range between -10~+10 dB of system input SNR as demonstrated by Table VII, the overall intelligibility gain remains relatively stable. A 3–4 dB net intelligibility improvement is usually obtained in our simulation.

It is important to note that the overall intelligibility improvements listed in Table VII for options #1 and #2 are not the largest possible values that can be achieved in terms of the intelligibility-weighted measure. This is because cancellation at frequencies where the noise component is larger than that of the desired signal will always lead to an additional increase in these values. However, cancellation under these circumstances will also lead to a loss in the desired signal, which is obviously undesired. Thus, pursuing a high value in the overall intelligibility-weight gain is not our goal. Rather, we pursue the highest possible value of this parameter on the premise that the loss in the desired signal is the least.

Informal listening agrees with this conclusion. A larger amount of noise cancellation can be perceived when the system input SNR is lowered, and the improvement in the speech quality is obvious. On the other hand, the loss of the desired signal increases as well.

III. SUMMARY AND CONCLUSION

Following our earlier work on the superdirective fixed-beamforming preprocessor for hearing aids (Liu and Sideman, 1996), the present study complements this directional targeting approach by adding an extracting step to the beamforming output so as to enhance hearing aids adaptability in the “cocktail party” environment. Consequently, we propose a combined targeting-and-extracting procedure: The fixed-beamforming approach is supplemented by noise cancellation by comb filtering or attenuation. This cancellation step is needed due to limitations in the array dimensions whereby the effectiveness of the beamforming in spatial noise reduction is largely reduced when applied to hearing aids. Consistent with the array size, it is assumed that the HFs (≥ 3 kHz) of the signal have become sufficiently clean of noise at the beamforming stage, hence the extracting step should provide additional cleaning only for the LF (< 3 kHz) range. Based on the determination of the voiced/unvoiced/silence state and the fundamental pitch frequency of the desired signal, either comb filtering or attenuation is applied.

The capability of the post-targeting extracting procedure to enhance noise cancellation has been clearly confirmed by our simulation experiments. Although a small amount of the desired signal is also inevitably canceled by this procedure, this loss is relatively small and practically insignificant. In contrast, cancellation of the noise by this method clearly im-

proves the overall intelligibility. Our experiments demonstrate 6–7 dB of further noise cancellation and 3–4 dB of further overall intelligibility improvement.

It is noteworthy that the extraction procedure which is applied here to supplement the fixed-beamforming approach can also be applied to the adaptive beamformers. However, adopting the fixed beamformers may be significant when a multiple speech interference is present, since the adaptive beamformer is considered to be much less effective for noise cancellation when the number of noise sources is larger than the number of microphones (Greenberg and Zurek, 1992).

Since this is a theoretical analysis aimed to give indications and possibly guidelines to future work, obviously, much more work is needed to improve the procedure suggested here. For example, we need to know how to recognize and apply noise cancellation for irregular frames, and how to determine and cancel noise in speech with jitters and shimmers. In addition, the present algorithms for voiced/unvoiced/silence decision and pitch determination are obviously imperfect, the correction rate is not very high, and work is needed to improve these limitations. It is reasonable that recognition of other speech features of the speaker may help improve the correction rate of the determination of the speaker's voicing state as well as the fundamental pitch frequency, and hence improve the performance of the extracting stage.

Although the performance of the system was evaluated through informal listening, formal real tests on human subjects should be done in our future work. Especially, it is interesting to evaluate the effect of the combined targeting-and-extracting on the intelligibility improvement of hearing-impaired people against the approach of using targeting alone. Since all our computer simulation experiments assumed an anechoic environment, we will also test the influence of the reverberation factor on the performance of the system in both our future simulation experiments and real-environment tests.

Last, simplified algorithms which allow to implement the extracting stage in real time are also needed. Limitations notwithstanding, the study successfully explores the potential improvement in intelligibility of signals from a single speaker in the "cocktail party" environment by combining the obvious advantages of directional beamformer targeting with residual noise suppression by comb filtering and attenuation.

ACKNOWLEDGMENTS

The authors wish to thank R. Beckers and E. Naheer for helpful discussions. This research has been supported by the Albert and Ethel Herzstein Foundation, Houston, TX, USA. The Gutwirth Prize to C. L. is noted with great pleasure. Part of this manuscript was written when one of the authors (S.S.) was a DFG Fellow in Karlsruhe University, Germany. The comments by two anonymous reviewers are also gratefully acknowledged.

spectrum in a deterministic way, since it is only when the processing length is very large that the main difference between the analysis of deterministic and stochastic signals arises.

- Allen, J. B. (1977). "Short-term spectral analysis, synthesis, and modification by discrete Fourier transform," *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-25**, 235–238.
- Allen, J. B., and Rabiner, L. R. (1977). "A unified approach to short-time Fourier analysis and synthesis," *Proc. IEEE* **65**, 1558–1564.
- Assmann, P. F., and Summerfield, Q. (1990). "Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies," *J. Acoust. Soc. Am.* **88**, 680–697.
- Atal, B. S., and Rabiner, L. R. (1976). "A pattern recognition approach to voiced-unvoiced-silence classification with applications to speech recognition," *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-24**, 201–212.
- Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., and Rzezchowski, C. (1984). "Standardization of a test of speech perception in noise," *J. Speech Hear. Res.* **27**, 32–48.
- Boll, S. F. (1979). "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-27**, 113–120.
- Brady, P. T. (1965). "A technique for investigating on-off patterns of speech," *Bell Syst. Tech. J.* **44**, 1–22.
- Cox, H., Zeskind, R. M., and Kooji, T. (1986). "Practical supergain," *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-34**, 393–398.
- Crochiere, R. E., and Rabiner, L. R. (1983). *Multirate Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ).
- Darwin, C. J. (1981). "Perceptual grouping of speech components differing in fundamental frequency and onset time," *Q. J. Exp. Psychol.* **33A**, 1636–1647.
- De Cheveigné, A. (1993). "Separation of concurrent harmonic sounds: Fundamental frequency estimation and a time-domain cancellation model of auditory processing," *J. Acoust. Soc. Am.* **93**, 3271–3290.
- 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.
- Flanagan, J. L. (1972). *Speech Analysis, Synthesis and Perception* (Springer-Verlag, Berlin).
- Frazier, R. H., Samsam, S., Braida, L. D., and Oppenheim, A. V. (1976). "Enhancement of speech by adaptive filtering," *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.* **ICASSP-76**, 251–253.
- Freeman, D. K., Cosier, G., Southcott, C. B., and Boyd, I. (1989). "The voice activity detector for the Pan-European digital cellular mobile telephone service," *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.* **ICASSP-89**, 369–372.
- Gardner, R. B., Gaskill, S. A., and Darwin, C. J. (1989). "Perceptual grouping of formants with static and dynamic differences in fundamental frequency," *J. Acoust. Soc. Am.* **85**, 1329–1337.
- Gordon-Salant, S. (1986). "Recognition of natural and time/intensity altered CVs by young and elderly hearing-impaired subjects," *J. Acoust. Soc. Am.* **80**, 1599–1607.
- Greenberg, J. E., and Zurek, P. M. (1992). "Evaluation of an adaptive beamforming method for hearing aids," *J. Acoust. Soc. Am.* **91**, 1662–1676.
- Hanson, B. A., and Wong, D. Y. (1984). "The harmonic magnitude suppression (HMS) technique for intelligibility enhancement in the presence of interfering speech," *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.* **ICASSP-84**, 18A.5.1–18A.5.4.
- Hecker, M. H. L. (1974). "A study of the relationships between consonant vowel ratios and speaker intelligibility," Ph.D. thesis, Stanford University.
- Hess, W. (1983). *Pitch Determination of Speech Signals: Algorithms and Devices* (Springer-Verlag, Berlin).
- Jenkins, G. M., and Watts, D. G. (1969). *Spectral Analysis and Its Applications* (Holden-Day, San Francisco).
- Kates, J. M. (1993). "Superdirective arrays for hearing aids," *J. Acoust. Soc. Am.* **94**, 1930–1933.
- Kobatake, H. (1987). "Optimization of voiced/unvoiced decisions in non-stationary noise environments," *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-35**, 9–18.
- Krubsack, D. A., and Niederjohn, R. J. (1991). "An autocorrelation pitch detector and voicing decision with confidence measures developed for noise-corrupted speech," *IEEE Trans. Signal Process.* **39**, 319–329.

¹The periodogram, also called Fourier line spectrum (Jenkins and Watts, 1969), is a measure of the power contribution at each frequency. Note that although the speech signal is a stochastic procedure, we calculate the power

- Lee, C. K., and Childers, D. G. (1988). "Cochannel speech separation," J. Acoust. Soc. Am. **83**, 274–280.
- Lee, H. H., and Un, C. K. (1986). "A study of on-off characteristics of conversational speech," IEEE Trans. Commun. **COM-34**, 630–637.
- Liu, C., and Sideman, S. (1995). "Digital frequency-domain implementation of optimum broadband arrays," J. Acoust. Soc. Am. **98**, 241–247.
- Liu, C., and Sideman, S. (1996). "Simulation of fixed microphone arrays for directional hearing aids," J. Acoust. Soc. Am. **100**, 848–856.
- Martin, P. (1982). "Comparison of pitch detection by cepstrum and spectral comb analysis," Proc. IEEE Int. Conf. Acoust. Speech Signal Process. **ICASSP-82**, 180–183.
- Meddis, R., and Hewitt, M. J. (1992). "Modeling the identification of concurrent vowels with different fundamental frequencies," J. Acoust. Soc. Am. **91**, 233–245.
- Montgomery, A. A., and Edge, R. A. (1988). "Evaluation of two speech enhancement techniques to improve intelligibility for hearing-impaired adults," J. Speech Hear. Res. **31**, 386–393.
- Monzingo, R. A., and Miller, T. W. (1980). *Introduction to Adaptive Arrays* (Wiley, New York).
- Naylor, J. A., and Boll, S. F. (1987). "Technique for suppression of an interfering talker in cochannel speech," Proc. IEEE Int. Conf. Acoust. Speech Signal Process. **ICASSP-87**, 6.12.1–6.12.4.
- Noll, A. M. (1970). "Pitch determination of human speech by the harmonic product spectrum, the harmonic sum spectrum, and a maximum likelihood estimate," Proc. Symposium Computer Process. in Commun., Polytechnic Institute of Brooklyn, pp. 779–797.
- Paliwal, K. K., and Rao, P. V. S. (1983). "A synthesis-based method for pitch extraction," Speech Commun. **2**, 37–45.
- Parsons, T. W. (1976). "Separation of speech from interfering speech by means of harmonic selection," J. Acoust. Soc. Am. **60**, 911–918.
- Peterson, P. M. (1989). "Adaptive array processing for multiple microphone hearing aids," D.Sc. Dissertation, Dept. Elect. Eng. and Comp. Sci., MIT; Res. Lab. Elect. Tech. Rept. 541, MIT.
- Preves, D. A., Fortune, T. W., Woodruff, B., and Newton, J. (1991). "Strategies for enhancing the consonant to vowel intensity ratio with in the ear hearing aids," Ear Hear. **12**, 139s–153s.
- Rabiner, L. R., Cheng, M. J., Rosenberg, A. E., and McGonegal, C. A. (1976). "A comparative study of several pitch detection algorithms," IEEE Trans. Acoust. Speech Signal Process. **ASSP-24**, 399–418.
- Saito, S., and Nakata, K. (1985). *Fundamentals of Speech Signal Processing* (Academic Press, Tokyo).
- 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.
- Schroeder, M. R. (1968). "Period histogram product spectrum: New methods for fundamental-frequency measurement," J. Acoust. Soc. Am. **43**, 829–834.
- Siegel, L. J. (1979). "A procedure for using pattern classification techniques to obtain a voiced/unvoiced classifier," IEEE Trans. Acoust. Speech Signal Process. **ASSP-27**, 83–89.
- Siegel, L. J., and Bessey, A. C. (1982). "Voiced/unvoiced/mixed excitation classification of speech," IEEE Trans. Acoust. Speech Signal Process. **ASSP-30**, 451–460.
- Simpson, A. M., Moore, B. C. J., and Glasberg, B. R. (1990). "Spectral enhancement to improve the intelligibility of speech in noise for hearing-impaired listeners," Acta Oto-Laryngol. (Stockh.) Suppl. **469**, 101–107.
- Soede, W., Berkhout, A. J., and Bilsen, F. A. (1993). "Development of a directional hearing instrument based on array technology," J. Acoust. Soc. Am. **94**, 785–798.
- Sreenivas, T. V., and Rao, P. V. S. (1981). "Functional demarcation of pitch," Signal Process. **3**, 277–284.
- Stadler, R. W., and Rabinowitz, W. M. (1993). "On the potential of fixed arrays for hearing aids," J. Acoust. Soc. Am. **94**, 1332–1342.
- Stone, M. A., and Moore, B. C. J. (1992). "Spectral feature enhancement for people with sensorineural hearing impairment: Effects on speech intelligibility and quality," J. Rehab. Res. Dev. **29**, 39–56.
- Stubbs, R. J., and Summerfield, Q. (1988). "Evaluation of two voice-separation algorithms using normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **84**, 1236–1249.
- Stubbs, R. J., and Summerfield, Q. (1991). "Effects of signal-to-noise ratio, signal periodicity, and degree of hearing impairment on the performance of voice-separation algorithms," J. Acoust. Soc. Am. **89**, 1383–1393.
- Terhardt, E. (1979). "Calculating virtual pitch," Hearing Res. **1**, 155–182.
- Tucker, R. (1992). "Voice activity detection using a periodicity measure," IEE Proc. I, Commun. Speech Vis. (UK) **139**, 377–380.
- Un, C. K., and Lee, H. H. (1980). "Voiced/unvoiced/silence discrimination of speech by data modulation," IEEE Trans. Acoust. Speech Signal Process. **ASSP-28**, 398–407.
- Wang, F.-M., Kabal, P., Ramachandran, R. P., and O'Shaughnessy, D. (1993). "Frequency domain adaptive postfiltering for enhancement of noisy speech," Speech Commun. **12**, 41–56.
- Weintraub, M. (1986). "A computational model for separating two simultaneous talkers," Proc. IEEE Int. Conf. Acoust. Speech Signal Process. **ICASSP-86**, 3.1.1–3.1.4.

A study on robust utterance verification for connected digits recognition

Mazin G. Rahim

AT&T Labs—Research, Murray Hill, New Jersey 07974

Chin-Hui Lee and Biing-Hwang Juang

Lucent Technologies, Bell Labs, Murray Hill, New Jersey 07974

(Received 21 October 1995; revised 20 May 1996; accepted 15 October 1996)

Utterance verification represents a key technology in the design of a user-friendly speech recognition system. One essential element when designing such a system is the ability to maintain a uniform performance over a wide range of acoustic conditions. An acoustic mismatch between training and testing conditions often results in an undesirable performance degradation. This paper addresses the issue of *robustness* in utterance verification of a speech recognition system. Two techniques, namely signal bias removal (SBR) and on-line adaptation, are studied. The SBR algorithm is used to deal with global mismatch conditions caused by handset and channel differences. The on-line adaptation algorithm is used to adjust verification threshold at runtime for achieving a desirable trade-off between false rejection and false alarm in new test conditions. Various on-line adaptation schemes are investigated. We show that both supervised or unsupervised adaptation can effectively adjust the verification threshold to achieve a desirable performance trade-off irrespective of the initial setting of the threshold. We report on connected digit recognition/verification results for *matched* and *mismatched* training and testing conditions. At a 5% digit string rejection rate, the proposed robust utterance verification system gives a reduction in string error rate between 32% and 35% over the conventional system, while still correctly rejects over 99.9% of nonvocabulary utterances. © 1997 American Institute of Physics. [S0001-4966(97)03005-1]

PACS numbers: 43.72.Ne [JS]

INTRODUCTION

During recent years, it has become increasingly essential to equip speech recognition systems with the ability to accommodate spontaneous speech input. Although providing this capability allows a friendly user-interface, it also poses a number of new problems due to the inclusion of out-of-vocabulary words, false starts, and disfluency. In these situations, a speech recognition system must be able to detect and recognize the “keywords” and reject the “nonkeywords.” Here keywords refer to known words in the application vocabulary while nonkeywords represent unspecified words or extraneous acoustic events. Recognizers equipped with a *keyword spotting* capability allow users the flexibility to speak naturally without the need to follow a rigid speaking format. Significant progress has been made in keyword spotting for unconstrained speech using hidden Markov models (HMMs). Conventional keyword spotting systems incorporate filler (or garbage) models for enhancing keyword detection and absorbing out-of-vocabulary events just like the way background models are used to absorb nonspeech events. Proper modeling of fillers using out-of-vocabulary events is essential for enhancing the performance of a general keyword spotting system. The issue of building appropriate filler models has been extensively studied by Rose and Paul,¹ Wilpon *et al.*,² Jeanrenaud *et al.*,³ and others. Single and multiple filler models with and without grammatical constraints, including modeling a filler with a structure of a whole word, a list of common non-keywords, a monophone

or triphone network, a broad phonemic class network, or a combination of these units, have been investigated.

The use of fillers improves keyword detection rate. However, in order to reduce false alarm rate, several studies (e.g., Refs. 4–6) have proposed the use of *keyword verification* following the detection and segmentation of input speech into keyword hypotheses. Keyword verification, similar to speaker verification (e.g., Ref. 7), is performed by computing a score for each detected keyword and comparing the score with a predetermined *verification threshold*. This process is often referred to as *testing hypothesis* in which a test is constructed to verify the *null hypothesis* that the keyword exists in the detected speech segment against the *alternative hypothesis* that no keyword exists in that segment of speech. Traditionally, likelihood scores of the speech segment corresponding to the keyword is used as the *test statistic*. Recent studies have suggested a *likelihood ratio test* which evaluates the ratio of the keyword likelihood score and some score of a filler model which measures the likelihood of the alternative hypothesis. Further enhancement of these methods have been reported using linear transformation of the feature space,^{8,9} discriminative feature analysis,¹⁰ and discriminative training.¹¹

As a generalization to keyword spotting and keyword verification, *utterance verification*, also referred to as UV in this study, attempts to accept part or all of an utterance as being *correctly* recognized or to reject it based on a computed confidence score (e.g., Ref. 10). This is particularly useful in situations where utterances are spoken without valid keywords or when significant confusion exists among

keywords which may result in a large number of substitution errors. The problem of verifying if an utterance has been correctly recognized is considerably more difficult than simply rejecting non-keyword utterances. In general, to deal with these types of problems, recognizers must be equipped with both a keyword spotting capability to correctly detect and recognize keywords embedded in the utterance, and with an utterance verification capability to reject utterances that do not contain valid keywords and utterances that have low confidence scores. In Ref. 12, utterance verification was formally defined in the framework of *statistical hypothesis testing*. The issues of selecting an optimal verification test, defining an alternative hypothesis and optimally estimating parameters of the test statistic were addressed.

An important part in the design of utterance verification systems is the selection of an appropriate *operating point* (or critical threshold) that can provide a desirable trade-off between *type I error* (false rejection of valid keywords or correctly recognized strings) and *type II error* (false alarm, or false acceptance of invalid nonkeywords or incorrectly recognized strings). In this study, we show that utterances recorded under different environmental conditions require different verification thresholds in order to satisfy some operational requirements in some optimal manner. This indicates that the verification statistics based on a particular training or evaluation set would not work consistently under mismatched training and testing conditions (see, e.g., Ref. 13). Further, it will be shown that a small deviation in the selected operating point could result in severe changes in the overall performance of the verification system.¹⁴ This critical issue raises the important question of how to maintain “robustness” in verification.

This paper provides a study on robust utterance verification for HMM-based automatic speech recognition (ASR) systems. Our objective is to design a robust verification system that maintains a uniform verification performance across a wide range of conditions. Two approaches are investigated. The first approach is a signal bias removal method based on maximum-likelihood estimation for the minimization of extraneous additive signal components.¹⁵ This technique has been merely applied for channel normalization. The second approach is an on-line adaptation technique which helps to achieve a desirable trade-off between detection rate and false alarm rate under changing test conditions. We report on experiments using various parametric and nonparametric adaptation methods, including maximum-likelihood (ML) adaptation, decision feedback adaptation,¹⁶ and maximum *a posteriori* (MAP) adaptation.^{17–19}

The rest of the paper is organized as follows. In Sec. I we describe the baseline utterance verification system and the connected digit databases used in our experiments. In Sec. II, we describe the problem of acoustic mismatch in UV that may exist between training and testing conditions. A set of preliminary results will be used to motivate the study of robust utterance verification. Achieving robustness through channel normalization is addressed in Sec. III. Once the UV system is in operation we can still improve the verification performance by adjusting the verification threshold on an utterance-by-utterance basis. This topic is discussed in detail

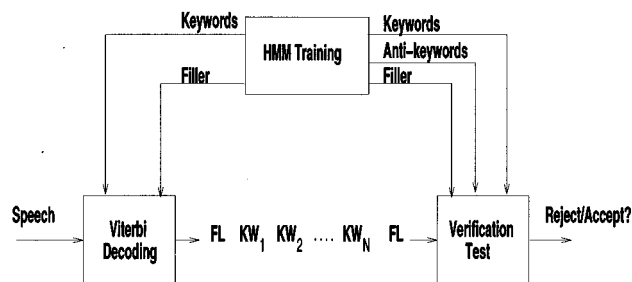


FIG. 1. A block diagram of the training/recognition/verification system.

in Sec. IV. Finally we summarize our findings and discuss some unresolved issues in Sec. V.

I. BASELINE SYSTEM AND DATABASES

A block diagram of the training/recognition/verification system is shown in Fig. 1. A two-pass strategy is adopted. It consists of recognition followed by verification. In the first pass, recognition is performed via a conventional beam search algorithm which segments the test utterance into a string of keyword hypotheses. In the second pass, an utterance-based confidence measure is constructed by combining the likelihood scores of all keywords and their corresponding antikeyword and filler models. A verification test is then performed and the utterance is either accepted or rejected (e.g., Ref. 12).

The front-end feature extraction process was conducted as follows. Input speech, sampled at 8 kHz, was initially preemphasized ($1 - 0.95z^{-1}$) and grouped into frames of 240 samples with a shift of 80 samples. For each frame, a Hamming window was applied followed by a 10th-order linear predictive coding (LPC) analysis.²⁰ A lifted 12-dimensional LPC-derived cepstral vector was then computed. The first and second time derivatives of the cepstrum, the so-called delta-cepstrum and delta-delta cepstrum, were also computed.²¹ Besides the cepstral-based features, the log of the energy, normalized by the peak sample, and its first- and second-order time derivatives were also computed. Thus, each speech frame was represented by a vector of 39 features consisting of 12 cepstrum, 12 delta-cepstrum, 12 delta-delta cepstrum, 1 log energy, 1 delta-energy, and 1 delta-delta energy.

Each keyword (i.e., digit in this study) was modeled by an N_{st} -state left-to-right HMM with M_{mix} mixture Gaussian state observation densities. The probability density function (PDF) for the observation vector O_t from state S_j of HMM λ_k was defined as

$$p(O_t | S_j, \lambda_k) = \sum_{m=1}^{M_{\text{mix}}} c_{kjm} \mathcal{N}(O_t; \mu_{kjm}, \Sigma_{kjm}), \quad (1)$$

where c_{kjm} is the mixture weight and $\mathcal{N}(\cdot)$ is a multivariate Gaussian distribution with mean μ_{kjm} and diagonal covariance Σ_{kjm} .

Training of each *keyword* model consisted of estimating the mean, covariance, and mixture weights for each state using ML estimation (e.g., Ref. 22). In this study, the state transition probabilities were fixed at 0.5. For each keyword model, an *antikeyword* model was also trained. The antikey-

word model can be considered as a digit-specific filler model. It is based on a similar concept to the speaker cohort in speaker verification.²³ An antikeyword model $\bar{\lambda}_k$, for example, for keyword, λ_k is generally trained on the data of all keywords except KW_k . Further explanation of this will be provided in the next section.

Aside from keywords and antikeywords, we also introduced a general acoustic filler model, λ_f , trained on nondigit speech data, and a background/silence model, λ_s , trained on the nonspeech segments of the signal. Therefore, a total of 24 models were used, each represented by a 10-state HMM with the exception of a single-state silence model. Up to 16 Gaussian mixtures were used in each state.

A. Connected digit databases

In order to assess the robustness of the various rejection techniques presented in this paper, three connected digits databases were considered. The databases were recorded over telephone lines, some by having individuals read digit strings from a predefined list and others from real applications where talkers spoke naturally without format constraints. The following databases have been investigated in this study.

(1) The first database was collected from two regions, namely, Long Island and Boston, over a digital T1 interface. Speech was recorded using four different microphone handsets, two electret and two carbon button. Digit strings of lengths 10, 14, 15, and 16 digits, corresponding to credit card numbers and long-distance telephone numbers, were collected from 250 adult talkers (125 males and 125 females). Approximately half of the speakers were used for training the HMMs and the other half for testing and evaluating the various verification techniques presented in this paper. The testing database which consists of 2842 strings will be referred to as DB1.

(2) The second database was collected from five regions within the United States, namely, Long Island, Chicago, Boston, Columbus, and Atlanta. Each region consisted of 100 adult talkers (50 males and 50 females), each speaking 66 connected digit strings from a predefined list (11 digit strings for each of the lengths two through seven). Half of the utterances were recorded using two electret microphone handsets, and the other half using two carbon button microphone handsets. Speech was transmitted over a long-distance telephone network which was either all analog, all digital, or a mix, depending on the region. A subset of this database consisting of 7073 strings was assigned for testing. This database will be referred to as DB2.

(3) The third database was collected in Jacksonville, Florida, over an analog tip-and-ring connection. Each caller was asked to speak his/her 16-digit credit card number. This database is different from the previous two in that each utterance was spontaneously spoken by a different person. All of the 3496 utterances in this database were selected for testing and will be referred to as DB3.

(4) A fourth database consisting of phonetically-rich sentences, modeled after the TIMIT sentences,²⁴ was also used in order to provide nonvocabulary (i.e., nondigit-like) words during training and recognition. This data set was recorded from two regions, namely, Boston and Chicago.

Speech was recorded using four different microphone handsets, two electret and two carbon button. This database was divided into three subsets. The first two were taken from Boston and the third from Chicago. The three subsets were used as follows. The first was used to train the filler model (about 3000 utterances). The second was added to DB1 (about 2900 utterances) and the third was added to both DB2 and DB3 (about 6000 utterances).

B. Utterance verification

Keyword verification can be considered as a *statistical hypothesis* testing problem in which the task is to test the *null hypothesis*, H_0 , that a given keyword, KW_k , exists within a segment of speech, O , against the *alternative hypothesis*, H_1 , that KW_k does not exist or is incorrectly recognized within that speech segment. Another scenario is to test H_0 , that a given keyword KW_k is correctly recognized within a segment of speech, O , against H_1 , that the given KW_k is incorrectly recognized. As discussed in Ref. 12, a verification test is performed to accept the null hypothesis, H_0 , if

$$\mathcal{LR}_k(O, \Lambda) = \frac{p_k(O|H_0)}{p_k(O|H_1)} > \tau_k. \quad (2)$$

This is referred to as the *likelihood ratio test*, where $p_k(O|H_0)$ and $p_k(O|H_1)$ are the PDFs of the null and the alternative hypotheses, respectively, Λ is the set of parameters characterizing both H_0 and H_1 , and τ_k is the *critical threshold* of the test.^{25,26}

For word (or digit) verification, several different formulations for the alternative hypothesis were evaluated in Ref. 12. It was shown that improved discrimination between keyword and nonkeyword models as well as reasonable detection of putative errors (i.e., misrecognition) is achieved by combining the likelihoods of digit-specific antidigit models with a general acoustic filler model using a type of geometric average. Antidigit models $\{\bar{\lambda}_k\}$ were trained using the same ML training procedure that was used for obtaining the digit models. Each model $\bar{\lambda}_k$ was trained on data corresponding to all digits except digit k . Maximum likelihood training was also applied to construct the filler model using the phonetically-rich sentences of the training data set.

For utterance verification, two techniques were investigated in Ref. 12 based on the likelihood ratio scores of the detected digits. The first was to reject the digit string if the likelihood ratio of any detected digit falls below a specified critical threshold. The second was to introduce a string-based score which is the geometric average of the individual digit likelihood ratio scores, such that

$$S(O; \Lambda) = -\log \left[\frac{1}{N} \sum_{q=1}^N \exp\{-\eta \cdot \mathcal{LR}_q(O; \Lambda)\} \right]^{1/\eta}, \quad (3)$$

where η is a positive constant and $\Lambda = \{\{\lambda_k\}, \{\bar{\lambda}_k\}, \lambda_s, \lambda_f\}$. Using the string-level based statistic was shown to give an improved verification performance for connected digits over using individual digit statistics.¹²

TABLE I. String recognition error rate for DB1 at different rejection rates.

Rejection rate (%)	String error (%)
0.0	9.0
1.0	8.2
3.0	7.1
5.0	6.4
7.0	5.9
9.0	5.2

When the proposed utterance verification system was trained and tested on DB1, over 99.9% of the nondigit word strings were correctly rejected. For valid digit strings, Table I shows the baseline string recognition error rate at different rejection rates. Note that error rate represents the percentage of strings that are incorrectly recognized (i.e., falsely accepted), and have failed to be detected by the verification process. At a 5% rejection rate, for example, the string error rate is reduced by about 29% over the baseline system (i.e., without rejection). It can be seen that rejecting nondigit utterances is rather easy in the case of connected digit recognition. For the remainder of this study we focus on the more *challenging* problem of testing the null hypothesis that *a given utterance is correctly recognized*. We use the single-string score defined in Eq. (3) to formulate the verification test. By treating the string score as a *confidence measure* for a test utterance, it offers a system the flexibility of incorporating an intelligent user interface to perform *voice repair*, *reprompt*, *confirmation*, or to accept the user input unconditionally based on the estimated confidence in recognition.

II. ROBUST UTTERANCE VERIFICATION

A speech signal transmitted through a telephone channel encounters various sources of degradation such as noise distortion and channel interference. Those effects as well as others, such as different sound pick-up equipment and articulatory effects are known to degrade the performance of an HMM-based speech recognition system. For utterance verification, these factors could affect the verification performance in several ways. For example, they could cause fluctuation of likelihood scores or variation of critical threshold. These types of effects will be investigated in this section.

The ability to improve recognition performance by introducing utterance verification was clearly demonstrated by the results in Table I. Although the testing database, namely DB1, was somewhat similar to the training corpus in that it was digitally recorded using the same types of microphones, this database still included different speakers and was recorded under different channel conditions. Nevertheless,

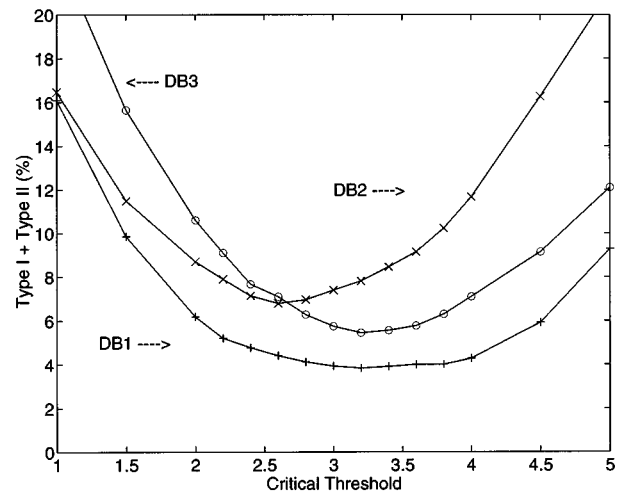


FIG. 2. Total verification error rate for DB1, DB2, and DB3 as a function of the critical threshold.

DB1 will be considered as a “matched” database in this study. In order to evaluate the robustness property of our rejection technique, further experiments have been conducted with databases DB2 and DB3. These databases were both recorded in completely different environments from the training corpus and will be considered as “mismatched” databases.

The baseline utterance verification system was evaluated on DB1, DB2, and DB3 using different critical thresholds. Figure 2 shows plots of the string verification error rate (i.e., combined type I and type II errors) as a function of the critical threshold. Three remarks can be made. First, the optimal threshold to achieve minimum verification error is different across the three databases. For example, DB2 has a minimum at an operating point of 2.6 whereas both DB1 and DB3 have their minima at about 3.2. Setting the critical threshold to 3.2 causes the verification error rate for DB2 to increase by about 15%. Second, both DB2 and DB3 have higher error rates than DB1. This is due to an environmental mismatch between the training model and the testing data. For example, the minimum verification error rates for DB1, DB2, and DB3 are 3.9%, 6.8%, and 5.5%, respectively. Third, it appears that DB2 and DB3 are more affected by changes in the critical threshold than DB1. For example, if an original threshold at 3.2 is raised by, say, 25% to 4.0 then this causes the total error rates for DB1, DB2, and DB3 to increase by 12% (to 4.3%), 50% (to 11.7%), and 32% (to 7.1%), respectively.

Table II presents the string recognition error rate before and after utterance verification for DB1, DB2, and DB3. These results are obtained when operating at a threshold of

TABLE II. String recognition error rate for DB1, DB2, and DB3 before and after UV using SBR with an operating point of 3.2.

Database	Baseline (%)	Rejection rate (%)	Digit string error (%)	Nondigit string error (%)
DB1	9.0	3.1	7.0	0.1
DB2	15.5	17.8	7.0	0.1
DB3	36.3	32.6	13.7	0.1

TABLE III. String recognition error rate for DB1, DB2, and DB3 before and after UV using SBR with an operating point of 3.2.

Database	Baseline (%)	Rejection rate (%)	Digit string error (%)	Nondigit string error (%)
DB1	8.4	2.8	6.6	0.1
DB2	13.8	13.9	7.0	0.1
DB3	29.7	26.4	11.9	0.1

3.2 which is the level where a minimum verification error rate was achieved for DB1. It is clear that for a given critical threshold, the verifier shows substantial performance fluctuations. For valid digit strings, it is shown that the performances on DB2 and DB3 are significantly increased after verification (see column 4), but at the expense of having a high rejection rate (see column 3). For example, the perceived string error rate for DB3 decreased by 63% (from 36.3% to 13.7%) at a rejection rate of 32.6% (see column 3). Certainly such a high rejection rate is undesirable in practice. It discourages users from utilizing the ASR system. Note that the recognition error rate for nondigit word strings fell below 0.1% for all the three databases (see column 5).

Ideally, one would like to have a verification system that is robust to different environments and shows minimal changes in performance when operating at various initial thresholds. In an effort to achieve this goal, two techniques have been investigated: (a) signal bias removal (SBR) to reduce channel mismatch between the training model and the testing data, (b) on-line adaptation schemes to enable sequential adaptation of the verification threshold in either a supervised or an unsupervised mode. The following two sections will discuss these techniques in detail.

III. ROBUST UV: SIGNAL BIAS REMOVAL (SBR)

The SBR method is an iterative procedure, which minimizes the effects of unknown adverse conditions that contaminate speech.¹⁵ It is based on a formulation that aims at separating two processes, one being the speech signal and the other is what we call a bias process. Assuming an additive uncorrelated bias term, b , to an observation sequence $X = \{x_1, x_2, \dots, x_t, \dots, x_T\}$ of T frames, then with a simplified notation

$$y_t = x_t + b \quad (4)$$

and $\{Y = y_1, y_2, \dots, y_t, \dots, y_T\}$, we have

$$p(Y|b) = p(Y - b). \quad (5)$$

Thus given a set of HMMs, $\Lambda = \{\lambda_i, i = 1, 2, \dots, M\}$, the likelihood function becomes

$$p(Y|b, \Lambda) = \prod_t \max_{\Lambda} p(y_t - b | \Lambda), \quad (6)$$

and for a given value of Λ , the ML bias estimator \bar{b} is the one that achieves

$$p(Y|\bar{b}, \Lambda) = \max_b p(Y|b, \Lambda). \quad (7)$$

This can be accomplished by an iterative procedure, reminiscent of the EM (estimation-maximization) algorithm.

Further details of the SBR method are provided in Ref. 15.

It has been shown that the SBR algorithm improves speech recognition performance for telephone speech collected over a wide range of handset and channel conditions.¹⁵ In this study, our objective is to investigate the impact of this algorithm when performing utterance verification. Table III presents the string recognition error rate before and after utterance verification for DB1, DB2, and DB3 when applying SBR with a critical threshold of 3.2 (determined from DB1). When applying SBR, a codebook of size 16 was employed which was generated from the training data by using the generalized Lloyd clustering algorithm.²⁷ For the baseline system (see column 2), the SBR method clearly reduces the string error rate reported in Table II especially for the mismatched databases. The string error rate drops by about 7%, 11%, and 18% for DB1, DB2, and DB3, respectively. Furthermore, when performing utterance verification, the SBR technique results in a lower rejection rate than that reported in Table II (see column 3). This is illustrated in Figs. 3–5. The plots show the string recognition error rate as a function of rejection rate when testing on DB1, DB2, and DB3 with and without the SBR method. At a rejection rate of 5%, for example, the string error rate is reduced from 6.4% to 6.1% for DB1, from 12.6% to 10.8% for DB2, and from 33.0% to 26.0% for DB3. It is interesting to note that the percentage error reduction for the three databases is maintained even up to a 15% rejection rate. On the other hand for nondigit strings, the verification system with SBR still accepts less than 0.1% of the utterances (see column 5).

It is seen that minimizing channel effects and other ex-

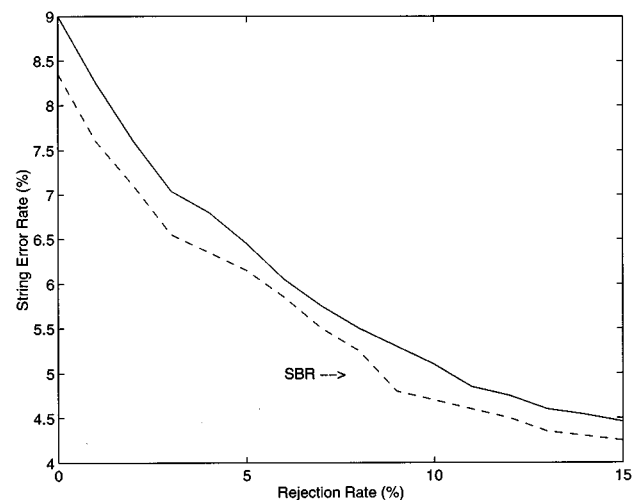


FIG. 3. String error rate for DB1 as a function of rejection rate with and without the introduction of SBR.

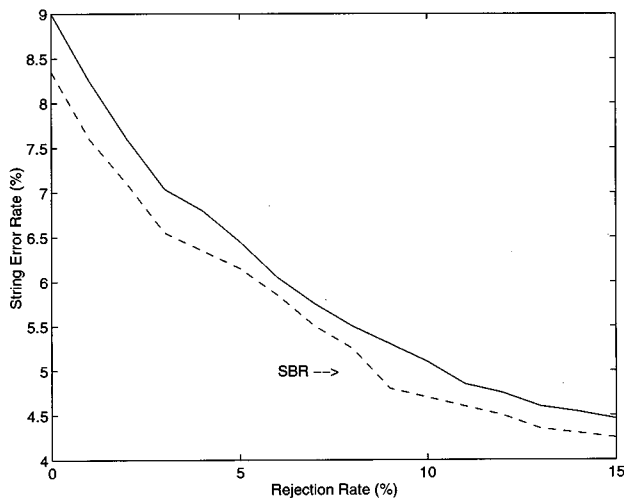


FIG. 4. String error rate for DB2 as a function of rejection rate with and without the introduction of SBR.

traneous signals has helped to produce a moderate improvement in recognition and verification. As expected, the improvement was more pronounced for the mismatched databases (i.e., DB2 and DB3). Note that the improvement for DB3 is higher than that for DB2. This observation is believed to be due to the longer digit strings in the former database which helps in computing a more accurate bias for each utterance. The SBR technique will thus be utilized in all subsequent experiments. (Detailed analysis of the recognition performance as a function of string length is provided in Ref. 15.)

IV. ROBUST UV: ON-LINE THRESHOLD ADAPTATION

The performance of utterance verification systems largely depends on the selection of an appropriate critical threshold. In previous sections, thresholds have been set according to a predefined criterion, such as to obtain a minimum verification error rate or to achieve an equal error rate. We demonstrated that based on a particular setting of the

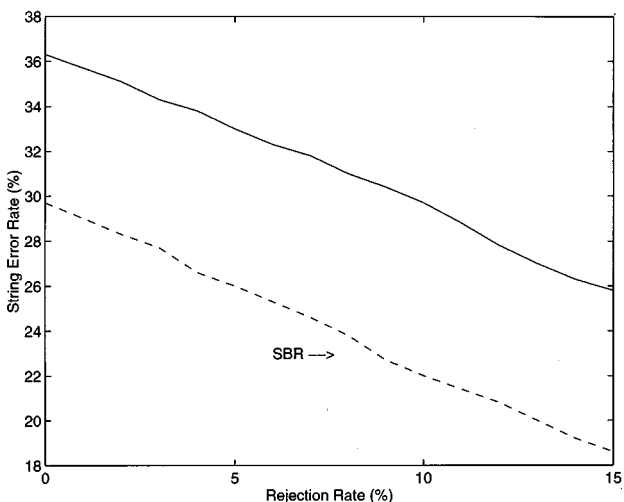


FIG. 5. String error rate for DB3 as a function of rejection rate with and without the introduction of SBR.

verification threshold a desired trade-off between type I and type II errors may not be maintained when dealing with an environmental mismatch between the training and the testing data, such as when testing on DB2 and DB3. For example, when considering nonvocabulary word strings as well as valid digit strings, the equal error rate for DB1 and DB2 are at about 2% and 3.2% for thresholds of 4.1 and 3.1, respectively (see Fig. 2). At a threshold of 4.1, the type I error for DB2 is a factor of 9 times higher than the type II error. Therefore, a critical threshold for a specific condition may not be appropriate for all different conditions. Ideally, one would like to have a mechanism that would continuously adjust the critical threshold in order to maintain a desirable trade-off between type I and type II errors according to a specific operating requirement.

Let S be the score function [or confidence measure, see Eq. (3)] used to compare with the critical threshold τ_s in the verification phase. Let $p_c(S)$ and $p_{\bar{c}}(S)$ be the densities of the score function for the correct and incorrect verification decisions, respectively. Then the type I and type II errors, $E_I(\tau_s)$ and $E_{II}(\tau_s)$, can be evaluated as

$$E_I(\tau_s) = \int_{-\infty}^{\tau_s} p_c(S) dS \quad (8)$$

and

$$E_{II}(\tau_s) = \int_{\tau_s}^{\infty} p_{\bar{c}}(S) dS. \quad (9)$$

The critical threshold can be selected to achieve an equal error rate, i.e., $E_I(\tau_s) = E_{II}(\tau_s)$. It can also be determined based on a desired operational requirement for a particular application. For the remainder of the study, we adopt the criterion of minimizing the total verification error, $E_T(\tau_s) = E_I(\tau_s) + E_{II}(\tau_s)$. Depending on the densities $p_c(S)$ and $p_{\bar{c}}(S)$, the critical threshold can usually be computed either directly or numerically given the parameters of the densities. Since these densities are often estimated from the training data and are not known exactly, their parameters are likely to be changed in operational environments. This usually results in a shift of the optimal threshold and an imbalance between type I and type II errors. One way to improve the robustness of the verification system is to adjust the critical threshold by adapting the parameters of the two score distributions using the actual operational data.

In the following subsections, we consider three different techniques for on-line threshold adaptation, namely maximum-likelihood (ML) adaptation, decision feedback adaptation, and maximum *a posteriori* (MAP) adaptation. These techniques utilize the string scores of a block of testing utterances to adapt the verification threshold so as to achieve a desired criterion which is to minimize the total verification error rate.

To illustrate that such procedures can be designed, the empirical distributions of the string confidence scores [from Eq. (3)] for DB1 are shown in Fig. 6. The right distribution corresponds to the confidence scores obtained when a digit string is both valid and recognized correctly. The distribution on the left corresponds to the scores of invalid digit strings and other strings that have been incorrectly recognized. Due

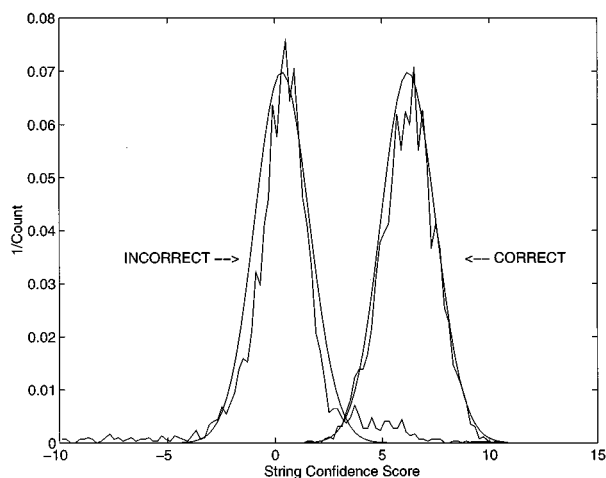


FIG. 6. Two histograms showing the distribution of the string confidence scores for the correct and the incorrect classes.

to the formulation of the geometric average in Eq. (3) which emphasizes low individual digit scores, the two distributions seem to have a similar shape with the exception of the long tails that characterize the left distribution. (Although the left distribution was constructed from invalid digit strings and other strings that were incorrectly recognized, it may also be formed from confidence scores of valid digit strings when passed through random digit models.) Both distributions can be approximated by two Gaussian densities having roughly the same variance as shown in Fig. 6.

A. Maximum-likelihood adaptation

A straightforward method for on-line adaptation is to assume that the two score distributions in Fig. 6 are of the same shape which is unimodal and symmetric around the mode of the distributions. The locations of the two distributions are estimated using test/adaptation data. Then in order to achieve a minimum total error rate (also an equal error rate), the critical threshold τ_s would need to be positioned at

$$\tau_s = \frac{\bar{X}_c + \bar{X}_{\bar{c}}}{2}, \quad (10)$$

where \bar{X}_c and $\bar{X}_{\bar{c}}$ are the modes of the correct class and the incorrect class distributions, respectively. Since the “true” means of the two distributions are unlikely to be known while the system is running in practice, the two modes can be estimated from operational data. When the two score distributions are assumed Gaussian, the means (or modes) are the maximum-likelihood estimates. If adaptation is super-

vised, then \bar{X}_c and $\bar{X}_{\bar{c}}$ can represent good estimates of the true means. However, if the system is running in an unsupervised mode, then the estimates of the means would be biased depending on the running value of τ_s . Here supervised adaptation assumes that a prompt is available to confirm whether or not the spoken digit string was recognized correctly. For example, a typical prompt is “Did you say 1 3 5 7 9?” which is followed by the user saying “yes” or “no.” In unsupervised adaptation, such confirmation is assumed unavailable and the verification results are used to determine if an utterance is correctly recognized.

In practice, whether running supervised or unsupervised adaptation, it is unlikely to have a sufficient amount of samples for estimating \bar{X}_c and $\bar{X}_{\bar{c}}$, and subsequently τ_s . An alternative, but approximate, solution would be to adapt τ_s for every R number of strings. In order to avoid biasing the averages due to the long tails of the distributions, scores exceeding predefined maximum and minimum thresholds are clipped.

In the following experiment, ML adaptation of the verification threshold was applied for DB1, DB2, and DB3. Both valid and invalid strings were presented for recognition/verification in a random order to ensure equal and sufficient information for estimating the averages. The critical threshold was initially set at the minimum error rate point for DB1 (i.e., 3.2) and was then adjusted for every 20 strings (i.e., $R=20$). Table IV presents the verification error rate before and after adaptation. The least possible error rate obtained when adjusting the verification threshold manually is shown under the column “Target.” As expected, since the initial setting of the critical threshold was based on DB1, the performance on this database did not change after adaptation. A small improvement was obtained when testing on DB2 and DB3. This was again expected since the optimum results on these databases (i.e., “Target”) were not significantly different than those obtained by either supervised or unsupervised adaptation.

To investigate the robustness of the ML adaptation technique to initial settings of the critical threshold, DB2 was selected for further analysis. Figure 7 shows the error rate of the verification system before adaptation (i.e., baseline) and following either supervised or unsupervised ML adaptation. The initial setting for τ_s was varied between 1 and 5. The results suggest that it is possible to reduce the verification error rate from anywhere between 0% and 15% for unsupervised adaptation, and between 0% and 30% for supervised adaptation. The inability to achieve further improvement is

TABLE IV. String error rate for DB1, DB2, and DB3 when applying ML adaptation of the verification threshold with an initial value of 3.2.

Database	Error before adapt (%)	Error after adapt (%)		
		Unsupervised	Supervised	Target (%)
DB1	3.5	3.5	3.5	3.5
DB2	6.8	6.5	6.5	6.1
DB3	5.8	5.7	5.6	5.6

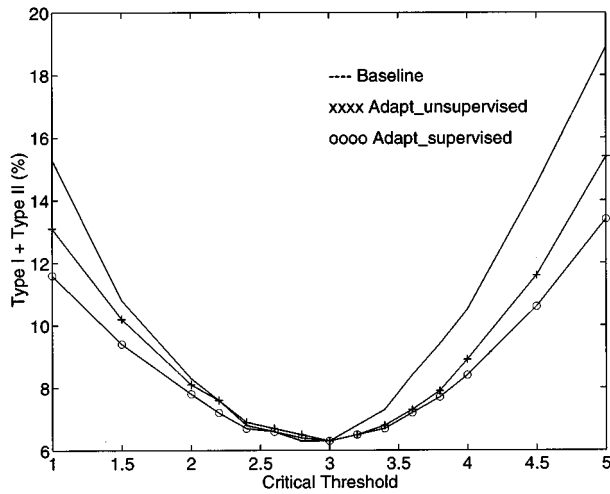


FIG. 7. Total verification error rate for DB2 at different initial critical thresholds before adaptation and after either supervised or unsupervised ML adaptation.

believed to be due to the slow learning rate, the insufficient number of strings used for computing the averages, and the underlying assumption that the optimum operating point lies midway between the means of the two distributions for all possible conditions.

B. Decision feedback adaptation

Since the equal variance assumption and maximum-likelihood adaptation produced unsatisfactory results when the initial threshold was far from the operating point needed to handle real operational data, we decided to experiment with a decision feedback approach in which the parameters of the two score distributions are embedded in a decision function. The error of the decision process is then minimized by a gradient decent algorithm which adjusts the parameters of the distributions so that to minimize the total verification error rate. Let the mean and variance of the left distribution in Fig. 6 (i.e., when $O \notin C_s$) be denoted by μ_c^- and σ_c^2 , respectively, and that for the right distribution (i.e., when $O \in C_s$) be denoted by μ_c and σ_c^2 , respectively. The goal is to adapt the means and variances as testing string confidence scores become available following recognition. Once these parameters are updated then the critical threshold, τ_s , can be estimated to minimize $E_T(\tau_s)$.

Given the string score $S(O; \Lambda)$ in Eq. (3), we define a decision function

$$E(O; \Lambda) = \frac{\Gamma(O)}{2} \left[\frac{(S(O; \Lambda) - \mu_c)^2}{2\sigma_c^2} - \frac{(S(O; \Lambda) - \mu_c^-)^2}{2\sigma_c^2} \right], \quad (11)$$

where

$$\Gamma(O) = \begin{cases} +1, & O \in C_s, \\ -1, & O \notin C_s. \end{cases} \quad (12)$$

Such a decision function, similar to the *misclassification measure* defined in minimum error-based discriminative training algorithms (e.g., Ref. 16), measures the normalized distance between the score $S(O; \Lambda)$ and the two means of the

two score distributions. Therefore for a string that is recognized correctly or incorrectly with a confidence score close to the desired mean, the value of $E(O; \Lambda)$ would be negative. To approximate an empirical error count, $E(O; \Lambda)$ is then smoothed using a 0-1 sigmoid function

$$\bar{E}(O; \Lambda) = \frac{1}{1 + \exp\{-\alpha E(O; \Lambda)\}}, \quad (13)$$

where α is a constant. If $v(n)$ denotes either μ_c , σ_c^2 , μ_c^- , or σ_c^2 after processing n strings, then using a gradient descent method

$$\Delta v(n) = -\beta \frac{\delta E}{\delta v(n)} + \gamma \Delta v(n-1), \quad (14)$$

where β is a learning step size and γ is a momentum term which is commonly used in the backpropagation training of neural networks to accelerate convergence and to avoid getting trapped in local minima of the error surface.^{28,30}

Experiments using decision feedback adaptation were carried out on all the three databases, namely, DB1, DB2, and DB3. Initial estimates for the means and variances of the distributions were calculated based on DB1 (see Fig. 6) so as to obtain a minimum verification error rate. In the experiments reported in this section, the variances were fixed and only the means were updated. Adaptation of the means was conducted upon recognizing every string, however, the critical threshold τ_s was only updated at every batch of 20 strings. This was found to be necessary to allow a more stable estimate of the means prior to updating the operating point. The results of the experiments using decision feedback adaptation were very similar to those obtained for ML threshold adaptation (see Table IV). Basically, only a small improvement was observed when testing on DB2 and DB3.

To test the robustness property of the proposed technique, an experiment was conducted in which the initial estimates of the means of each distribution were deviated within $\pm 40\%$. This caused the value of τ_s to vary between 1 and 5. The goal was to investigate whether decision feedback adaptation would be capable of retaining the target performance when varying the initial conditions of the system. Figure 8 shows the variation of the verification error rate for DB2 as a function of τ_s . Results before adaptation and after either supervised or unsupervised adaptation are presented. Clearly the performance is better than that reported for ML threshold adaptation (see Fig. 7). For example, when τ_s is initially set to 4.0 (i.e., 25% increase from the optimal point of 3.2), the error rate increases from 6.8% to 10.8% (i.e., an increase by 59%). After ML threshold adaptation, the error rate drops to 9.0% and 8.5% for unsupervised and supervised adaptation, respectively. When using decision feedback adaptation, the error rate drops further to 8.0% and 6.8% for unsupervised and supervised adaptation, respectively. Although the improvement in performance when using the gradient descent algorithm is higher in this case, ML threshold adaptation offers a lower computational complexity by a factor of 2 to 3 times.

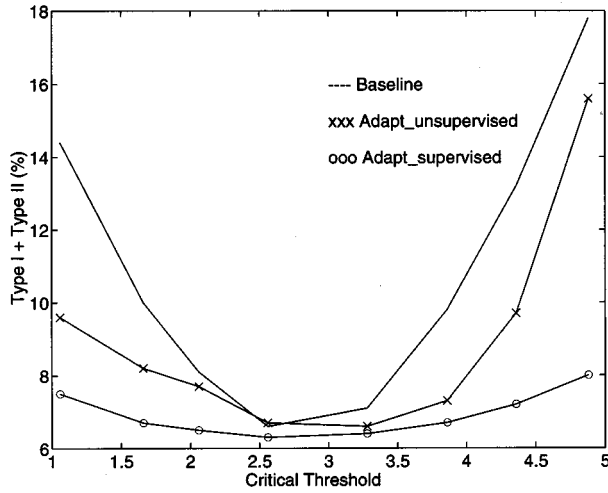


FIG. 8. Total verification error rate for DB2 at different initial critical thresholds before adaptation and after either supervised or unsupervised decision feedback adaptation.

C. Maximum *a posteriori* (MAP) adaptation

Instead of using the decision feedback approach which requires the use of a gradient decent algorithm, we can also improve the ML approach used in Sec. IV A by incorporating prior information in the estimation of the parameters of the two score distributions. A Bayesian learning approach based on MAP adaptation (e.g., Refs. 17–19) is studied in the following. It will be shown that the MAP estimate of the Gaussian parameters can be simply evaluated avoiding the need for an iterative gradient search. The difference between an ML approach and a MAP estimate lies in the assumption of an appropriate prior distribution. When estimating a set of parameters, the ML approach aims at maximizing the conditional density

$$\Lambda_{\text{ML}} = \arg \max_{\Lambda} p(O|\Lambda), \quad (15)$$

whereas the MAP estimation assumes the parameters to be random and aims at maximizing the posterior density.

$$\Lambda_{\text{MAP}} = \arg \max_{\Lambda} p(\Lambda|O) = \arg \max_{\Lambda} p(O|\Lambda)p(\Lambda). \quad (16)$$

The prior density, $p(\Lambda)$, is typically computed using a small set of training data.

MAP estimation has been applied successfully to a number of adaptation problems in speech and speaker recognition.^{29,17} In this study, the Bayesian learning approach is used to adapt the parameters of the two Gaussian score densities. Assuming the prior densities for the two mean parameters are also Gaussian with prior means, μ_{p_c} and $\mu_{p_{\bar{c}}}$, and prior variances, $\sigma_{p_c}^2$ and $\sigma_{p_{\bar{c}}}^2$, respectively. Then the MAP estimates for the mean parameters are simply¹⁹

$$\begin{aligned} \tilde{\mu}_c &= \frac{n_c \sigma_c^2}{\sigma_{p_c}^2 + n_c \sigma_c^2} \bar{X}_c + \frac{\sigma_{p_c}^2}{\sigma_{p_c}^2 + n_c \sigma_c^2} \mu_{p_c}, \\ \tilde{\mu}_{\bar{c}} &= \frac{n_{\bar{c}} \sigma_{\bar{c}}^2}{\sigma_{p_{\bar{c}}}^2 + n_{\bar{c}} \sigma_{\bar{c}}^2} \bar{X}_{\bar{c}} + \frac{\sigma_{p_{\bar{c}}}^2}{\sigma_{p_{\bar{c}}}^2 + n_{\bar{c}} \sigma_{\bar{c}}^2} \mu_{p_{\bar{c}}}, \end{aligned} \quad (17)$$

where n_c and $n_{\bar{c}}$ are the number of observations from the correct and incorrect classes, respectively. This formulation raises two interesting thoughts. One is that the MAP estimate of the mean is a weighted average of the prior mean and the sample mean. Two, when a large number of samples are provided then the MAP estimate approaches the ML estimate.

In the following experiment, the initial estimates of the means and variances of the densities were computed from Fig. 6. The distribution parameters were updated in a batch of 40 strings. It was established that about 30–50 strings were necessary in order to obtain stable parameter estimates. It was also found that applying the same formulations presented in Eq. (17) for updating the variances helped in improving the overall performance of the verification system. Therefore besides the mean, the variance of each distribution was also adapted using a weighted average of the prior variance and the sample variance. Clearly, updating the means and variances in this manner is computationally less expensive than the gradient descent approach. Following every cycle update of the distribution parameters, the critical threshold was recomputed to minimize the total verification error rate. The verification results when employing MAP estimation were quite similar to those previously obtained with maximum-likelihood adaptation and decision feedback adaptation.

A further experiment was then conducted to investigate the robustness of MAP estimation to varying initial conditions. Database DB2 was again selected for this experiment. The initial estimates of the means of the prior distributions were varied to allow a changing verification threshold between 1 and 5. The means and variances of the underlying distributions were updated every 40 utterances. The critical threshold was then adjusted to achieve a minimum verification error rate. Figure 9 shows the variation of the verification error rate before adaptation and after either supervised or unsupervised adaptation. Two interesting features can be concluded from this plot. First, the difference between supervised and unsupervised adaptation is small. Second, the verification error rate is unaffected by the setting of the initial parameters thus resulting in a robust verification system. The results using unsupervised MAP adaptation is the best we have had so far. Although MAP adaptation did provide robustness towards initial parameter settings, the performance at the optimal operating point was found to be slightly worse than that achieved by the previous techniques for on-line adaptation.

V. DISCUSSION AND SUMMARY

As the demand for speech recognition technologies keeps on the rise, the need for the development of systems that are robust to speaking style, accents, environmental mis-

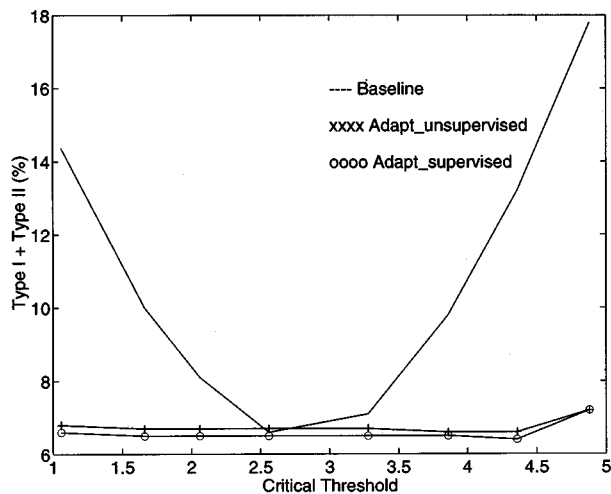


FIG. 9. Total verification error rate for DB2 at different initial critical thresholds before adaptation and after either supervised or unsupervised MAP adaptation.

match, disfluency, etc., becomes increasingly essential. In these circumstances, utterance verification plays an important role in providing the flexibility for designing an intelligent user interface. Given the changing testing conditions, a robust utterance verification is then needed to maintain an acceptable error rate and to provide a desirable trade-off between false alarm rate and false rejection rate.

This paper described a robust utterance verification system for connected digits recognition. A digit-based likelihood ratio combining the scores of keyword, antikeyword, and filler models was found to be effective in detecting non-vocabulary words and, in some instances, putative errors (i.e., valid but misrecognized digit strings). Three key features of this system have been discussed, namely: (1) robust string-based likelihood measure for detecting incorrectly recognized keywords; (2) SBR for channel normalization; and (3) on-line adaptation for achieving a desired trade-off between false alarm rate and detection rate during actual system operations.

Although the SBR method has been traditionally used to improve the acoustic match between a training model and a test utterance, our experiments have shown that the same technique can help lower the absolute rejection rate between 2% and 10% while maintaining a uniform verification error rate. However, we still have experienced fluctuations in likelihoods when testing on mismatched databases especially when selecting an inappropriate verification threshold. Consequently, we proposed three methods for adapting the verification threshold on-the-fly, namely by ML, decision feedback, and MAP. The three proposed methods varied in their complexity and in the amount of data required for adaptation. Further investigation of their robustness to the initial setting of the verification threshold revealed that the MAP framework whether running in a supervised or an unsupervised mode resulted in the best performance.

When combining the SBR algorithm and the unsupervised MAP threshold adaptation technique, the verification system correctly rejected over 99.9% of the nonvocabulary

TABLE V. String error rate for databases DB1, DB2, and DB3 before and after verification at a rejection rate of 5%. The utterance verification system includes SBR and unsupervised MAP adaptation with an initial operating point of $\tau_s = 3.2$.

Database	Error before rejection (%)	Error after rejection (%)
DB1	9.0	6.1
DB2	15.5	10.1
DB3	36.3	24.0

utterances. For testing digit strings, the performance is summarized in Table V. It can be seen that the system effectively improves rejection of putative errors (i.e., incorrectly recognized utterances). At a rejection rate of 5%, the experimental results indicate a reduction in the string error rate by about 32%, 35%, and 34% for DB1, DB2, and DB3, respectively. It is our belief that this improvement is somewhat conservative since the databases used in this study included only valid digit strings with no embedded speech. Digit strings with out-of-vocabulary words would typically represent 10% to 20% of each database, and when considered, further-improvement is expected. This is reasonable due to the nature of the utterance verification measure in Eq. (3) which provides good detection of segments or words with low confidence score.

Although our results are still preliminary and more research is needed, we believe this is the first study to illustrate the important concept of robust utterance verification. The techniques presented in this paper are rather general and they can be extended to other similar problems like speaker verification. However, a number of unresolved issues in MAP threshold adaptation still requires some further investigation. For instance, in many applications, the channel conditions may change rapidly and therefore a slow adaptation algorithm like MAP cannot accommodate well with the change. It is important to find ways to improve adaptation efficiency while maintaining a low number of adaptation samples. Another problem is how to reduce the performance gap between supervised and unsupervised adaptation especially when only a small number of samples is available.

ACKNOWLEDGMENTS

The authors acknowledge useful discussions with W. Chou, R. Sukkar, and A. Setlur. We would like to thank Eric Brill and the other anonymous reviewer for their helpful comments.

¹R. Rose and D. Paul, "A Hidden Markov Model Based Keyword Recognition System," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing* (IEEE, New York, 1990), Vol. I, pp. 129–132.

²J. Wilpon, L. Rabiner, C.-H. Lee, and E. Goldman, "Automatic recognition of keywords in unconstrained speech using hidden Markov models," *IEEE Trans. Acoust. Speech Signal Process.* **38**, 1870–1990 (1990).

³P. Jeanrenaud, M. Siu, J. Rohlicek, M. Meteer, and H. Gish, "Spotting events in continuous speech," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing* (IEEE, New York, 1994), Vol. I, pp. 381–384.

⁴P. Moreno, D. Roe, and P. Ramesh, "Rejection techniques in continuous speech recognition using hidden Markov models," in *Proceedings European Conference on Signal Processing* (1990), pp. 1383–1386.

- ⁵B. Chigier, "Rejection and keyword spotting algorithms for a directory assistance city name recognition application," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing* (IEEE, New York (Alberta, Canada, 1992), Vol. II, pp. 93–96.
- ⁶M.-W. Feng and B. Mazar, "Continuous word spotting for applications in telecommunications," in *Proceedings of the International Conference on Spoken Language Processing* (Alberta, Canada, 1992), pp. 21–24.
- ⁷C.-S. Liu, C.-H. Lee, W. Chou, A. E. Rosenberg, and B.-H. Juang, "A study on minimum error discriminative training for speaker recognition," *J. Acoust. Soc. Am.* **97**, 637–648 (1995).
- ⁸R. Sukkar and J. Wilpon, "A two pass classifier for utterance rejection in keyword spotting," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing* (IEEE, New York, 1993), Vol. II, pp. 451–454.
- ⁹J. Sorensen and M. Savic, "Hierarchical pattern classification for high performance text-independent speaker verification systems," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing* (IEEE, New York, 1994), Vol. I, pp. 157–160.
- ¹⁰R. Sukkar, "Rejection for connected digit recognition based on GPD segmental discrimination," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing* (IEEE, New York, 1994), Vol. I, pp. 393–396.
- ¹¹R. Rose, "Discriminant wordspotting techniques for rejecting non-vocabulary utterances in unconstrained speech," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing* (IEEE, New York, 1992), Vol. II, pp. 105–108.
- ¹²M. Rahim, C.-H. Lee, and B.-H. Juang, "Discriminative utterance verification for connected digits recognition," in *Proceedings European Conference on Speech Communications*, Madrid (1995), Vol. I, pp. 529–532.
- ¹³R. Rose and E. Hofstetter, "Techniques for robust wordspotting in continuous speech messages," in *Proceedings European Conference on Speech Communications* (Elsevier Science, Amsterdam, The Netherlands, 1991), pp. 1183–1186.
- ¹⁴M. Rahim, C.-H. Lee, and B.-H. Juang, "Robust utterance verification for connected digits recognition," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing* (IEEE, New York, 1995), Vol. I, pp. 285–288.
- ¹⁵M. Rahim and B.-H. Juang, "Signal bias removal by maximum likelihood estimation for robust telephone speech recognition," *IEEE Trans. Speech Audio Process.* **IV**, 19–30 (1996).
- ¹⁶B.-H. Juang and S. Katagiri, "Discriminative learning for minimum error classification," *IEEE Trans. Signal Process.* **40**, 3043–3054 (1992).
- ¹⁷J.-L. Gauvain and C.-H. Lee, "Bayesian learning for hidden Markov model with Gaussian mixture state observation densities," *Proceedings European Conference on Speech Communications* (Elsevier Science, Amsterdam, The Netherlands, 1992), Vol. II, pp. 205–213.
- ¹⁸J.-L. Gauvain and C.-H. Lee, "Maximum a posteriori estimation for multivariate Gaussian mixture observation of Markov chains," *IEEE Trans. Speech Audio Process.* **II**, 291–298 (1994).
- ¹⁹C.-H. Lee, C.-H. Lin, and B.-H. Juang, "A study on speaker adaptation of the parameters of continuous density hidden Markov models," *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-39**, 806–814 (1991).
- ²⁰L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals* (Prentice-Hall, Englewood Cliffs, NJ, 1978).
- ²¹C.-H. Lee, E. Giachin, L. R. Rabiner, R. Pieraccini, and A. E. Rosenberg, "Improved acoustic modeling for large vocabulary continuous speech recognition," *Comput. Speech Lang.* **6**, 103–127 (1992).
- ²²L. R. Rabiner and B.-H. Juang, *Fundamentals of Speech Recognition* (Prentice-Hall, Englewood Cliffs, NJ, 1993).
- ²³A. Rosenberg, J. Delong, C.-H. Lee, B.-H. Juang, and F. Soong, "The use of cohort normalized scores for speaker recognition," in *Proceedings of the International Conference on Spoken Language Processing* (Alberta, Canada, 1992), pp. 599–602.
- ²⁴L. F. Lamel, R. H. Kessel, and S. Seneff, "Speech database development: design and analysis of the acoustic-phonetic corpus," in *Proceedings of the Speech Recognition Workshop (DARPA)* (IEEE, New York, 1986).
- ²⁵P. Bickel and K. Doksum, in *Mathematical Statistics: Basic Ideas and Selected Topics* (Prentice-Hall, Englewood Cliffs, NJ, 1977).
- ²⁶K. Fukunaga, *Statistical Pattern Recognition* (Academic, New York, 1990), 2nd ed.
- ²⁷Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.* **28**, 84–95 (1980).
- ²⁸R. P. Lippmann, "An introduction to computing with neural nets," *IEEE Acoust. Speech Signal Process. Mag.* **4**, 4–22 (1987).
- ²⁹C.-H. Lee and J.-L. Gauvain, "Speaker adaptation based on MAP estimation of HMM parameters," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing* (IEEE, New York, 1993), Vol. II, pp. 588–591.
- ³⁰D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, edited by D. E. Rumelhart and J. L. McClelland (MIT, Cambridge, MA, 1986), Vol. 1.

Acceptance limits for the duration of pre-Helmholtz transients in bowed string attacks

Knut Guettler

Norwegian State Academy of Music, P.O. Box 5190, Majorstua, N-0302 Oslo, Norway

Anders Askenfelt

Department of Speech, Music and Hearing, Royal Institute of Technology, P.O. Box 70014, S-100 44 Stockholm, Sweden

(Received 31 May 1996; revised 9 December 1996; accepted 16 December 1996)

The attack of most bowed notes shows an initial part before Helmholtz triggering occurs (the *pre-Helmholtz transient*), during which the stick-slip interaction promotes frequencies other than that of the string's fundamental. Depending on the particular combination of bowing parameters, this state is characterized either by periods that are *prolonged*, or by a division of the period into two or more parts, *multiple flyback*. An onset with perfectly periodic motion (Helmholtz triggering) directly from the very start is also possible. A sample of violin tones representing these three classes of attacks, and with different duration of the pre-Helmholtz transient, has been collected by the use of a computer-controlled bowing machine. The tones were evaluated by 20 advanced string students and professionals in a listening test, judging the acceptance and quality of the attacks. The maximum acceptable duration of the pre-Helmholtz transient was estimated to 50 ms (≤ 10 nominal periods, open G string, violin) for attacks with prolonged periods, and 90 ms (≤ 18 periods) for multiple-flyback attacks. These values refer to a neutral start in a neutral context, such as when practicing a scale. A playing test, in which the performances of two professional violinists were analyzed, confirmed these results, and showed that the same limits apply to a larger group of bowing styles as well. © 1997 Acoustical Society of America. [S0001-4966(97)05504-5]

PACS numbers: 43.75.De, 43.75.Cd, 43.66.Lj [WJS]

INTRODUCTION

The attack of a note on a bowed instrument shows a transition phase during which the “steady-state” motion develops. In particular, the stick-slip interaction often starts with an aperiodic initial part before Helmholtz triggering occurs and the fundamental periodicity is established. This is so because the conditions for a “perfect attack,” where the string slips under the bow hair only once per fundamental period from the very beginning of the note, are too demanding to be met in practice at each stroke.

Depending on the actual combination of bowing parameters, the initial part of the attack is characterized either by periods which are *prolonged* (sometimes referred to as “delayed triggering”), or by a division of the period into two or several parts, *multiple flyback*,¹ or “multiple slip” (see Fig. 1). As will be shown, roughly half of the attacks in a normal string performance will be either one of these two types. An attack with perfectly periodic motion (Helmholtz triggering) directly from the very start is also possible, being the standard reference of an ideal attack in string playing.

An attack with prolonged periods will normally be perceived as “choked/creaky,” while a multiple-flyback attack will be characterized as “loose/slipping.” However, provided that the aperiodic states do not last too long, such attacks could very well pass as acceptable. *How long* has been a matter of discussion, but professional string players claim that the transient must be short, the attack should sound “clean.”

The main question addressed in this study was: “How

long aperiodic attack transients (before the triggering of the periodic Helmholtz motion) do professional string players accept for notes with a neutral attack in a neutral context?” Such a reference would be useful when analyzing the bowed string, either through computer modeling, or from observations on real strings.

It is to be noted that even notes with a perfect Helmholtz motion from the very beginning show a transient state during which the amplitude develops and subfundamental frequencies are excited, among other phenomena. In particular, the time for reaching final “steady-state” amplitude, the *amplitude buildup time*, is an important characteristic of a bowed attack. The main aspect of the transient state treated in this study is, however, the duration of the aperiodic part before the triggering of the Helmholtz motion. In the following, this state will be referred to as the *pre-Helmholtz transient*.

The study was divided into two parts: (1) a listening test, in which a large group of string players judged the acceptance of recorded bowed attacks, and (2) a playing test, in which the performances of two professional violinists were analyzed.

I. LISTENING TEST

A. Recordings

A series of bow strokes on the open G string of a violin ($G_3 = 196$ Hz) was recorded on a DAT recorder, using a computer controlled bowing machine.² One channel recorded the string velocity at the bowing point (later used for evaluating the stick-slip action),³ while the other channel was fed with

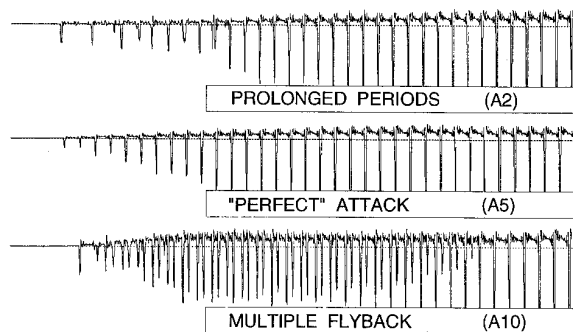


FIG. 1. Measurements of the string motion on a violin illustrating the three principal types of bowed string attacks. String velocity at the bowing point for prolonged (*top*), “perfect” (*middle*), and multiple-flyback attacks (*bottom*). The notes were played with a bowing machine using a normal bow. The time window shown is 200 ms and the nominal period 5.1 ms (open G string). Notice the accented onset of the multiple-flyback attack (A10).

the signal from a small electret microphone inside the violin. All strokes were run with a bow-bridge distance of 28 mm (relative bowing position $\beta=1/11.6$), and a constant bow force of 800 mN. The amount of rosin on the bow (and thus the frictional properties) was varied from stroke to stroke. Different strokes were programmed on the bowing machine, each starting from rest and accelerating according to a predetermined value up to a steady velocity of 20 cm/s.

A selection of attacks was made, resulting in a series of bow strokes representing *perfect* attacks, and *prolonged* and *multiple-flyback* attacks of different durations. The microphone recordings of these notes were transferred to a digital sound editing program,⁴ by which they were normalized to equal durations (1000 ms) and decay times (300 ms). A modest reverberation and equalization were then added in order

to mimic a normal violin sound (as perceived by the performer), as closely as possible.

B. The series of attacks

Eleven attacks were chosen to represent a reasonable spread in transient behavior, ranging from 160 ms of prolonged periods to 200 ms of multiple flyback periods before the occurrence of Helmholtz triggering (see Table I, Series A, violin). In order to study whether or not an accent would influence the perceived quality of an attack, copies of two of the selected attacks (A5 and A11) were artificially accented and added to the series (with an X added to the name), giving 13 attacks in all. The envelope of the artificial accents boosted the initial period by 12 dB, followed by a smooth decay covering 5–6 periods down to the original level. Of the six multiple-flyback attacks, two had pronounced natural accents (A8 and A10, the latter shown in Fig. 1).

A second series was derived from the first by changing the sampling rate (see Table I, Series B, “viola”). For series B, the pitch was lowered a fifth from 196 Hz corresponding to the open G string of the violin, to 131 Hz, which would compare with the open C string on the viola. This series was referred to as “another string instrument” when presenting the sound examples in the listening test, avoiding a direct reference to the viola. The second series was normalized and equalized similarly to series A.

The buildup time for reaching final string amplitude was measured with a sound file editor,⁴ observing both the string velocity and the sound track. The buildup time was measured from the first flyback to a point where the amplitude had reached 90% (−1 dB) of the “steady-state” level. The attacks in the series were chosen primarily on the basis of the

TABLE I. Data for the two series of attacks in the listening test. Series B (“viola”) was obtained from Series A (violin) by lowering the sampling rate to 2/3. **Attack characteristics** indicates the type of stick-slip pattern. **Bow acc.** gives the programmed acceleration of the bow. Each stroke started with the bow at rest and accelerated to a steady velocity of 20 cm/s. **Buildup time** indicates the time for reaching 90% (−1 dB) of final “steady-state” amplitude, counting from the first flyback. Different amounts of rosin were used on the bow, which explains the inconsistency between bow acceleration and buildup time. **Pre-Helmholtz** gives the duration of the pre-Helmholtz transient, counting measured in absolute time [ms] and nominal periods [T_0] (violin 5.1 ms; “viola” 7.7 ms). Asterisks indicate attacks with pronounced accents. An X added to the name indicates that the attack was accented artificially (see text). **Answer Yes** refers to the question: “Does the attack sound acceptable?” in the listening test. Attacks achieving more than 50% acceptance are indicated by bold print and shading. The values are based on the answers from 20 subjects \times 3 repetitions=60 answers per attack.

			Series A violin					Series B “viola”				
Attack characteristics		Bow acc.	Attack name	Buildup time	Pre-Helmholtz		Answer Yes	Attack name	Buildup time	Pre-Helmholtz		Answer Yes
Type	Accent	cm/s ²		ms	ms	T_0	%		ms	ms	T_0	%
Prolonged		40	A1	420	160	31	0	B1	630	240	31	0
Prolonged		100	A2	180	55	11	35	B2	270	83	11	8
Prolonged		125	A3	130	30	6	77	B3	200	45	6	47
Prolonged		100	A4	170	20	4	37	B4	260	30	4	23
Perfect		150	A5	120	0	0	92	B5	180	0	0	78
Perfect	*	150	A5X	120	0	0	92	B5X	180	0	0	82
Mult. flyback		150	A6	130	15	3	93	B6	200	23	3	78
Mult. flyback		300	A7	70	40	8	90	B7	110	60	8	87
Mult. flyback	*	200	A8	90	60	12	80	B8	140	90	12	38
Mult. flyback		150	A9	130	90	18	52	B9	200	135	18	48
Mult. flyback	*	500	A10	140	140	27	30	B10	210	210	27	38
Mult. flyback		150	A11	200	200	39	12	B11	300	300	39	7
Mult. flyback	*	150	A11X	200	200	39	17	B11X	300	300	39	33

duration of the pre-Helmholtz transient, but as will be shown later, also accents and amplitude buildup time played important roles in the judgment of attack quality.

C. Design and procedure

The two series of attacks were evaluated in a listening test by a panel of 20 advanced string students and professionals at the Norwegian State Academy of Music. The recorded attacks were presented to the test panel in a medium sized auditorium at a moderate sound level, using a single loudspeaker.

Three sequences were played, each consisting of 26 attacks. Each attack was repeated three times in quick succession, followed by 10 s of silence, which allowed for writing down the scores. In the first sequence, the 13 attacks of series A appeared twice in random order. The next sequence was composed in an equal manner from series B, while in sequence three, series A and B were scrambled. The subjects thus judged each attack three times during the test session, which allowed an estimation of the intraindividual reliability. For each attack, three scores were given:

(1) **Acceptance.** The acceptance of the attack was indicated by answering the question: “Does the attack sound acceptable?” with “Yes” or “No.”

(2) **Quality.** The quality of the attack was indicated on a line (100-mm long) with the end points defined as: “The attack is catastrophic” and “The attack is perfect.”

(3) **Character.** The character of the attack was indicated on a line (100-mm long) with the end points defined as: “The attack sounds extremely choked/creaky” and “The attack sounds extremely loose/slipping” with the midpoint indicated by “without creaks or slips.”

The participants were instructed to judge the attacks with reference to a “neutral attack,” for example as when practicing a scale. For Quality and Character, the positions of the marks (measured in millimeters) were taken as the score values, thus ranging from 0 to 100.

Before the actual listening test started, two extreme examples of attacks with very long pre-Helmholtz transients were presented, in order to illustrate the two contrasting types of attacks (prolonged and multiple flyback) and establish the terminology. Being professional string players and students, the subjects could be assumed to be familiar with the phenomena *per se*. Following, all 13 attacks of series A were played in quick succession. Finally, the subjects judged four other attacks in a training session, in order to become familiar with their task.

D. Results

1. Acceptance

A first estimation of the acceptance limits for the duration of the pre-Helmholtz transient was obtained from the answers to the Acceptance question: “Does the attack sound acceptable?” Table I gives the percentages of approvals with a total of 60 answers per attack. Using a lower limit of 50% “Yes” as a reasonable *ad hoc* criterion of acceptance, the maximum acceptable duration of the pre-Helmholtz transient for the violin attacks can be estimated to 30–50 ms on the

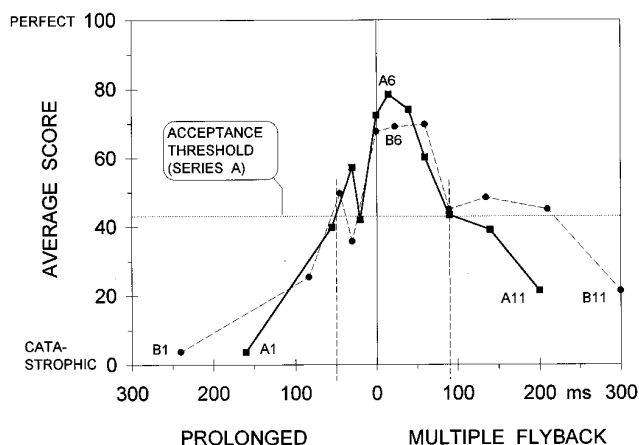


FIG. 2. Judgment of the Quality of the attacks. Averaged scores across 20 subjects for two series of attacks (series A, violin, solid line; series B, “viola,” dashed line). The scores for the artificially accented X versions are not included. The figure shows that attacks with a short duration of the pre-Helmholtz transient are preferred. A calculated threshold of acceptance for series A (43 units) is included for reference (horizontal dotted line). Estimated acceptance limits for series A are shown at 50 ms of prolonged periods and 90 ms of multiple flyback (vertical dashed lines). Interestingly, the data indicate a greater tolerance for multiple-flyback attacks than for attacks with prolonged periods (see text).

prolonged side, and around 90 ms on the multiple flyback side. (The poor rating of attack A4 will be discussed in Sec. I D 3.)

For the “viola” attacks, the acceptance limits seemed shorter, 60 ms on the multiple flyback side, and with no tolerance at all for prolonged-period attacks (0 ms). A strict application of the rough acceptance limit may, however, be slightly misleading in this case, taking into account that attacks B3 (prolonged, 45 ms) and B9 (multiple flyback, 135 ms) just barely fell below 50% acceptance.

2. Quality

The Quality scores for series A and B, averaged across subjects, are shown in Fig. 2, with the prolonged period and multiple-flyback cases arranged mirror symmetrical around the perfect case. When plotted in this manner, the abscissa corresponds (qualitatively) to the continuous progression in bowing parameters from high bow force and low acceleration (prolonged, left), to low bow force and high acceleration (multiple flyback, right).

The score values form a bell-shaped distribution. As expected, the attacks with the highest acceptance are those with a short pre-Helmholtz transient. The peak in the score distribution was, however, not located at zero duration (perfect attack) but rather at about 15–25 ms (3–5 nominal periods) of multiple flyback. This bias, as well as the influence of artificially introduced accents, will be discussed below (Secs. I D 3 and I D 4).

The attacks were classified as accepted or rejected by introducing a threshold of acceptance in the score distribution (see Fig. 2). The threshold was estimated to be 43 units for series A (violin) on the rating scale for quality ranging from 0 (“catastrophic”) to 100 (“perfect”). This value was found by arranging all quality scores for series A in a de-

scending array and selecting the M th element, where M is the total number of *accepted* attacks for series A according to the acceptance judgment above (Sec. I D 1). It would then be expected that a majority of the listeners would judge a violin attack with an average score of less than 43 units as “not acceptable,” and vice versa. Using this definition, all prolonged-period attacks shorter than approximately 50 ms (10 nominal periods) were judged as acceptable (excluding attack A4). The corresponding limit for the multiple-flyback attacks was 90 ms (18 periods).

The same approximate limits in absolute terms (ms) seem to apply to series B (“viola”), taking into account that the acceptance threshold for this series was 51 units. In this case, the 50- and 90-ms limits correspond to only 7 and 12 nominal periods for the prolonged and multiple-flyback attacks, respectively. Some precaution in the interpretation of these results may be necessary, however, as the “viola” examples were artificially derived from the violin attacks. It can be noted in passing that the scores outside the acceptance limits dropped steeper for the violin than for the “viola,” possibly indicating a lower sensitivity to poor “viola” attacks.

As expected, the classification of the attacks as accepted or rejected with the aid of the computed threshold gave essentially the same result as the answers to the Acceptance question, except a slight widening of the acceptance range for series B.

The acceptance limits for the duration of the pre-Helmholtz transient given above form a major result of the study. Following, some particular conditions will be discussed which may have influenced the judgments.

3. Influence of bow acceleration and amplitude buildup time

The bow acceleration and the resulting time for reaching final “steady-state” amplitude (amplitude buildup time) are important for the rating of the quality of an attack. A short buildup time is rated favorable. Two observations illustrate this dependence.

The shift of the peak in the score distribution towards attacks with some multiple-flyback periods (see Fig. 2) is most likely explained by the amplitude buildup time. Two effects seem to combine in this case: (1) A small amount of multiple flyback (15–25 ms) is probably not perceived as disturbing, being masked by the following stronger part of the note; (2) The relatively high acceleration of the bow (which is the primary cause of the multiple flybacks) will force the string to accelerate fast as soon as it has been securely caught by the bow and a regular Helmholtz triggering sets in, giving a short buildup time. Altogether, these two effects can give the impression of a fast and clean attack, which probably explains why the short multiple-flyback attacks A6 and A7 were rated the highest, in preference to the perfect attack.

Also the dips in the averaged score distribution in Fig. 2 are most likely connected with the amplitude buildup time. In particular attack A4 (20 ms, prolonged), and the corresponding “viola” attack B4 (30 ms, prolonged) obtained low ratings. The explanation here would be that attack A4

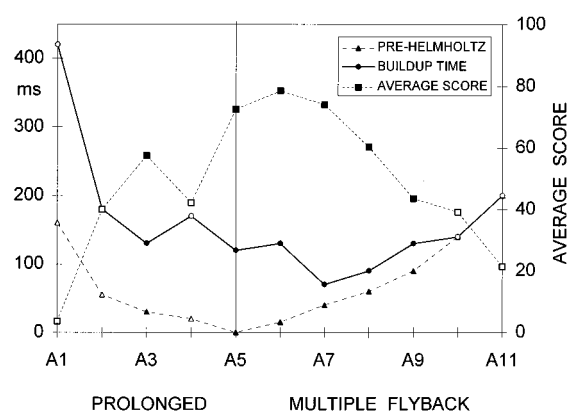


FIG. 3. Comparison between the duration of the amplitude buildup time (solid line) and the pre-Helmholtz transient (dashed line). The average scores for series A from Fig. 2 are included for reference (dotted line). Accepted attacks (above the acceptance threshold in Fig. 2) are shown with filled symbols.

was played with lower bow acceleration than the preceding attack in the series (A3), thus giving longer buildup time (200 ms compared to 160 ms, see Table I). Logically, A4 should have been recorded with a higher bow acceleration than A3, but as mentioned, attacks of suitable duration of the pre-Helmholtz transient were collected by changing the bow acceleration, or friction characteristics, or both.

At this point it is appropriate to ask whether the subjects actually rated the duration of the pre-Helmholtz transient, or rather the amplitude buildup time. A comparison between the buildup time and the duration of the pre-Helmholtz transient shows, as expected, that they are related (see Fig. 3). On the prolonged side, the buildup time was long due to a low bow acceleration (between 420 and 130 ms), and always much longer than the pre-Helmholtz transient. For the perfect attack with no pre-Helmholtz transient (A5) the buildup time was 120 ms. The minimum in buildup time (70 ms) was found a little in on the multiple-flyback side (A7, 40 ms multiple flyback), next to the attack with peak scores (A6). For the longest multiple-flyback attacks (A10, A11) the buildup time equaled the duration of the pre-Helmholtz transient (140 and 200 ms), simply because the bow reached its final velocity before a safe grip of the string had been established.

A comparison between the Quality-score distribution and the buildup times shows clearly that the buildup time alone cannot have served as the basis for the judgment (see Fig. 3). If that were the case, multiple-flyback attacks would have obtained much more favorable ratings in general, as the shortest buildup times were found in this region. In particular, there are no reasons to expect that an acceptance limit in buildup time would be different on the prolonged and multiple flyback side. But in this test, attacks with the same or comparable buildup times were rated quite differently. Three attacks with the same buildup time (130 ms) give a striking illustration. The prolonged-period attack (A3) was gladly accepted, the attack with a short multiple-slip transient obtained peak scores (A6), while the long multiple-flyback attack (A9) was at the threshold of being rejected.

From this digression it is clear that it is the duration of

the pre-Helmholtz transient which serves as the main parameter when judging the acceptance of bowed attacks. A desire for short buildup times bias the judgments towards attacks with a slight amount of multiple flyback (peak scores for attack A6 with 15 ms of multiple flyback). A lengthening in buildup time relative to adjacent cases is immediately reflected in lowered scores (e.g., A4 and A9). Even the attack with peak scores (A6) would probably have been rated still higher, had the buildup time been more in line with the adjacent attacks.

From the player's point of view, a desire for short buildup times means that some of the perfection in the initial Helmholtz triggering may be sacrificed in favor of a shorter buildup time. With this in mind, somewhat excess bow acceleration is safer than too little. All attacks above the acceptance threshold in Fig. 2 had amplitude buildup times shorter than 130 ms (series A), but this value should not be interpreted as an indication of the lower limit in general. In real music performance, a variety of bow accelerations are required and a correspondingly large variation in buildup times will be accepted. When practicing scales and clean attacks, however, the musician is in general striving for the quickest possible response from the instrument, which would explain the preference for short buildup times in the listening test.

4. Effect of accents

The effect of (artificially introduced) accents seemed to be dependent on the quality of the attack. The two artificially accented attacks A5X and B5X, both being perfect attacks with no pre-Helmholtz transient, were rated very similar to their unboosted originals as regards Acceptance (see Table I). The difference in average rating of Quality was only one score unit (1% of rating range) approximately, also indicating a very small perceptual difference.

On the other hand, for attacks with a long multiple-flyback transient as A/B11X (200/300 ms) the introduced accents were perceived as a major improvement by a majority of the listeners. For these attacks, the averaged quality scores increased by 10 and 20 units, respectively, compared to the originals. A possible explanation of this effect might be that an accent gives the impression that the bow has a better grip of the string than is actually the case.

5. Character

The scores for Character for series A and B, averaged across subjects, are shown in Fig. 4. The figure shows that there is a clear relation between the type of string motion during the attack and the perceived character. Attacks with prolonged periods are perceived as "choked/creaky," while multiple-flyback attacks are referred to as "loose/slipping" (with a few exceptions).

Again, accents seem to play an important role, now for the classification of an attack as being of one or the other type (cf. Sec. I D 4). The data in Fig. 4 show clearly that any accent is interpreted in the direction of "choked/creaky," as illustrated by the lowered score values of A/B 5X and 11X, to which artificial accents had been added. In particular, a remarkable shift in rating was observed when an accent was

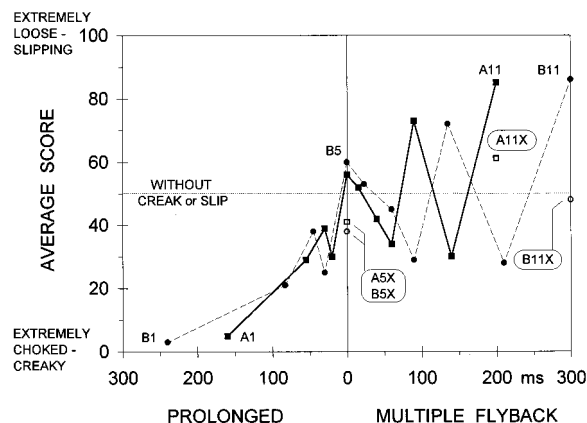


FIG. 4. Judgment of the Character of the attacks. Averaged scores across 20 subjects for the same two series of attacks as in Fig. 2 (series A, violin, *solid line*; series B, "viola," *dashed line*). The scores for the artificially accented X versions are marked separately. The figure shows that attacks with prolonged periods are generally characterized as "loose/slipping," while multiple-flyback attacks are judged as "choked/creaky." Accented attacks tend to be perceived as choked/creaky (see text).

added to the long multiple-flyback attack B11. The average score changed from 85 (corresponding to something like "very loose/slipping") to 48 for B11X ("without creak or slip"). This shift in perceived quality is even more surprising when considering that the duration of the accent corresponded to a few nominal periods only.

The scores on the multiple flyback side in Fig. 4 appears to be less systematic than those on the prolonged period side. In particular, the low scores for attacks A/B 8 and 10 (indicating a "choked/creaky" character) upset the picture. Here too, the explanation is probably connected with accents. In both series, attacks A/B 8 and 10 had (natural) accents, and attacks A/B 6 and 7 sounded accented as well, although by far not as pronounced.

The multiple-flyback examples discussed above and the profound influence of accents clearly illustrate that even advanced string players may have difficulties in classifying bow attacks into the proper categories. A correct classification is necessary, however, in order to be able to adjust the bow acceleration or bow force, if the player is striving for a perfect attack. Although the task is difficult, a player will be in a slightly better position than the subjects in the listening test (or a teacher) when deciding which corrections should be made. Experience tells that, in real playing, a good deal of information on the bow-string interaction can be perceived through the bow and fingertips as a result of the high and rapidly changing frictional forces during the attack.

E. Statistical analysis

1. Analysis of variance

The ratings from the listening test were analyzed by means of analysis of variance, using a mixed model with "Attacks" as a fixed factor and "Subjects" as a random factor (see Table II). This choice of model would allow a generalization of the result of the listening test to represent the opinion of a large population of string students and professionals.

TABLE II. Summary of analysis of variance of the pooled scores for Quality and Character for series A (violin) and B (“viola”). The artificially accented attacks (X attacks) were not included in the analysis.

QUALITY			
Source of variation	df	MS	F value
Attacks (A)	21	28210	55.97 ^a
Subjects (S)	19	6004	20.49 ^a
Interaction (A×S)	399	504	1.72 ^b
Within cell (w)	880	293	
CHARACTER			
Source of variation	df	MS	F value
Attacks (A)	21	31611	117.51 ^a
Subjects (S)	19	12570	7.22 ^a
Interaction (A×S)	399	269	1.55 ^b
Within cell (w)	880	174	

^a $p < 0.001$.

^b $p < 0.01$.

A preliminary inspection of data showed that the artificially accented attacks (X attacks) confused the subjects in their judgments. A plausible reason would be that these attacks sounded artificial in the word’s literal meaning (“lacking in natural quality”), and hence was not possible to rate with reference to the subjects’ experiences as string players. Whatever the cause, these attacks were discarded from the subsequent analyses.

The analysis of variance showed, as expected, that there were significant differences in average scores between the attacks, both in Quality and Character (see Table II, factor A). A nonsignificant result in this factor would have been very disappointing as the attacks were carefully chosen to represent a wide spread in these respects. Further, there were significant differences between the ratings given by different subjects (S), meaning that the subjects used different parts of the rating scale.

A more interesting question is whether the *differences* (in average scores) between the attacks were the same or not for different subjects, regardless of the subjects’ mean position on the rating scale. As seen in Table II, the (A×S) interaction was significant, which means that the differences between the mean ratings of individual attacks were not similar for all subjects. This fact is by no means alarming in itself, as will be discussed in Sec. I E 3 (“percent of variance accounted for”), but it illustrates that the subjects did not quite agree upon how much the attacks differed.

The subjects did agree on the *ranking* of the attacks, however, as shown by a rank correlation test.⁵ This test checked whether the subjects ranked the attacks in the same order, with no attention paid to the score magnitudes. The result of this test was highly significant ($p < 0.001$).

All results and conclusions for Quality scores given in this section apply to the Character scores as well (see Table II).

2. Intra-individual reliability

The intra-individual reliability, or the consistency of the ratings within each subject, was checked by calculating a

TABLE III. Percent of variance accounted for by different sources.

Source	Quality	Character
Attacks (A)	50	70
Subjects (S)	9	2
Interaction (A×S)	7	4
Error (“within cell”)	34	24

“Reliability index for individual subjects” (r_w).^{6,7} An interpretation of this measure would be that if the listening test was repeated with the same subject under the same conditions, the correlation between the mean ratings for the individual attacks in the two tests would be approximately r_w . As a rule of thumb, subjects with r_w above 0.50 can be considered as acceptable in a listening test.

In addition, the error variance for each subject $MS_{\text{within cell}}$ was calculated. This measure indicates how much a subject varies in repeated ratings of the same attack during the test. An approximate upper limit for MS_w would be 150 units squared $\approx (12 \text{ units})^2$ (Ref. 6).

All subjects showed a satisfactory reliability in their judgments according to the two measures above. For the Quality (Character) scores, 11 (15) subjects out of 20 had a reliability index above 0.90, the lowest being 0.75 (0.85). The median of the MS_w values was 59 (25).

3. Inter-individual reliability

The inter-individual reliability, or the agreement between the ratings given by different subjects, was estimated by a “Reliability index between subjects” (r_B).⁷ This index has a powerful interpretation which makes it particularly informative. If the listening test was to be repeated with another random sample of 20 subjects from the same population (advanced string students and professionals), the correlation between the mean ratings for the individual attacks in the two tests would be approximately r_B . In the present case, r_B turned out to be as high as 0.99 both for Quality and Character, which indicates an excellent agreement between subjects, and consequently a high reliability of the mean ratings presented in Fig. 2. A very similar result would have been obtained with another sample of subjects.

Still more information on the reliability can be obtained from a calculation of how much of the variance in the subjects’ data is accounted for by different sources (see Table III).^{7,8} These calculations revealed that by far the largest part of the total variance in the ratings, 50% for Quality (70% for Character), can be attributed to differences between the attacks. A sequential analysis of the averaged ratings showed no appreciable learning effects between the first and third rating of the same attack.

A closing remark on the subjects’ task might be appropriate. It can be argued that the sound generated in such a complex process as a bowed attack must be related to several perceptual dimensions, and a straightforward judgment in one or two parameters may not be meaningful. Before the listening test started, some subjects actually expressed concerns about summarizing the quality of an attack in a single rating. Reference was given to the many different ways of



FIG. 5. Music used in the playing test, representing different styles of bowing. (a) A simple tune “Twinkle, twinkle, little star,” *tenuto-détaché-martellato*, performed in tempo M.M.=100–110 beats/min; (b) Theme from Beethoven’s violin concerto, Op. 61, *tenuto*, M.M.=93; (c) Excerpt from Preludio V from Das Wohltemperierte Klavier by J. S. Bach, *détaché*, M.M.=60; (d) Theme from L’Arlésienne suite by Bizet, *martellato*, M.M.=110; (e) Theme from Wilhelm Tell overture by Rossini, *ricochet*, M.M.=125. All examples were performed both on the G string as shown, and on the E string transposed an octave plus a major sixth.

starting notes which are used in contrasting musical contexts. Their concerns faded away, however, as soon as the introductory sound examples were presented, indicating that a single perceptual dimension, in this experiment referred to as Quality, was possible to rate on a scale between “Catastrophic” and “Perfect.”

II. PLAYING TEST

A. Method

A playing test was performed in order to study the pre-Helmholtz transients in real violin playing, in particular the distribution of transient durations. Two professional violinists performed a simple tune (the first phrase of “Twinkle, twinkle little star”) on the G string (G major), and on the E string (E major), respectively [see Fig. 5(a)]. Three bowing styles were used; *tenuto* (“held,” broad strokes, fully sounding through the entire note value), *détaché* (“separated,” articulated attacks with no or little release of bow force between notes), and *martellato* (“hammered,” very pronounced attacks, anticipating each stroke with extra bow force and releasing the force rapidly after the onset). The 14 notes in the phrase were played at a tempo of 100–110 beats/min, using alternating up and down-bows, with one exception at which two up-bows followed in succession [see Fig. 5(a)]. Each manner of bowing was repeated four times (three times by the second performer) and at three dynamic levels (p – mf – f), resulting in a total of $2 \times 3 \times 4(3) \times 3 = 72$ (54) renderings by the two performers, respectively, corresponding to 1008 (756) attacks.

A two-channel recording was made on a DAT recorder, documenting the string velocity at the bowing point,³ and the

sound pressure close to the violin (measured by a dedicated violin microphone mounted to the instrument).⁹

A fine violin of professional quality was used,¹⁰ and both players were most satisfied with the instrument. They also assured that the minor preparations of the violin for the recording did not disturb their playing. Both players used their own bows. The players were not informed about the purpose of the experiment, but were only given the music with the instructions of bowing styles and dynamic levels and asked to perform accordingly. If a player was not satisfied with a certain rendering, this version was deleted and replaced with a new recording.

The recordings were analyzed using a sound file editor.⁴ The duration of the transient state before Helmholtz triggering was measured with the aid of the string velocity signal, and the attacks were classified into three categories (*prolonged*, *perfect*, and *multiple flyback*). The criteria for classifying an attack as “perfect” were: (1) the string velocity signal should show a smooth buildup in flyback velocity amplitude, and (2) only occasional drops in velocity (down to zero or crossing the zero line) should take place during the nominal sticking part of the period. Minor partial slips with no dynamic consequences for the development of the tone may thus have occurred also in attacks classified as perfect.

In order to avoid ambiguous classification of “almost perfect attacks,” an short initial buffer zone was introduced. All attacks with an initial aperiodic transient shorter than 5 ms were classified as perfect. This leeway was particularly apt for the classification of attacks at p level. A total of 1694 out of the 1764 recorded attacks (96%) were successfully analyzed and classified.

B. Results

1. A simple tune

a. G string. The results for the simple tune played by the first subject is shown in Fig. 6, in which distributions of the pre-Helmholtz transient durations are displayed separately for the three dynamic levels, and also summarized in a pooled plot.

The figure shows, first of all, a good agreement between the actual durations of the pre-Helmholtz transients which occur in normal playing, and the limits for acceptable durations obtained in the listening test. Very few prolonged-period attacks exceed 50 ms, and attacks with multiple flyback are normally not longer than 90 ms. Between 80% and 90% of the attacks (depending on bowing style) stayed between these limits. Such an agreement is not obvious as the players in no case were instructed to play the notes with “neutral” attacks, which, in contrast, was the rating reference in the listening test.

Between 20% and 50% of the attacks were perfect, depending on the dynamic level and bowing style. The highest rate of perfect attacks which occurred in the experiment (79%), was observed for the second subject in *martellato* at p level. It is to be noted, however, that out of the 121 analyzed recordings of the tune, only *one* had all 14 consecutive attacks classified as perfect. Although such a perfect series thus seems to be rare in normal performance, informal ex-

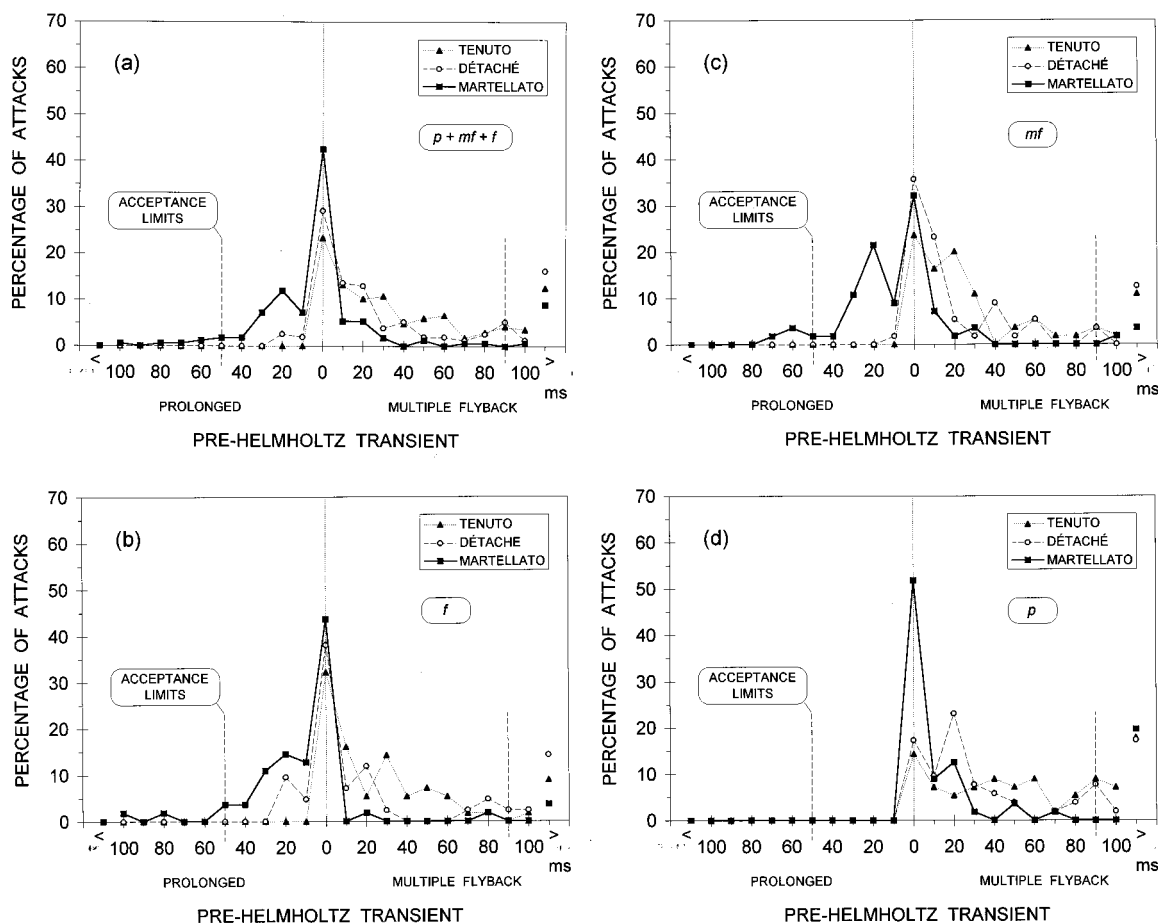


FIG. 6. Distributions of the durations of pre-Helmholtz transients in a simple tune ("Twinkle, twinkle, little star") played on the G string using three different bowing styles (*tenuto*, *détaché*, and *martellato*) by one performer. (a) Pooled plot showing three dynamic levels, $p-mf-f$. Each bowing style was repeated four times at each dynamic level (with one exception) giving a total of 35 repeats, corresponding to 490 analyzed attacks. The class width is 10 ms centered around the tick mark values. Transients with durations longer than 105 ms are summarized in the outermost data points (indicated by < and >). The acceptance limits obtained in the listening test are included for reference (vertical dashed lines). Separate plots of the distributions at different dynamic levels are shown in: (b) *forte* (168 attacks); (c) *mezzo forte* (154 attacks); and (d) *piano* (168 attacks).

periments showed that a professional string player can easily reproduce a sizable series of consecutive perfect attacks when concentrating on this task, for example, during a practice session.

As regards the multiple-flyback attacks, a certain amount of very long pre-Helmholtz transients were observed. These ranged between 100 and 600 ms, and would in some cases last the entire note. Even so, the perceived pitch did not deviate from the intended (by an octave or some other integer ratio), but the string spectrum showed a weak fundamental and sometimes also weak odd partials. Such notes with long multiple-flyback attacks are perceived as having a "surface" quality, giving associations to cautious string playing with light bow pressure rather than to a poor attack. These notes with long multiple-flyback attacks were observed preferably in *tenuto* and *détaché*, being most sustained at p level.

The influence of dynamic level and bowing style was striking [see Fig. 6(b)–(d)]. The perfect attacks showed the highest rate for all bowing styles at all dynamic levels. The distributions on either side of the perfect class differed, however. At p level only perfect and multiple-flyback attacks were observed, some of them being very long, as discussed

above. At mf and f levels, multiple-flyback attacks were still the dominating type for *tenuto* and *détaché*, while the *martellato* strokes were played with a pronounced rate of prolonged-period attacks. For the second subject, being more prone to high bow pressures,¹¹ almost all attacks were of the prolonged period type at f level, also for *tenuto* and *détaché* bowing.

The analogy between a *martellato* stroke and a consonant in an initial position in speech is evident. Such a "transient" in speech with a typical duration of 70 ms, is, however, slightly longer than most pre-Helmholtz transients in *martellato*.

Some notes are easier to start than others. The second of the two notes played in two successive up-bows [see Fig. 5(a)] showed markedly shorter attack durations, suggesting that this note was particularly easy to start. Both players played this note with a perfect attack in 12 successive recordings. The explanation would be that the Helmholtz corner is already circulating on the string in the same direction as the second bow stroke will induce. Accordingly, no phase reversal of the string motion needs to take place, and the frictional forces during the attack of the second note will be within the

allowed range for a perfect attack in a larger number of cases.

b. E string. The attack durations for the E string were distributed within the same limits as for the G string (<50 ms, prolonged periods; <90 ms, multiple flyback), but with a clear shift towards prolonged periods. This was in particular the case for the martellato strokes for which the multiple-flyback attacks gave way for perfect attacks at *p* level, and where essentially all attacks were found on the prolonged side at *mf* and *f* levels. The peak in the pooled martello distribution ($p + mf + f$) was located at 10–20 ms of prolongation instead of at zero duration (perfect attack), which was the case for the G string.

Possibly, the explanation could be sought in the need for safety margins in bowing. String players often attack the open top string with excess bow force (and/or lower velocity) in order to establish a safety margin against long multiple-flyback attacks. Thin strings, like the top E string on the violin, are much more inclined to “whistling” (lasting multiple-flyback motion) than thicker strings such as the low G string. The reason is that the reflections returning to the bow are better defined on a thin string. The need for larger safety margins in the playing of the open string might have infected the bowing of the stopped E string as well, although much less needed. The kinks are now reflected at the soft finger pad and loose in definition, and an excessive bow pressure only results in prolonged periods, as observed.

A comparison between the accepted attacks for the G and the E string, respectively, shows that fewer periods were spent before the triggering of the Helmholtz motion when playing on the G string. As the acceptance limits were the same in absolute terms, the number of periods in accepted attacks was only 30% (196 : 660 Hz) for the G string compared to the E string. An extension of this result to the lower string instruments, assuming an approximately constant upper limit for the attack duration in absolute terms (50 and 90 ms), would imply a need for very high precision in the control of the bowing. In these terms, an acceptable attack on the lowest string on the cello and double bass would correspond to a few nominal periods only (3–6 periods for cello; 2–4 periods for double bass).

2. Beethoven, Bach, and Bizet

As a complement to the simple tune, the first subject was asked to perform some music from the violin literature of his own choosing, representative of the three types of bowing. Excerpts from the following pieces were selected; Beethoven violin concerto (*tenuto*), a Bach prelude (*détaché*), and L’Arlésienne suite by Bizet (*martellato*), all performed on the G string [see Fig. 5(b)–(d)].

The distributions of attack durations for these excerpts were rather similar to the corresponding versions of the simple tune, taking into account that the Bizet example was performed fairly loud (about *f*), and the Beethoven and Bach examples much softer (*mf*) (cf. Figs. 6 and 7). This result suggests that the particular choice of music, given a certain dynamic level and bowing style, does not affect the distribution of the durations of the pre-Helmholtz transients dramati-

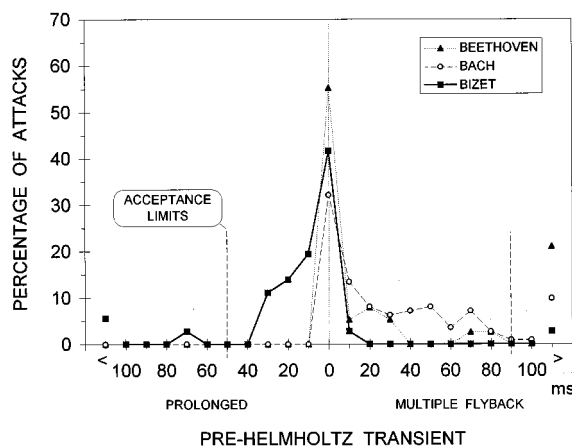


FIG. 7. Distributions of the durations of pre-Helmholtz transients in performances of excerpts from the classical violin literature representing different bowing styles, including theme from Beethoven’s violin concerto, *tenuto* (39 attacks), excerpt from Preludio V from Das Wohltemperierte Klavier by J. S. Bach, *détaché* (112 attacks), and theme from L’Arlésienne suite by Bizet, *martellato* (36 attacks). All examples played on the G string.

cally. Within the “archetype” distributions, variations will, of course, occur, reflecting the mood and characteristics of the piece among other things.

3. Spiccato, sautillé, and ricochet

Other types of bowings were tried as well, including *spiccato* (“cut off,” with a springing bow) in eighth notes, a faster version (*sautillé*) in sixteenth notes, and a very rhythmic *ricochet* (“rebounding”) bowing à la Wilhelm Tell overture, combining several bouncing strokes in one direction [see Fig. 5(e)].

In spiccato, very few multiple-flyback attacks were observed (see Fig. 8). The prolonged-period attacks stayed well within the 50-ms limit found earlier, with few exceptions. The perfect attacks were particularly numerous for spiccato on one string (played in groups of four with three consecutive notes of the same pitch) reaching 64%. As soon as a second string was interleaved every other note, the “perfect” rate lowered to 33%.

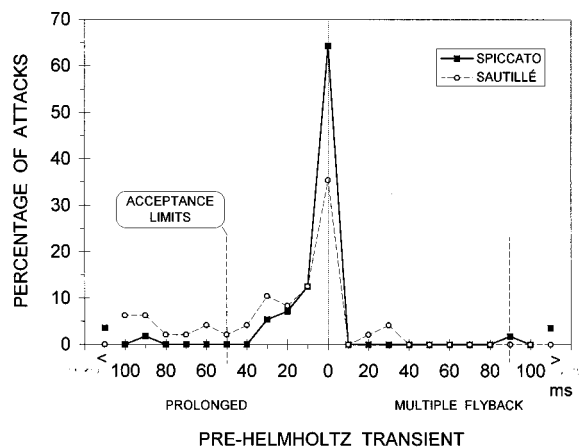


FIG. 8. Distributions of the durations of pre-Helmholtz transients in *spiccato* on one string (56 attacks), and *sautillé* (48 attacks), both examples played on the G string.

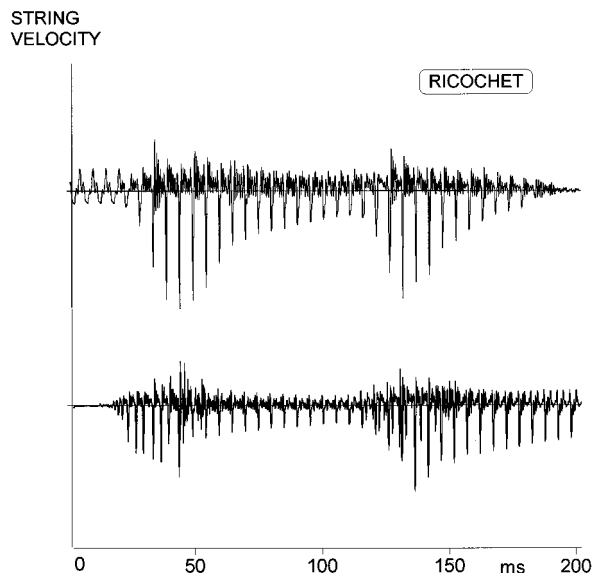


FIG. 9. Examples of excellent (*top*) and poor attacks (*bottom*) of sixteenth notes in the ricochet example, showing the string velocity close to the bowing point. Note that in the upper example a decaying string vibration (mainly fundamental) is present at the onset of the first note.

Sautillé bowing gave a higher rate of long prolonged-period attacks, up to 100-ms duration, and a “perfect” rate of 35% (see Fig. 8). Short multiple-flyback attacks occurred as well, altogether giving a much less clear perception of each individual note compared to the spiccato.

The rapid ricochet bowing gave no multiple flybacks, a high rate of perfect attacks (45%), and the main part of the prolonged-period attacks within the 50 ms limit. A certain amount of longer prolonged attacks, up to 100 ms, was observed as well. For a sixteenth note, which in this example lasted between 70 and 100 ms (corresponding to 14–20 nominal periods only), this means that a major part of the note was spent on the pre-Helmholtz transient. This “failure” occurred in almost 20% of the sixteenth notes in the ricochet example. Examples of excellent and poor ricochet attacks are shown in Fig. 9.

In addition to the results reported above, the playing test offered several interesting comparisons between the actual string motion and the perceived tone quality, which would motivate a larger separate study.

III. CONCLUSIONS

The study shows that string players are very sensitive to the quality of bowed attacks, and that they agree in their preferences. A narrow span defines the allowed duration of transients in acceptable attacks. For the main parameter investigated, that is, the duration of the aperiodic part of the attack before the triggering of the Helmholtz motion (the *pre-Helmholtz transient*), the acceptance limits were 50 ms (prolonged periods) and 90 ms (multiple flyback). In terms of nominal periods these limits correspond to 10 and 18 periods, respectively (violin, open G string). Attacks with pro-

longed periods are perceived as being “choked/creaky,” while multiple-flyback attacks are characterized as “loose/slipping.”

The same approximate limits in absolute terms (ms) were obtained for “viola” attacks (obtained from transposed violin attacks), the number of nominal periods now being reduced to only 7 for prolonged-period attacks, and 12 for multiple-flyback attacks.

An analysis of variance using a mixed model with subjects as a random factor showed a high reliability in the data, which indicates that the results of the listening test can be generalized to represent the opinion of a large population of string students and professionals.

The acceptance limits were first obtained in a listening test in which advanced string students and professionals judged the acceptance of recorded attacks with reference to a neutral attack. Following, a playing test with two professional violinists verified that the same approximate limits apply to the pre-Helmholtz transients in normal violin playing as well. As a rule of thumb, only 20%–50% of the notes in a performance can be expected to start with a perfect attack, showing a periodic Helmholtz triggering from the very beginning. On the other hand, as much as 80%–90% of the attacks will stay within the acceptance limits (50/90 ms).

A closer examination of the data indicates that several parameters other than the duration of the pre-Helmholtz transient play important roles in the perception of bowed attacks.

Accents. Any initial accent would shift the perceived character in the direction of “choking the string.” This tendency is, however, probably strongly related to the musical context. In the listening test the attacks were judged with reference to a “neutral” context, like when practicing a scale.

Amplitude buildup time. A low bow acceleration which results in a slow buildup to final “steady-state” amplitude, after the Helmholtz triggering has been established, adds to the impression of the string being choked.

Nominal fundamental frequency. The results of the listening test as well as the playing test suggest that the acceptance limits for the duration of the pre-Helmholtz transient is about equal for different fundamental frequencies when expressed in absolute terms (ms). In the listening test the fundamental frequencies were limited to 196 and 130 Hz (open violin G string and “viola” C string), while a larger range was sampled in the playing test (196–940 Hz). If this finding was verified in a formal study, it would mean that instruments operating in a lower fundamental frequency range (viola, cello, double bass) cannot afford as many periods as the violin before reaching Helmholtz triggering. This would imply a narrower range in bowing parameters for these instruments, closer to the perfect conditions.

It may well be, however, that the upper frequency limit in the spectra and the rate at which the spectral changes take place—parameters which both are dependent on the pitch range—play a role for the perceived attack quality. A hint in this direction is found in Fig. 2 where the rejected attacks (below the acceptance threshold) were rated lower for the violin than for the “viola.” This result was verified in a smaller preliminary test, in which the fundamental frequency

of series B was lowered to 65 Hz, comparable to the open C string of a cello. The reason may be that a poor attack sounds less erroneous when the spectrum changes more slowly. A consequence of that would be that bowing errors are less dramatic for the lower-pitched bowed instruments than for the violin. A dedicated study would be needed to settle these questions.

Pre-Helmholtz transients with durations well outside the limits found in this study certainly do occur in professional performances. Although all string players spend a considerable part of their studies learning to master the bow-string control, not all professionals are able to perform attacks at the same quality level. Nevertheless, the ideal remains the same, a perfect attack with periodic Helmholtz triggering from the very start. Another aspect of the pre-Helmholtz transient is the occasional musical desire to create tone colors outside of the *beau idéal*. For instance, a series of soloistic sforzandi may consist of prolonged periods only, while a soft (“from nothing”) entrance of a string section may show several seconds of nothing but multiple-flyback attacks.

ACKNOWLEDGMENTS

The authors are indebted to a large group of string students and teachers at the Norwegian State Academy of Music, Oslo, and to violinists Semmy Lazaroff, Royal Philharmonic Orchestra, Stockholm, and Per Sandklef, Swedish Radio Symphony Orchestra, for their kind participation in

the experiments. The guiding expertise of Alf Gabrielsson, Uppsala University in the statistical analysis is gratefully acknowledged. Several thoughtful comments by an anonymous reviewer helped us improve the manuscript. This study was supported by the Swedish Natural Science Foundation and the Nordic Research Council.

¹M. E. McIntyre and J. Woodhouse, “A parametric study of the bowed string: The violinist’s menagerie,” *J. Catgut Acoust. Soc.*, No. 42, 18–21 (1984).

²A. Cronhjort, “A computer-controlled bowing machine,” *Speech Transmission Laboratory Quarterly Progress and Status Report*, Dept. of Speech Communication and Music Acoustics, Royal Institute of Technology, Stockholm, STL-QPSR 2-3/92, 61–66 (1992).

³The string velocity was measured by applying a small magnet under the string at the bowing point and recording the voltage across the string. The setup gave a measurement range of 10 kHz (−4 dB).

⁴SWELL Soundfile Editor from AB Nyvalla DSP, Teknikhögden, S-104 05 Stockholm, Sweden.

⁵G. K. Kanji, *100 Statistical Tests* (Sage, London, 1994), p. 115.

⁶B. J. Winer, *Statistical Principles in Experimental Design* (McGraw-Hill, New York, 1971), 2nd ed.

⁷A. Gabrielsson, “Statistical treatment of data from listening tests on sound-reproducing systems,” Karolinska Institute, Dept. of Technical Audiology, Report No. TA 92, 1979.

⁸R. E. Kirk, *Experimental Design: Procedures for the Behavioral Sciences* (Brooks-Cole, Belmont, CA, 1982), 2nd ed.

⁹Brüel & Kjaer 4021 with mounting VH4000.

¹⁰Léon Bernardel, Paris, 1909.

¹¹A. Askenfelt, “Measurement of the bowing parameters in violin playing. II: Bow-bridge distance, dynamic range, and limits of bow force,” *J. Acoust. Soc. Am.* **86**, 503–516 (1989).

Sound production in recorderlike instruments. I. Dimensionless amplitude of the internal acoustic field

Marc-Pierre Verge^{a)} and Benoit Fabre

Laboratoire d'Acoustique Musicale, CNRS, Université de Paris VI, Ministère de la Culture, 4, Place Jussieu, case 161, 75252 Paris Cedex 05, France

A. Hirschberg and A. P. J. Wijnands

Technische Universiteit Eindhoven, W & S 0.54, Postbus 513, 5600 MB, Eindhoven, The Netherlands

(Received 14 June 1996; revised 22 November 1996; accepted 17 December 1996)

Data on the internal acoustic pressure signals in an experimental recorderlike flue organ pipe are presented. A dimensionless representation appears to be a powerful basis for the analysis of these data. The dimensionless amplitude of the fundamental is, for a given geometry, a function of the Strouhal number only. For the first hydrodynamic mode of the jet, this amplitude is independent of the acoustic mode involved. The dimensionless amplitude of the second harmonic displays two different behaviors, depending on whether the jet is laminar or turbulent. This specific Strouhal number dependency implies a strong influence of the distance W from the flue exit to the labium on the timbre of recorderlike instruments. In recorders the ratio W/h , where h is the height of the flue exit, is adjusted by craftsmen to a value close to 4. This specific ratio ensures an optimal harmonic content to noise ratio in the produced sound. Increasing the distance W yields a more powerful sound but at the expense of additional turbulence noise. For these large ratios W/h , transitions between different hydrodynamic modes of the jet are observed at low blowing pressures. This phenomenon is avoided by the choice of a recorderlike geometry. © 1997 Acoustical Society of America. [S0001-4966(97)05904-3]

PACS numbers: 43.75.Np, 43.28.Ra [WJS]

INTRODUCTION

Self-sustained oscillations of recorderlike instruments involve a combination of very complicated hydrodynamic phenomena occurring in the embouchure of the instrument. In this tiny region one can observe simultaneously an hydrodynamically unstable jet, vortex shedding, and turbulence, all interacting with the acoustic field from a resonator. The modeling of the functioning of recorderlike instruments is further complicated by their great variety, characterized by their own geometry and blowing pressure range,¹ which results in a wide diversity of timbres and tone qualities. The behavior of recorderlike instruments appears, however, to be governed by common mechanisms which are grasped by a dimensionless representation of measurements of internal acoustic signals. We propose the use of the dimensionless ratio obtained by dividing the mean acoustic velocity through the embouchure of the instrument with the jet velocity. This dimensionless ratio indeed appears to obey a specific Strouhal number dependency for a wide variety of geometries and playing conditions. This type of representation is therefore helpful in understanding the important phenomena involved in the sound production and in determining the critical parameters that should be considered when modeling this type of instruments. It also provides quantitative criteria to evaluate numerical simulation results, as is shown in the companion to this paper.

The different mechanisms involved in the sound produc-

tion of recorderlike instruments are first reviewed. Experimental data on the internal pressure signal are then presented, for a specific instrument geometry, in a dimensionless representation. This representation enables one to gain some insights in the amplitude limiting mechanisms involved in the functioning of the instrument. Similar data obtained with different mouth geometries are then presented in order to demonstrate the generality of the results. The effects of the change in the mouth geometry on the tone produced by the instrument are also discussed. In the following section the importance of the different dissipation mechanisms identified is evaluated quantitatively and the behavior of our experimental flue pipe is compared to that of other aeroacoustic systems based on self-sustained flow oscillations. Finally, a simple model of the effects of acoustically induced vortex shedding at the edge of the labium is proposed and discussed.

I. SOUNDING MECHANISMS OF RECORDERLIKE INSTRUMENTS

Recorderlike instruments produce sound as a result of the coupling between a hydrodynamically unstable flow and a resonant acoustic field in a pipe. The geometry of a typical recorderlike instrument is shown in Fig. 1. The flow is generated by blowing through a narrow windway called the *flue*. At the flue exit, a free jet is formed by flow separation. This jet flows across the *mouth* or *window* of the instrument and is directed toward a sharp edge called the *labium*. At the flue exit, the jet is submitted to the transverse acoustic flow due to the oscillation in the pipe of the instrument. Because of its intrinsic instability, the jet is very sensitive to these pertur-

^{a)}Part of this work was done while the author was at IRCAM and at the Technische Universiteit Eindhoven.

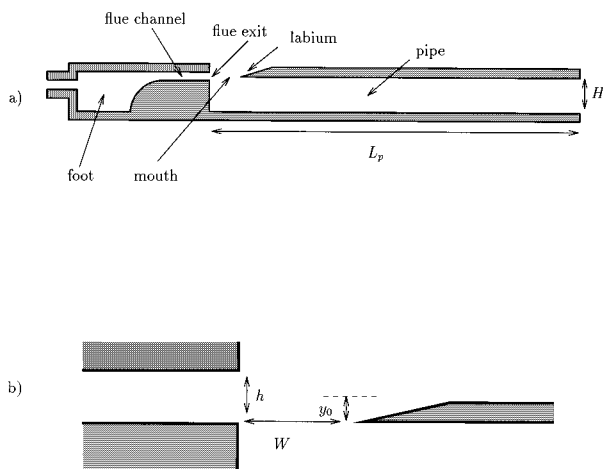


FIG. 1. (a) Typical geometry of a recorderlike instrument and (b) geometry of the mouth of the instrument. For the experimental flue organ pipe used: $L_p = 0.283$ m, $H = 0.020$ m, $h = 0.001$ 08 m, and W is adjustable. The origin of the transverse position y_0 of the labium is in the middle of the flue exit.

bations. They propagate on the jet and are amplified as they travel toward the labium. This results in a flipping of the jet on each side of the labium at the same frequency as that of the acoustic field. This motion of the jet around the labium provides energy to the acoustic field which enables its maintenance. This energy can accumulate in standing waves corresponding to the resonances of the pipe of the instrument which favors oscillations at specific frequencies. The problem of the modeling of the jet oscillation and its coupling with the acoustic field from the pipe was discussed by Verge *et al.*² The driving effect of this motion of the jet was also evaluated, in the same paper, in the potential flow approximation.

The functioning of recorderlike instruments is therefore usually represented by a feedback loop³ where the jet-edge system, the excitator, is nonlinearly coupled to the pipe of the instrument represented as a linear acoustic filter. In this representation, a steady-state amplitude is reached by saturation of the jet flow on each side of the labium when the amplitude of the jet oscillations at the labium becomes greater than the half-width of the jet. The amplitude of the steady-state internal acoustic field of the resonator is then determined by a balance between the energy supplied by the jet oscillations and the loss mechanisms of the system namely the visco-thermal losses in the resonator and sound radiation at the mouth of the instrument and at the passive extremity of the pipe.

Early models of the functioning of flue instruments by Cremer and Ising,⁴ Coltman,^{5,6} Elder,⁷ and Fletcher^{8,9-11} are based on variants of this representation. Although these models have considerably contributed to the understanding of the sounding mechanisms of flue instruments, they do not enable one to correctly predict basic features of the tone of flue instruments such as the steady-state amplitude and spectrum. This type of representation indeed neglects details of the flow at the labium which appear to be fundamental in the functioning of the instrument.

The importance of vortex shedding at the edge of the

labium has been pinpointed by Howe.¹² Using a simplified model in which only a single acoustic pipe mode is taken into account, Fabre^{13,14} showed that during steady-state oscillation of the instrument, the shedding of vortices at the edge of the labium, due to separation of the acoustic flow from the pipe, constitutes the dominating nonlinear saturating mechanism for the fundamental of the acoustic oscillations of the instrument. It is therefore essential, when trying to determine the playing amplitude of flue instruments, to take into account the sound source associated with vortex shedding. Furthermore, while the separation of the acoustic flow at the labium represents a damping mechanism for the fundamental, Fabre also showed that it can represent a source mechanism for the second and higher harmonics. Vortex shedding therefore also contributes to determine the tone of the instrument. These effects of vortex shedding had already been pointed out by Coltman⁵ when he found out that the mouth impedance of a flute behaves nonlinearly at acoustic amplitudes corresponding to playing conditions. The modeling of this vortex shedding at the labium will be discussed in more detail in Sec. V. Vortex shedding at the labium has also been shown by Verge *et al.*¹⁵ to be crucial in the triggering of the attack transient.

For high enough jet velocities, the jet flow becomes turbulent which provides another sound source. Turbulence generates a broadband noise which corresponds to the whistling and windy sounds that are very typical of the timbre of flue instruments. The production of sound by turbulence is further enhanced in this kind of instruments by the presence of the labium in the jet flow.¹⁶ This aspect of sound production will be discussed in the companion to this paper.

An interesting and powerful way to study the tone produced by flue instruments is to perform a dimensionless analysis of the internal pressure signal measured in the pipe of the instrument. This kind of analysis clearly brings out the importance of the aeroacoustic sources that have just been discussed and also helps to understand some effects of the instrument geometry on the tone it produces.

II. DIMENSIONLESS PLAYING AMPLITUDE OF RECORDERLIKE INSTRUMENTS

The fundamental frequency f_1 of the internal pressure signal $p'(x_m)$ measured in a small experimental pipe is shown in Fig. 2 as a function of the blowing pressure p_f . The instrument used for these measurements is the recorderlike experimental flue organ pipe, having a length L_p of 0.283 m and a square cross section $H^2 = (0.02 \text{ m})^2$, described in detail by Verge *et al.*¹⁵ For these measurements, the ratio W/h between the distance W from the flue exit to the edge of the labium and the height h of the flue exit was adjusted to a value close to 4, which is typical of the geometry of recorders. The measurements were performed at position $x_m = 50.25$ mm (where x_m is measured from the passive end of the pipe). Operation of the system in the different acoustic modes of the pipe ($L_p \approx i\lambda/2$, $i = 1, 2, 3, \dots$, where λ is the wavelength of the acoustic oscillation in the pipe) is clearly observed. The corresponding amplitude of the pressure signal $p'(x_m)$ is shown in Fig. 3 as a function of the driving pressure p_f . Note the gap between operation on the first and

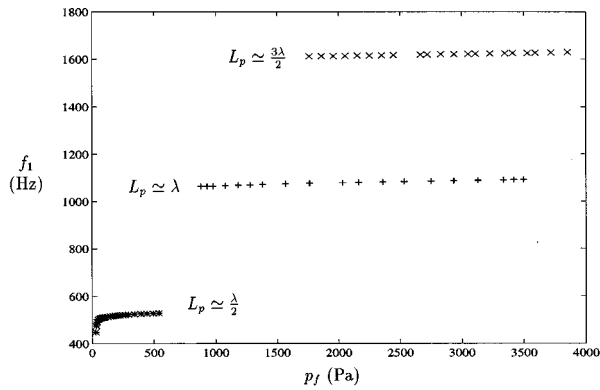


FIG. 2. Operating fundamental frequency f_1 of the experimental flue organ pipe as a function of the driving pressure p_f . The edge of the labium was placed 4.02 mm in front of the flue exit ($W/h=3.72$) and its transverse position y_0 was adjusted with the lower wall of the flue exit ($y_0/h=0.5$, the origin of y_0 is in the middle of the flue exit).

second mode of the pipe and the overlapping of the curves corresponding to the second and third modes of the pipe due to an hysteresis in the behavior of the system.

In Fig. 4(a), the same information is displayed in a dimensionless representation. The data are presented as the ratio

$$\frac{u_i}{U_j} = \frac{\alpha_i}{\rho_0 c_0 U_j} \frac{H}{W} \quad (1)$$

of the mean acoustic velocity $u_i = \alpha_i / \rho_0 c_0 (H/W)$ through the mouth of the instrument and the maximum jet velocity $U_j = \sqrt{2p_f / \rho_0}$, where α_i is the amplitude of the i th mode and ρ_0 is the air density. Note that in calculating the acoustic velocity u_i in the mouth of the instrument, the acoustic flow at the mouth was implicitly assumed to have the same amplitude as at the passive end of the pipe, which is not exactly true. The amplitude α_i of the different pressure modes can be obtained from the amplitude $p_i(\omega_i)$ of the harmonics of the

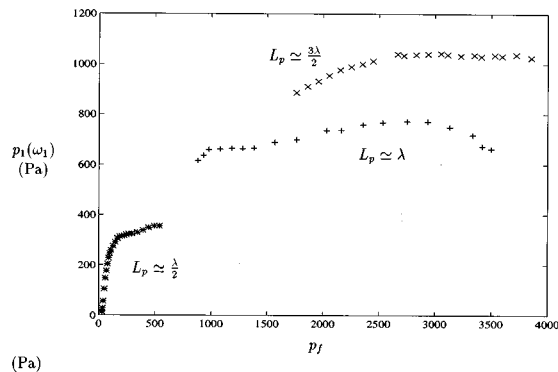


FIG. 3. Peak amplitude $p_1(\omega_1)$ of the fundamental as a function of the driving pressure p_f for the small experimental flue organ pipe. The microphone was placed at $x_m=50.25$ mm from the passive end of the pipe. The edge of the labium was placed 4.02 mm in front of the flue exit ($w/h=3.72$) and its transverse position y_0 was adjusted with the lower wall of the flue exit ($y_0/h=0.5$, the origin of y_0 is in the middle of the flue exit). The amplitude of the fundamental is isolated from the pressure signal by using a bandpass filter.

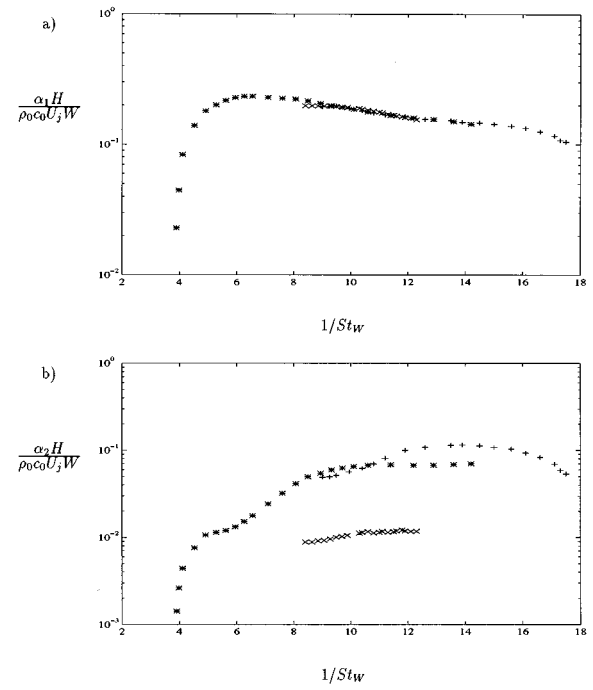


FIG. 4. (a) Dimensionless amplitude α_1 of the fundamental and (b) α_2 of the second harmonic as a function of the Strouhal number $St_W = f_1 W / U_j$ for operation of the experimental flue organ pipe in the first three acoustic modes of the pipe: *, first mode; +, second mode; x, third mode). The edge of the labium was placed 4.02 mm in front of the flue exit $W/h=3.72$ and its transverse position y_0 was adjusted with the lower wall of the flue exit ($y_0/h=0.5$, the zero of the transverse position is in the middle of the flue exit).

measured pressure signal $p'(x_m)$ by means of the following expression

$$\alpha_i = \frac{p_i(\omega_i)}{\sin\left(\frac{\omega_i(x_m + \delta_p)}{c_0}\right)}, \quad (2)$$

for the i th harmonic, having an angular frequency $\omega_i = i\omega_1$ ($i=1,2,3,\dots$), δ_p is the end correction of the pipe, and c_0 the speed of sound. Data on the value of δ_p are given in an earlier paper.¹⁵ The ratio u_i / U_j is plotted as a function of the Strouhal number St_W

$$St_W = \frac{f_1 W}{U_j}, \quad (3)$$

where f_1 is the frequency of the fundamental, which is a ratio of the travel time of a fluid particle along the jet (W/U_j) and the oscillation period ($1/f_1$).

In this dimensionless representation, the different curves associated with the fundamental ω_1 of each acoustic oscillation mode have now collapsed onto a single curve. This indicates a common behavior of the fundamental oscillation over the entire range of the functioning of the instrument. The maximum ratio between the transverse acoustic flow velocity in the mouth and the jet velocity is approximately equal to $\frac{1}{5}$ and is reached at a value $1/St_W \approx 6$. Similar data are shown in Figs. 5(a) and 6(a) for different transverse positions y_0 of the labium with respect to the flue channel axis. A similar collapse is observed.

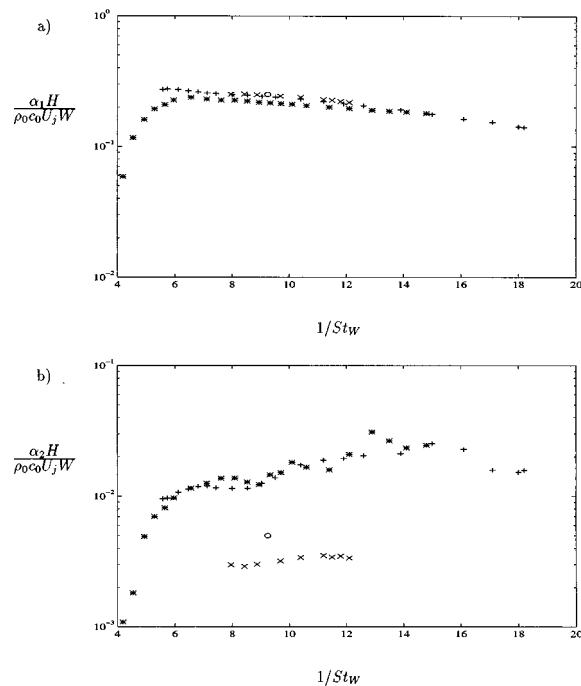


FIG. 5. (a) Dimensionless amplitude α_1 of the fundamental and (b) α_2 of the second harmonic as a function of the Strouhal number $St_W = f_1 W / U_j$ for operation of the experimental flue organ pipe in the first three acoustic modes of the pipe: *, first mode, +, second mode, x, third mode; O, fourth mode). The edge of the labium was placed 4.02 mm in front of the flue exit $W/h = 3.72$ and its transverse position y_0 was adjusted 0.25 mm above the lower wall of the flue exit ($y_0/h = 0.27$, the zero of the transverse position is in the middle of the flue exit).

This result is quite interesting because it shows that, for a given geometry, the dimensionless representation of the amplitude α_1 of the fundamental is a function of the Strouhal number St_W alone and no other dimensionless parameter. In the data displayed in Fig. 4(a), for a fixed Strouhal number, the jet velocity and therefore the Reynolds number $Re_h = h U_j / \nu$, where ν is the kinematic viscosity, increases by a factor of 3 as the pipe is overblown. This indicates that the dimensionless amplitude u_i / U_j does not depend on the viscosity in the bulk of the fluid and therefore on the structure of the jet which goes from laminar on the first pipe mode to turbulent on the higher modes. Furthermore, the similar increase of the Helmholtz number $He = H \omega / c_0$ clearly implies that the playing amplitude can not be determined, as assumed in most models,¹¹ by a balance between acoustic energy production by a jet drive and energy losses due to frequency-dependent damping mechanisms such as radiation and visco-thermal dissipation at walls. These observations and their generalization to other configurations will be discussed in Sec. IV. The key of this surprising result is that at typical amplitude, the energy losses due to vortex shedding at the edge of the labium are crucial. We will discuss this in more detail in Sec. V. This example shows that dimensionless representation of experimental data is of considerable help in tracking basic physical mechanisms.

Similar data are presented in Figs. 4(b), 5(b), and 6(b) for the dimensionless amplitude α_2 of the second harmonic. When the instrument oscillates in the first and second acous-

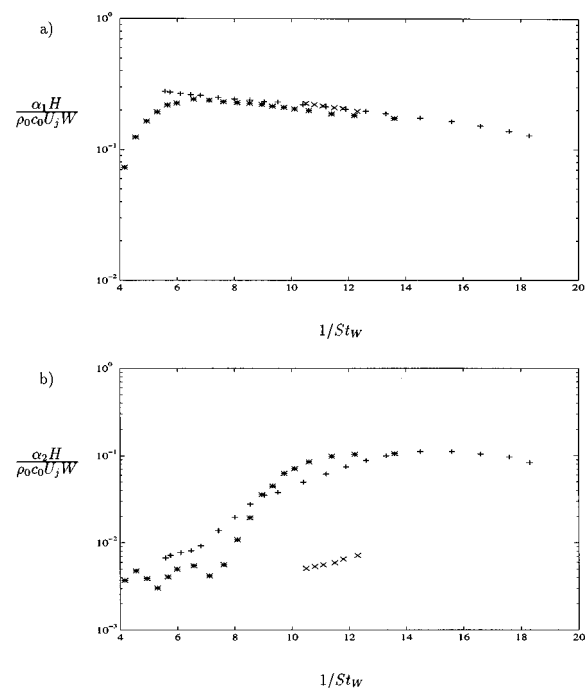


FIG. 6. (a) Dimensionless amplitude α_1 of the fundamental and (b) α_2 of the second harmonic as a function of the Strouhal number $St_W = f_1 W / U_j$ for operation of the experimental flue organ in the first three acoustic modes of the pipe: *, first mode; +, second mode; x, third mode). The edge of the labium was placed 4.02 mm in front of the flue exit $W/h = 3.72$ and its transverse position was adjusted 0.39 mm above the lower wall of the flue exit ($y_0/h = 0.14$, the zero of the transverse position is in the middle of the flue exit).

tic mode of the pipe, the dimensionless amplitudes α_2 display a common Strouhal number dependency. The amplitude α_2 appears to be lower than that of the fundamental but increases strongly as the jet velocity is increased. When the jet velocity is increased and the third mode of the pipe is reached, the dimensionless amplitude α_2 drops by an order of magnitude. This reduction of the second harmonic correlates with the disappearance of the “acoustically” induced vortex shedding at the edge of the labium when the jet becomes turbulent enough to prevent the formation of these vortices at the edge of the labium.¹⁴ When the jet is laminar, its displacement around the labium is large. As these displacements are larger than the jet width, the flow around the labium is, during part of the oscillation period, dominated by the acoustic flow from the pipe. When the jet becomes turbulent, it also becomes wider and for the particular ratio $W/h = 4$ it seems to be attracted by the labium (Coanda effect). This reduces the amplitude of the jet displacement to values smaller than the width of the jet which prevents the formation of acoustically induced vortices. We conclude that the behavior of the second harmonic, in contrast with the behavior of the fundamental, depends strongly on the structure of the jet. The dimensionless amplitude α_2 is governed by a common Strouhal number dependency while the jet is laminar and displays a completely different behavior when the jet becomes turbulent. As will be seen later, for larger values of the ratio W/h (one commonly finds values of the order of 10 in organ pipes) this dramatic difference in oscil-

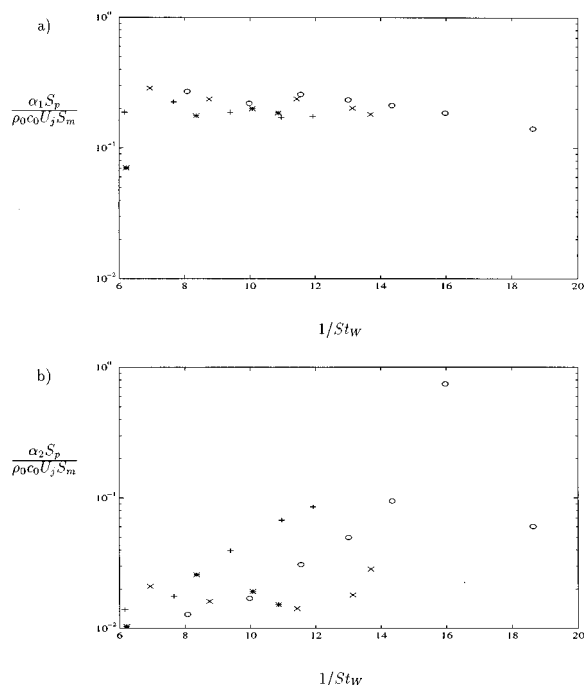


FIG. 7. (a) Dimensionless amplitude of the fundamental and (b) the second harmonic as a function of the Strouhal number $St_W = f_1 W/U_j$ for operation of an alto recorder (mouthpiece and cylindrical bore) in the first four acoustic modes of the pipe: (*, first mode; +, second mode; x, third mode; o, fourth mode). S_m and S_p are the mouth and pipe cross section, respectively.

lation behavior depending on whether the jet is laminar or turbulent disappears. This is due to the fact that in these conditions, the displacement of the turbulent jet around the labium is also large in comparison with its width.

In order to demonstrate that these features are general and not due to special properties of the experimental pipe, similar data measured on an alto recorder are shown in Fig. 7. In these measurements, the body of the recorder was replaced by a cylinder in order to facilitate the acoustic interpretation of the data. This instrument behaves similarly to the experimental flue organ pipe. Experiments with the original pipe of the instrument provided similar results but these data are less accurate because of the uncertainty in the acoustic model of the pipe.

III. INFLUENCE OF THE MOUTH GEOMETRY

A. The horizontal position of the labium

The ratio W/h is an important parameter since it determines, for a given value of h , the distance traveled by the hydrodynamic perturbation on the jet and hence its oscillating mode. It also determines the window surface to pipe cross-section ratio and consequently the acoustic velocity in the mouth.

In recorders of all sizes, the ratio W/h is adjusted by instrument makers to a value close to 4. In other instruments, such as large flue organ pipes, the ratio W/h can be up to three times as large (in flutes, it is adjusted by the player). Changing this ratio modifies the operating characteristics of the instrument. [Figures 10 and 11 show the frequency f_1 and amplitude $p_1(\omega_1)$ of the fundamental of the pressure



FIG. 8. Flow visualization of the mouth of the experimental flue pipe for a ratio W/h equal to 8.0 and a driving pressure of 40 Pa. The labium was placed 0.2 mm above the lower wall of the flue exit ($y_0/h=0.3$, the zero of the transverse position is in the middle of the flue exit). In these conditions, the jet operates in its *second hydrodynamic mode*.

signal measured in the experimental flue pipe as a function of the driving pressure p_f for a ratio $W/h=8.0$.] Comparison with Fig. 2 and Fig. 3 shows that transitions between the functioning of the instrument in the different acoustic modes of the resonator appear at higher blowing pressures. This can be understood from the fact that the perturbations on the jet have a longer distance to travel before they reach the labium; a higher jet velocity is therefore required in order to fulfill the operating conditions of the feedback loop of the system. At very low pressures the operating frequency is strongly dependent on the driving pressure p_f , displaying jumps as the driving pressure is changed. This is due to jet oscillations on higher hydrodynamic modes as can be observed in the flow visualizations of Fig. 8. At the transition between two hydrodynamic modes of the jet, the response of the instrument is very unstable. Increasing the jet velocity enables the system to reach the first hydrodynamic mode of the jet as shown in Fig. 9. Note in this hydrodynamic mode, the large amplitude of the jet oscillation and the turbulence in the jet flow around the labium. Nolle¹⁷ observed similar shifts in the functioning of flue organ pipes for different mouth geom-

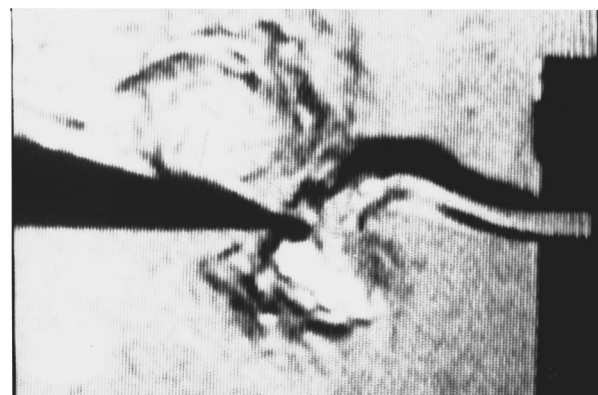


FIG. 9. Flow visualization of the mouth of the experimental flue pipe for a ratio W/h equal to 8.0 and a driving pressure of 270 Pa. The labium was placed 0.2 mm above the lower wall of the flue exit ($y_0/h=0.3$, the zero of the transverse position is in the middle of the flue exit). In these conditions, the jet operates in its *first hydrodynamic mode*.

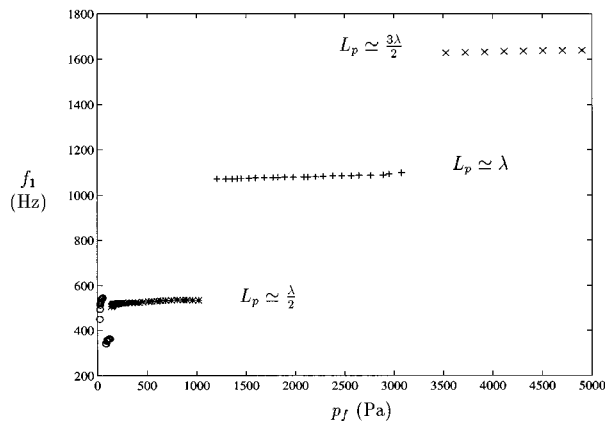


FIG. 10. Operating fundamental frequency f_1 of the experimental flue organ pipe as a function of the driving pressure p_f (Pa). The edge of the labium was placed 8.0 mm ($W/h=8.0$) in front of the flue exit and its transverse position y_0 was adjusted 0.2 mm above the lower wall of the flue exit ($y_0/h=0.3$, the zero of the transverse position is in the middle of the flue exit). The “○” corresponds to operation of the system in the second hydrodynamic mode of the jet.

etries. Further increasing the jet velocity brings the system to operate on higher acoustic modes of the pipe.

The effects of the ratio W/h can be evaluated again by using a dimensionless representation of experimental data. In Fig. 12 the data of Figs. 10 and 11 are made dimensionless by using equations 1 and 2 and are presented as a function of the Strouhal number $St_W=fW/U_j$.

Operation of the instrument in the second hydrodynamic mode of the jet is clearly observed at low blowing pressures. In this mode, the dimensionless amplitude is weak compared to that of the first hydrodynamic mode. This could be due to the fact that, as can be observed in Fig. 8, the jet oscillation is saturated by a breakdown of the jet into discrete vortices. The vortices appear on each side of the jet, as a result of the roll-up of the unstable shear layers of the jet. These vortices are fully developed before the interaction of the jet with the labium.

At higher blowing pressures, the data corresponding to operation in the first hydrodynamic mode of the jet again

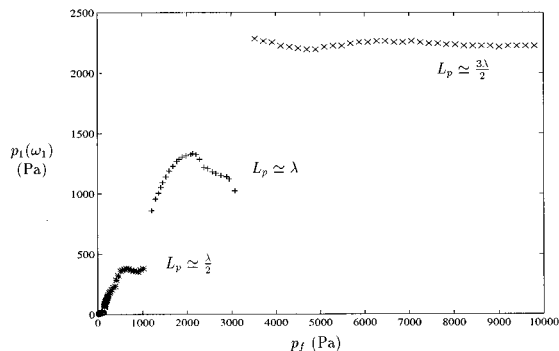


FIG. 11. Amplitude $p_1(\omega_1)$ of the fundamental as a function of the driving pressure p_f for the small experimental flue pipe. The microphone was placed 277.8 mm from the passive end of the pipe. The distance between the flue exit and the labium was adjusted to 8.0 mm ($W/h=8.0$) and its transverse position y_0 was adjusted 0.2 mm above the lower wall of the flue exit ($y_0/h=0.3$, the origin of y_0 is in the middle of the flue exit).

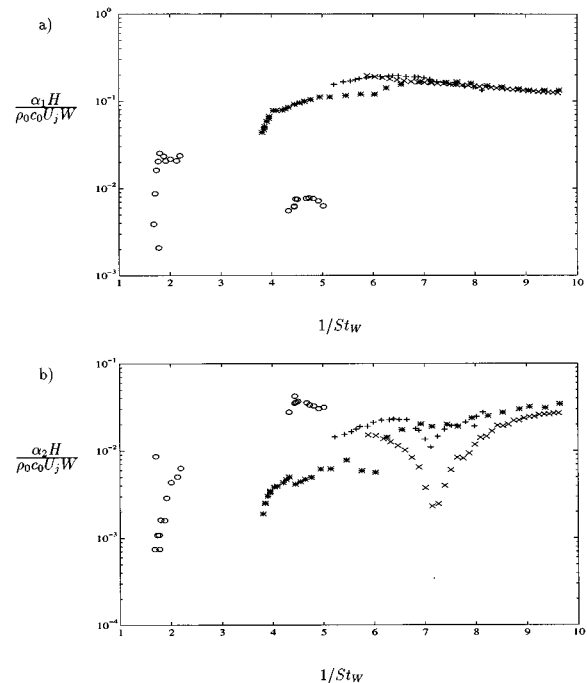


FIG. 12. (a) Dimensionless amplitude of the fundamental and (b) the second harmonic as a function of the Strouhal number $St_W=f_1W/U_j$ for operation of the experimental flue organ in the first three acoustic modes of the pipe: *, first mode; +, second mode; x, third mode). The edge of the labium was placed 8.0 mm in front of the flue exit $W/h=8.0$ and its transverse positions y_0 was adjusted 0.2 mm above the lower wall of the flue exit ($y_0/h=0.3$, the zero of the transverse position is in the middle of the flue exit). The “○” corresponds to operation of the system in the second hydrodynamic mode of the jet.

collapse on a single curve, regardless of the acoustic mode of the resonator. In this hydrodynamic mode, the oscillations of the jet have again a high amplitude compared to the jet width h , as can be observed in Fig. 9. It is interesting to note that this curve roughly has the same shape and corresponds to amplitudes close (about 30% lower for $6<1/St_W<10$) to that presented in Figs. 4(a), 5(a), and 6(a) for a ratio $W/h=3.72$. It is fascinating to note that the measured maximum ratio u_i/U_j between the transverse acoustic velocity in the mouth and the jet velocity is again close to $\frac{1}{5}$, even though the mouth geometry has been changed. This indicates that the dimensionless representation grasps a quite fundamental characteristic of the functioning of recorderlike instruments.

The dimensionless amplitude of the second harmonic α_2 is presented in Fig. 12(b) as a function of the Strouhal number for a ratio $W/h=8.0$. Its evolution with the jet velocity is quite similar to that obtained for a ratio $W/h=3.72$ during operation in the first hydrodynamic mode of the jet and the first acoustic mode of the pipe. The behavior in the second and third acoustic mode of the pipe are quite different, however, displaying strong minima for values of the Strouhal number $1/St_W=f_1W/U_j$ close to 7. We have no explanation for these minima. Note that the shift in the amplitude α_2 observed for a lower ratio W/h (see Fig. 4) between the second and third acoustic mode does not occur here. As discussed previously this could be due to the fact that, as is observed in Fig. 9 for large ratios W/h , the am-

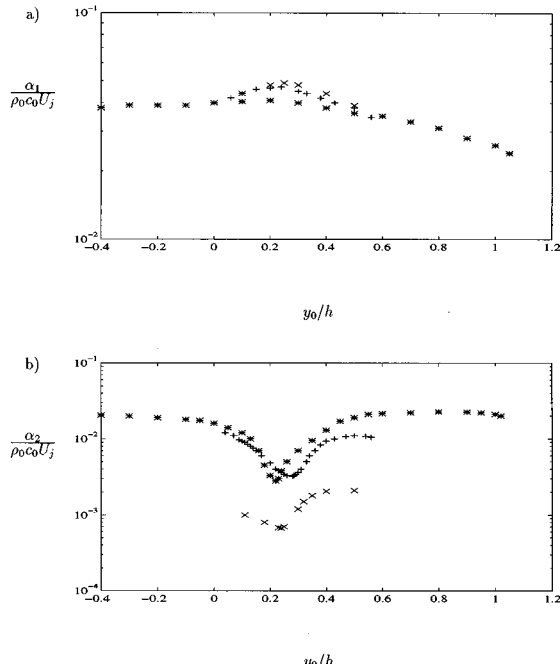


FIG. 13. (a) Dimensionless amplitude of the fundamental and (b) the second harmonic as a function of the transverse position y_0 of the labium for a Strouhal number $St_W = f_1 W/U_j = 0.099$: (*, first acoustic mode of the pipe; +, second acoustic mode; ×, third acoustic mode). The edge of the labium was placed 4.02 mm in front of the flue exit ($W/h = 3.7$). A ratio y_0/h of zero corresponds to the middle of the flue exit and positive values are oriented towards the interior of the pipe.

plitude of the jet displacement at the labium is large in comparison with its width even when the jet is turbulent.

The maximum efficiency of recorderlike instruments for the fundamental is obtained for values of the Strouhal number in the range $0.14 < St_W = f_1 W/U_j < 0.2$. This range of Strouhal number also enables one to obtain a strong second harmonic. The corresponding jet velocities necessary to obtain these Strouhal numbers are determined by the distance W . For small distances such as found in recorderlike instruments, the jet is laminar for this range of Strouhal numbers while it is turbulent for larger distances such as in large flue organ pipes. A large distance between the flue exit and the labium further enhances this effect because the jet has more time to develop into turbulence before it interacts with the labium. The higher jet velocities found in flue organ pipes enable one to obtain a much more powerful sound than in recorders but at the expense of turbulence noise. In our measurements we observed the same behavior for a ratio $W/h = 12$ as when a ratio $W/h = 8$ was used. The dimensionless amplitude was then decreased, however, by a factor of 2 [$(u_1/U_j)_{\max} = 0.11$ at $1/St_W = 6$] compared with the case $W/h = 4$.

B. The transverse position of the labium

Another parameter which is important for the timbre of the instrument is the transverse position y_0 of the labium with respect to the flue channel axis. In flue organ pipes and recorders this parameter is fixed by the craftsmen, but in flutes it can be varied by the musician.

Figure 13 shows the effects of the transverse position

y_0 (where y_0 is positive toward the interior of the pipe) of the labium on the amplitude of the first and second harmonic of the pipe for a typical Strouhal number ($St_W = f_1 W/U_j = 0.099$) which corresponds to optimal sound production. This parameter appears to strongly affect the amplitude of the second harmonic. The same behavior can also be observed by comparing Figs. 4(b), 5(b), and 6(b). This effect has been studied by Nolle and Fletcher in flue organ pipes^{18,19} and could be linked to a change in duty cycle of the flow waveform on each side of the labium which determines the drive of the acoustic oscillations in the resonator by the jet oscillations. The use of a linear jet oscillation model however implies that a minimum in the generation of second harmonic should appear when the edge of the labium is aligned with the middle of the flue exit ($y_0 = 0$) because of the symmetry of the jet displacement at the labium. In the measurements, however, this minimum of the second harmonic is observed when the labium is slightly decentred with respect to the flue exit ($y_0 = 0.2h$). This shift in the “effective” zero transverse position of the labium could be associated with a nonlinear effect due to the interaction of the jet with the labium. Such an effect can be expected as a result of air entrainment by the jet which can create a static vacuum in the pipe (Coanda effect) and deflect the jet. In a closed pipe, a pressure build up which would induce an outward static deflection of the jet would be expected. Such effects are usually more pronounced with turbulent than with laminar jets.

IV. HIGH-AMPLITUDE BEHAVIOR OF SELF-SUSTAINED FLOW OSCILLATIONS

The measurements presented in the previous sections show that, for a specific flue pipe geometry, the ratio u_1/U_j appears to be a function of the Strouhal number St_W which for the first hydrodynamic mode of the jet is independent of the acoustic pipe mode excited. This implies that the amplitude of the fundamental is not determined by a balance between energy supplied by the jet motion and energy losses due to a frequency dependent loss mechanism such as radiation. This statement can be quantified by comparing the power $\langle P \rangle_{\text{rad}}$ dissipated by radiation

$$\langle P \rangle_{\text{rad}} \sim \frac{\rho_0 (HW)^2 f_1^2 u_1^2}{c_0}, \quad (4)$$

with the power $\langle P \rangle_{\text{jet-drive}}$ provided by the jet-drive model of Coltman⁶ modified by Verge *et al.*²

$$\langle P \rangle_{\text{jet-drive}} \sim \rho_0 f_1 \delta (h H U_j) u_1. \quad (5)$$

In this jet-drive model δ is a distance accounting for the impedance difference experienced by sources placed on each side of the labium. Verge *et al.*² show that by assuming the jet flow injection points to be located at a distance ϵ from the edge of the labium, the distance δ can be estimated by

$$\delta/W \sim \sqrt{\epsilon/W}. \quad (6)$$

As a first guess ϵ is assumed to be constant and of the order of the jet width h . A balance between $\langle P \rangle_{\text{rad}}$ and $\langle P \rangle_{\text{jet-drive}}$ would imply

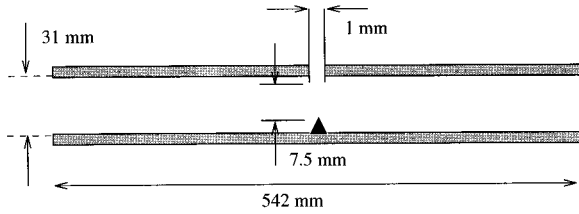


FIG. 14. Confined jet-edge system used by Powell²¹ ($h=1$ mm, $W=7.5$ mm).

$$\frac{u_1}{U_j} \sim \frac{\delta h}{WHSt_W M_j}, \quad (7)$$

where $M_j = U_j/c_0$ is the Mach number. The Mach number increases every time the pipe is overblown. For a specific Strouhal number, reaching the fourth mode of the pipe implies an increase of the Mach number by a factor of 4 and consequently a decrease in the dimensionless ratio u_1/U_j by a factor of 4. Taking into account the visco-thermal losses at the walls of the pipe would imply a less spectacular decrease of this ratio but certainly not a constant value. We can now conclude that either the jet-drive model is wrong or either there exist another dissipative mechanism which dominates radiation and friction. As explained by Fabre,^{13,14} this suspicion is reinforced by the fact that the Fletcher's model predicts a much higher value of the amplitude u_i/U_j than what we actually measure. This is further confirmed by simulation results presented in the companion to this paper. Fabre suggests that dissipation by flow separation at the edge of the labium is the missing link. This effect was already apparent in the mouth impedance measurements by Coltman⁶ which displayed a significant nonlinearity at amplitudes corresponding to normal playing conditions. Coltman associated these observations to the nonlinear behavior of diaphragms with sharp edges which have been systematically studied by Ingard and Ising²⁰ and will be discussed in more details in the next section.

Although our measurements were performed on a specific acoustic system, we expect that our conclusions can be generalized to various jet-edge driven acoustic systems. A first confirmation of this affirmation comes from the measurements of Powell²¹ on the sound radiated by the confined jet-edge system illustrated in Fig. 14. The jet-edge system, characterized by a ratio $W/h=7.5$ was placed in the middle of a pipe open at both ends normal to the axis of the pipe. Powell overblew the system eight times and found that the power $\langle P \rangle_{\text{rad}}$ radiated outside the system scaled, as shown in Fig. 15, with U_j^4 over two decades of the jet velocity. As shown in Fig. 16, pulsations of the system on the first hydrodynamic mode of the jet are located within a relatively narrow range of the Strouhal number centered around $St_W=0.18$ as in our measurements. Since the Strouhal number does not vary much, we know that the oscillating frequency must roughly be proportional to the jet velocity. This implies, as can be deduced from Eq. (4), that $\langle P \rangle_{\text{rad}} \sim f_1^2 u_1^2 \sim U_j^2 u_1^2$. The data of Powell can therefore be explained if the ratio u_1/U_j is assumed to be constant. Since he used a wide range of jet velocities, up to $M_j=0.8$, we can conclude that

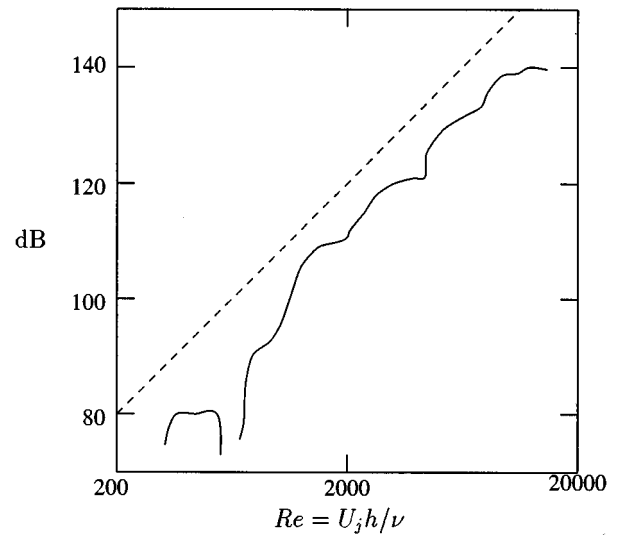


FIG. 15. Intensity of the sound radiated by the system of Fig. 14 as a function of the Reynolds number $Re = U_j h / \nu$ and comparison with a U_j^4 scaling (dotted line).²¹ $(Re)_{\text{max}} = 2 \times 10^4$, $(U_j)_{\text{max}} = 300$, $(U_j/c_0)_{\text{max}} = 0.85$.

this behavior is independent, for a given geometry, of the Mach and Reynolds numbers.

Saturation of pulsations at a fixed value of the ratio u_1/U_j depending only on the value of the Strouhal number and on the geometry of the source regions is a behavior that we also observed in our study of oscillations of a flow in a pipe having a tandem of closed side branches.²² When the distance between the branches is equal to twice the length L_b of the branches, the tandem acts as a perfect nonradiating resonator for frequencies $f_1 = (2m+1)c_0/4L_b$ ($m=0,1,2,\dots$). By varying the side branches length L_b and the static pressure p_0 , measurements were carried out in the range $10^5 \leq Re_D \leq 10^7$ and $3 \times 10^{-2} \leq M_0 \leq 0.5$, where $Re_D = U_0 D / \nu$ and $M_0 = U_0 / c_0$ and D is the diameter of the pipe. The visco-thermal losses could be reduced by increasing the pressure p_0 and decreasing the length L_b of the side branches. It appears that, as visco-thermal dissipation is re-

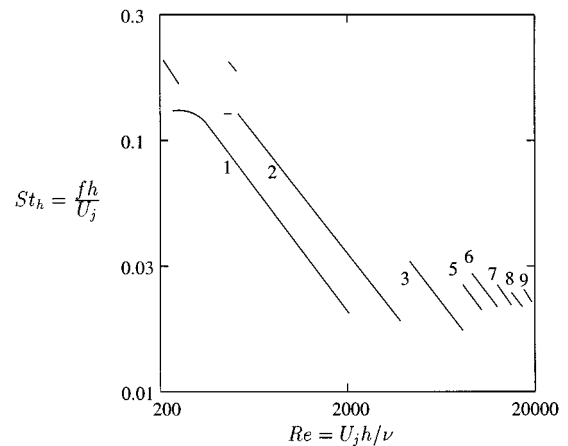


FIG. 16. Strouhal number $St_h = f_h / U_j$ as a function of the Reynolds number $Re = U_j h / \nu$ of the confined jet-edge system shown in Fig. 14 as a function of the jet velocity.²¹ The number of the harmonic is indicated on the graphic.

duced, the value of the dimensionless amplitude of the fundamental increases until a maximum $(u_1/U_j)_{\max}$ is reached. The ratio u_1/U_j is then independent of visco-thermal dissipation. The value of this maximum is determined by the geometry of the junction between the side branches and the main pipe. For junctions having sharp edges, $(u_1/U_j)_{\max} = 0.2$. For edges with radius of curvature greater than $D/10$, $(u_1/U_j)_{\max} = 0.6$. The amplitude of the pulsations was particularly sensitive to the shape of the upstream edge of the downstream branch junction. Placing a thin plate of length $D/5$ along the direction of the flow at this edge reduced the maximum amplitude of the pulsations by two orders of magnitude. This effect was explained by Bruggeman²² in terms of the vortex sound theory of Powell²³ and Howe.¹² It is interesting to note that Kriesels²⁴ obtained similar results in a configuration where the two side branches are placed opposite to each other along the main pipe. The results of these scale model experiments were confirmed by measurements at compressor stations of the Dutch natural gas transport system²⁵ for $Re_D = 10^8$. These studies reinforce our assumption that the amplitude limiting mechanism in recorderlike instruments is independent of the Reynolds and Mach numbers.

Recently we have started a study of the pulsation of grazing pipe flow along a diaphragm opening of a volume V . This flow triggers Helmholtz resonances similar to that studied by Nelson.²⁶ In a series of preliminary experiments, Bailliet²⁷ found a constant maximum pulsation amplitude $(u_1/U_j)_{\max} = 0.2$ independently of the ratio $D/V^{1/3}$ of the diaphragm diameter D and the characteristic length $V^{1/3}$ of the volume. This ratio was varied within the range $10^{-1} \leq D/V^{1/3} \leq 1$ which seems to confirm that the shape of the resonator is not critical for the amplitude saturating mechanism which determines the maximum amplitude of the pulsations.

We conclude from these observations that the behavior displayed by our flue organ pipe is typical of high amplitude pulsation in aeroacoustic systems in which saturation is not determined by radiation nor visco-thermal dissipation. Following Howe,²⁸ we therefore expect the pulsation amplitude to be determined by vortex damping which constitutes a strong nonlinearity in the mouth flow itself.

V. QUASI-STATIONARY MODEL OF VORTEX DAMPING

Visualizations of the flow around the labium presented by Fabre^{13,14} for $W/h = 4$ show that injection of vorticity into the main flow at the edge of the labium results from two distinct phenomena. When the jet passes the labium, the shedding of vorticity clearly occurs as the result of the separation of the jet flow. In addition, for relatively low blowing pressures (first and second pipe mode), the shedding of a vortex occurs once the jet has passed the labium. This vortex is therefore expected to be mainly induced by separation of the acoustic flow at the edge of the labium. Visualizations show, furthermore, that this acoustically induced vortex is present during less than half the fundamental oscillation period. This implies that it can only extract energy from the fundamental of the acoustic field. In order to supply the fun-

damental with energy, it would have to remain at least until the opposite phase of the acoustic velocity which is not the case. The life span of this vortex is however long enough to transfer energy to the higher frequencies, which correlates with the presence of high harmonics in measured spectra of the acoustic field in the pipe.

In order to model the effects of the acoustically induced vortex shedding at the labium, a simple quasi-stationary free-jet model is proposed. For reasons of simplicity, the effects of the separation of the jet flow are neglected since no model of this complex phenomenon is available at the present time. The formation of the jet is represented by a fluctuating pressure jump Δp_a across the mouth of the instrument

$$\Delta p_a = -\frac{1}{2} \rho_0 \left(\frac{Q_p}{\alpha_v S_m} \right)^2 \text{sign}(Q_p), \quad (8)$$

where ρ_0 is the density, α_v is the vena-contracta factor of the jet, which, depending on the geometry, can vary within the range $0.6 \leq \alpha_v \leq 1$. We used later a value $\alpha_v = 0.6$. This pressure drop opposes the acoustic flow in the mouth and corresponds to the dissipation of the kinetic energy of the free jet by turbulence with negligible pressure recovery. Power dissipation $\langle P \rangle_{\text{free-jet}}$ induced by this model scales with the third power of the acoustic flow

$$\langle P \rangle_{\text{free-jet}} = \frac{1}{T} \int_0^T \Delta p_a \frac{Q_p}{S_m} dt, \quad (9)$$

where T is the period of the acoustic signal. This crude model should constitute an upper limit approximation of the effects of the vortex shedding since it implicitly assumes that dissipation occurs during the entire period of the fundamental. It is obvious that this model has only been chosen for its simplicity. While it is not a realistic flow model, it appears to predict some essential aspects of the behavior of flue instruments.

Assuming an energy balance between the jet-drive model discussed in the previous section and this simple vortex model, the following relationship is obtained:

$$\left(\frac{u_1}{U_j} \right)^2 \sim St_W \frac{\delta h}{W^2} \sim St_W \left(\frac{h}{W} \right)^{3/2}, \quad (10)$$

which gives an estimate of the influence of the ratio W/h on the dimensionless amplitude. Dividing $St_W^{1/2} (h/W)^{3/4}$ with the maximum $(u_1/U_j)_{\max}$ found in our measurements at $1/St_W = 6$, we obtain 0.70 for $W/h = 3.72$, 0.85 for $W/h = 8$ and 0.71 for $W/h = 12$. The almost constant result shows that the model accounts for the shift in the value of $(u_1/U_j)_{\max}$ observed when the ratio W/h is varied. This is a surprisingly good agreement in view of the drastic changes observed in the flow as the ratio W/h is varied.

This result does not, however, prove that the value $(u_1/U_j)_{\max}$ is correctly predicted by this model. The validity of the model was confirmed by Fabre,¹⁴ using the same jet-drive and vortex models and also including radiation and visco-thermal dissipation, for a sinusoidal time dependency of the acoustic flow. We will show in the companion to this

paper that extending the model to an arbitrary number of harmonics (below the cutoff frequency for nonplanar propagation) does not modify the conclusions.

The assumption of quasi-stationarity should only be valid when the Strouhal number $St_a = \omega W^2 H / Q_p \ll 1$, that is, when the amplitude of the acoustic displacement is much larger than the characteristic dimensions of the mouth. Measurements indicate, however, that in playing conditions, the Strouhal number is rather of the order of 1. As mentioned by Fabre,¹⁴ using an unsteady point vortex model as proposed by Howe,¹² valid when $St_a \gg 1$, gave the same order of magnitude for the power dissipation due to vortex shedding in the limit $St_a = \mathcal{O}(1)$ and therefore did not change the results for the fundamental. For time-domain simulation purposes, we therefore chose to use, for simplicity the simple quasi-stationary free-jet model.

VI. CONCLUSION

Dimensionless representation of data measured on different recorderlike instruments shows that the dimensionless amplitude of the fundamental of the acoustic response of the instrument is, for the first hydrodynamic mode of the jet, a function of the Strouhal number (only), regardless of the acoustic mode excited and the pipe and mouth geometry of the instrument. The maximum ratio between the acoustic flow velocity in the mouth and the jet velocity appears to be of the order of 10^{-1} and to scale with the ratio $(h/W)^{3/4}$. This general behavior indicates that amplitude saturation is due to a nonlinear interaction of the acoustic flow with the jet flow. The amplitude of the fundamental can therefore not be simply determined by a balance between acoustic energy production by a jet drive and energy losses due to friction and radiation. Vortex shedding at the edge of the labium is therefore expected to induce significant energy losses for the fundamental. The optimal efficiency of the instrument is obtained at a Strouhal number $St_w = fW/U_j \approx 0.17$. This value also enables one to obtain a strong second harmonic and the presence of high-frequency components in the spectrum. The dimensionless amplitude α_2 of the second harmonic displays two distinct kinds of behavior depending on whether the jet is laminar or turbulent. These observations have been correlated by Fabre¹⁴ with visualizations of the flow in the mouth of the instrument.

Although the dimensionless amplitude of the acoustic response of flue instruments appears to be determined by a physical mechanism common to different instruments, the choice of a particular mouth geometry has important consequences on the tone they produce. Recorder makers have favored a ratio $W/h \approx 4$ and a small window. This results, in normal playing conditions, in low jet velocities compared to flue organ pipes. Recorders therefore require a minimum of air flow in order to produce a rich timbre which is a reasonable requirement for an instrument driven by human power. The range of Reynolds numbers obtained in a recorder geometry, $1000 < Re = U_j h / \nu < 2000$, ensures that except for the highest notes, the jet is laminar up to the labium which prevents the generation of turbulence noise. As the jet remains laminar a sharp edge will enhance the generation of higher harmonics. A small window therefore implies that

transitions between operation in the different acoustic modes of the pipe are reached at low blowing pressures which is convenient to play the instrument over a large range. Low blowing pressures might be convenient for the musician but also results in modest playing amplitudes. A consequence of this particular choice of ratio W/h is that recorders always operate on the same (the first) hydrodynamic mode of the jet. This avoids unstable transitions between different jet modes which probably results in a better control of attacks by musicians. Recorders therefore seem to have been optimized by craftsmen to yield an optimal harmonic content to noise ratio and clear attacks.

In large organ pipes, craftsmen wish to obtain a powerful sound which requires a high jet velocity and therefore a larger distance W . This is obtained at the expense of a noisy sound, which is very characteristic of some flue organ pipes. The noise production at the labium can be reduced, however, by using a square or round labium edge instead of a very sharp one as found in recorders. This is indeed common practice in organ building. A large distance W also prevents the collapse of the generation of even harmonics (α_2) as observed in recorders at the transition between a laminar and a turbulent jet. More quantitative analysis of these effects could only be obtained by a detailed study of the flow around the labium. A drawback of this type of mouth geometry, for the stability of the instrument, is the fact that at low blowing pressures, it may allow oscillation of the jet in its second hydrodynamic mode. When a single note is played on some organ pipes, one can clearly hear the transition from the first to the second mode of the jet after the release of a key; the sound reappears after it has died out a first time. This is probably not so important for flue organ pipes, since they are adjusted to play only one note in a limited range of driving pressures in contrast with recorders which must respond well over more than 2 octaves.

¹M. Castellengo, "Contribution à l'étude expérimentale des tuyaux à bouche," Ph.D. thesis, Université de Paris VI, Paris, France, 1976.

²M. P. Verge, R. Causse, B. Fabre, A. Hirschberg, A. P. J. Wijnands, and A. van Steenberghe, "Jet oscillations and jet drive in recorder-like instruments," *Acustica* **2**, 403–419 (1994).

³M. E. McIntyre, R. T. Schumacher, and J. Woodhouse, "On the oscillations of musical instruments," *J. Acoust. Soc. Am.* **74**, 1325–1345 (1983).

⁴L. Cremer and H. Ising, "Die selbsterregten Schwingungen von Orgelpfeifen," *Acustica* **19**, 143–153 (1967–1968).

⁵J. W. Coltman, "Sounding mechanism of the flute and organ pipe," *J. Acoust. Soc. Am.* **44**, 983–992 (1968).

⁶J. W. Coltman, "Jet drive mechanism in edge tones and organ pipes," *J. Acoust. Soc. Am.* **60**, 725–733 (1976).

⁷S. A. Elder, "On the mechanism of sound production in organ pipes," *J. Acoust. Soc. Am.* **54**, 1554–1564 (1973).

⁸N. H. Fletcher, "Jet-drive mechanism in organ pipes," *J. Acoust. Soc. Am.* **60**, 481–483 (1976).

⁹N. H. Fletcher, "Sound production by organ flue pipes," *J. Acoust. Soc. Am.* **60**, 926–936 (1976).

¹⁰N. H. Fletcher and S. Thwaites, "Wave propagation on a perturbed jet," *Acustica* **42**, 323–334 (1979).

¹¹N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments* (Springer-Verlag, New York, 1991).

¹²M. S. Howe, "Contribution to the theory of aerodynamic sound, with application to excess jet noise and the theory of the flute," *J. Fluid Mech.* **71**, 625–673 (1975).

¹³B. Fabre, "La Production de Son dans les Instruments à Embouchure de

- Flûte: Modèle Aéro-Acoustique pour la Simulation Temporelle," Ph.D. thesis, Université du Maine, Le Mans, France, 1992.
- ¹⁴B. Fabre, A. Hirschberg, and A. P. J. Wijnands, "Vortex shedding in steady oscillations of a flue organ pipe," *Acustica* **82**, 863–877 (1996).
 - ¹⁵M. P. Verge, B. Fabre, W. E. A. Mahu, A. Hirschberg, R. van Hassel, A. P. J. Wijnands, J. J. de Vries, and C. J. Hogendoorn, "Jet formation and jet velocity fluctuations in a flue organ pipe," *J. Acoust. Soc. Am.* **95**, 1119–1132 (1994).
 - ¹⁶M. P. Verge, "Aeroacoustics of Confined jets, with Applications to the Physical Modeling of Recorder-Like Instruments," Ph.D. thesis, Eindhoven University of Technology, The Netherlands (1995).
 - ¹⁷A. W. Nolle, "Some voicing adjustments in flue organ pipes," *J. Acoust. Soc. Am.* **66**, 1612–1626 (1979).
 - ¹⁸A. W. Nolle, "Flue organ pipes: Adjustments affecting steady waveform," *J. Acoust. Soc. Am.* **73**, 1821–1832 (1983).
 - ¹⁹N. H. Fletcher and L. M. Douglas, "Harmonic generation in organ pipes, recorders and flutes," *J. Acoust. Soc. Am.* **68**, 767–771 (1980).
 - ²⁰U. Ingard and H. Ising, "Acoustic nonlinearity of an orifice," *J. Acoust. Soc. Am.* **42**, 6–17 (1967).
 - ²¹W. K. Blake and A. Powell, "The development of contemporary views of flow-tone generation," in *Recent Advances in Aeroacoustics* (Springer-Verlag, New York, 1983).
 - ²²J. C. Bruggeman, A. Hirschberg, M. E. H. van Dongen, A. P. J. Wijnands, and J. Gorter, "Self-sustained aeroacoustic pulsations in gas transport systems: experimental study of the influence of closed side branches," *J. Sound Vib.* **150**, 371–393 (1991).
 - ²³A. Powell, "Theory of vortex sound," *J. Acoust. Soc. Am.* **36**, 177–195 (1964).
 - ²⁴P. C. Kriesels, M. C. A. M. Peters, A. Hirschberg, A. P. J. Wijnands, A. Iafrati, G. Riccardi, R. Piva, and J. C. Bruggeman, "High amplitude vortex-induced pulsations in a gas transport system," *J. Sound Vib.* **184**, 343–368 (1995).
 - ²⁵J. Gorter, A. Hirschberg, A. P. J. Wijnands, and J. C. Bruggeman, "Flow induced pulsations in gas transport systems," in *International Gas Research Conference*, Tokyo, November 1989.
 - ²⁶P. A. Nelson, N. A. Halliwell, and P. E. Doak, "Fluid dynamics of a flow excited resonance, Part 1: Experiment," *J. Sound Vib.* **78**, 15–38 (1981).
 - ²⁷H. Bailliet, 1996 (private communication).
 - ²⁸M. S. Howe, "On the absorption of sound by turbulence and other hydrodynamic flows," *Inst. Math. Appl., J. Appl. Math.* **32**, 187–209 (1984).

Sound production in recorderlike instruments.

II. A simulation model

M. P. Verge^{a)}

Laboratoire d'Acoustique Musicale, CNRS, Université de Paris VI, Ministère de la Culture, 4, Place Jussieu, case 161, 75252 Paris Cedex 05, France

A. Hirschberg

Technische Universiteit Eindhoven, W & S 0.54, Postbus 513, 5600 MB, Eindhoven, The Netherlands

R. Causse

IRCAM, 1 Place Igor Stravinsky, 75004, Paris, France

(Received 14 June 1996; accepted 17 December 1996)

A simple one-dimensional representation of recorderlike instruments, that can be used for sound synthesis by physical modeling of flutelike instruments, is presented. This model combines the effects on the sound production by the instrument of the jet oscillations, vortex shedding at the edge of the labium, and turbulence in the mouth of the instrument. The jet oscillation model used is a modification of the semi-empirical model by Fletcher [J. Acoust. Soc. Am. **60**, 926–936 (1976)]. The steady-state drive of the acoustical oscillations in the pipe by the jet motion is represented by a pressure jump in the mouth of the instrument. Vortex shedding at the edge of the labium during steady-state operation is taken into account by the use of a free-jet model. The combined effects of this nonlinearity and the jet-drive model enable one to correctly predict the steady-state amplitude of the fundamental. The turbulence noise source is represented by an additional pressure jump across the mouth of the instrument having an amplitude scaling with the square of the jet velocity. This simple model appears to correctly predict the noise level and its spectrum, both with and without oscillations of the jet. The transient response of the model is triggered by the initial volume injection into the mouth of the instrument and is dependent on the steepness of the driving pressure rise. © 1997 Acoustical Society of America. [S0001-4966(97)06004-9]

PACS numbers: 43.75.Np, 43.28.Ra, 43.75.Wx [WJS]

INTRODUCTION

The objective of this paper is to present a simple model of recorderlike instruments that could be used for sound synthesis by physical modeling. Despite the apparent simplicity of flue instruments, their sounding mechanisms are not yet clearly understood. The difficulties arise from the complex nature of the flow in the embouchure of the instrument. In this small region, one can observe simultaneously a hydrodynamically unstable jet interacting with an acoustic field, vortex shedding, and turbulence. The exact simulation of these fluid dynamics phenomena requires the use of direct simulation techniques in order to solve the Navier–Stokes equations. This kind of work has already been undertaken by Skordos¹ for low Reynolds number flows. This approach involves tremendous computing power and is certainly not adapted for time-domain sound synthesis. Instead, it is preferred to rely on simplified representations of the hydrodynamical phenomena involved in the sound production that yet enable one to simulate the basic features of measured acoustic signals. The model was developed from extensive sets of measurements and flow visualizations performed on a small experimental flue organ pipe.^{2–5} The main concern was to develop a model that could reproduce the main features of measurements with as little arbitrary parameter adjustment as

possible and in particular the Strouhal number dependency of the dimensionless amplitude of the fundamental discussed in the companion to this paper. This dependency indeed appears to be a fundamental characteristic of the functioning of flutelike instruments.

I. THE BASIC MODEL

The simple one-dimensional model shown in Fig. 1 is used to simulate sound production by recorderlike instruments. This simple low-frequency approximation representation was obtained by “unfolding” the two-dimensional geometry of a recorderlike instrument shown in Fig. 2. The resonator is represented by a large tube of cross section $S_p = H^2$ and length $L_p + \delta_p$, where L_p is the length of the pipe and δ_p an end correction associated with the open pipe termination. The mouth of the instrument is represented by a smaller tube of cross section $S_m = WH$, where W is the distance between the flue exit and the labium. Because the mouth region is small compared to the wavelengths considered, the flow in this region can be assumed to be incompressible. The complicated two-dimensional geometry of the mouth of the instrument can therefore be represented by an equivalent pipe segment of length δ_m . The length of this “end correction” is determined by inertia effects related to the acoustic flow at the embouchure of the instrument and to the constriction in the mouth of the instrument resulting from the presence of the labium.

^{a)}Part of this work was done while the author was at IRCAM and the Technische Universiteit Eindhoven.

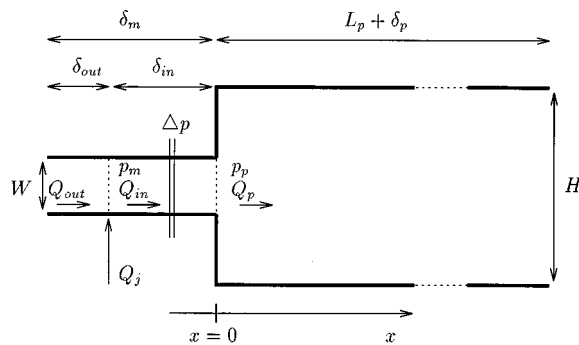


FIG. 1. Simple one-dimensional model used to simulate sound production by recorderlike instruments.

In this one-dimensional representation of the instrument, the flue exit, where the jet is formed, is located at an “acoustic” distance δ_{out} from the “outside world” and δ_{in} from the entrance of the resonator ($\delta_m = \delta_{in} + \delta_{out}$). The details of the determination of these parameters were discussed by Verge *et al.*³ In the same paper, the flue exit was shown to be a volume source which acts as a sound source when the jet velocity U_j fluctuates. This effect appears to be very important during the attack transient since it provides, in combination with the initial interaction of the jet with the labium, the signal that triggers the attack transient.

During steady-state operation, the jet is submitted to the transverse acoustic volume flux Q_m at the flue exit which controls the oscillating motion of the jet in the mouth of the instrument. At the labium, this movement results in a complex flow which is lumped, as proposed by Powell,⁶ Elder,⁷ and Coltman,⁸ into two complementary flow sources Q_1 and Q_2 , placed on each side of the labium, as shown in Fig. 3. As discussed by Verge *et al.*,² these two sources induce a pressure difference across the mouth of the instrument which constitutes the “jet-drive” mechanism for the acoustic oscillations in the pipe. In a one-dimensional representation, the complementary sources can be represented by a dipole placed in the end correction δ_m . For a given distance of the injection point from the edge of the labium, the “acoustic” distance δ_d between the sources, determining the strength of the dipole, can be calculated from the two-dimensional geometry of the instrument, in the potential flow approximation, by using conformal mapping techniques.² In analogy with Powell,⁹ Elder,⁷ and Coltman,⁸ it is assumed that the injection points are located at a fixed position independent of the jet flow velocity. The distance between these injection points and the edge of the labium is estimated to be of the order of the width h of the jet. These arbitrary assumptions

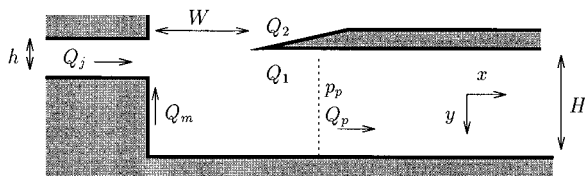


FIG. 2. Two-dimensional representation of recorderlike instruments.

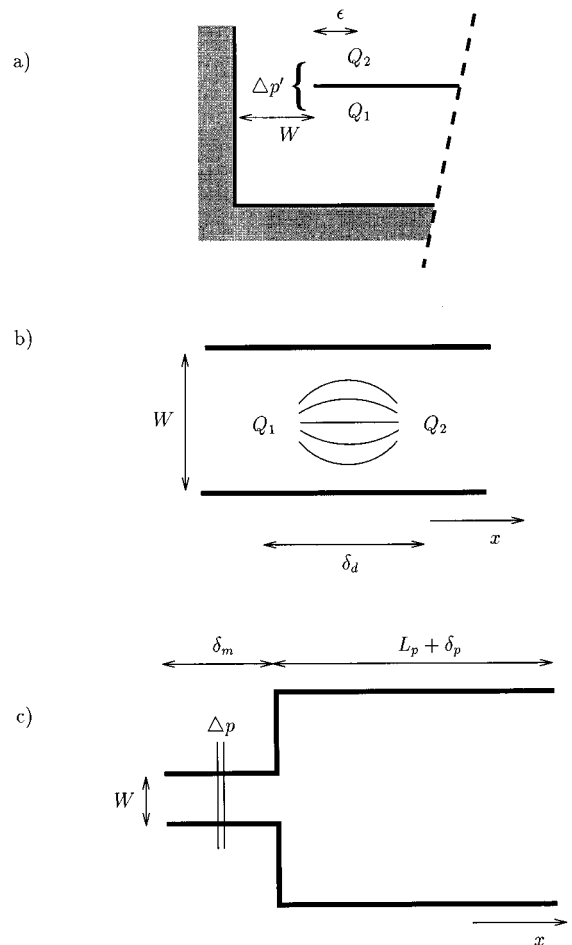


FIG. 3. Pressure source Δp at the labium; (a) real mouth geometry with the sources Q_1 and Q_2 a distance ϵ from the edge of the labium; (b) equivalent one-dimensional geometry: The complementary sources are placed a distance δ_d apart in a tube having a cross section equal to that of the mouth; (c) equivalent pressure source of strength Δp in a one-dimensional representation of the instrument.

provide reasonable results. Such a dipole induces a fluctuating pressure jump Δp across the pipe segment representing the mouth of the instrument, such as shown in Figs. 1 and 3.

The model is driven by a pressure signal p_f representing the pressure in the foot of a flue organ pipe or in the mouth of the musician. Neglecting friction, a relationship between the pressure drop $p_f - p_m$ across the flue canal, where p_m is the acoustic pressure at the flue exit, and the resulting jet velocity U_j at the flue exit is obtained from the Bernoulli equation¹⁰

$$\rho_0 l_c \frac{dU_j}{dt} + \frac{1}{2} \rho_0 U_j^2 = p_f - p_m, \quad (1)$$

where ρ_0 is the air density, l_c is the length of the flue canal, and t is time. As was discussed by Verge *et al.*,³ this representation enables one to take into account the effects of the jet velocity variations due to pressure fluctuations at the flue exit. At the flue exit, the mass conservation law can be applied

$$Q_j + Q_{out} = Q_{in}, \quad (2)$$

where Q_j , Q_{out} , and Q_{in} are the jet flow and the flow in the portion δ_{out} and δ_{in} of the mouth end correction, respectively. At the flue exit ($x = -\delta_{in}$), the pressure p_m is determined by the value of the radiation impedance Z_{out} . In a low-frequency approximation (with the sign convention of Fig. 1)¹¹

$$Z_{out} = \frac{p_m}{Q_{out}} = -\frac{\rho_0 c_0}{S_m} \left[\frac{1}{4} \left(\frac{\omega r_m}{c_0} \right)^2 + i \frac{\omega}{c_0} \delta_{out} \right], \quad (3)$$

where ω is the angular frequency, r_m the radius of a circle having the same cross section as that of the mouth ($\pi r_m^2 = S_m$), and c_0 is the speed of sound. By inverse Laplace transformation, a time-domain relationship between the pressure p_m at the flue exit and the flow Q_{out} is obtained:

$$p_m = \frac{\rho_0 c_0}{S_m} \left(\frac{1}{4} \frac{r_m^2}{c_0^2} \frac{d^2 Q_{out}}{dt^2} - \frac{\delta_{out}}{c_0} \frac{dQ_{out}}{dt} \right). \quad (4)$$

The pressure p_p at the entrance of the resonator can be related to the pressure p_m at the flue exit by applying the Bernoulli equation in the pipe representing the mouth of the instrument

$$p_p - p_m = -\frac{\rho_0 \delta_{in}}{S_m} \frac{dQ_{in}}{dt} + \Delta p, \quad (5)$$

where continuity of pressure at the entrance of the resonator ($x=0$) was assumed and the jet-drive mechanism was represented by a pressure jump Δp across the end correction δ_{in} . Note that the other sound production mechanisms (vortex shedding at the labium and turbulence noise) can also be included, as will be discussed later, in this pressure jump. At the interface between the mouth of the instrument and the pipe, flow continuity is also assumed:

$$Q_p(0) = Q_{in}(0), \quad (6)$$

where Q_p is the volume flux in the resonator.

The amplitude of the pressure jump Δp_{jd} associated with the jet-drive is determined, as discussed by Verge *et al.*,² by the time derivative of the flow source Q_1 corresponding to the portion of the jet flow entering the pipe at the labium:

$$\Delta p_{jd} = -\frac{\rho_0 \delta_d}{S_m} \frac{dQ_1}{dt}, \quad (7)$$

where δ_d is the ‘‘acoustic’’ distance between the sources Q_1 and Q_2 in the one-dimensional representation of the instrument. The flow Q_1 is calculated, as suggested by Cremer and Ising,¹² from the estimated jet position $\eta(W, t)$ at the labium

$$Q_1 = H \int_{(y_0 - \eta)}^{\infty} U_j(y) dy, \quad (8)$$

where H is the jet width, y_0 the labium position with respect to the flue exit axis, and where the y axis is positive toward the interior of the pipe. By assuming that the jet has a Bickley velocity profile [$U_j(y) = U_0 \operatorname{sech}^2(y/b)$], as proposed by Powell,⁶ the following expression is obtained:

$$Q_1 = b H U_0 \left[1 + \tanh \left(\frac{\eta - y_0}{b} \right) \right], \quad (9)$$

where U_0 is the jet central velocity and b is the jet width parameter which determines the velocity profile at the flue exit and the stability properties of the jet.²

Equations (1), (2), (4), (5), (6), (7), and (8) constitute the basic recorder model. The details of the implementation of the different elements of this model are now discussed.

II. THE JET OSCILLATION MODEL

The problem of the jet oscillation is very complex. There is, at the present time, no available physical model that enables an accurate description of the jet motion and the determination of the transverse position $\eta(W, t)$ of the jet at the labium. Verge *et al.*² have proposed to use a modification of the semi-empirical formula by Fletcher and Thwaites¹³ for laminar jets which is based on the model proposed by Powell;⁶

$$\eta(W, \omega) = \frac{1}{i\omega} \left(\frac{2}{\pi} \frac{Q_p}{S_m} - \frac{0.38 Q_1'}{S_m} \right) \times \left[1 - \cosh(\mu W) \exp \left(\frac{-i\omega W}{u} \right) \right], \quad (10)$$

where W is the distance between the flue exit and the edge of the labium, $\mu(\text{St})$ is an amplification coefficient, $u(\text{St})$ the speed of the hydrodynamic perturbation on the jet, St the Strouhal number, and where Q_1' is defined as

$$Q_1' = Q_1 - \frac{1}{2} b H U_0. \quad (11)$$

The terms in the parentheses of Eq. (10) represent the transverse acoustic velocity v_m' at the flue exit. As discussed by Verge *et al.*,² this velocity is the sum of a contribution due to the acoustic field from the pipe and a direct hydrodynamic feedback from the edge of the labium, which may account for an edgetone behavior of the instrument. The hyperbolic cosine in the second factor of Eq. (10) represents the amplification of the perturbation by the jet, and the complex exponential the phase delay due to its convection towards the labium. The effect of these two terms depends strongly on the value of the parameters $\mu(\text{St}_b)$ and $u(\text{St}_b)$, and it is proposed to use the data, based on a spatial stability analysis, calculated by Mattingly and Criminale¹⁴ for an infinite parallel frictionless flow having a Bickley velocity profile

$$U_j(y) = U_0 \operatorname{sech}^2 \left(\frac{y}{b} \right), \quad (12)$$

where y is the distance from the center of the jet, U_0 the centerline velocity, and b a jet width parameter. In the earlier works of Sato¹⁵ and Fletcher,¹³ data corresponding to a temporal stability analysis are used. There is some indication that a spatial stability analysis is more appropriate to describe the response of a jet.¹⁶ Since the value of these parameters used in the simulations depends on the Strouhal number St_b based on the parameter b , the estimation of this latter parameter is crucial. Calculations based on the conservation of the jet momentum flux show that this parameter should be

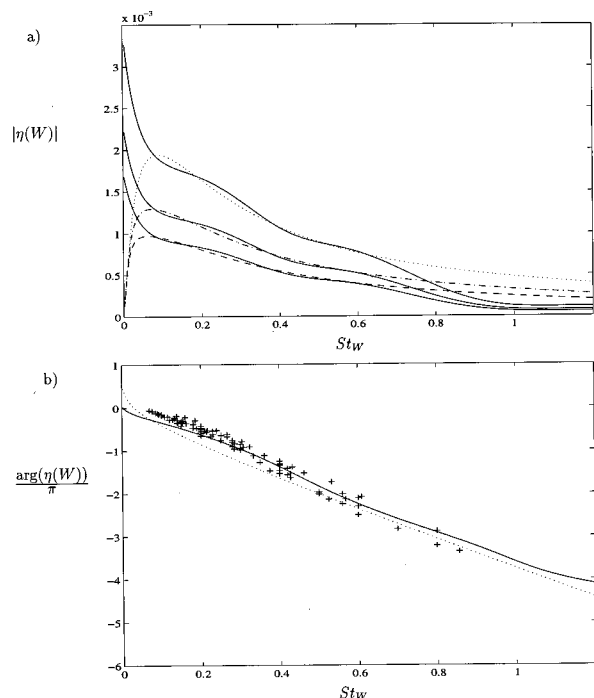


FIG. 4. (a) Modulus of the frequency response of a second-order Butterworth passband filter and comparison with the modulus of Eq. (10) (bold line) for different jet velocities: 10 m/s (dotted), 15 m/s (dashdot), and 20 m/s (dashed) as a function of the Strouhal number $St_W = fW/U_j$; (b) corresponding phase of the filter (dotted) and a delay line of length $0.6U_j$ and comparison with the phase of Eq. (10) (bold line) and experimental data by Coltman.⁴⁰

equal to $0.4h$ for a Poiseuille flow at the flue exit and to $0.75h$ for a uniform velocity profile at the flue exit.² Since the jet velocity profile depends on the convergence of the flue canal, this parameter should be characteristic of a given instrument. Note that Eq. (10) displays an unrealistic behavior at zero frequency because of the factor $1/i\omega$. Another difficulty in the implementation of the model is the fact that both μ and u depend on the Strouhal number St , which means that no simple analytical representation can be found in the time domain. It is therefore proposed to implement the model as a digital filter.

An order two Butterworth bandpass filter enables one to obtain, as can be seen from Fig. 4(a), a reasonable fit of the modulus of Eq. (10). Using a bandpass filter is necessary in order to get rid of the singularity of the model at zero frequency. Using a low-order filter does not, of course, yield a good fit of the jet model at high frequencies but is convenient for computation efficiency. One can argue anyway that the jet model is certainly not exact and that it therefore makes no sense to try to fit Eq. (10) exactly. Simulation results will be affected by the bandwidth chosen for the filters. It is therefore possible to change the behavior of the model by varying the sensitivity of the jet model at different frequencies. In this way one could account, in an arbitrary manner, for the effects of parameters, such as the geometry of the flue exit, which are certainly not taken into account by Eq. (10).

The convection of the perturbations toward the labium can be simulated by using a delay line. In a first approximation, the propagation velocity u of the perturbation on the jet

can be assumed to depend on the jet velocity only which allows to represent this effect by a simple delay line. Experimental data by Powell,¹⁷ Sato,¹⁵ and Maurel¹⁸ show that indeed the velocity u does not depend strongly on the Strouhal number. The total phase shift introduced by the jet filter and a delay line of length $0.6U_j/W$ gives, as shown in Fig. 4(b), a good fit of the phase response of Eq. (10) and of experimental data by Coltman^{2,5,40} for Strouhal numbers $St_W = fW/U_j > 0.15$.

In the simulations that will be presented, the filter coefficients were determined, for a given jet velocity, by selecting a bandwidth, which enabled to fit well the jet model response for the fundamental frequencies considered and the first harmonics. The transfer function of the filter was multiplied by an arbitrary constant in order to match the amplitude of the jet model. These steps must be repeated for the whole range of jet velocities that will be used during the simulations. Different sets of data must also be calculated for different flue exit to labium distance W and width parameter b . The filter coefficients were stored in a look-up table.

In implementing the jet model presented in this section, one has to be careful in choosing the value of the jet velocity used to determine the filter coefficients. The jet velocity U_j is determined by the pressure difference between the foot of the instrument (or the mouth of the musician) and the flue exit. Because of large pressure fluctuations at the flue exit, the jet velocity fluctuates at the same frequency as that of the acoustic oscillations in the resonator. This implies that the filter coefficients should vary, and consequently the amplitude of the filter response, within one period of oscillation. Originally, the fluctuating value of the jet velocity U_j was used in the calculation of the filter coefficients, but it was found that the model then displays a rather unrealistic behavior. The change in the response of the filter within an oscillation period results in asymmetrical oscillations of the jet. After a certain number of oscillations a mean displacement of the jet brings the jet entirely on one side of the labium. This awkward behavior was in fact introduced artificially by the jet model and has no physical foundations. It is due to the fact that the mean flow effects, which determine the average jet position, were not described.

In order to avoid these problems, it is proposed to use, when determining the jet coefficients, not the actual jet velocity, but the velocity obtained by assuming a zero pressure at the flue exit and using the stationary approximation ($U_j = \sqrt{2p_f/\rho_0}$). This drastic simplification of the model provides reasonable results. Note that U_j still displays a time dependence due to the variations of the driving pressure p_f .

The jet oscillation model that has just been discussed represents the motion of a laminar jet and a linear amplification of the perturbations acting on it. As was discussed in the companion to this paper, for flue instruments characterized by a large ratio W/h between the distance W from the flue exit to the labium and the height h of the flue exit, the operating conditions are such that the jet is usually turbulent. For turbulent jets, a model has been proposed by Bechert¹⁶ which should be compared to the works of Fletcher and Thwaites.¹³ Furthermore, for very low driving pressures, the

jet may oscillate on higher hydrodynamical mode, which is characterized by a nonlinear roll-up of the jet which limits its oscillation amplitude. A simplified jet model based on the assumption of a break down of the jet into a von Karman vortex street has been proposed by Holger *et al.*¹⁹ The jet model described by Eq. (10) is therefore most likely to be valid only for describing flue instruments having a small ratio W/h as is the case for recorders and must therefore be used with care.

III. THE RESONATOR

The well-known D'Alembert's solution of the wave equation enables one to represent the pressure signal p_p in the pipe of the instrument as the sum of two traveling waves, p_o and p_i :

$$p_p = p_o(x - c_0 t) + p_i(x + c_0 t). \quad (13)$$

For an observer placed at the entrance of the pipe, the incoming wave p_i is equal to the outgoing wave p_o delayed by the time of a roundtrip in the resonator and filtered by the effects of the visco-thermal losses and sound radiation at the end of the resonator. Using Euler's equation and this expression, the flow Q_p in the pipe can be expressed as

$$Q_p = \frac{S_p}{\rho_0 c_0} (p_o - p_i). \quad (14)$$

An efficient implementation method of the wave equation consists in the use of so-called digital waveguides.²⁰ This representation is based on the traveling-wave solution (13) of the wave equation and can be implemented with the use of delay lines.

The effects of sound radiation can be estimated by considering the value of the radiation impedance at the open extremities of the instrument. Radiation was already taken into account at the active extremity of the pipe with the use of Eq. (4). Similarly, the following expression is obtained at the end of the resonator with the help of Eqs. (13) and (14):

$$p_o + p_i = \frac{-r_p^2}{4c_0^2} \frac{\partial^2(p_o - p_i)}{\partial t^2} + \delta_p \frac{\partial(p_o - p_i)}{\partial t}. \quad (15)$$

By using a finite difference scheme to evaluate this expression, one can obtain a relationship between the outgoing and ingoing pressures p_o and p_i . Radiation losses can then be implemented by filtering the outgoing signal p_o before reflecting it back into the resonator. For the time-domain simulations that will be presented later, we used a second-order filter obtained from a backward scheme. The frequency response of this filter is shown in Fig. 5 and is compared to the value of the reflection coefficient given by Levine and Schwinger.²¹ The model we used in the simulation overestimates the radiation losses for frequencies below 12 kHz.

Visco-thermal damping induces a frequency-dependent dispersion and exponential decay of traveling waves. The effects of visco-thermal damping in a tube can be represented, in the frequency domain, by the use of a complex wave number k .²²

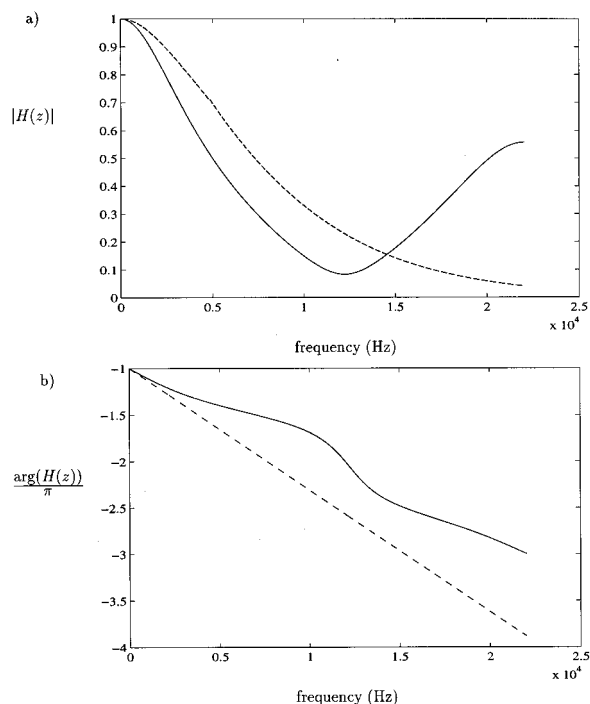


FIG. 5. (a) Modulus of the frequency response of the second-order filter used in the simulations to take into account radiation losses and comparison with theoretical values of the modulus of the reflection coefficient at the end of an unflanged circular pipe given by Levine and Schwinger (1948) (dashed line) and (b) corresponding phase (bold line) compared to the phase shift introduced by an end correction of length $\delta_p = 0.82r_p$ (dashed line).

$$p_p = (p_o e^{-ik^+ x} + p_i e^{ik^- x}) e^{i\omega t}, \quad (16)$$

where

$$k^\pm = \frac{\omega}{c_0} \pm \sqrt{\frac{\omega}{2}} \frac{\alpha}{r_p} (1 - i), \quad (17)$$

with α a damping coefficient given by

$$\alpha = \frac{\sqrt{l_v} + (\gamma - 1) \sqrt{l_t}}{\sqrt{c_0}}, \quad (18)$$

where l_v and l_t are the viscous and thermal characteristic lengths (in air at 20 °C, $l_v = 4.0 \times 10^{-8}$ m, and $l_t = 5.6 \times 10^{-8}$ m) and γ is the Poisson constant, i.e., the ratio of specific heats at constant pressure and volume, respectively ($\gamma = 1.4$).

An exact time-domain transformation of Eq. (16) involves fractional derivatives which are introduced by the $\sqrt{\omega}$ term in Eq. (17). This complex problem is beyond the scope of this work, but has been discussed at length by Matignon.²³ A low-order IIR filter was used in the simulation model. For efficiency, instead of filtering the pressure signal at every tube cell, the effects of damping have been lumped at both extremities of the tube. The frequency response of an order 3 IIR filter representing the effects of damping on a wave having traveled a whole length $L_p = 0.283$ m of pipe is shown in Fig. 6 and is compared to the theoretical value obtained by using the complex wave number of Eq. (17). The filter coefficients were found numerically by using a method based on minimizing the weighted sum of the

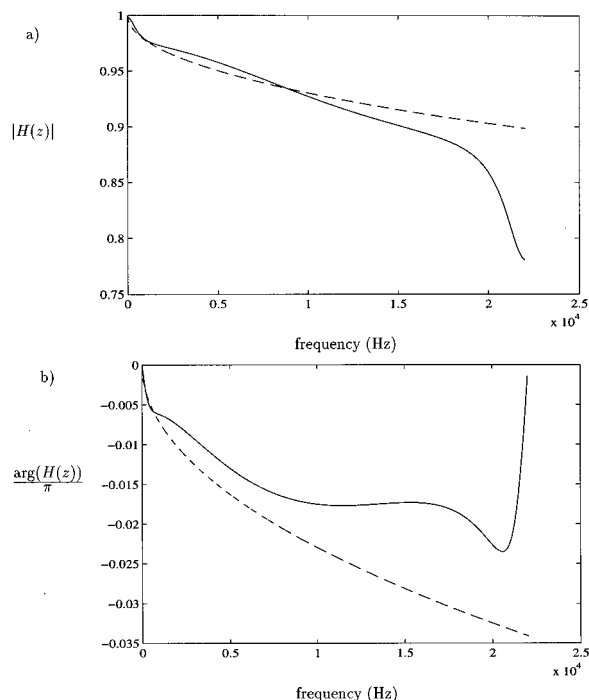


FIG. 6. (a) Magnitude and (b) phase of the frequency response of the filter representing the effects of visco-thermal damping: results obtained from Eqs. (16) and (17) (bold line) and an order 3 IIR filter (dashed line).

squared error between the response of the filter and the desired one. The frequency response of this filter is shown in Fig. 6.

IV. VORTEX SHEDDING AT THE LABIUM

The jet-drive model discussed in the previous section appears to predict reasonably well the linear behavior of a recorder during a smooth attack transient.² In order to explain the nonlinear behavior of the instrument, however, vortex shedding at the edge of the labium must be taken into account and should constitute a crucial element of a simulation model. This has already been pointed out by Coltman²⁴ when he found that the mouth impedance of a flute behaves nonlinearly at acoustic amplitudes corresponding to playing conditions.

Vortex shedding in flue instruments has been discussed by Fabre *et al.*^{25,4} During steady-state, vortex shedding at the edge of the labium appears to be an important nonlinear amplitude limiting mechanism. It is also correlated with the generation of high harmonics in the tone of the instrument. Flow separation at the edge of the labium, resulting in the formation of vortices, appears to be due to the interaction of both the jet flow and the acoustic flow Q_p from the resonator with the edge of the labium. At the present time, an accurate model of the periodic vortex shedding induced by the jet is not available and therefore, only the vortex shedding associated with the quasi-steady separation of the acoustic flow will be taken into account in the simulation model. Vortex shedding also appears to play a crucial role in the triggering of the attack transient.³ A crude model of this effect will be proposed here.

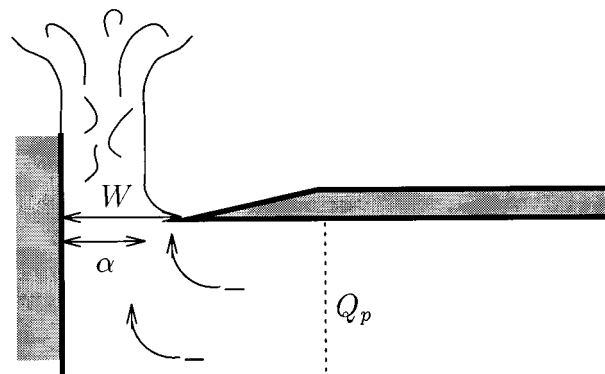


FIG. 7. Flow separation of the acoustic flow from the resonator, resulting in the formation of a jet in the mouth of the instrument.

The vortex shedding induced at the labium by the transverse acoustic flow from the pipe can be modeled by assuming that flow separation of the acoustic flow Q_p at the labium results in the formation of a free jet in the mouth of the instrument as shown in Fig. 7. In this quasi-stationary free-jet model, proposed by Ingard and Ising,²⁶ the effects of vortices are represented by a fluctuating pressure jump Δp_a across the mouth of the instrument

$$\Delta p_a = -\frac{1}{2} \rho_0 \left(\frac{Q_p}{\alpha_v S_m} \right)^2 \text{sign}(Q_p), \quad (19)$$

where α_v is the vena-contracta factor of the jet, which, depending on the geometry, can vary within the range $0.6 \leq \alpha_v \leq 1$. A value $\alpha_v = 0.6$, which should correspond to a sharp labium, was later used. This pressure drop opposes the acoustical flow in the mouth and corresponds to the dissipation of the kinetic energy of the free jet by turbulence. Energy losses E_a induced by this model scale with the third power of the acoustic flow

$$E_a = \int_0^T \Delta p_a \frac{Q_p}{S_m} dt, \quad (20)$$

where T is the period of the acoustic signal. This crude model should constitute an upper limit approximation of the effects of this vortex shedding because it implicitly assumes that a vortex exists during the entire period of the fundamental. The generation of high harmonics by this simple vortex model is hard to estimate, and beyond the scope of this paper, because of the nonlinear nature of the model. It can be suggested, however, that breaking the symmetry of the vortex model, by assuming, for example, that more vorticity is shed on one side of the labium than on the other one, could strongly affect the amplitude of the higher harmonics. This could be a way to qualitatively account for the effects of the shape of the labium which is not taken into account in the free-jet model.

During the attack transient, the vortex shedding due to the initial interaction of the jet with the labium is crucial.³ When the jet hits the labium, a vortex is shed which appears to be important in the triggering of the attack transient and the generation of high frequencies during the transient. In order to determine the jet position during its formation, it is

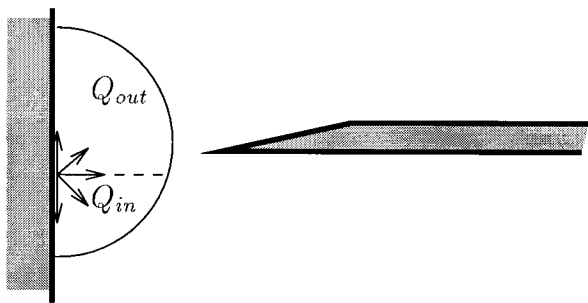


FIG. 8. Vertical position of the half-circle representing the forming jet such as determined by the ratio of the flow Q_{in} and Q_{out} in the mouth of the instrument. In this model, the flue exit is represented as a volume point source. Its vertical position is indicated by the dashed line.

proposed to represent the flue exit, in a two-dimensional potential flow approximation, as a point source $Q_j(t)$ flowing into a semi-infinite space. The forming jet therefore appears as a growing semi-circle. The vertical position of this semi-circle is determined by the ratio of the flow Q_{out} and Q_{in} in the mouth of the instrument, as shown in Fig. 8. This model should constitute a lower limit approximation.³ When the semi-circle hits the labium, a pressure pulse Δp_{tr} having an amplitude proportional to the square of the jet front velocity U_{fr} is sent in the pipe:

$$\Delta p_{tr} \propto \frac{1}{2} \rho_0 U_{fr}^2. \quad (21)$$

The proportionality constant in this formula was determined by fitting experimental data.

V. NOISE PRODUCTION BY TURBULENCE

Turbulence noise constitutes an essential component of the tone of flutelike instruments. Although it does not contribute fundamentally to the functioning of the instrument, a noise module is important to include in a simulation model in order to obtain a natural sounding synthesis. Figure 9 shows a power spectrum measured in the small recorderlike organ pipe, described by Fabre and Verge,^{4,3} upon steady blowing conditions. The harmonically related frequency components associated with the feedback mechanism of the instrument clearly dominate the spectrum. However, the tur-

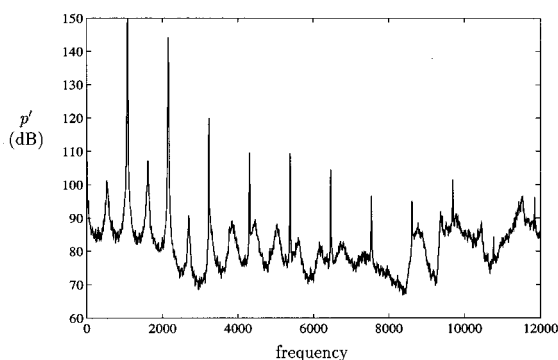


FIG. 9. Power spectrum of the pressure signal p' (in dB re: 20 μ Pa) measured in a small experimental recorderlike organ pipe at the end of the resonator (50.25 mm from the passive end of a 0.283-m-long pipe) for a driving pressure p_f of 2500 Pa.

bulent structure of the jet flow also produces a strong background noise. This broadband sound source is filtered by the pipe resulting in the presence of little bumps at frequencies corresponding to passive resonances of the pipe. As will be shown, a simple turbulence noise source can be used, in the one-dimensional representation of the instrument, to reproduce this type of background noise.

The Lighthill analogy^{27,28} enables one to describe the sound production from so-called *aero-acoustical* sources induced by unsteady flows. The noise produced by turbulent jets can only be characterized analytically for simple geometries since noise production depends on the jet shape and is strongly affected by the presence of bodies or walls in the flow. This implies that in order to go beyond generalities, one has to rely on empirical data to describe noise production in complex geometries such as that of flue instruments.

Turbulence in free field is dominated by quadrupoles. Lighthill showed that for a subsonic flow as in flue instruments, it is a very inefficient sound production mechanism with a radiated sound power scaling with the eighth power of the Mach number $(U_j/c_0)^8$. The efficiency of the noise source is enhanced, however, by the presence of objects in the flow. In the case of the interaction of a compact edge with a turbulent flow, the sound production is dominated by the force provided by the edge and therefore has a dipolar nature. This was proved by Powell⁹ in a beautiful experiment. Taking as an upper limit the force necessary to bring the jet to rest on the body in a quasi-stationary process, one can show that this force is proportional to the square of the jet velocity $(\frac{1}{2}\rho U_j^2)$. In free-field conditions, this results in a sound intensity scaling with the sixth power of the Mach number $(U_j/c_0)^6$. When placed in a pipe of infinite length, such a source can be shown, following Ffowcs-Williams,²⁹ to radiate at low frequencies like the fourth power of the jet Mach number $(U_j/c_0)^4$.⁵

Because of the presence of the edge of the labium, the turbulence noise source in flue instruments can be expected to display a dipolar nature and therefore to have an intensity that scales with the square of the jet velocity $(\frac{1}{2}\rho U_j^2)$. Because of the presence of the edge of the labium, it can be assumed that, in analogy to the free-field edgetone, the force term corresponding to the jet-labium interaction dominates over the volume integral term associated with the quadrupolar turbulence sound sources. This assumption is subject to uncertainties, however, because, unlike the case of an infinite pipe, there is no simple analytical theoretical argument to support such a statement. For high Reynolds numbers one can expect that this force exerted on the labium scales with the jet stagnation pressure $\frac{1}{2}\rho U_j^2$ as was found by Powell in the case of the edgetone. This of course should be checked because, in playing conditions, the Reynolds number of the flow is rather low and can therefore exhibit an additional velocity dependency due to the evolution of the jet turbulence with the Reynolds number. It is at least expected that there exists a critical Reynolds number R_c below which turbulence will be negligible.

In a preliminary study of turbulence noise in flue instruments,⁵ noise sources were deduced from measurements on an experimental pipe. The measured noise spectra

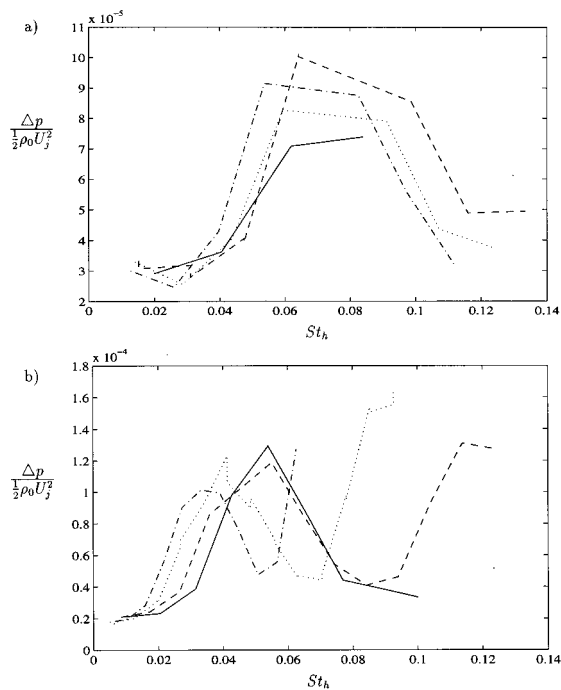


FIG. 10. Amplitude of the turbulence noise source Δp calculated from acoustic pressure measurements made at the end of the pipe as a function of the Strouhal number $St = fh/U_j$: (a) driving pressure $p_f = 400$ Pa (solid line), 690 Pa (dashed), 800 Pa (dotted), and 980 Pa (dashdot); (b) $p_f = 1560$ Pa (solid), 2200 Pa (dashed), 3920 Pa (dotted), and 6080 Pa (dashdot).

were obtained by damping the auto-oscillations of the pipe by using a muffler or lifting the labium above the flue exit. The spectra of the noise sources were then deduced from these measurements and the one-dimensional model of the instrument shown in Fig. 1. In this representation, a dipolar sound source is equivalent to the pressure jump of amplitude Δp across the mouth of the instrument, similar to the other sound generating mechanisms. The amplitude of the sound source at a given frequency and jet velocity was assumed to be equal to the amplitude of the pressure jump that would result in the measurements. This operation was repeated for different frequencies and jet velocities in order to obtain the noise sources shown in Fig. 10.

The noise sources appear to have a bell-shaped spectrum. Their dimensionless amplitude of the different sources is relatively constant over a wide range of driving pressures, which indicates that the assumption of a source having a dipolar character and scaling with the square of the jet velocity is reasonable. The calculated noise sources display a maximum amplitude for values of the Strouhal number ($St = fh/U_j$) shifting from approximately 0.08 to 0.03 as the blowing pressure is increased which can be linked to the maximum instability of a planar jet.^{15,30} For a circular jet, one would expect a maximum of noise production by the jet at a Strouhal number $St_D = fD/U_j$ (where D is the jet diameter) of the order of one,³¹ which implies a maximum of sound production at much higher frequencies than for a planar jet. The planar shape of the jet in recorderlike instruments is therefore important perceptively since it implies that noise will be produced at characteristic frequencies where the ear is more sensitive. This difference in behavior between a pla-

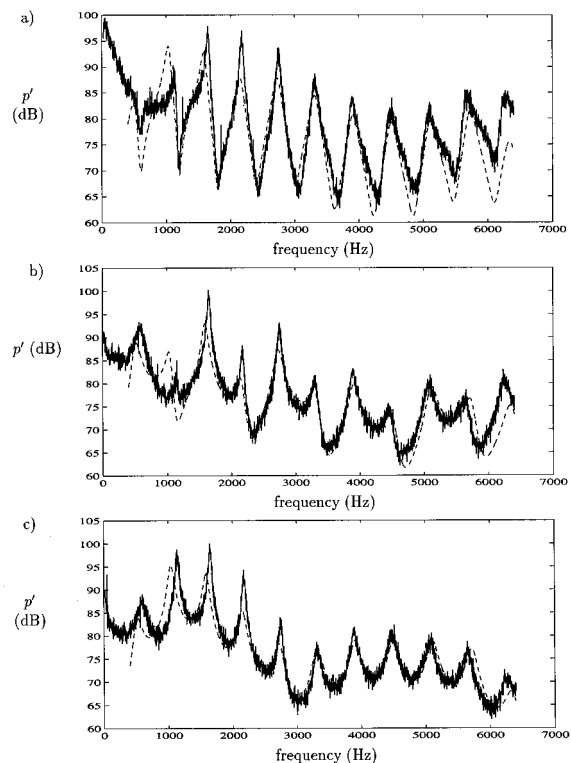


FIG. 11. Comparison between a power spectrum of the pressure p' (in dB re: 20 μ Pa) measured (bold line) in the pipe with a muffler placed at the end of the pipe (no pipe tone) at a driving pressure of 7250 Pa and the corresponding predicted (dashed line) spectrum. (a) Under the labium; (b) in the middle of the pipe; (c) at the end of the pipe.

nar and a circular jet can be explained by the fact that the most unstable modes of a circular jet correspond to varicose (symmetrical) rather than sinusoidal (antisymmetrical) oscillations as is the case for a planar jet.

The validity of these sources can be checked by reintroducing them into the flute model. The amplitude of the source for different jet velocities is obtained by modulating the amplitude of the source shown in Fig. 10 by the square of the jet velocity U_j . Figure 11 presents a comparison between a measured and an estimated noise spectrum at different points in the resonator. In Fig. 12, the same estimated signals are compared to power spectra measured with auto-oscillations of the pipe. The general noise level is well predicted in both cases, which indicates that the motion of the jet in playing conditions does not appear to modify substantially the nature of the turbulence noise source. In these examples, simulations were compared to measurements performed at fairly high driving pressures, certainly higher than typical playing conditions. These measurements were chosen because they enable one to clearly see the filtering effect on the noise source of the longitudinal resonances of the pipe in the frequency range of interest. For lower jet velocities, components of the measured noise spectra, which can be associated with edgetonelike oscillations of the jet, appear at lower frequencies and “corrupt” the noise measurements,⁵ which does not enable one to compare our model with the measurements as clearly as in Figs. 11 and 12. These edgetonelike oscillations of the jet, which can be observed above the cut-

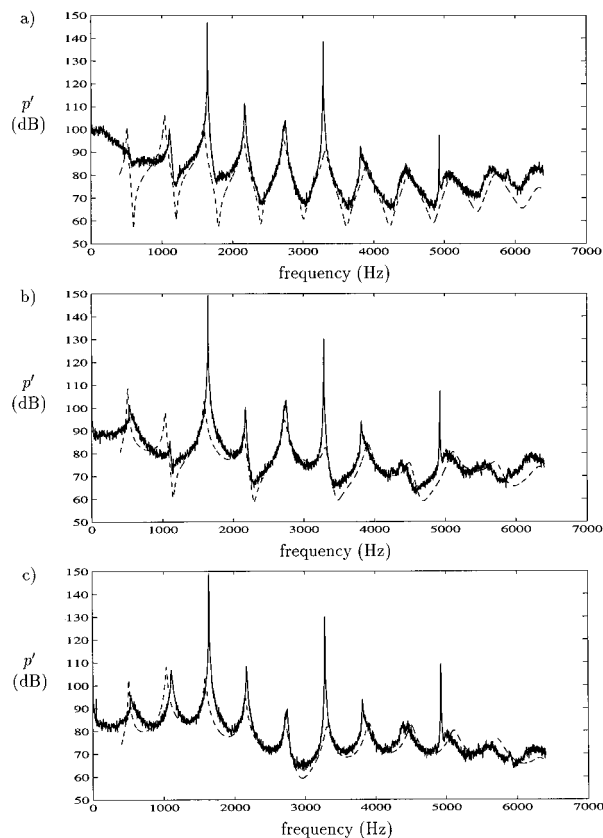


FIG. 12. Comparison between a power spectrum of the pressure p' (in dB re: 20 μ Pa) measured in the pipe without muffler (with a pipe tone) at a driving pressure of 7250 Pa (solid line) and the corresponding predicted spectrum (dashed line). (a) Under the labium; (b) in the middle of the pipe; (c) at the end of the pipe.

off frequency of the pipe in the spectrum of Fig. 13, appear as a result of random feedback loops which are not directly associated with noise production by turbulence and do not appear in normal playing conditions because of the jet motion. Simulations performed for lower driving pressures gave results similar to that presented in Figs. 11 and 12.

Flow visualizations of the jet in a recorderlike experimental flue pipe show that the nature of the flow in the mouth of the instrument is strongly dependent on the driving

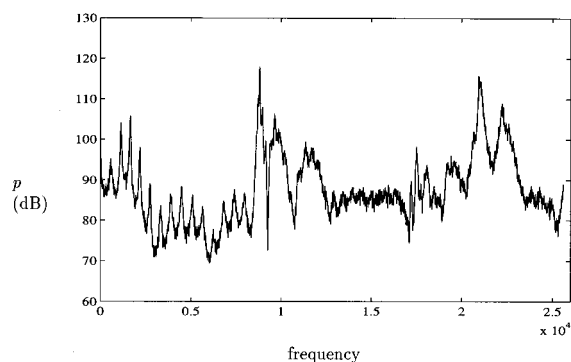


FIG. 13. Power spectrum of the pressure signal p (in dB re: 20 μ Pa) measured in a small experimental recorderlike organ pipe at the end of the resonator (50.25 mm from the passive end of a 0.283-m long pipe) for a driving pressure p_f of 7250 Pa with a muffler placed at the end of the pipe.

pressure p_f .⁵ For driving pressures below 60 Pa the flow is laminar. This driving pressure corresponds to a jet velocity U_j of 10 m/s and a Reynolds number (based on the flue exit height h) of approximately 700. When the jet velocity is increased, the coherent two-dimensional structure of the jet is unstable after it has reached the labium and the flow then becomes turbulent. The Reynolds number is therefore high enough for the jet to develop into turbulence but only after a certain delay or having interacted with the edge of the labium. Because of the jet oscillations, puffs of turbulence appear periodically on each side of the labium and the noise is therefore expected to display a pulsating (intermittent) character. When the Reynolds number reaches a value of approximately 3000 (a jet velocity of about 45 m/s corresponding to a driving pressure of 1200 Pa), the jet is already fully turbulent at the flue exit. The turbulence noise is then expected to be less dependent on the acoustically driven jet oscillations. In real instruments, the importance of the noisy component in the tone is therefore expected to vary greatly depending on the playing conditions. Consequently, a thorough study of the behavior of the noise source at low blowing pressure should include an analysis of the dynamics of the noise production phenomena.

For simulations purposes, it is proposed to use a bell-shaped broadband source having an amplitude given by

$$\Delta p_t \propto \frac{1}{2} \rho_0 U_j^2. \quad (22)$$

In the simulations that will be presented, we used a fit of the sources shown in Fig. 10. This source should be triggered for Reynolds numbers $Re > 700$. In the range $700 < Re < 3000$ the turbulence noise is expected to be intermittent as a result of the high amplitude of the jet oscillations. Its amplitude should therefore be modulated with the jet oscillations. It can furthermore be speculated that the phase between these two signals, introduced by the time taken by the jet flow to become turbulent, might be perceptively important. In speech synthesis, similar phenomena are observed because of the pulsed opening and closing of the vocal folds and variations of the phase between the noise and the periodic part of the signal appear to strongly affect the quality of synthetic voice.³² Similar ideas were explored by Chafe³³ for adding breath noise to synthesized flute sound.

Finally, it should be noted that this model is limited and is not adapted to reproduce the measured spectra in all their complexity. Figure 13 shows a spectrum obtained by damping the auto-oscillations of the system by placing a muffler at the end of the resonator. At low frequencies, the filtering effect of the noise source by the longitudinal passive resonances of the pipe are clearly observed. The spectrum is complicated, however, by a spectacular jump of more than 30 dB at 8.1 kHz which corresponds to the cutoff frequency of the pipe. Clearly, the turbulence noise source also excites transverse modes of the pipe that would require a two-dimensional description of the resonator. This resonance explains the second increase in the amplitude of the deduced sources presented in Fig. 10(b). For large resonators, this jump might occur at a frequency low enough to be perceptually significant. For example, in typical bass recorders, the resonator has a diameter of the order of 35 mm, which results

in a cut-off frequency around 5 kHz. In large organ pipes, this is certainly quite relevant. High-frequency resonances are also observed in the spectrum of Fig. 13, which appear to be linked to the maintenance of complex edgetone feedback loops.⁵ Note that this strong influence of the pipe acoustics on the one-dimensional source model results into a frequency dependency of the source amplitude which is not described by the Strouhal number.⁵

VI. SIMULATION RESULTS

Equations (1), (2), (4), (5), (6), (7), (9), (10), (13), and (14) constitute the complete recorder model. The driving pressure signal p_f is considered to be a known parameter as well as the incoming pressure wave $p_i(0)$ at the entrance of the resonator since it can be determined from past values of the outgoing signal $p_o(0)$. One therefore has a set of ten equations with ten unknowns which must be solved simultaneously. By using Eqs. (2), (4), (5), and (6), a relationship between the pressure p_p and the flow Q_p at the entrance of the resonator can be obtained. Then, by using Eqs. (13) and (14), one finds an expression with a single unknown:

$$p_o = \frac{r_m^2 S_p}{4c_0^2 S_m} \frac{\partial^2 (p_o - p_i)}{\partial t^2} - \frac{\rho_0 r_m^2}{4c_0 S_m} \frac{d^2 Q_j}{dt^2} - \frac{\delta m S_p}{c_0 S_m} \frac{\partial (p_o - p_i)}{\partial t} + \frac{\rho_0 \delta_{out}}{S_m} \frac{dQ_j}{dt} + \Delta p - p_i. \quad (23)$$

The outgoing wave p_o is determined by the effects of inertia in the mouth of the pipe, feedback from the resonator, the pressure source in the mouth of the instrument, and the jet velocity fluctuations. During steady-state operation, the pressure source Δp is the main driving mechanism, while during the early stages of the attack transient the sound production is dominated by the effects of the jet velocity fluctuations. Note that the different sound producing mechanisms at the edge of the labium (jet-drive, vortex shedding at the edge of the labium and turbulence) are included in the pressure jump Δp :

$$\Delta p = \Delta p_{jd} + \Delta p_a + \Delta p_{ir} + \Delta p_t. \quad (24)$$

Simulation can now be obtained by solving numerically Eq. (23) by using finite differences schemes. This was implemented in "C" language on a DEC Alpha station for real-time sound synthesis. The numerical version of the model was validated by checking that the program could reproduce results obtained in the linear analysis of smooth transients presented by Verge *et al.*²

A. Attack transient

During the attack transient, Eq. (23) is dominated by the terms with time derivatives of the jet flow Q_j . Taking into account the acceleration of the jet flow generates a pressure pulse during the initial volume injection into the mouth of the instrument. This and the simple modeling of the early interaction of the jet with the labium enable the simulation of the differences between a slow and a fast attack which were discussed in detail by Verge *et al.*³

The transient behavior of the model appears to be very sensitive to the value of the different parameters of the

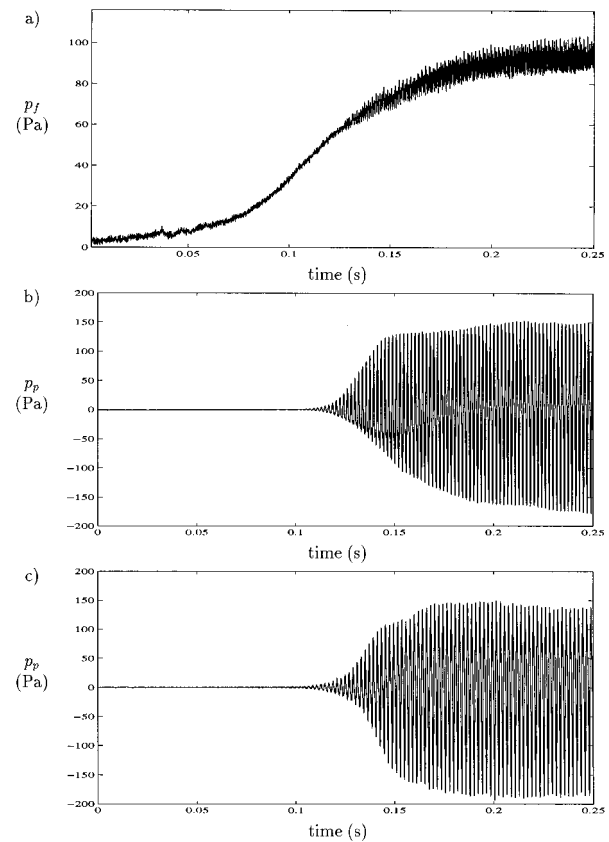


FIG. 14. (a) Driving pressure rise measured in the mouth of a musician playing a *smooth* transient on an alto recorder (Zen-on, Pol Bressan model). Steady-state oscillations correspond to the first mode of the pipe; (b) corresponding response of the instrument measured under the labium; (c) simulated response at the entrance of the pipe using the driving pressure rise shown in (a). In the simulations, the distance $\delta_d = 4.0$ mm and the width parameter $b = 0.5h$.

model. Different combinations of the model parameters can lead to very different transients and steady-state waveforms. This behavior is, after all, not so surprising in view of the fact that very different signals can be obtained on real instruments with very similar driving signals, almost identical instruments, or slight modifications of the same instrument. The simulation model appears to be especially sensitive to the value of the mouth and constriction end corrections, of the distance δ_d , the b parameter of the jet model, the length of the delay line representing propagation on the jet, and the shape of the jet filters used to simulate the jet oscillations. The strategy adopted in order to fit the measurements was to keep all the geometric parameters fixed and to only adjust the distance δ_d and the b parameter, within the limits calculated from physical arguments,² whose values are more arbitrary. The length of the delay line was chosen, for a given set of filter coefficients, so that it best reproduced the experimental data by Coltman presented in Fig. 4(b) and was then kept fixed.

1. Slow transients

Figure 14 shows a smooth attack transient, on the first mode of the pipe, measured on a plastic alto recorder (Zen-on, Pol Bressan model) and the corresponding driving pres-

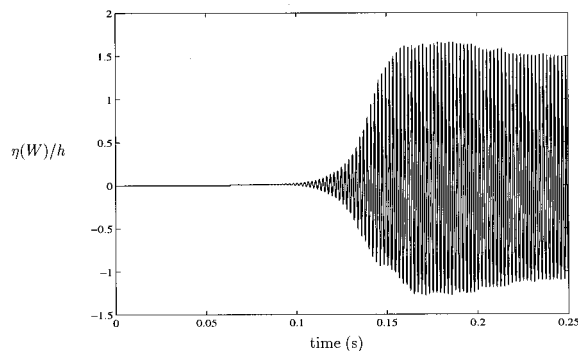


FIG. 15. Jet displacement $\eta(W)$ at the labium corresponding to the simulation of Fig. 14(c).

sure rise. The response $p_p(0)$ obtained from the simulation model with the measured driving pressure signal p_f of Fig. 14(a) as input is shown in Fig. 14(c). In this simulation, the width parameter b of the jet model was set to $0.5h$ and the distance δ_d between the sources Q_1 and Q_2 of the flow dipole in the mouth of the instrument was adjusted to 4 mm. This roughly corresponds to the value of δ_d calculated by Verge *et al.*² for the experimental pipe used when the sources are assumed to be located at a distance h from the edge of the labium. The value of the b parameter is very close to the optimal one found, in the linear analysis of smooth transients carried out on the same instrument and presented in the same paper. As expected, the oscillations appear to grow exponentially while the jet displacement at the labium, shown in Fig. 15, is smaller than its half-width. With these values, the threshold of the oscillations as well as the rate of increase of their amplitude appear to be well predicted which enables one to simulate the measured delay between the triggering of the driving pressure signal and the onset of the oscillations in the instrument. It is interesting to note that the pressure pulse due to the initial volume injection into the mouth of the instrument, although very faint, is sufficient to trigger the oscillations. The steady-state playing amplitude as well as the playing frequencies are well predicted.

2. Fast transients

Figure 16 shows a comparison between a fast transient response measured on the same alto recorder as in the previous section and a simulated signal. The measured driving pressure signal is presented in Fig. 16(a) and was used as the input signal of the simulation program. The measured response of the instrument is shown in Fig. 16(b). The response of the model, presented in Fig. 16(c), is triggered by the initial volume injection into the mouth of the instrument, resulting in the pressure pulse appearing at approximately 9 ms. The amplitude of this initial pulse depends on the time derivative of the jet flow Q_j and is well predicted by the model. This pulse is followed by a complex response due to the early interaction of the jet with the labium. This early interaction was very crudely modeled and its details are therefore not reproduced in the simulation signal of Fig. 16(c). The growing of the oscillation is quantitatively well reproduced and the oscillations reach the correct steady-state

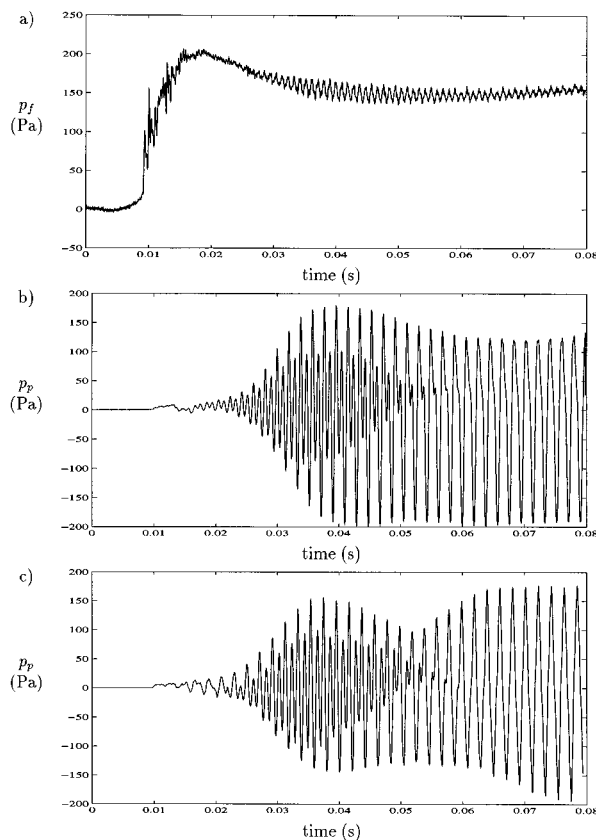


FIG. 16. (a) Driving pressure rise measured in the mouth of a musician playing a *fast* transient on an alto recorder (Zen-on, Pol Bressan model). Steady-state oscillations correspond to the first mode of the pipe; (b) corresponding response of the instrument measured under the labium; (c) simulated response at the entrance of the pipe using the driving pressure rise shown in (a). In the simulations, the distance $\delta_d = 4.0$ mm and the width parameter $b = 0.5h$.

amplitude at about the same time as in the measurements. The model also reproduces qualitatively the amplitude and evolution of the transient burst on the second mode of the pipe observed before stabilization of the oscillations on the fundamental.

B. Steady state

A typical simulation result is presented in Fig. 17(a). As expected, the playing amplitude is of the same order of magnitude as that of the driving pressure signal. In this simulation, separation of the acoustic flow at the edge of the labium is taken into account by the model given by Eq. (19).

It is interesting to evaluate the importance of the choice of the distance δ_d and of the nonlinearity associated with the flow separation at the labium since they are the main features that distinguish the model from the widely accepted model by Fletcher.³⁴ In Fletcher's model, the flow separation at the labium is ignored. Furthermore, only the flow source Q_1 at the labium is considered. This is equivalent to assuming that the acoustic oscillations in the pipe are driven by a monopole rather than by a dipole. In the one-dimensional representation of Fig. 3 this is equivalent to using a distance δ_d between the flow sources Q_1 and Q_2 equal to the total end correction δ_m . This brings the source Q_2 "outside" the in-

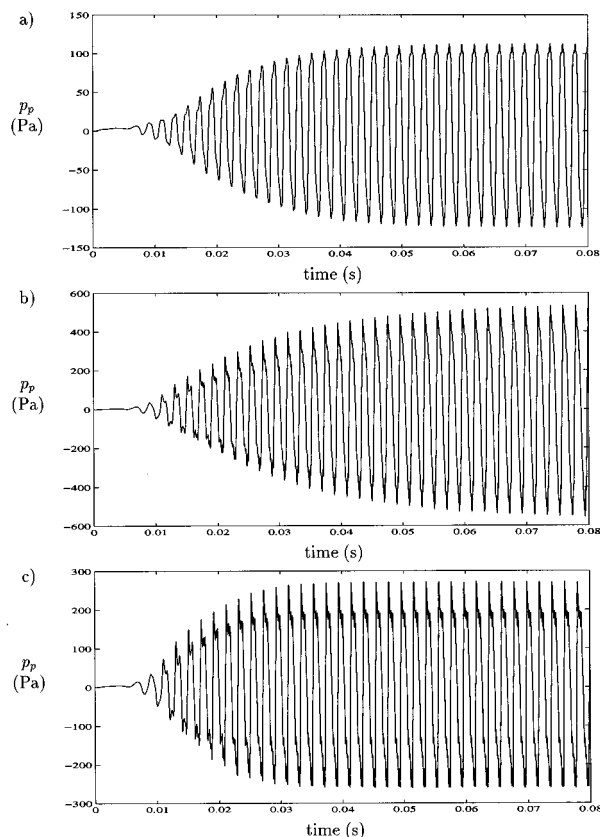


FIG. 17. Simulated signal obtained, for the first mode of the pipe, by using a steady-state driving pressure $p_f = 100 \text{ Pa}$. (a) response obtained with the model given by Eq. (23); (b) response obtained when neglecting the vortex shedding at the labium and using a distance $\delta_d = \delta_m$; and (c) response obtained when taking into account vortex shedding at the labium and using a distance $\delta_d = \delta_m$.

strument and prevents it from performing acoustic work. As can be observed from Eq. (7), this increases the efficiency of the jet-drive mechanisms. Simulation results obtained with a distance $\delta_d = \delta_m$ and by neglecting flow separation at the labium ($\Delta p_a = 0$) are presented in Fig. 17(b) for oscillation of the system on the first mode of the pipe. By comparison with the results of Fig. 17(a), the playing amplitude has been increased by about a factor of 5. Taking into account the vortex shedding reduces, as is seen in Fig. 17(c), the amplitude by a factor of 2. The effects of these two parameters on the playing amplitude is spectacular. Note that in the simulations of Fig. 17(b) and (c), the jet velocity U_j was assumed to be independent of the value of the pressure p_m at the flue exit ($U_j = \sqrt{2p_f/\rho_0}$) in order to neglect losses associated with the jet velocity fluctuations, discussed by Verge *et al.*³ which are not taken into account by Fletcher. It is anyway only sensible to take this effect into account when the playing amplitude is correctly predicted. When the fluctuations of the pressure p_m are too large they may result in jet velocity variations allowing the jet to have a negative velocity.

The results presented in the companion to this paper showed that a fundamental characteristic of flue instruments is the dependency of the dimensionless amplitude of the fundamental of the oscillation on the Strouhal number only. In order to validate the simulation model, a similar dimension-

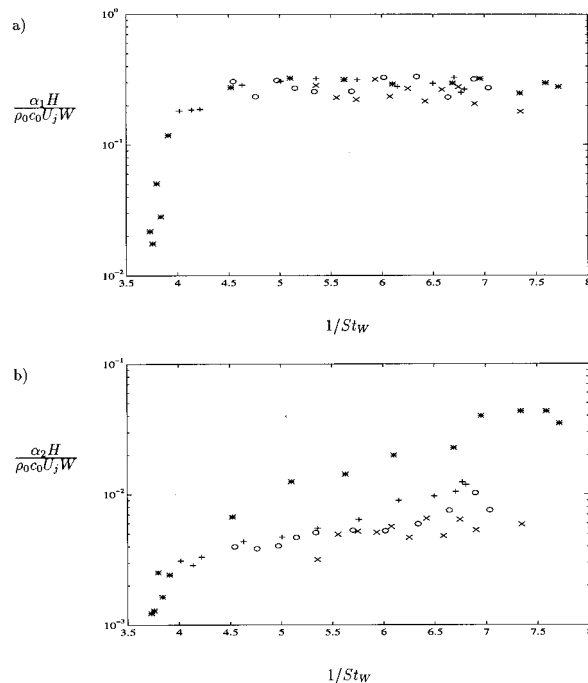


FIG. 18. (a) Dimensionless amplitude α_1 of the fundamental and (b) α_2 of the second harmonic as a function of the Strouhal number $St_W = f_1 W / U_j$ obtained with the one-dimensional simulation model of Fig. 1: *, first mode; +, second mode; O, third mode). In the simulations, flow separation at the labium was taken into account, the horizontal position of the labium was adjusted 0.1 mm below the middle of the flue exit, the ratio $W/h = 4$, the width parameter $b = 0.4h$, and the distance $\delta_d = 3.5 \text{ mm}$.

less analysis of the simulation results was performed for a wide range of driving pressures and is shown in Fig. 18 as a function of the Strouhal number $St_W = f_1 W / U_j$. In the simulations model, we used the same geometric parameters as that of the small organ pipe on which the measurements were performed. As was observed with experimental data, the curves associated with the different acoustic modes of the pipe collapse onto a single curve in a dimensionless representation. The general shape of this function as well as the value of its maximum are globally very similar to that corresponding to measurements. This good agreement between the simulations and the measurements over most of the operating range of the model gives us confidence in its ability to determine the amplitude of the fundamental during steady-state oscillations.

The oscillation threshold of the simulations is slightly underestimated compared to the measurements while the dimensionless amplitude of the oscillations is overestimated near this point. This can be explained by the fact that in the jet model used, the perturbations are assumed to grow exponentially with the distance traveled. In flow visualizations (see companion paper or Ref. 5) the jet appears to break down into vortices at very low blowing pressures. This nonlinear effect clearly limits the amplitude of the jet oscillations and is not taken into account by the model. Note that in the simulations, for the same reasons, one can observe oscillations of the jet on its second hydrodynamic mode even for short distances W . For clarity, these points were not plotted in Fig. 18.

The main difference between these results and the measurements presented in the companion to this paper is that the transition points between operation on the different modes of the pipe appear at blowing pressures that are too low. This could again be due to the linear amplification of the jet model. Another explanation may be the fact that the filters used to simulate the jet motion do not, as can be seen from Fig. 4(a), amplify enough at low Strouhal numbers. Experiments with the simulation program showed that shifting the position of the maximum of the frequency response of these filters strongly affected the position of the transition points. Slight variations in the length of the delay line, representing propagation on the jet, also appeared to strongly influence the position of the transition points. In the simulations it was adjusted in order to fit at best the data of Coltman presented in Fig. 4 and calculated by assuming the propagation velocity u to be proportional to the jet velocity. The optimal value of this proportionality constant used in the simulations depends on the phase shift introduced by the jet filters and is therefore a little arbitrary. It can therefore be argued that the value of this constant could be used for “fine tuning.” Finally, the choice of the value of the distance δ_d between the sources Q_1 and Q_2 of the jet-drive model also appears to affect the behavior of the model near transition points. A larger distance δ_d will favor oscillations at lower frequencies compared to a smaller one because of the effects of greater inertia in the mouth. The value used in the simulations was kept constant and calculated by assuming the flow sources to be located on each side of the labium at a distance h from the edge of the labium.² It is not impossible that these sources could “move” which suggests that the simulations could be improved by having a Strouhal number-dependent jet-drive model. A simpler solution is to also use this parameter for “fine tuning.”

The amplitude α_2 of the second harmonic appears to be of the order of the experimental data measured when the transverse position of the labium is set to $y_0=0.27$ (see companion paper). It is interesting to note that the measured drop of the amplitude α_2 between the second and third acoustical mode of the pipe does not appear in the simulations. This is due to the fact that the transition from a laminar to a turbulent jet was not taken into account in the jet oscillation model. This stresses the importance of the structure of the jet.

The effects of the flow separation model at the edge of the labium and of the dipolar nature of the jet drive can be estimated from the results presented in Fig. 19. In these simulations, vortex shedding at the edge of the labium were neglected and the flow source Q_2 was moved “outside” the instrument ($\delta_d=\delta_m$). The dimensionless amplitude α_1 of the fundamental is again greatly overestimated. This confirms the importance of the flow source Q_2 located on the outer side of the labium. Clearly, the steady-state amplitude of the fundamental appears to be determined by the combination of the effects of the dipolar nature of the jet drive and of flow separation at the edge of the labium. Furthermore, in these results, the amplitude α_1 is not a function of the Strouhal number only. This can be explained by the fact that the only loss mechanisms considered in these simulations (friction

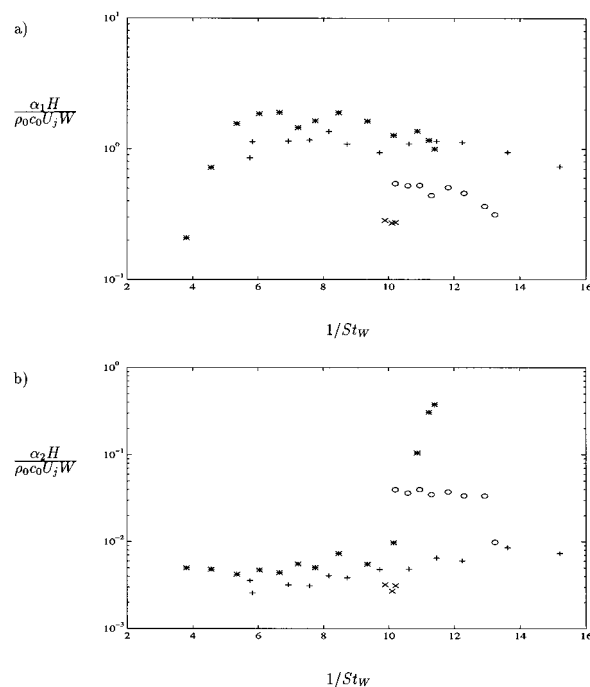


FIG. 19. (a) Dimensionless amplitude α_1 of the fundamental and (b) α_2 of the second harmonic as a function of the Strouhal number $St_W=f_1 W/U_j$ obtained with the one-dimensional simulation model of Fig. 1: *, first mode; +, second mode; \circ , third mode). In the simulations, flow separation at the labium was not taken into account, the jet velocity was assumed to be constant ($U_j = \sqrt{2p_f/\rho_0}$), the horizontal position of the labium was adjusted 0.1 mm below the middle of the flue exit, the ratio $W/h=4$, the width parameter $b=0.4h$, the distance $\delta_d=\delta_m=9.7$ mm.

and radiation) are frequency dependent and therefore depend on the operating mode. In these simulations the transition points between the different modes of the pipe appear at values of the Strouhal number close to the measurements. They appear at higher driving pressures than in the previous simulations, even though the same filters were used to simulate the jet motion, because of the higher value of the distance δ_d .

The absolute value of the dimensionless amplitude α_2 does not appear to be affected by the strength of the jet drive. In relation to the first harmonic (α_2/α_1), however, the second harmonic is now an order of magnitude lower than in

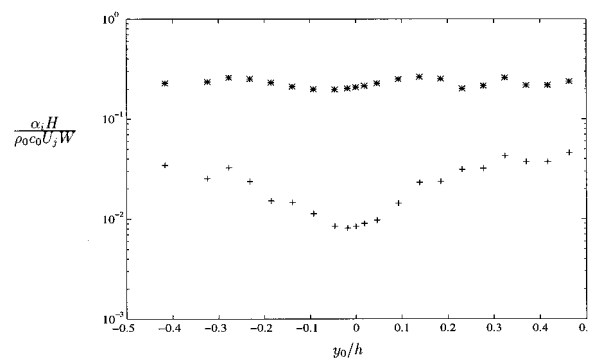


FIG. 20. Amplitude α_1 of the fundamental (*) and of the second harmonic (+) as a function of the transverse position y_0 of the labium for a Strouhal number $St_W=f_1 W/U_j=0.2$.

the experiments. This implies that the model of the vortex shedding at the labium provides a significant production of higher harmonics. The exact shape of the labium, which is not considered in the modeling, is of course the parameter which determines the details of the flow around the labium and is therefore crucial for the generation of high harmonics. It is, however, difficult to deduce the effects of the vortex shedding model on the generation of high harmonics from these measurements. Indeed, the very high amplitude predicted by this model also results in an overestimate of the jet displacement at the labium. This results in a sharpening of the waveform of the flow sources Q_1 and Q_2 which “artificially” generates high frequencies.

1. Effects of the transverse position of the labium

The linear jet oscillation model used in the simulations implies that a minimum of even harmonics should be maintained when the edge of the labium is aligned with the middle of the flue exit $y_0=0$ as can be observed in Fig. 20. This behavior was explained by Fletcher and Nolle^{35,36} and is due to the fact that in this configuration, the jet displacement at the labium is symmetric. In the measurements presented in the companion to this paper, however, the position of the “effective zero” of the transverse position y_0 of the labium, where the amplitude α_2 has a minimum, is located approximately a distance $0.2h$ below the middle of the flue exit. This shift in the effective transverse position of the labium, probably due to a nonlinear interaction between the jet flow and the labium, has important implications since this parameter determines the ratio α_2/α_1 of the amplitude of the first and second harmonic. It can also strongly affect the attack transient. This suggests that when trying to match measurements, the transverse position of the labium should be shifted by $0.2h$ but also that it is not impossible to match both the attack transient and steady-state sound without a physical model of the mean jet position.

VII. CONCLUSION

Despite its simplicity, the one-dimensional model described in this paper enables one to understand and simulate a surprisingly large amount of observations on the functioning of recorderlike instruments. The model reproduces quantitatively many of the results obtained from analytical calculations as well as the measurements performed on an experimental flue pipe. The model is also adapted for sound synthesis since it can be implemented for real-time simulations.

During the attack transient, the model is triggered by the initial volume injection into the mouth of the instrument and a simple model of the initial impact of the jet with the labium. The model is also sensitive to the steepness of the driving pressure rise which enables one to simulate the differences between a fast and a slow attack observed in the measurements. The response of the model is triggered by its reaction to the driving pressure signal. This may seem straightforward, but simulation models, such as that proposed by McIntyre *et al.*,³⁷ usually need to be triggered by an arbitrary additional impulse.

During steady-state operation, the dimensionless amplitude of the fundamental displays the expected Strouhal number dependency. The simulated response is very close to experimental results which implies that it grasps some basic features of recorderlike instruments. This Strouhal number dependency of the dimensionless amplitude is not predicted by the widely used model of Fletcher.³⁴ This appears to be the result of the combination of the use of the jet-drive mechanism described by Verge *et al.*² and of a simple quasi-stationary model of the separation of the acoustic flow at the edge of the labium. The distance δ_d between the flow sources Q_1 and Q_2 at the labium (in the one-dimensional representation of the instrument, see Fig. 3) determined by Verge *et al.*² appears to be of the right order of magnitude. This stresses the fact that the flow source Q_2 located on the outer side of the labium belongs to the “acoustic world” of the instrument and must be taken into account. Flow separation at the edge of the labium introduces a crucial nonlinear damping of the fundamental which is essential in order to obtain the particular Strouhal number dependency of its dimensionless amplitude (independent of the acoustic mode). Note that in the simulations, the “momentum drive” corresponding to the injection of momentum by the oscillation of the jet flow was not taken into account. The quasi-stationary model of this effect used in the literature is quite doubtful since it assumes instantaneous mixing of the jet with the pipe flow across a distance (pipe diameter) for which the Strouhal number is of the order of the unity or larger. Furthermore, the simulations confirm Coltman’s experimental results³⁸ which show that it is a secondary effect.

The jet oscillation model is one of the main weaknesses of the global model. In the simulations, a modification of the original model was used in order to obtain a simple filter representation and to remove its singularity at zero frequency. The behavior of the model is very sensitive to the choice of the width parameter b of the jet model as well as the shape of the filters chosen to implement the jet oscillation model. This parameter is critical during the attack transient since it determines the hydrodynamic amplification of the perturbations by the jet and hence the swiftness of the response of the model. During steady-state operation, this parameter b slightly affects the playing amplitude, since it determines the amplitude of the source Q_1 . The optimal value found in matching measurements performed on the experimental flue pipe described by Verge *et al.*² varied from $b=0.39h$ for slow transients to $b=0.35h$ for fast transients. These values are very close to those found in the linear analysis performed in the same paper on measured transients. This range of values is also very close to the theoretical value ($b=0.4h$) obtained by assuming a Poiseuille jet velocity profile at the flue exit. For a convergent canal, such as found in recorders, the optimal value is expected to be higher because the convergence flattens the velocity profile. Another weak point of the model is that it does not take into account the effects of the shape of the labium on the sound quality. Such an effect could, however, be taken into account in a separate “sound quality” module which would be driven by the model.

The simulation model is globally very satisfying in view

of the number of approximations that it involves. Nevertheless, the range of uncertainty of its different parameters still leaves more latitude in the model response than one would wish. The determination of the optimal value of these parameters is complicated by the fact that different combinations of these parameters can be used to obtain a given response. The range of variation of these parameters has been reduced enough to reproduce the main characteristics of recorderlike instruments. However, for a “fine tuning” of the model, adjustment of the parameters is still necessary. This might, in fact, be inevitable. Craftsmen also have to do the fine tuning of a recorder by trial and error. In the simulations it was preferred to keep all the parameters constant and to adjust the parameter b . In order to account for the subtleties of the work of the instrument maker which determines the quality of the instrument further research would be necessary on three main points. First the separation of the jet and the acoustic flow at the labium would have to be described in more detail. The sharpness of the labium indeed appears experimentally to be crucial for the generation of high harmonics.^{39,36} The effect of the convergence of the flue channel, which determines the velocity profile at the flue exit and its influence on the jet motion, appears to be important as can be concluded from the great sensitivity of this simple model to the parameter b . Finally, the interaction of the transverse acoustic field from the pipe with the jet flow at the edges of the flue exit, the receptivity problem, would have to be reformulated in order to understand and model the effects of chamfers (or the lips of a flute player) on the behavior of the jet.

¹P. Skordos, “Modeling Flue Pipes: Subsonic Flow Lattice Boltzmann, and Parallel Distributed Computers,” Ph.D. thesis, Department of Electrical Engineering, MIT, Cambridge, MA (January, 1995).

²M. P. Verge, R. Caussé, B. Fabre, A. Hirschberg, A. P. J. Wijnands, and A. van Steenberghe, “Jet oscillations and jet drive in recorder-like instruments,” *Acta Acustica* **2**, 403–419 (1994).

³M. P. Verge, B. Fabre, W. E. A. Mahu, A. Hirschberg, R. van Hassel, A. P. J. Wijnands, J. J. de Vries, and C. J. Hogendoorn, “Jet formation and jet velocity fluctuations in a flue organ pipe,” *J. Acoust. Soc. Am.* **95**, 1119–1132 (1994).

⁴B. Fabre, “La Production de Son dans les Instruments à Embouchure de Flûte: Modèle Aéro-Acoustique pour la Simulation Temporelle,” Ph.D. thesis, Université du Maine, Le Mans, France (1992).

⁵M. P. Verge, “Aeroacoustics of Confined jets, with Applications to the Physical Modeling of Recorder-Like Instruments,” Ph.D. thesis, Eindhoven University of Technology, The Netherlands (1995).

⁶A. Powell, “On edge tones and associated phenomena,” *Acustica* **3**, 233–243 (1953).

⁷S. A. Elder, “On the mechanism of sound production in organ pipes,” *J. Acoust. Soc. Am.* **54**, 1554–1564 (1973).

⁸J. W. Coltman, “Jet drive mechanism in edge tones and organ pipes,” *J. Acoust. Soc. Am.* **60**, 725–733 (1976).

⁹A. Powell, “On the edgetone,” *J. Acoust. Soc. Am.* **33**, 395–409 (1961).

¹⁰A. R. Paterson, *A First Course in Fluid Dynamics* (Cambridge U.P., Cambridge, UK, 1983).

¹¹Lord Rayleigh, *The Theory of Sound* (Dover, New York, 1894).

¹²L. Cremer and H. Ising, “Die selbsterregten Schwingungen von Orgelpfeifen,” *Acustica* **19**, 143–153 (1967–1968).

¹³N. H. Fletcher and S. Thwaites, “Wave propagation on a perturbed jet,” *Acustica* **42**, 323–334 (1979).

¹⁴G. E. Mattingly and W. O. Criminale, “Disturbance characteristics in a plane jet,” *Phys. Fluids* **14**, 2258–2264 (1971).

¹⁵H. Sato, “The stability and transition of a two-dimensional jet,” *J. Fluid Mech.* **7**, 53–80 (1960).

¹⁶D. Bechert, “Die Steuerung eines ebenen turbulenten Freistrahls durch eine seitliche Wechselströmung, erzeugt in einem Schallfeld,” *Z. Flugwiss.* **24**, 25–33 (1976).

¹⁷W. K. Blake and A. Powell, “The development of contemporary views of flow-tone generation,” in *Recent Advances in Aeroacoustics* (Springer-Verlag, New York, 1983).

¹⁸A. Maurel, “Instabilité d’un Jet Confiné,” Ph.D. thesis, Université de Paris VI, Paris, France (July 1994).

¹⁹D. K. Holger, T. A. Wilson, and G. S. Beavers, “The amplitude of edgetone sound,” *J. Acoust. Soc. Am.* **67**, 1507–1511 (1980).

²⁰J. O. Smith, “Physical modeling using digital waveguides,” *Comput. Music J.* **16**, 75–87 (1992).

²¹H. Levine and J. Schwinger, “On the radiation of sound from an unflanged circular pipe,” *Phys. Rev.* **73**, 383 (1948).

²²J. D. Polack, X. Menial, J. Kergomard, C. Cosnard, and M. Bruneau, “Reflection function of a plane wave in a cylindrical tube,” *Rev. Phys. Appl.* **22**, 331–337 (1987).

²³D. Matignon, “Représentation en Variables d’Etat de Modèles de Guide d’Ondes avec Dérivation Fractionnaire,” Ph.D. thesis, Université de Paris-Sud, Centre d’Orsay, France (1994).

²⁴J. W. Coltman, “Sounding mechanism of the flute and organ pipe,” *J. Acoust. Soc. Am.* **44**, 983–992 (1968).

²⁵B. Fabre, A. Hirschberg, and A. P. J. Wijnands, “Vortex shedding in steady oscillations of a flue organ pipe,” *Acustica* **82**, 863–877 (1996).

²⁶U. Ingard and H. Ising, “Acoustic nonlinearity of an orifice,” *J. Acoust. Soc. Am.* **42**, 6–17 (1967).

²⁷M. J. Lighthill, “On sound generated aerodynamically 1: General theory,” *Proc. R. Soc. London, Ser. A* **211**, 564–587 (1952).

²⁸M. J. Lighthill, “On sound generated aerodynamically 2: Turbulence as a source of sound,” *Proc. R. Soc. London, Ser. A* **222**, 1–32 (1954).

²⁹J. E. Ffowcs Williams, “Hydrodynamic noise,” *Annu. Rev. Fluid Mech.* **1**, 197–222 (1969).

³⁰L. Bjørnø, and P. N. Larsen, “Noise of air jets from rectangular slits,” *Acustica* **54**, 247–256 (1984).

³¹M. Goldstein, *Aeroacoustics* (McGraw-Hill, New York, 1976).

³²D. J. Hermes, “Synthesis of breathy vowels: Some research methods,” *Speech Commun.* **10**, 497–502 (1991).

³³C. Chafe, “Adding vortex noise to wind instrument physical models,” in *International Computer Music Conference* (Banff, Canada, 1995), pp. 57–60.

³⁴N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments* (Springer-Verlag, New York, 1991).

³⁵N. H. Fletcher and L. M. Douglas, “Harmonic generation in organ pipes, recorders, and flutes,” *J. Acoust. Soc. Am.* **68**, 767–771 (1980).

³⁶A. W. Nolle, “Flue organ pipes: Adjustments affecting steady waveform,” *J. Acoust. Soc. Am.* **73**, 1821–1832 (1983).

³⁷M. E. McIntyre, R. T. Schumacher, and J. Woodhouse, “On the oscillations of musical instruments,” *J. Acoust. Soc. Am.* **74**, 1325–1345 (1983).

³⁸J. W. Coltman, “Momentum transfer in jet excitation of flutelike instruments,” *J. Acoust. Soc. Am.* **69**, 1164–1168 (1981).

³⁹M. Castellengo, “Contribution à l’étude expérimentale des tuyaux à bouche,” Ph.D. thesis, Université de Paris VI, Paris, France (1976).

⁴⁰J. W. Coltman, (private communication), 1994.

Inertial cavitation and associated acoustic emission produced during electrohydraulic shock wave lithotripsy

Pei Zhong

Department of Mechanical Engineering and Materials Science and Division of Urology/Department of Surgery, Duke University, Durham, North Carolina 27708

Iulian Cioanta and Franklin H. Cocks

Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27708

Glenn M. Preminger

Division of Urology/Department of Surgery, Duke University, Durham, North Carolina 27708

(Received 14 August 1996; accepted for publication 26 December 1996)

The inertial cavitation and associated acoustic emission generated during electrohydraulic shock wave lithotripsy were studied using high-speed photography and acoustic pressure measurements. The dynamics of cavitation bubble clusters, induced in vitro by an experimental laboratory lithotripter, were recorded using a high-speed rotating drum camera at 20 000 frames/s. The acoustic emission, generated by the rapid initial expansion and subsequent violent collapse of the cavitation bubbles, was measured simultaneously using a 1-MHz focused hydrophone. The expansion duration of the cavitation bubble cluster was found to correlate closely with the time delay between the first two groups of pressure spikes in the acoustic emission signal. This correlation provides an essential physical basis to assess the inertial cavitation produced by a clinical Dornier HM-3 shock wave lithotripter, both in water and in renal parenchyma of a swine model. In the clinical output voltage range (16–24 kV), the expansion duration of the primary cavitation bubble cluster generated by the HM-3 lithotripter in water increases from 158 to 254 μs , whereas the corresponding values in renal parenchyma are much smaller and remain almost unchanged (from 71 to 72 μs). In contrast, subsequent oscillation of the bubble following its primary collapse is significantly prolonged (from 158–235 μs in water to 1364–1373 μs in renal parenchyma). These distinctive differences between lithotripsy-induced inertial cavitation in vitro and that in vivo are presumably due to the constraining effect of renal tissue on bubble expansion. © 1997 Acoustical Society of America. [S0001-4966(97)02305-9]

PACS numbers: 43.80.Ev, 43.80.Gx, 43.25.Yw, 43.80.Sh [FD]

INTRODUCTION

Evidence from recent studies has suggested that inertial cavitation, induced by the tensile stress of a lithotripter shock wave pulse, plays an important role in the process of stone comminution and tissue injury during shock wave lithotripsy (SWL).^{1–9} Inertial cavitation is referred generally to the oscillation of cavitation bubbles in response to microsecond acoustic pressure pulses of short duty cycles, in which the inertial effect dictates the dynamics of bubble oscillation. Cavitation and associated microjet impingements, produced by the asymmetric collapse of cavitation bubbles near a stone surface, have been identified as the primary mechanism for stone fragmentation.^{1–5} The mechanical stresses generated by shock wave–cavitation bubble interaction and subsequent microjet impingement, as well as cavitation-induced formation of free radicals have been attributed to various tissue and cellular damage during SWL.^{6–9}

Despite these findings, current commercial lithotripters are not equipped with any means to assess quantitatively the transient cavitation activity in patients during clinical lithotripsy. The treatment procedure is largely empirical and 2000–4000 high-intensity shock pulses are often used during one treatment. Following clinical lithotripsy, acute renal in-

jury including gross hematuria and renal edema is frequently observed, although the long-term effects of SWL on renal tissue and function are still yet to be completely determined.¹⁰ Clearly, a technique that can assess quantitatively the lithotripsy-induced cavitation in vivo would be clinically valuable to ensure the effectiveness and safety of SWL treatment. Such a technique may be used to identify the correlation between lithotripsy-induced cavitation and resultant stone comminution and tissue injury, and thus providing treatment guidelines to maximize stone fragmentation while minimizing potential renal tissue injury and functional alterations.

Using a focused hydrophone, Coleman and associates have measured the acoustic emission in water emanating from the beam focus of an electrohydraulic shock wave lithotripter.¹¹ Qualitative agreements were obtained between the measurement results and bubble dynamics predicted theoretically based on the Gilmore–Akulichev model.¹¹ In addition, measurements of sonoluminescence were made separately to provide indirect evidence for the presence of cavitation. Although a general correlation between the spatial and temporal distribution of acoustic emission and sonoluminescence signals,¹¹ as well as cell lysis in vitro¹²

has been observed, a clear interpretation of the results in association with the corresponding bubble activities is still difficult. Others have studied lithotripsy-induced cavitation¹³ and shock wave-bubble interaction¹⁴ using high-speed photography, yet in those studies the correlation between the dynamics of cavitation bubbles and associated acoustic emission was not determined.

In addition, several animal and clinical studies have provided evidence of lithotripsy-induced inertial cavitation in vivo, identified as transient echoes in ultrasound *B*-scan imaging.^{9,15–18} Histology examinations following SWL treatment have revealed a close correlation between regions of hyperechogenicity and tissue injury, and based on such a correlation, the appearance of the first echo has been used to define the threshold for lithotripsy-induced tissue injury.^{9,16–18} However, in order to accurately detect the onset of inertial cavitation, the acquisition of *B*-scan imaging must be synchronized with the cavitation event, a condition that is not provided by current commercial lithotripters. Moreover, *B*-scan imaging cannot be used to determine the intensity of inertial cavitation bubble expansion and collapse, which may correlate directly with tissue injury.

The present study aims to explore the feasibility of assessing inertial cavitation in vivo during SWL based on acoustic emission measurements. First, simultaneous high-speed photography and acoustic emission measurements were used to characterize the inertial cavitation and associated acoustic emission in vitro induced by lithotripsy shock waves. This approach was used to clearly identify the correlation between the dynamics of lithotripsy-induced cavitation bubbles and resultant acoustic emission, and to provide a physical basis for assessing cavitation activity using acoustic emission measurements alone. Second, acoustic emission due to inertial cavitation produced around the beam focus of a commercial shock wave lithotripter (Dornier HM-3) was measured both in water (in vitro) and from renal parenchyma of a swine model (in vivo). The results have shown that inertial cavitation in vivo during SWL can be detected using acoustic emission measurements. Distinctive differences between lithotripsy-induced inertial cavitation in vitro and that in vivo presumably due to the constraining effect of tissue on bubble expansion have also been identified.

I. MATERIALS AND METHODS

A. Shock wave lithotripters

In this study, an experimental laboratory lithotripter and a clinical electrohydraulic shock wave lithotripter (HM-3, Dornier Medical Systems) were used. The experimental lithotripter, described in detail previously,¹⁹ is similar in design principle to a Dornier HM-3 unit, except that a 120-nF capacitor and a semiellipsoidal aluminum reflector (semimajor axis $a=93$ mm, semiminor axis $b=80$ mm and half-focal length $c=47$ mm) is used. In comparison, the Dornier HM-3 lithotripter uses a 40-nF capacitor and a truncated ellipsoidal brass reflector with a larger aperture and a longer focal length ($a=143$ mm, $b=86$ mm, and $c=114$ mm). In addition, the shock wave generator of the experimental lithotripter is housed in a transparent water tank (100×50

×50 cm) filled with degassed water (25 °C), and thus simultaneous high-speed photography of the dynamics of inertial cavitation bubbles and measurements of concomitant acoustic emission can be carried out. The Dornier HM-3 lithotripter, with its shock wave generator mounted on the base of a 1000 liter stainless-steel water tub, is used to characterize the acoustic emission produced around its beam focus both in vitro and in vivo under clinically relevant lithotripsy conditions.

The typical output voltage for both lithotripters can be varied between 16 and 24 kV. For the Dornier HM-3 lithotripter, the pressure waveform at the beam focus can be characterized by a compressive component with a leading shock front of less than 30 ns in rise time and a phase duration between 1 and 2 μ s, followed by a tensile component with a longer phase duration of 3 to 6 μ s.²⁰ The peak compressive and peak tensile pressure of the shock wave have been measured to be between 33–50, and 7.1–9.5 MPa, respectively, with the lithotripter output setting varying from 15 to 25 kV.²⁰

B. High-speed photograph

The experimental setup for high-speed photography and acoustic emission measurements performed on the experimental lithotripter is shown in Fig. 1. A high-speed rotating drum camera (Model 350, Cordin), equipped with a FUJINON 17.5–105-mm TV zoom lens (FUJI Photo Inc.) and a 2×teleconverter (Cosmicar, ASAHI Precision), was used to capture the dynamics of inertial cavitation bubbles induced during SWL. This high-speed camera can record a total of 224 frames in one sequence, with a maximum framing rate of 35 000 frames per second (FPS). The camera operation was controlled, and the framing rate of each high-speed recording was set preoperatively, using a camera controller (model 447A, Cordin). Synchronization between the firing of the experimental lithotripter and triggering of the high-speed camera was realized with the aid of a digital delay generator (DS535, Stanford Research Systems). Because of the extremely short exposure time (e.g., 3.3 μ s at 20 000 FPS) on each frame, highly sensitive black and white film (Kodak T-Max 3200) was used. Homogeneous backlighting illumination for the high-speed camera was provided by using a 1000-W halogen lamp together with a 4-mm-thick ground glass plate. The scale of the high-speed images was calibrated in a test run with reference to an object of known dimension under the same magnification.

C. Acoustic emission measurements

Acoustic emission, emanating from the beam focus of the lithotripters, was measured using a broadband focused hydrophone (V392-SU, Panametrics), having a 38-mm diameter sensing element with a 101.6-mm focal length and 1-MHz resonant frequency (−6-dB frequency bandwidth: 0.54–1.38 MHz). A similar focused hydrophone of 1-MHz resonant frequency has been used previously by Coleman and associates to detect the acoustic emission signals generated by the oscillation of cavitation bubbles during electrohydraulic SWL.¹¹ The advantage of using a focused hydro-

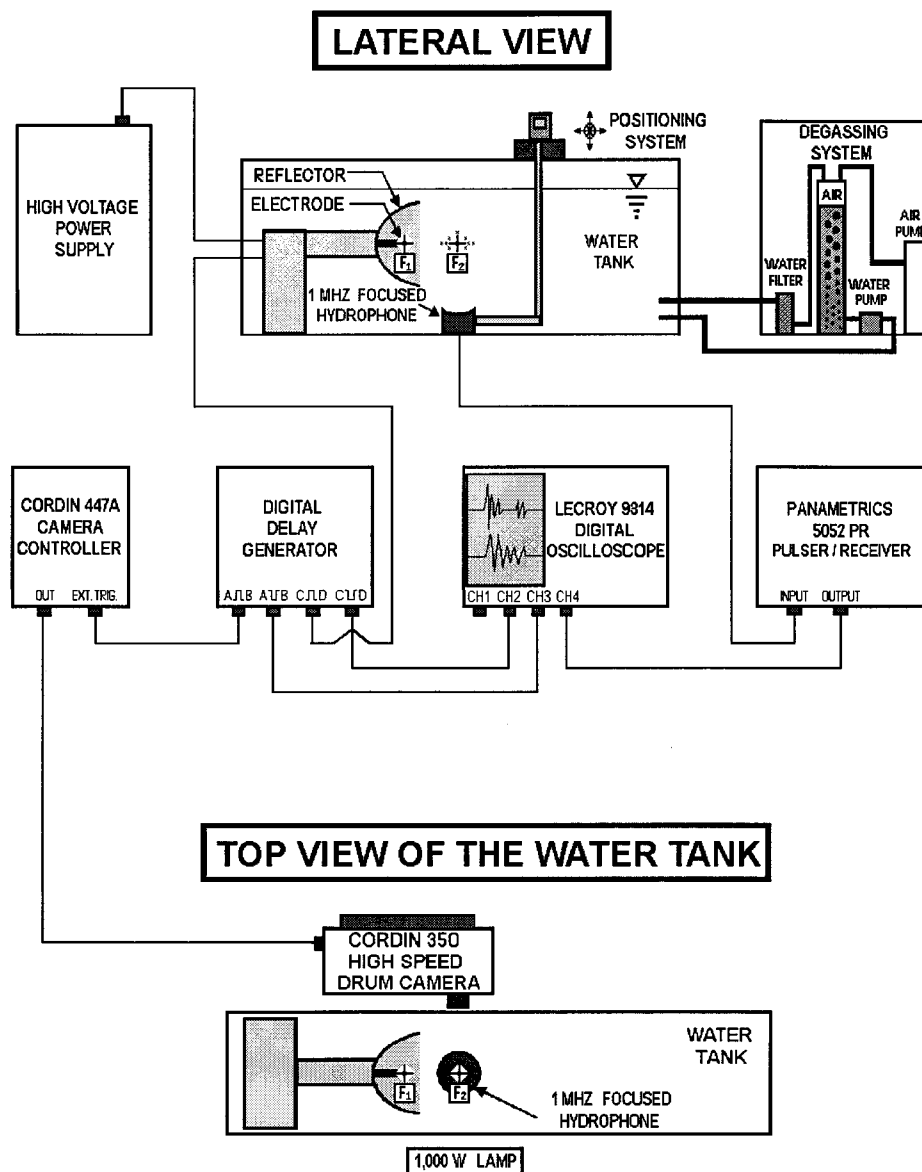


FIG. 1. A schematic diagram of the experimental setup for high-speed photography and acoustic emission measurements in vitro using an experimental laboratory lithotripter.

phone (a passive detector) for assessing lithotripsy-induced inertial cavitation, compared to using ultrasound B-scan imaging or an active detector, is that the passive detector can monitor continuously the whole cavitation event, and thus is better suited for characterizing the dynamics of overall bubble activities. Being focused, the hydrophone also provides adequate sensitivity and spatial resolution (i.e., the focused hydrophone used by Coleman and associates has a focusing gain of the order of 1000, and a cigar shape reception zone of 5 mm long and 3 mm in diameter).¹¹

To measure the acoustic emission generated by the experimental lithotripter, the focused hydrophone was connected to an x - y - z translational stage (0.01-mm precision) and placed near the bottom of the water tank (Fig. 1). The symmetric axis of the focused hydrophone was aligned perpendicular to that of the shock wave reflector and confocally with the focus of the lithotripter (F_2). The output signal of

the hydrophone was fed into a high-pass filter with 0.3-MHz cut-off frequency and a broadband amplifier (5052 PR Pulser/Receiver, Panametrics), and then recorded on a digital oscilloscope (LeCroy 9314) with a maximum sampling rate of 100×10^6 samples/s. The high-pass filter was used to eliminate signals due to the lateral vibration of the hydrophone sensing element in the acoustic field of the lithotripter.¹⁰ To ensure that the acoustic emission is measured from the reception zone of the hydrophone, an acoustic mask, made of foam rubber and having a conical-shape channel at its center, was constructed and attached to the front surface of the hydrophone.

For acoustic emission measurements carried out on the Dornier HM-3 lithotripter, the focused hydrophone, connected to an x - y - z translational stage, was placed on the lateral side of the shock wave reflector, and immersed in a water tub filled with degassed and deionized water (37 °C),

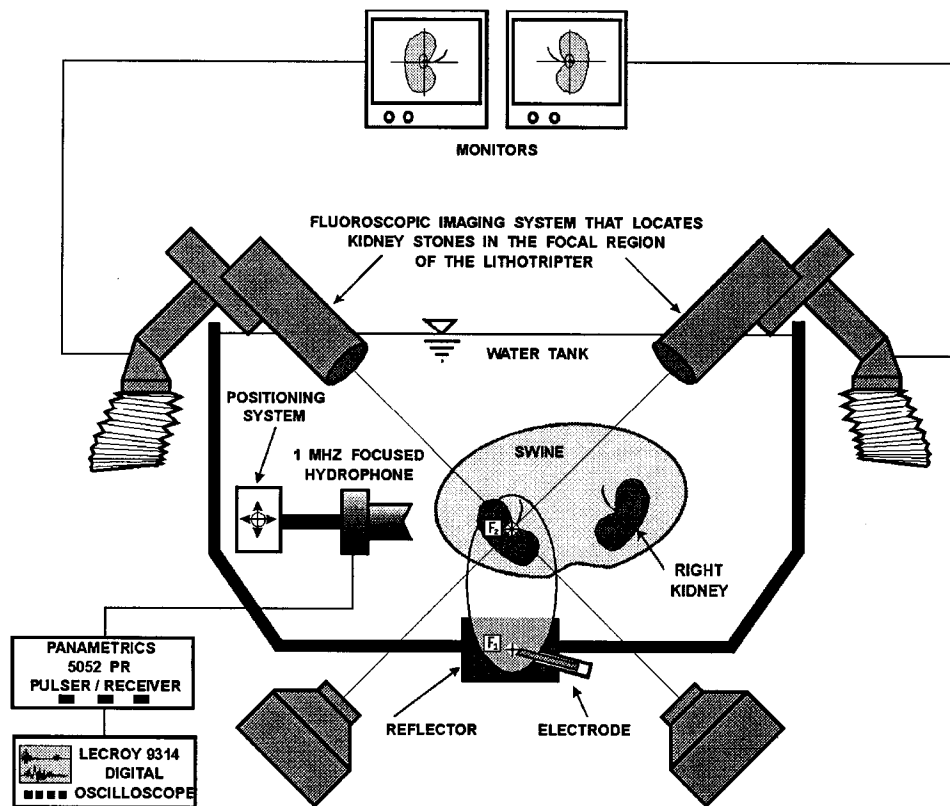


FIG. 2. A schematic diagram of the experimental arrangement for acoustic emission measurements in vivo using a Dornier HM-3 lithotripter.

as shown in Fig. 2. The hydrophone was aligned confocally with the focus (F_2) of the Dornier lithotripter, with the aid of a mechanical pointer that highlights the position of F_2 on two orthogonally aligned fluoroscopic imaging cameras. This alignment was carried out by scanning the focused hydrophone, operated in a transmitter/receiver mode via the Panametric Pulser/Receiver, around the tip of the mechanical pointer until the maximum echo signal was detected at the focal length of the hydrophone. Before each experiment, a new electrode was installed and 100 shocks were delivered at 20 kV, for the purpose of preconditioning the electrode. In the first series of experiments, acoustic emission emanating from the beam focus of the lithotripter was measured in water. A total of 15 measurements were made at 16, 18, 20, and 24 kV, respectively, and the signals were saved on a floppy disk for later analysis.

In the second series of experiments, two 20-month juvenile swine, weighing approximately 18 kg, were anesthetized with 5% Pentothal (10 mg/kg), and placed on a motorized stretcher and lowered into the water tub of the HM-3 lithotripter. The focused hydrophone was again aligned confocally with F_2 using the mechanical pointer, before the anesthetized animal was placed into the water tub. Under bi-plane fluoroscopic guidance, the lower pole of the left kidney of the animal was then aligned with F_2 . Shock waves generated at 16, 18, 20, and 24 kV were delivered in a randomized sequence, and a total of 15 measurements were taken at each output voltage setting.

D. Calibration and data analysis

Calibration of the measurement system was carried out using a spherically divergent shock wave source produced by an AUTOLITH IEHLTM (Northgate Technologies Inc.) electrohydraulic lithotripter.²¹ Acoustic emission from the shock source was measured simultaneously using both the focused hydrophone and a calibrated needle PVDF transducer (MHA9-6, FORCE INSTITUTE) with a flat frequency response up to 20 MHz. By comparing the amplitudes of 1-MHz components in the frequency spectrum of the acoustic emission signals measured by both transducers, the effective sensitivity of the focused hydrophone system (including high-pass filtered and amplifier) was determined to be 6.46 $\mu\text{V}/\text{Pa}$, with a -6-dB reception zone of 20-mm long (in the beam axis) and 9-mm wide (in the focal plane).

After the experiments, acoustic emission signals were retrieved back to the LeCroy oscilloscope, and the peak value, duration, and time delay of each burst were measured. Using the mathematical analysis facility of the LeCroy oscilloscope, the effective acoustic emission energy, E_{AE} in the frequency domain of the hydrophone was calculated using the following equation:

$$E_{\text{AE}} = \frac{4\pi r^2}{\rho c} \int_0^T P_{\text{AE}}^2 dt, \quad (1)$$

where r is the distance from the source to the transducer (101.6 mm), ρ (1000 kg/m³), and c (1498 m/s) are the density and sound speed in water, respectively, P_{AE} is the pres-

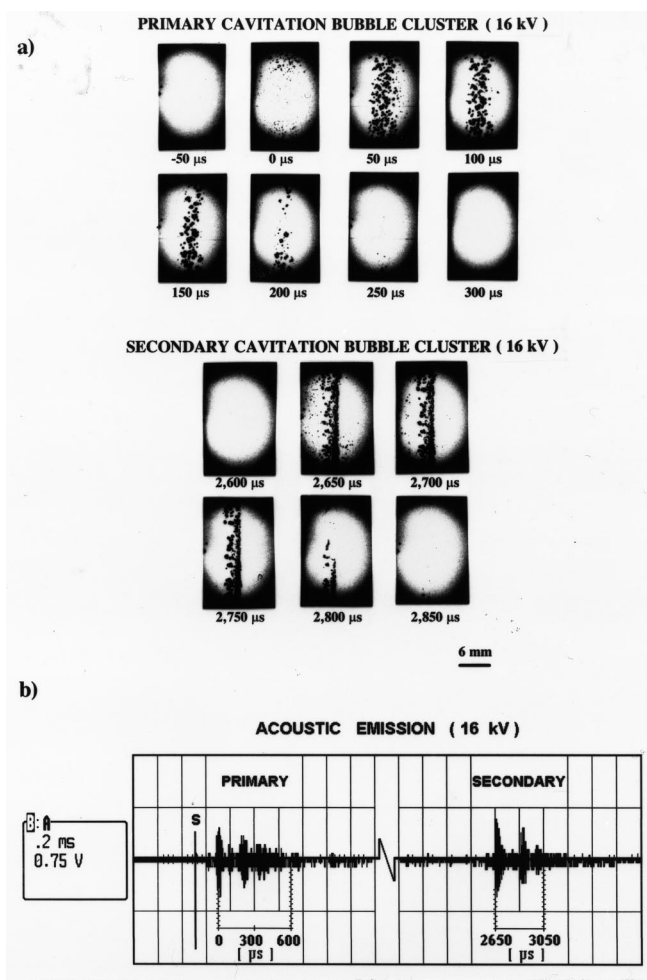


FIG. 3. A representative high-speed sequence of the inertial cavitation bubble clusters (a) and associated acoustic emission bursts (b), produced by an experimental laboratory lithotripter at 16 kV.

sure, and T is the duration of the acoustic emission burst.

II. RESULTS

A. Experimental lithotripter

1. At output voltage of 16 kV

Figure 3 shows a representative high-speed photographic sequence of the inertial cavitation bubbles, recorded at a framing rate of 20 000 FPS, and concomitant acoustic emission signal measured simultaneously after a single discharge of the experimental lithotripter at 16 kV. In this and subsequent high-speed photographic sequences, the incident shock wave was propagating upwards with its focal point coinciding with the center of each frame. For convenience, time zero in each sequence was selected to be the frame when cavitation bubbles were first observed, and the primary incident shock wave arrived at the beam focus shortly before $t=0$ μ s.

Following the discharge, two distinctive clusters of cavitation bubbles, separated in time by approximately 2.65 ms, were observed [Fig. 3(a)]. These two cavitation bubble clusters were induced, respectively, by a primary and a secondary shock wave pulse, generated by the initial plasma expansion and the rapid collapse/rebound of an ensuing cavitation

bubble originated from the spark discharge at F_1 .¹¹ In the primary cluster [$t=0-300$ μ s, Fig. 3(a)], cavitation bubbles of different sizes, distributed in a cylindrical volume around the beam axis of the lithotripter, were observed. These bubbles were found to decrease in size with radial distance from the beam axis and F_2 , corresponding presumably to the pressure gradient around the focal point of the lithotripter.²⁰ This finding suggests that the primary shock wave has encountered a cavitation nucleus population of nonuniform sizes. While small cavitation bubbles formed in the periphery of the cylindrical volume appeared to last for a short period (<100 μ s), large cavitation bubbles, formed within a 5-mm diameter around the beam axis, were seen to first expand rapidly ($t=0-50$ μ s), followed by a long period ($t=50-150$ μ s) in which they stabilized near their maximum sizes (ranging from 0.5 to 1.1 mm), before collapsing quickly within 50 μ s ($t=150-200$ μ s). After this primary collapse, a few bubbles near the beam axis were seen to rebound ($t=250$ μ s), but with much reduced size and shortened oscillation duration.

Bubbles in the secondary cavitation cluster [$t=2650-2950$ μ s, Fig. 3(a)], observed in 2.65 ms after the primary one, were formed within a 3-mm diameter cylindrical volume around the beam axis. Although a similar pattern of bubble oscillation was observed, including a rapid initial expansion to maximum size, followed by a prolonged period of stabilization (greater than 100 μ s) before the final collapse, two distinctive differences have been identified. First, cavitation bubbles in the secondary cluster were more uniform in size; and secondly, they appeared to coagulate together around the beam axis. The uniform bubble size indicates that the secondary shock wave pulse has encountered a narrowly distributed bubble population, as first suggested by Church's theoretical work²² and later confirmed experimentally by Coleman *et al.*, based on interpretation of acoustic emission signals associated with cavitation bubbles induced by an electrohydraulic shock wave lithotripter.¹¹ The observation of bubble coagulation also suggests that these cavitation bubbles were oscillating in phase, and thus attracting each other by mutual Bjerknes forces,²³ evidence which further confirms that the secondary shock wave has encountered an uniform bubble population.¹¹

Correspondingly, the acoustic emission (AE) measured simultaneously revealed two distinctive pressure bursts [Fig. 3(b)]. The primary burst occurred about 192 μ s after the interference signal from the spark discharge ["s" in Fig. 3(b)], corresponding to a time delay (124 μ s) for the shock wave to propagate from the discharge source (F_1) to the beam focus (F_2), plus the time (67.7 μ s) for the AE signal to reach the focused hydrophone. The secondary AE burst was measured in 2618 μ s following the primary burst, matching fairly well with the time delay of 2650 μ s between the formation of the primary and the secondary bubble clusters [Fig. 3(a)]. In addition, both the primary and the secondary AE signals were found to comprise of an initial and a delayed burst of pressure spikes, corresponding to the initial expansion of shock wave-induced cavitation bubbles and their primary collapses. The duration of these AE signals, measured by the time delay between peak initial and peak

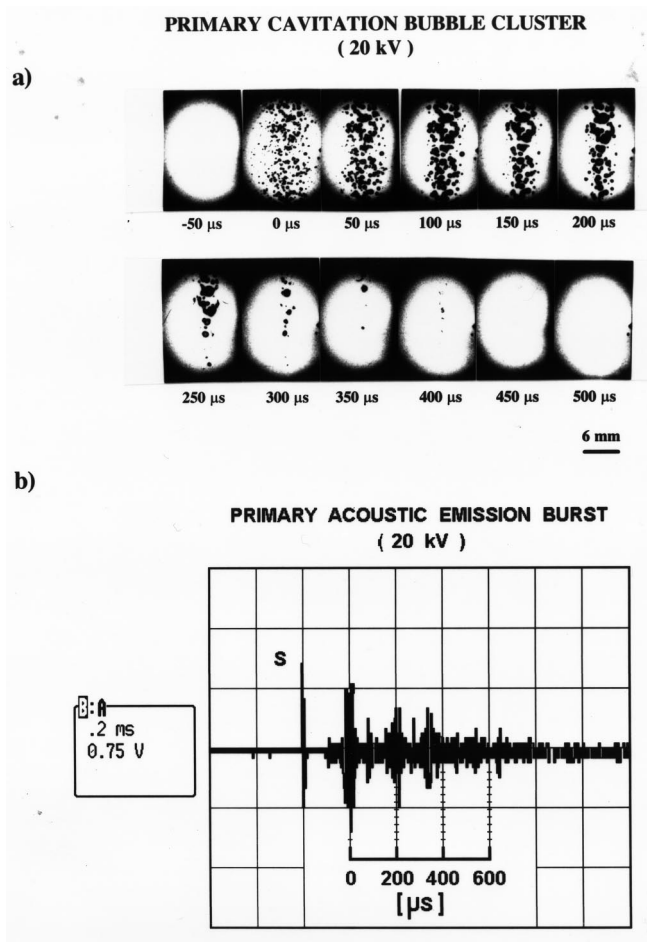


FIG. 4. Typical high-speed sequence of the primary cavitation bubble cluster (a) and the associated acoustic emission burst (b), produced by an experimental laboratory lithotripter at 20 kV.

delayed pressure spikes, were found to be 218 and 198 μs for the primary and the secondary bursts, respectively. These values correlate well with the duration of the two cavitation bubble clusters (200 and 200 μs), determined photographically.

In the primary AE signal, a small intermediate burst of pressure spikes was observed between the initial and the delayed bursts, which was found to correlate with the collapse of small bubbles on the periphery of the cavitation cluster formed around the beam axis [$t=0-100 \mu\text{s}$, Fig. 3(a)]. Following the delayed burst, additional pressure spikes of reduced amplitudes were observed in both the primary and the secondary AE signals, which were likely to be generated by the rebounds of the cavitation bubbles after their primary collapse.^{11,22} However, limited by the temporal and spatial resolution of the high-speed camera system, cavitation bubbles with an oscillating duration of less than 50 μs cannot be resolved.

2. At output voltage of 20 kV

Typical inertial cavitation bubbles and concomitant acoustic emission induced, respectively, by the primary and the secondary shock waves after a single discharge of the experimental lithotripter at 20 kV, are shown in Figs. 4 and

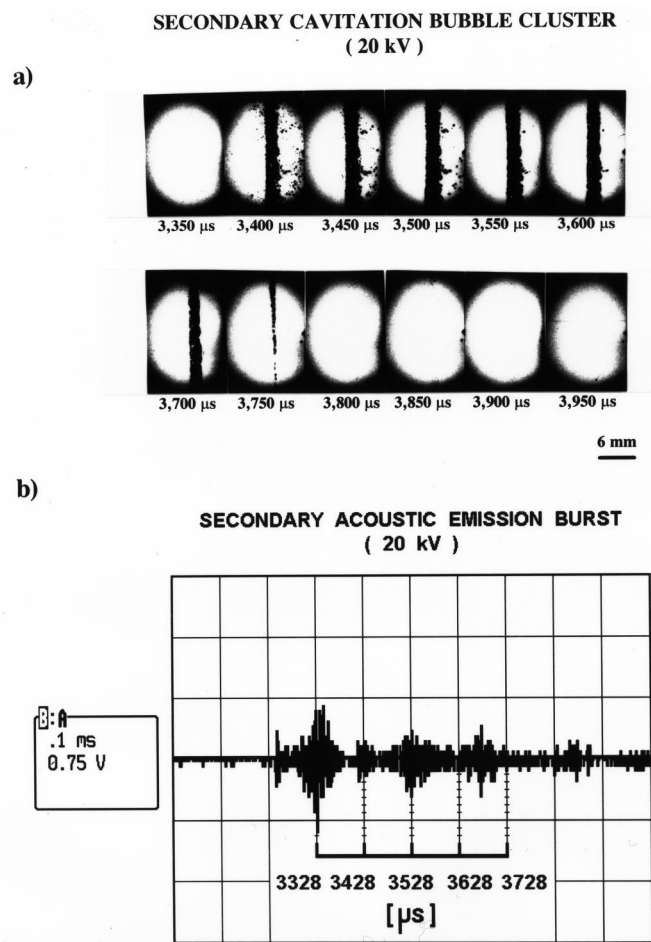


FIG. 5. Typical high-speed sequence of the secondary cavitation bubble cluster (a) and the associated acoustic emission bursts (b), produced by an experimental laboratory lithotripter at 20 kV.

5. Following the incidence of the primary shock wave, cavitation bubbles of different sizes were observed in a cylindrical volume of 8-mm diameter surrounding the beam axis [Fig. 4(a)]. Again, small cavitation bubbles, which lasted less than 100 μs before collapsing, were induced around the periphery of the cylindrical volume. In contrast, large cavitation bubbles formed near the beam axis appeared to be stabilized at their maximum size (1.0–1.7 mm) for more than 150 μs before collapsing. Several large-size bubbles, possibly formed by coalescence of individual bubbles around their maximum expansion, were observed along the beam axis beyond the focal point. The primary collapse of this cavitation bubble cluster appeared to occur slightly before $t=250 \mu\text{s}$ and continued until $t=350 \mu\text{s}$. After the primary collapse, several small bubbles were seen to rebound along the beam axis at about $t=400 \mu\text{s}$. However, because of the resolution of the high-speed camera, subsequent oscillation of cavitation bubbles of much reduced size cannot be identified.

Simultaneously, AE measurements revealed three distinctive bursts of pressure spikes [Fig. 4(b)]. The first peak pressure was again observed in about 192 μs following the interference signal from the spark discharge ["s" in Fig. 4(b)], whereas the peak pressure spikes of the second and

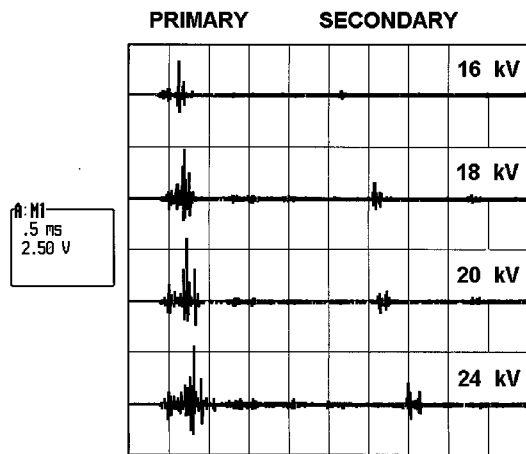


FIG. 6. A series of representative acoustic emission signals generated in vitro by a Dornier HM-3 lithotripter at 16, 18, 20, and 24 kV, respectively.

third bursts were observed in a time delay of 220 and 360 μ s, respectively. Comparison with the dynamics of the primary cavitation cluster showed that the first AE burst correlated with the initial expansion, while the second and the third AE bursts were generated by the subsequent collapses of the inertial cavitation bubbles. Similar to the AE signals produced at 16 kV, a small intermediate pressure spike was observed between the first and second bursts, corresponding to the collapse of small cavitation bubbles surrounding the beam axis. In addition, several pressure spikes of reduced amplitude were observed following the third AE burst. These pressure spikes are presumably generated by subsequent oscillation of small, residual cavitation bubbles along the beam axis.

For the secondary cluster, as shown in Fig. 5(a), the most significant feature is the strong coagulation of cavitation bubbles along the beam axis. Density of the bubble population is much higher than that induced at 16 kV, such that individual bubbles inside the cluster cannot be distinguished. After the initial expansion, coalescence of bubbles had led to the formation of a bubble column of 5 mm in diameter. This bubble column appeared to be stabilized at its maximum size for about 200 μ s, followed by a rapid collapse starting at $t=250$ μ s, and a complete collapse of the cavitation cluster between $t=300$ –350 μ s. Correspondingly, acoustic emission measured simultaneously [Fig. 5(b)] revealed an initial pressure burst that started at time 3328 μ s, followed by two additional bursts of pressure spikes occurring about 200 and 330 μ s later, respectively. In comparison, these AE bursts were found to correspond in time with the initial expansion, and subsequent collapse of the secondary cavitation bubble cluster. In addition, the time delay between the primary and the secondary cavitation bubble clusters was measured to be 3328 μ s from the AE signal. This value correlates fairly well with the time delay of 3350 (± 50) μ s, measured directly from high-speed photographic sequence.

It is clear from Figs. 3–5 that the duration of the primary and the secondary bubble clusters, as well as the time delay between them, increase as the output voltage of the lithotripter increases in the range of 16 to 20 kV. Compared with

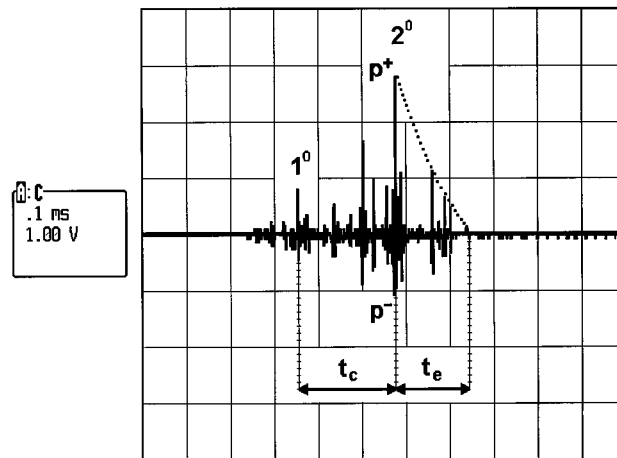


FIG. 7. The primary acoustic emission burst produced in vitro by a Dornier HM-3 lithotripter at 20 kV.

results obtained at 16 kV, cavitation activities induced at 20 kV were found to be much stronger, with increased bubble sizes, a larger distribution volume around the beam focus, and a prolonged duration of bubble activities. This increased cavitation activity is consistent with Church's theoretical prediction that lithotripsy-induced cavitation will increase in bubble size and expansion duration at higher shock wave amplitudes.²² Further, simultaneous high-speed photography and AE measurements have clearly shown a close correlation between dynamics of the inertial cavitation bubbles and resultant acoustic emission, generated at various voltage settings during SWL. Such a correlation provides an essential physical basis for assessing inertial cavitation during SWL using acoustic emission measurements, when a direct observation of cavitation bubbles is impossible as is generally the case in vivo.

B. Dornier HM-3 lithotripter

1. Acoustic emission: In vitro

Figure 6 shows a series of representative acoustic emission signals produced by inertial cavitation bubbles generated in water by the HM-3 lithotripter, at an output setting of 16, 18, 20, and 24 kV, respectively. To facilitate comparison, the same oscilloscope time base (500 μ s/div) and vertical sensitivity (2.5 V/div) were used for all measurements. In each trace, at least two (a primary and a secondary) acoustic emission bursts were identified, while a third burst with much reduced intensity was observed occasionally. It is clear from Fig. 6 that the intensity, duration of each acoustic burst, and time delay between the primary and the secondary acoustic bursts all increase with the discharge voltage of the lithotripter.

Figure 7 shows the detailed structure of a primary acoustic emission burst produced by the HM-3 lithotripter at 20 kV. Similar to results obtained from the experimental lithotripter, each burst was found to be comprised of an initial (1°) and delayed (2°) group of pressure spikes, which correspond, respectively, to the initial expansion and primary collapse/rebound of the cavitation bubble cluster induced around the beam focus. Two time-scale parameters were in-

TABLE I. *In vitro* acoustic emission data.

Output (kV)	Primary burst						Secondary burst					
	t_c (μ s)	t_e (μ s)	p^+ (kPa)	p^- (kPa)	E_{p1} (mJ)	E_{p2} (mJ)	t_c (μ s)	t_e (μ s)	p^+ (kPa)	p^- (kPa)	E_{s1} (mJ)	E_{s2} (mJ)
16	158 \pm 20	158 \pm 16	313 \pm 87	186 \pm 25	6.0 \pm 0.4	16.8 \pm 4.1	48 \pm 19	22 \pm 17	32 \pm 15	54 \pm 8	1.8 \pm 0.3	2.1 \pm 0.4
18	194 \pm 23	188 \pm 34	328 \pm 76	180 \pm 25	8.8 \pm 1.5	20.9 \pm 5.4	79 \pm 13	27 \pm 16	53 \pm 23	78 \pm 16	2.4 \pm 2.4	3.1 \pm 0.6
20	220 \pm 16	217 \pm 26	427 \pm 77	220 \pm 56	10.0 \pm 1.7	24.1 \pm 5.5	101 \pm 23	76 \pm 32	135 \pm 69	128 \pm 59	4.0 \pm 1.1	5.4 \pm 2.6
24	276 \pm 29	235 \pm 30	427 \pm 97	241 \pm 38	14.7 \pm 2.7	36.4 \pm 8.9	149 \pm 35	105 \pm 52	145 \pm 60	142 \pm 28	6.5 \pm 1.1	8.8 \pm 1.9

t_c : Duration of primary bubble expansion and collapse.
 t_e : Duration of subsequent bubbles oscillation following the primary collapse.
 p^+ : Positive peak pressure of the acoustic emission burst generated by the primary collapse of bubble cluster.
 p^- : Negative peak pressure of the acoustic emission burst generated by the primary collapse of bubble cluster.
 $E_{p1(s1)}$: Acoustic energy associated with the initial group of pressure spikes of the primary (secondary) AE burst.
 $E_{p2(s2)}$: Acoustic energy associated with the delayed group of pressure spikes of the primary (secondary) AE burst.

roduced: t_c which measures the time delay between peak pressure spikes of initial and delayed signals, and t_e which measures the exponential decay of the AE signal following the primary collapse of the cavitation bubble cluster. Physically, t_c represents the duration of primary bubble expansion and collapse, while t_e measures the subsequent oscillation period of cavitation bubbles following their primary collapse. In addition, both positive (p^+) and negative (p^-) peak pressure of the acoustic emission signals corresponding to primary collapse of the bubble cluster were measured. Because of the negative polarity of the focused hydrophone, p^+ and p^- correspond to the tensile and compressive acoustic peak pressure, respectively. The compressive pressure waves are generated by the rapid collapse and rebound of cavitation bubbles. The tensile pressure waves may be produced by scattering of the compressive pressure waves propagating outward through the cavitation cluster. The acoustic energy associated with each group of pressure spikes was also determined, using Eq. (1).

Table I summarizes results for the acoustic emission signals measured *in vitro*. The value of t_c (mean \pm s.d.) for the primary and secondary acoustic emission bursts was found to increase linearly from 158 \pm 20 to 276 \pm 29 μ s and from 48 \pm 19 to 149 \pm 35 μ s, respectively, as output voltage of the lithotripter increases from 16 to 24 kV. Correspondingly, the

value for t_e was also found to increase almost linearly from 158 \pm 16 to 235 \pm 30 μ s and from 22 \pm 17 to 105 \pm 52 μ s, respectively, in the same output voltage range. Because of increased initial expansion and subsequent collapse of cavitation bubbles at higher lithotripter output voltages, the resultant peak pressure spikes and associated acoustic energy of the AE burst were found to increase accordingly. In the lithotripter output voltage range of 16 to 24 kV, values of p^+ and p^- for the primary and secondary acoustic emission bursts were found to vary from 313 \pm 87 to 427 \pm 97, 186 \pm 25 to 241 \pm 38 kPa, and from 32 \pm 15 to 145 \pm 69, 54 \pm 8 to 142 \pm 28 kPa, respectively. Corresponding values of acoustic energy associated with the initial and delayed pressure spikes of the primary and secondary bursts were found to vary from 6.0 \pm 0.4 to 14.7 \pm 2.7 (E_{p1}), 16.8 \pm 4.1 to 36.4 \pm 8.9 mJ (E_{p2}), and from 1.8 \pm 0.3 to 6.5 \pm 1.1 (E_{s1}), 2.1 \pm 0.4 to 8.8 \pm 1.9 mJ (E_{s2}), respectively. In summary, lithotripsy-induced inertial cavitation bubble expansion and collapse *in vitro* clearly show a voltage-dependence, and more than 80% of the acoustic emission energy is produced by the first cavitation bubble cluster.

2. Acoustic emission: *In vivo*

Representative acoustic emission signals measured from the renal parenchyma of a swine model treated by the HM-3 lithotripter at an output setting of 16, 18, 20, and 24 kV, respectively, are shown in Fig. 8. Again, the same oscilloscope time base (500 μ s/div) and vertical sensitivity (1.5 V/div) were used for all measurements for easy comparison. In each trace, three acoustic emission bursts were identified, with intensity, burst duration and time delay between adjacent bursts all increasing with the output voltage of the lithotripter.

Table II summarizes results for the acoustic emission signals measured *in vivo*. When output voltage of the lithotripter increased from 16 to 24 kV, the value of t_c for primary acoustic emission bursts was found to be almost unchanged (from 71 \pm 2 to 72 \pm 3 μ s), while the value for t_e increased slightly from 1364 \pm 48 to 1412 \pm 65 μ s. For the secondary acoustic emission burst, the value of t_c was found to increase from 120 \pm 16 to 163 \pm 22 μ s, while the value for t_e increased from 99 \pm 30 to 161 \pm 29 μ s in the same output voltage range.

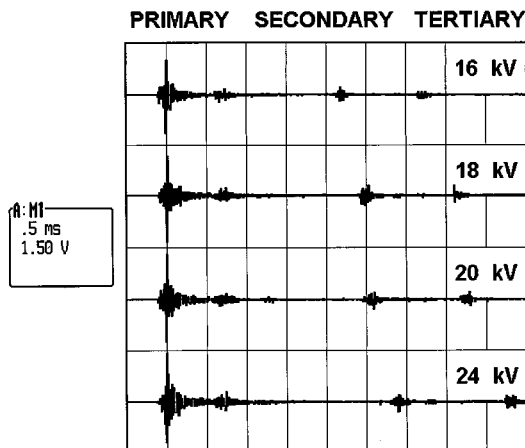


FIG. 8. A series of representative acoustic emission signals generated *in vivo* by a Dornier HM-3 lithotripter at 16, 18, 20, and 24 kV, respectively.

TABLE II. *In vivo* acoustic emission data.

Output (kV)	t_c (μ s)	t_e (μ s)	Primary burst				t_c (μ s)	t_e (μ s)	Secondary burst			
			p^+ (kPa)	p^- (kPa)	E_{p1} (mJ)	E_{p2} (mJ)			p^+ (kPa)	p^- (kPa)	E_{s1} (mJ)	E_{s2} (mJ)
16	71 \pm 2	1,364 \pm 48	178 \pm 33	157 \pm 25	1.5 \pm 0.2	19.6 \pm 2.4	120 \pm 16	99 \pm 30	27 \pm 8	41 \pm 7	0.6 \pm 0.1	1.5 \pm 0.2
18	71 \pm 2	1,390 \pm 44	165 \pm 39	161 \pm 30	1.7 \pm 0.3	20.8 \pm 2.3	135 \pm 22	140 \pm 33	29 \pm 10	46 \pm 10	1.0 \pm 0.2	2.2 \pm 0.5
20	71 \pm 2	1,412 \pm 65	175 \pm 33	154 \pm 30	2.0 \pm 0.2	25.0 \pm 4.1	145 \pm 24	166 \pm 29	33 \pm 11	51 \pm 12	1.1 \pm 0.2	2.4 \pm 0.4
24	72 \pm 3	1,373 \pm 52	196 \pm 29	173 \pm 27	2.2 \pm 0.4	26.4 \pm 2.9	163 \pm 22	161 \pm 29	47 \pm 30	55 \pm 19	1.2 \pm 0.3	2.7 \pm 0.9

t_c : Duration of primary bubble expansion and collapse.
 t_e : Duration of subsequent bubbles oscillation following the primary collapse.
 p^+ : Positive peak pressure of the acoustic emission burst generated by the primary collapse of bubble cluster.
 p^- : Negative peak pressure of the acoustic emission burst generated by the primary collapse of bubble cluster.
 $E_{p1(s1)}$: Acoustic energy associated with the initial group of pressure spikes of the primary (secondary) AE burst.
 $E_{p2(s2)}$: Acoustic energy associated with the delayed group of pressure spikes of the primary (secondary) AE burst.

In addition, when output voltage of the lithotripter increased from 16 to 24 kV, values of p^+ and p^- for the primary and secondary acoustic emission bursts were found to vary from 178 \pm 33 to 196 \pm 29, 157 \pm 25 to 173 \pm 27 kPa, and from 27 \pm 8 to 47 \pm 30, 41 \pm 7 to 55 \pm 19 kPa, respectively. Corresponding values of acoustic energy associated with the initial and delayed pressure spikes of each bursts were found to vary from 1.5 \pm 0.2 to 2.2 \pm 0.4 (E_{p1}), 19.6 \pm 2.4 to 26.4 \pm 2.9 mJ (E_{p2}), and from 0.6 \pm 0.1 to 1.2 \pm 0.3 (E_{s1}), 1.5 \pm 0.2 to 2.7 \pm 0.9 mJ (E_{s2}), respectively. In tissue, more than 87% of acoustic emission energy is produced by the first cavitation bubble cluster.

3. Acoustic emission: Comparison between *in vitro* and *in vivo* signals

Figure 9 shows the primary acoustic emission burst measured *in vivo* at 20 kV, together with the corresponding acoustic emission signal measured *in vitro*. In comparison, the *in vivo* acoustic emission signal was found to have a significantly shortened bubble expansion duration, but a substantially increased oscillation period following primary bubble collapse. At 20 kV, the average duration of the primary bubble expansion and collapse (t_c) is reduced from 220 to 71 μ s, yet the subsequent oscillation period of the bubble cluster (t_e) is increased from 217 to 1412 μ s, when

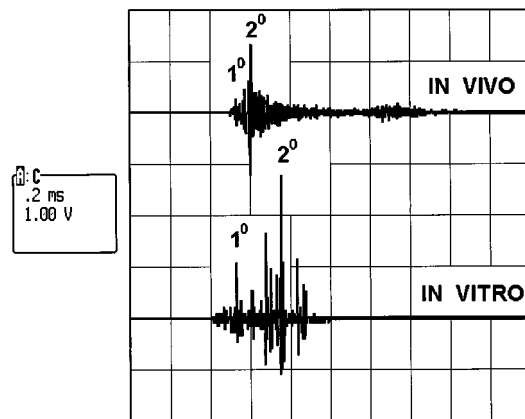


FIG. 9. The primary acoustic emission burst produced *in vivo* by a Dornier HM-3 lithotripter at 20 kV, together with the corresponding acoustic emission signal measured *in vitro*.

AE is measured in water and from renal parenchyma, respectively. Such characteristic changes in t_c and t_e of AE signals from *in vitro* to *in vivo* are also observed at other output voltage settings.

In addition, when the lithotripter output voltage increases from 16 to 24 kV, there is approximately linear increment in t_c in water (up to 75% for the primary cavitation cluster), corresponding to the increased bubble expansion and growth as shown photographically in Figs. 3–5. In contrast, in the same output voltage range, the corresponding value for lithotripsy-induced cavitation bubbles in renal parenchyma is increased by less than 2%. This finding indicates clearly that bubble expansion in renal parenchyma is constrained. Because it has been estimated that tissue attenuation reduces the tensile component of lithotripter shock waves by less than 6%,¹¹ this constrained bubble expansion may be primarily due to the confined space available in renal parenchyma for bubble expansion. In the same output voltage range, the rate of increment in t_c for the second cavitation cluster in renal parenchyma, though slightly higher, is still less than 50% of the corresponding increment in water. This result again indicates a tissue-constraining effect on bubble expansion, although factors attributing to the differences in bubble expansion in renal parenchyma between the primary and secondary clusters are not clear.

Similarly, the rate of increment in peak pressure and acoustic energy associated with lithotripsy-induced cavitation clusters were found to be significantly reduced when comparing AE signals *in vivo* with corresponding ones *in vitro*.

III. DISCUSSION

Using simultaneous high-speed photography and pressure wave measurements, the inertial cavitation bubbles and associated acoustic emission produced *in vitro* by electrohydraulic shock wave lithotripters have been characterized. Our results demonstrate that following each spark discharge, two distinct cavitation bubble clusters, separated in time by about 3 ms, are generated. The dynamics of these cavitation bubble clusters can be described as an initial rapid expansion, followed by a long period of stabilization around the maximum bubble size, and a subsequent rapid collapse. Correspondingly, two distinct bursts of acoustic emission signals are

detected, which match very closely in time with the expansion duration of cavitation bubble clusters and the time delay between them. The results of these simultaneous measurements have shown clearly a close correlation between the dynamics of cavitation bubbles and concomitant acoustic emission produced during SWL. This correlation provides a physically essential, as well as clinically significant, basis for the use of acoustic emission measurements to assess quantitatively cavitation activity induced in vivo during lithotripsy. Such a noninvasive technique may potentially be used to quantify the degree and violence of cavitation in vivo, an information that may help to ensure the effectiveness and safety of lithotripsy treatment.

Previous theoretical work by Church has predicted that following the primary collapse, SWL-induced cavitation bubbles will achieve the same final size, independent of the initial radii of cavitation nuclei (within 1–10 μm), provided that shock wave amplitude is sufficiently high.²² This prediction implies that the secondary shock wave, generated by the collapse/rebound of a large single cavitation bubble at the electrode tip, may encounter cavitation nuclei population of similar sizes. Consequently, inertial cavitation bubbles induced by the secondary shock waves in electrohydraulic SWL are expected to be more uniform in size and coherent in oscillation dynamics.^{11,22} In this study, high-speed photography has revealed that cavitation bubbles induced by the primary shock wave pulse vary significantly in size, with the largest bubbles produced along the symmetrical axis of the shock wave reflector. In contrast, cavitation bubbles induced by the secondary shock wave pulse are much more uniform in size, and are coalesced around the symmetrical axis of the shock wave reflector, presumably due to mutual attraction (Bjerknes forces) between cavitation bubbles oscillating in phase. This experimental observation has confirmed Church's theoretical prediction.²²

The clear correlation between the output voltage (or energy) of a shock wave lithotripter and resultant cavitation bubble activity both in vitro and in vivo, as identified in this study, supports the prevailing general consensus that lithotripsy-induced cavitation is a primary mechanism for both stone fragmentation as well as tissue injury.²⁴ The increased cavitation activity at higher output voltages is consistent with previous experimental findings that stone comminution²⁵ and tissue injury²⁶ increase significantly with lithotripter output voltage. However, the relationship between cavitation activity, resultant stone fragmentation, and concomitant tissue injury in SWL has yet to be determined.

Although the temporal pattern of acoustic emission produced during SWL is similar, there are distinct differences between signals measured in vivo and that detected in vitro. Most importantly, when cavitation bubble clusters are induced in renal parenchyma, the duration of the primary bubble expansion is greatly reduced, yet the subsequent oscillation period of the bubble is substantially increased as compared to the corresponding values in water. In addition, although the values of t_c (except for the primary cavitation cluster), t_e , p^+ , p^- , E_1 , and E_2 associated with the cavitation clusters induced in renal parenchyma all increase with the output voltage of the lithotripter, the rate of increment in

these parameters are significantly reduced compared to the corresponding values in water. These characteristic changes in the dynamics of lithotripsy-induced cavitation bubbles from in vitro to in vivo cannot be explained simply by accounting for tissue attenuation of incident shock waves. It has been estimated that tissue attenuation (4–6-cm thick) will only reduce the tensile components of lithotripter shock waves, which is responsible for inducing cavitation bubbles, by less than 6%.¹¹ Such a small percentage of attenuation should only reduce proportionally the duration of bubble expansion-collapse and subsequent oscillation in tissue from that observed in water, but should not greatly shorten bubble expansion duration while substantially prolonging subsequent bubble oscillation period. Further, a small percentage of tissue attenuation should not alter significantly the rate of increment in bubble expansion-collapse duration with regard to lithotripter output voltage.

Therefore, we hypothesize that the characteristic changes in bubble dynamics from in vitro to in vivo, as observed in this study, are due to primarily tissue constraining effect based on following speculations. It is well-known that nephron is the functional and structural unit of renal parenchyma. Each nephron consists of a glomerulus and a urinary tubule, filled with blood and urine, in which cavitation nuclei can exist. The maximum dimension of these fluid-filled lumens is in the range of 10 to 100 μm .²⁷ Therefore, when a cavitation bubble is induced within a nephron, the bubble cannot expand freely to the maximum size of 1 to 2 mm as expected in water. Instead, the bubble will expand in a confined space against the surrounding blood vessel or tubule. As the tissue is stretched, it will consume a portion of the kinetic energy of the expanding bubble, and simultaneously increase the ambient pressure outside the bubble, leading to a shortened expansion and a sooner collapse of the bubble. Moreover, because the potential energy of the bubble is reduced, it will collapse less violently, resulting in lower energy dissipation in heat, viscous loss, and acoustic emission (a lightly damped bubble oscillation), which would increase the oscillation period of the bubble after its primary collapse. In contrast, a lithotripsy-induced cavitation bubble in water can expand over 100 times its initial size,²² thus it accumulates much higher potential energy at its maximum size. Consequently, the bubble collapses more violently, leading to higher energy dissipation in heat, viscous loss, and acoustic emission (a heavily damped bubble oscillation), which shortens the oscillation period of the bubble after its primary collapse. These speculations are consistent with the dynamics of bubble oscillation in water and in renal parenchyma observed experimentally.

In summary, results of this study indicate that the oscillation of lithotripsy-induced cavitation bubbles in renal parenchyma is fundamentally different from that induced in water. Yet, it is clear that in order to facilitate the interpretation of acoustic emission signals in vivo, and to better understand the mechanism of tissue injury during SWL, a theoretical model for lithotripsy-induced cavitation in a tissue constraining medium needs to be developed.

IV. CONCLUSIONS

The inertial cavitation and associated acoustic emission generated during electrohydraulic shock wave lithotripsy have been characterized using high-speed photography and acoustic pressure measurements. Simultaneous recording of bubble oscillation and concomitant acoustic emission during lithotripsy in vitro have revealed a close correlation between the expansion duration of cavitation bubbles and the time delay between the first two pressure spikes in the acoustic emission signal. This correlation allowed us to assess the inertial cavitation produced by a clinical Dornier HM-3 shock wave lithotripter, using acoustic emission measurements only. In the clinical output voltage range (16–24 kV), similar patterns of acoustic emission, generated by the HM-3 lithotripter both in water and in renal parenchyma, were measured. However, comparing to in vitro results, lithotripsy-induced cavitation in vivo was found to have a significantly shortened expansion duration, but a substantially prolonged oscillation period following the primary collapse of the bubble. These characteristic changes in bubble dynamics from in vitro to in vivo are presumably due to constraining effect of renal tissue on bubble expansion. In addition, although a clear correlation between lithotripter output voltage (or energy) and resultant cavitation bubble activity both in vitro and in vivo has been observed, the relationship between cavitation activity, resultant stone fragmentation, and concomitant tissue injury in SWL has yet to be determined. Understanding this relationship is essential for the potential use of acoustic emission measurements to guide the safe and effective use of high-intensity shock waves during lithotripsy treatment.

ACKNOWLEDGMENT

This work was supported in part by the Whitaker Foundation and by NIH through Grant No. PO1-DK20543.

- ¹A. J. Coleman, J. E. Saunders, L. A. Crum, and M. Dyson, "Acoustic cavitation generated by an extracorporeal shockwave lithotripter," *Ultrasound Med. Biol.* **13**, 69–76 (1987).
- ²L. A. Crum, "Cavitation microjets as a contributory mechanism for renal calculi disintegration in ESWL," *J. Urol.* **140**, 1587–1590 (1988).
- ³M. Delius, W. Brendel, and G. Heine, "A mechanism of gallstone destruction by extracorporeal shock waves," *Naturwissenschaften* **75**, 200–201 (1988).
- ⁴P. Zhong and C. J. Chuong, "Propagation of shock waves in elastic solids caused by the impact of cavitation microjets: Part I. Theoretical formulation," *J. Acoust. Soc. Am.* **94**, 19–28 (1993).
- ⁵P. Zhong, C. J. Chuong, and G. M. Preminger, "Propagation of shock waves in elastic solids caused by the impact of cavitation microjets: Part II. Application to extracorporeal shock wave lithotripsy," *J. Acoust. Soc. Am.* **94**, 29–36 (1993).
- ⁶F. Brummer, T. Brauner, and D. F. Hulser, "Biological effects of shock waves," *World J. Urol.* **8**, 224–232 (1990).
- ⁷D. Suhr, F. Brummer, and D. F. Hulser, "Cavitation-generated free radicals during shock wave exposure: investigations with cell-free solutions

- and suspended cells," *Ultrasound Med. Biol.* **17**, 761–768 (1991).
- ⁸M. Delius, M. Jordan, H. Liebich, and W. Brendel, "Biological effects of shock waves: kidney haemorrhage by shock waves in dogs—administration rate dependence," *Ultrasound Med. Biol.* **14**, 689–694 (1988).
- ⁹M. Delius, R. Denk, C. Berding, H. Liebich, M. Jordan, and W. Brendel, "Biological effects of shock waves: cavitation by shock waves in piglet liver," *Ultrasound Med. Biol.* **16**, 467–472 (1990).
- ¹⁰A. P. Evan, L. R. Willis, B. Connors, G. Reed, J. A. McAteer, and J. E. Lingeman, "Shock wave lithotripsy-induced renal injury," *Am. J. Kidney Dis.* **17**, 445–450 (1991).
- ¹¹A. J. Coleman, M. J. Choi, J. E. Saunders, and T. G. Leighton, "Acoustic emission and sonoluminescence due to cavitation at the beam focus of an electrohydraulic shock wave lithotripter," *Ultrasound Med. Biol.* **18**, 267–281 (1992).
- ¹²A. J. Coleman, M. Whitlock, T. G. Leighton, and J. E. Saunders, "The spatial distribution of cavitation induced acoustic emission, sonoluminescence and cell lysis in the field of a shock wave lithotripter," *Phys. Med. Biol.* **38**, 1545–1560 (1993).
- ¹³W. Sass, M. Braunlich, H. Dreyer, E. Matura, W. Folberth, H. Priesmeyer, and J. Seifert, "The mechanism of stone disintegration by shock waves," *Ultrasound Med. Biol.* **17**, 239–243 (1991).
- ¹⁴A. Phillip, M. Delius, C. Scheffczyk, A. Vogel, and W. Lauterborn, "Interaction of lithotripter-generated shock waves with air bubbles," *J. Acoust. Soc. Am.* **93**, 2496–2509 (1993).
- ¹⁵M. Kuwahara, N. Ioritani, K. Kamke, S. Shirai, K. Taguchi, T. Saitoh, S. Orikasa, K. Takayama, S. Aida, and N. Iwama, "Hyperechoic region induced by focused shock waves in vitro and in vivo: Possibility of acoustic cavitation bubbles," *J. Lithotripsy Stone Dis.* **1**, 218–288 (1989).
- ¹⁶M. Delius and S. Gambihler, "Sonographic imaging of extracorporeal shock wave effects in the liver and gallbladder of dogs," *Digestion* **52**, 55–60 (1992).
- ¹⁷R. K. Seman, W. J. Davros, J. A. Goldberg, B. S. Garra, W. S. Hayes, E. L. Cattau, S. C. Horii, C. J. Cooper, and P. M. Silverman, "Cavitation effects during lithotripsy, Part II. Clinical observations," *Radiology* **177**, 163–166 (1990).
- ¹⁸A. J. Coleman, T. Kodama, M. J. Choi, T. Adams, and J. E. Saunders, "The cavitation threshold of human tissue exposed to 0.2 MHz pulsed ultrasound: Preliminary measurements based on a study of clinical lithotripsy," *Ultrasound Med. Biol.* **21**, 405–417 (1995).
- ¹⁹S. R. Akers and F. H. Cocks, "Experimental laboratory lithotripter: Design construction and operation," *Biomed. Instrument. Technol.* **23**, 50–53 (1989).
- ²⁰A. J. Coleman and J. E. Saunders, "A survey of the acoustic output of commercial extracorporeal shock wave lithotripters," *Ultrasound Med. Biol.* **15**, 213–227 (1989).
- ²¹P. Zhong, H. Tong, F. H. Cocks, and G. M. Preminger, "Asymmetric cavitation bubble collapse near stone surfaces," *J. Endourology* **9**, Suppl. **1**, S62 (1996).
- ²²C. C. Church, "A theoretical study of cavitation generated by an extracorporeal shock wave lithotripter," *J. Acoust. Soc. Am.* **86**, 215–227 (1989).
- ²³T. G. Leighton, *The Acoustic Bubbles* (Academic, London, 1994).
- ²⁴A. J. Coleman and J. E. Saunders, "A review of the physical properties and biological effects of the high amplitude acoustic field used in extracorporeal lithotripsy," *Ultrasonics* **31**, 75–89 (1993).
- ²⁵C. J. Chuong, P. Zhong, and G. M. Preminger, "A comparison of stone damage caused by different modes of shock wave generation," *J. Urol.* **148**, 200–205 (1992).
- ²⁶M. Delius, G. Enders, Z. Xuan, H. Liebich, and W. Brendel, "Biological effects of shock waves: Kidney damage by shock waves in dogs—Dose dependence," *Ultrasound Med. Biol.* **14**, 117–122 (1988).
- ²⁷J. Hamburger, G. Richet, and J. P. Grunfeld, *Organ Physiology—Structure and Function of the Kidney* (W.B. Saunders, Philadelphia, 1971).

The acoustic features of vowel-like *grunt* calls in chacma baboons (*Papio cyncephalus ursinus*): Implications for production processes and functions

Michael J. Owren

Department of Psychology, Reed College, 3203 SE Woodstock Boulevard, Portland, Oregon 97202

Robert M. Seyfarth and Dorothy L. Cheney

Departments of Psychology and Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

(Received 9 July 1996; accepted for publication 26 December 1996)

The acoustic features of 216 baboon *grunts* were investigated through analysis of field-recorded calls produced by identified females in known contexts. Analyses addressed two distinct questions: whether the acoustic features of these tonal sounds could be characterized using a source-filter approach and whether the acoustic features of grunts varied by individual caller and social context. Converging evidence indicated that grunts were produced through a combination of periodic laryngeal vibration and a stable vocal tract filter. Their acoustic properties closely resembled those of prototypical human vowel sounds. In general, variation in the acoustic features of the grunts was more strongly related to caller identity than to the social contexts of calling. However, two acoustic parameters, second formant frequency and overall spectral tilt, did vary consistently depending on whether the caller was interacting with an infant or participating in a group move. Nonetheless, in accordance with the general view that identity cueing is a compelling function in animal communication, it can be concluded that much of the observed variability in grunt acoustics is likely to be related to this aspect of signaling. Further, cues related to vocal tract filtering appear particularly likely to play an important role in identifying individual calling animals. © 1997 Acoustical Society of America. [S0001-4966(97)01605-6]

PACS numbers: 43.80.Ka, 43.70.Aj [FD]

INTRODUCTION

Studies of nonhuman primates (hereafter *primates*) have demonstrated many essential anatomical similarities in the sound-production mechanisms of these animals and humans (e.g., Geist, 1933; Negus, 1949; Hill and Booth, 1957; Kelemen, 1969; Hilloowala, 1975; Schon Ybarra, 1995). Because many primate vocalizations, like human speech sounds, are produced using vocal-fold vibration and one or more contiguous supralaryngeal vocal tract chambers (see Fig. 1 for schematic representations of the human and baboon vocal tract, respectively), researchers have long noted that these sounds are necessarily formed through the interaction of source energy input with resonances (*formants*) due to vocal tract cavities (e.g., Kelemen, 1938, 1949; Negus, 1949; Zhinkin, 1963). These concepts, well-developed in speech applications (e.g., Chiba and Kajiyama, 1941; Stevens and House, 1955; Fant, 1960), have significantly influenced both analysis and synthesis of speech [see Rubin and Vatikotis-Bateson (in press) for a recent review]. In contrast, with the notable exception of work by Lieberman and his colleagues (e.g., Lieberman, 1968, 1969; Lieberman *et al.*, 1969, 1972) and Zhinkin's (1963) largely unacknowledged work, production-based perspectives have historically not influenced acoustic analysis of primate calls. More recently, however, both empirical and tutorial papers have appeared that argue for the application of source-filter theory and other speech-related, production-based perspectives in acoustic primatology (see reviews by Fitch and Hauser, 1995; Owren and Linker, 1995).

Briefly summarizing the available evidence concerning the contribution of vocal-fold-related source energy in primates, many species appear to employ both humanlike vocal-fold action and vibration modes that humans can probably not emulate. For instance, a number of species can produce highly tonal vocalizations involving stable fundamental frequencies and rich harmonic spectra [e.g., baboons (Andrew, 1976); Japanese macaques (Green, 1975); gelada monkeys (Richman, 1976); rhesus macaques (Hauser *et al.*, 1993)]. Fundamentals in these and more pure-tone vocalizations show both significant intra- and inter-species variation, for instance, being somewhat lower [e.g., baboons (Owren *et al.*, 1993)], moderately higher [e.g., rhesus and Japanese macaques (Owren *et al.*, 1992)], or many times higher [e.g., bushbabies and slow loris (Zimmermann, 1981); chimpanzees (Clark and Wrangham, 1993); and rhesus macaques (Hauser *et al.*, 1993)] than corresponding values in human phonation.

However, while laryngographic measurements by Brown and colleagues (personal communication) showed that normative variability in the periodicity of vocal-fold vibration were quite similar in a Sykes's monkey and a human subject, macaques and other species also routinely produce calls based on vibration modes that are quite unlike those underlying human speech sounds. Such vocalizations can include high-amplitude, noise-based sounds, virtually pure-tone sinusoids, frequency sweeps that can cover many octaves in a fraction of a second, and apparent simultaneous combinations of periodic and aperiodic vocal fold vibration

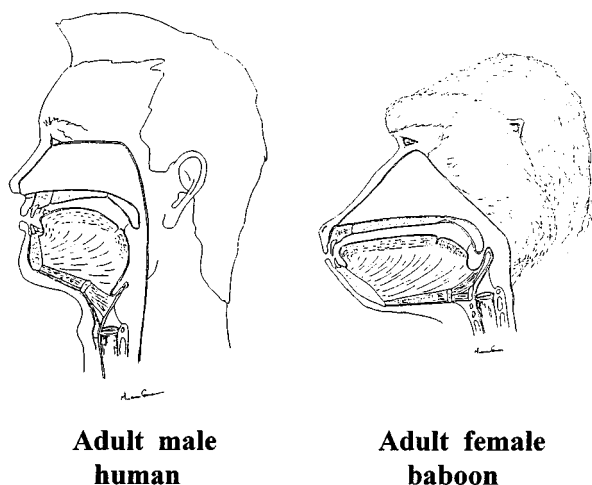


FIG. 1. Schematic, midsagittal views of a human and a baboon vocal tract. The relative headsize shown is approximately correct, although the drawings were not created to exact specifications. From Owren and Linker (1995), adapted with permission.

patterns [see, for example, spectrograms shown in Peters (1983) and Green (1975) for rhesus and Japanese macaques, respectively]. Anatomical studies by Schön Ybarra (1995) and others have indicated a probable basis for these capabilities in the occurrence of a relatively rigid “lip” on the medial extremity of each vocal fold in a number of primate species.

Sound spectrography has also revealed strong evidence of nonlaryngeal influences on vocalizations in many primates. Some, for instance, have been found to have extralaryngeal vocal sacs proposed to act as resonators in sounds specialized for long-distance transmission (see Gautier and Gautier, 1977). More commonly, however, spectrographic evidence shows apparent resonance effects that are analogous to those of the human supralaryngeal vocal tract, for instance in both tonal *coo* calls and other vocalizations in macaques (e.g., Grimm, 1967; Lieberman, 1968; Green, 1975; Peters, 1983; Hohmann and Herzog, 1985; Hohmann, 1989; Masataka and Thierry, 1993), in gelada monkey calls (Richman, 1976), in *double-grunts* produced by wild mountain gorillas (Seyfarth *et al.*, 1994), and in the baboon *grunts* (Andrew, 1976) that are the subject of this article.

Modification of vocal tract resonances has also been both proposed (e.g., Lieberman, 1975; Richman, 1976; Hauser and Fowler, 1991; Hauser, 1992) and indirectly demonstrated (e.g., Hauser *et al.*, 1993; Hauser and Schön Ybarra, 1994; Brown and Cannito, 1995). Such modification involves changes not only in laryngeal, lip, and mandible positioning, but may also include tongue movement. As illustrated in Fig. 1, primates differ from humans in having thinner tongues, larynges that are positioned higher in the neck, and a relative lack of soft tissues in the supralaryngeal cavities. Thus the former appear to have significantly less opportunity to alter the resonance properties of their vocal tract through flexible modification of articulator position than do the latter. While the functional significance of the spectral changes that can occur has not been tested, the perceptual salience of static spectral patterning has been demonstrated

both in vervets classifying synthetic versions of noisy species-typical alarm calls (Owren, 1990) and in baboons and Japanese macaques discriminating among human vowel sounds (Hienz and Brady, 1988; Sinnott, 1989).

Overall then, it is clear that the source-filter perspective can be usefully applied to a number of calls produced by primates, and that both vocal-fold action and extra-laryngeal resonance may play significant roles in shaping the functionally important acoustic characteristics of such calls. On the other hand, insufficient evidence is available to specify many details of either the source or filter components of production for any given call in any particular primate species, or concerning possible interaction effects between the two. Calls produced by baboons (*Papio cynocephalus*) may be of particular interest in this regard, due to the resemblance of these calls to human vowel sounds (Zhinkin, 1963; Andrew, 1976; Owren *et al.*, 1993, 1995). These similarities set the stage for detailed examination of grunts both for evidence of basic sound-producing processes and for the role that particular acoustic features, whether speechlike or not, play in the use of these sounds as species-typical social signals. In general, although a voluminous body of work exists on baboon social behavior, ecology, and reproduction, relatively little is known concerning either the acoustics or social functions of baboon calls. However, a brief description of the baboon vocal repertoire is provided by Bolwig (1959), while more detailed analyses of the acoustics and functions of grunts (Zhinkin, 1963; Andrew, 1976; Waser, 1982; Cheney *et al.*, 1995a), *contact barks* (Byrne, 1981; Cheney *et al.*, 1996), *copulatory* calls (Hamilton and Arrowood, 1978; O’Connell and Cowlshaw, 1994), and *loud calls* (Waser, 1982; Brown *et al.*, 1995) are also available.

The work reported here is an acoustic analysis of a sample of grunts recorded from a group of chacma baboons (*P. c. ursinus*) living under natural conditions in Botswana. There were two overriding goals for these analyses. First, we wanted to examine as well as possible the apparent contributions of laryngeally based sound generation and vocal tract filtering in the calls. To do so, we first characterized the apparent fundamental frequency of each vocalization through three different methods: (1) automated pitch extraction from the time waveform; (2) measuring the lowest visible harmonic in spectrographic representations; and (3) ascertaining the frequency intervals between partials that appeared above the apparent fundamental frequency. Then we examined the overall patterning of the frequency spectrum for evidence of formant-related spectral energy peaks and compared the relationship between derived values for fundamental frequencies and formant frequencies in each call. Finally, we used formant frequency measurement to estimate the baboon vocal tract length (based on an idealized, uniform-tube model) for comparison to values from humans.

The second overriding goal was to examine possible acoustic differences between grunts produced in varying behavioral circumstances, paying particular attention to the possibility that baboons might produce calls with varying spectral properties through apparent flexible modification of articulator positioning in the supralaryngeal vocal tract. As reported by Cheney *et al.* (1995a), baboons produce grunts

under a variety of social conditions, including situations in which these vocalizations appear to be directed toward particular receivers and act to facilitate further social interaction. As a first test of the relation between acoustic features and particular behavioral contexts, calls were scored according to the identity and relative social rank of the callers and apparent receivers (specific receivers were not always evident), whether an infant was involved, and whether the vocalization appeared as a component of initiating or participating in group movement. A variety of acoustic measures were made from each call and mean values for each were compared according to these context variables. Discriminant function analysis was then used to measure the classification accuracy achieved using each of these acoustic and social contexts.

I. METHOD

A. Study area and subjects

The study site was in the Okavango Delta, a vast seasonal swamp created by the floodplain of the Okavango River in northwestern Botswana. The habitat consisted of flat, open grasslands interspersed by slightly elevated wooded patches. Subjects were nine sexually mature, adult females that were part of a group of approximately 70 baboons. The group has been observed continuously since 1977 by Hamilton and colleagues (e.g., Bulger and Hamilton, 1988; Hamilton and Bulger, 1992) and ourselves (e.g., Cheney *et al.*, 1995a, b). All animals can be identified individually and maternal relatedness is known for natal individuals. Relative dominance ranks are also known for each animal. The group is fully habituated to the presence of human observers on foot.

B. Observational protocol and behavioral-context definitions

Data collection occurred continuously between July 1992 and July 1993, beginning each day as the group moved from the sleeping site in the morning and continuing for a predetermined interval of up to two and one-half hours. During data collection, two to four observers spread themselves throughout the group and noted the identities of as many calling females as possible. For each calling event or bout, the apparent receiver and salient behaviors by any individual occurring before and after the calling event were also noted. Other potentially significant circumstances (e.g., whether another group was nearby or whether a predator had recently been seen) were also available in the call records, allowing us to restrict the sample unambiguously to the variables of greatest interest.

Three primary variables were examined in detail in testing for possible differentiation in grunt call acoustics. First, we checked for individual variation in the parameters measured in each grunt based on the identity of the calling animal (variable *Caller*). Second, the behavioral context (variable *Context*) associated with each grunt was noted according to whether the call was produced when the calling

individual (and often others as well) were moving from one area of their range to another (*Move*) or whether the calling individual was engaged in interactions with an infant or the infant's mother (*Infant*). The *Move* and *Infant* contexts represented over 80% of all circumstances in which grunts were recorded. Both prior research on vervet monkeys (Cheney and Seyfarth, 1982; Seyfarth and Cheney, 1984) and impressions from field observations led us to hypothesize that, if the acoustic features of grunts were differentially associated with various social contexts, these two particular variables were the most likely to reveal such relationships. Other analyses have examined grunts given when animals are reconciling with one another following an agonistic interaction (Cheney *et al.*, 1995; Cheney and Seyfarth, in press). These grunts are also relatively common, and appear to be acoustically indistinguishable from those classified as *Infant* grunts in the current analyses (Seyfarth and Cheney, 1997).

The identity of the receiver was also recorded whenever evident, and calls to females were scored as occurring either to subordinate or to dominant animals (variable *Receiver-Rank*). This variable was also examined in separate analyses, due to the evident importance of relative dominance rank in many nonhuman primate species, including baboons (e.g., Smuts *et al.*, 1987), and previous evidence from other species that calls to dominant animals can be acoustically distinguishable from those directed to subordinate individuals (e.g., Seyfarth and Cheney, 1984). Classification on basis of *Receiver-Rank* was orthogonal to sorting based on the *Infant* and *Move* variables, neither of which required that a given call be unambiguously associated with a particular receiver. Accordingly, the sample tested using the *Receiver-Rank* variable constituted a subset of the calls identified for analysis using the primary context variables.

C. Apparatus

Calls were tape-recorded using Sennheiser directional microphones (ME80 heads with K3U power modules), Sony WC5M Professional Walkman cassette tape-recorders, and chromium-oxide metal tape cassettes. Tape-recorded grunts were subsequently digitized using the CSRE 4.0 software package (Avaaz Innovations, London, Ontario, Canada), implemented on a 486-type personal computer. Calls were sampled at 10.0 kHz using a DT2801A A/D board (Data Translation, Marborough, MA), a 16-channel audio interface, and 30-band, 1/4-octave, 20-kHz equalizer for low-pass filtering at approximately 4.8 kHz (preliminary analysis revealed only trace energy above 5.0 kHz). The bulk of acoustic analysis was conducted using the CSpeech version 4 software package (Paul Milenkovic, Dept. of Electrical and Computer Engineering, University of Wisconsin-Madison, 1415 Johnson Drive, Madison, WI 53706), also implemented on a 486-type personal computer. However, some measures were derived using the SIGNAL bioacoustics system (Engineering Design, Belmont, MA) on the same machine. The Number Cruncher Statistical System (NCSS; Jerry L. Hintze, 329 North 1000 East, Kaysville, UT 84037) was used for statistical analysis.

D. Acoustic analysis

1. Call selection

Baboons typically produce grunts in readily identifiable bouts, consisting of a series of two or more calls given in rapid succession (although individual grunts are occasionally produced). Here, a bout was defined as a set of one or more grunts given by a single individual that was both preceded and followed by at least 10 s during which it did not grunt. The database consisted of a total of 1361 bouts of grunts recorded over the entire, yearlong observation period, all of which were digitized prior to further call selection. Nine-hundred and thirty-four of these bouts were produced by 16 adult females that could potentially serve as subjects in acoustic analyses. One grunt was randomly selected from each bout by a research assistant (if the bout consisted of only a single call, that call was selected). In each case, the assistant was aware of the age-sex class of the caller and its two-letter identity code, but was unaware of any other aspect of the circumstances of call production. Nine females, who accounted for a total of 464 of the recorded grunt bouts, were found to be represented by the largest number of calls produced in the greatest variety of social contexts, and were therefore designated as subjects.

Each call was then classified according to the context that had been noted for the corresponding bout during field observations, and the potential sample was then restricted to exemplars that were either Infant or Move related, as defined above. We included only calls that could be unambiguously assigned to either the Infant or Move contexts, for instance excluding vocalizations that were given to an infant while the group was simultaneously on the move. Following context-based classification, each of the nine female's calls were examined in the chronological order of their original occurrence to further screen out vocalizations showing recording-related acoustic artifacts. The calls eliminated were those whose onset and offset points were unmeasurably obscured by background noise and exemplars that included extraneous sounds. In some instances, digital high-pass filtering at 50 Hz was used to improve the quality of the calls. The signal-to-noise ratio (variable SNR) of each call was then measured by subtracting the average rms background noise level immediately preceding or following the vocalization from the average rms energy over the duration of the vocalization. The call-selection process continued until a final sample of 216 grunts had been identified. This set included 24 grunts recorded from each of the 9 subjects, with 12 calls from the Infant context and 12 calls from the Move context in each case.

2. F_0 -related measures

Three methods were used to extract the fundamental frequency of each call. First, F_0 was determined from the call's time waveform using the CSpeech program (variable $F_{0_60\%}$). This algorithm is based on a least-mean-square, autocorrelation method described by Milenkovic (1987). It was applied over the middle 60% of each call, so as to maximize the probability of successful pitch extraction by excluding the lowest-amplitude cycles in each waveform. A repre-

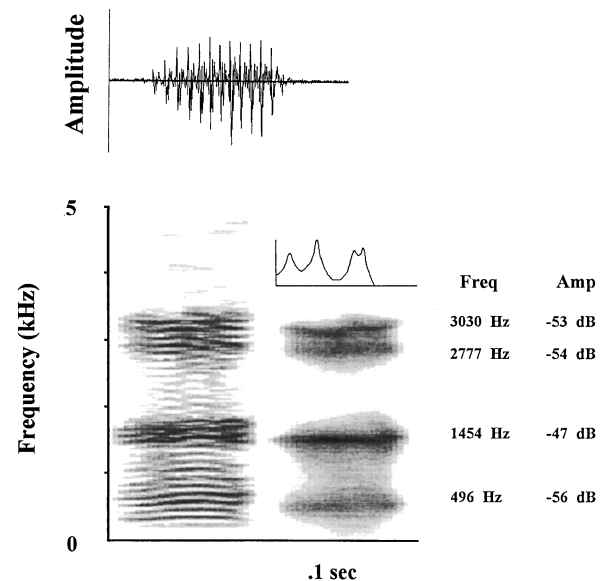


FIG. 2. A typical adult female chacma baboon grunt waveform (left) and corresponding digital spectrograms based on 256-point FFT (middle-left) and 14-coefficient, 256-point LPC computations (middle-right). A single LPC-based smoothed spectral envelope derived from the midpoint of the call is shown above the spectrograms; values listed on the right represent the frequency and amplitude values for each peak in this function. From Owren and Linker (1995), adapted with permission.

sentative baboon grunt waveform is shown in Fig. 2. F_0 values derived using this approach were spot checked for each waveform through direct cursor measurement of one or more cycles.

The second approach, using the SIGNAL bioacoustics package, derived both the F_0 and a number of other related measures from digital spectrograms. Each spectrogram was based on 2048-point FFTs computed using a Hanning analysis window, providing 9.8-Hz resolution. The cursor was used to measure frequency values of adjacent partials in the power spectrum, which was derived by centering the analysis window on the temporal midpoint of each call. Measurements were made of 3–5 partials in each call (as many as was practicable) in the portion of the frequency range showing greatest overall energy levels. These values were used to calculate a mean interval between the partials, which were hypothesized to be evenly spaced at multiples of the F_0 (variable $F_{0_Harmonics}$).

The third approach also made use of the SIGNAL-derived digital spectrograms, but involved deriving a fundamental frequency contour directly from that representation. This approach (described in Owren and Casale, 1994) involved first drawing time and frequency boundaries around the visually identified fundamental frequency or a higher-amplitude partial using cursors or the mouse, and then automatically extracting the highest-amplitude frequency from within each column of the spectrogram's underlying matrix using an automated peak-detection function in SIGNAL. The mean overall F_0 was then calculated from these contours (variable $F_{0_Contour}$).

3. Measurements from the time waveform

In addition to providing the basis for the $F_{0_60\%}$ variable described above, the time waveform was also used for measurement of overall call duration (variable *Dur*) and SNR.

4. Spectral measurements

Frequency and amplitude measurements of four major spectral peaks were derived from spectral slices computed through covariance-based linear predictive coding [LPC; see Owren and Linker (1995) and Owren and Bernacki (in press) for reviews of LPC approaches and applications in bioacoustics]. For each call, a single slice was computed using 15 coefficients and a 512-point Hamming analysis window centered on the temporal midpoint of the sound. Each LPC spectrum was checked by superimposing it on a corresponding 512-point fast Fourier transform (FFT) and displaying these characterizations simultaneously with a spectrographic representation computed using a 300-Hz analysis filter bandwidth. In a small number of cases, the LPC function was recomputed with a slightly higher or lower coefficient number to provide a more accurate match to the FFT spectrum. Cursor-based measurements of center frequency (variables $F1f$ – $F4f$) and amplitude (variables $F1a$ – $F4a$) were made from the spectral peaks. In order to characterize frequency and amplitude relationships among the peaks, a series of relational measures were computed from these absolute values. Using values from all possible pairs of peaks, relationships were computed for both frequency (e.g., variable $F2f$ – $F1f$) and amplitude values (e.g., variable $F1a$ – $F2a$). Finally, a measure of the overall slope of the frequency spectrum (variable *LPC_Tilt*) was derived by exporting the spectral slice to a spreadsheet program and calculating its least-squares regression line.

II. RESULTS

A. General characteristics of the vocalization sample

The number of calls occurring in the 216 bouts from which the grunts were selected varied substantially ($M = 4.38$ calls, $s.d.=4.52$, range 1–18). Using analysis of variance (ANOVA) with Context and Caller as factors, this variation was found to be statistically significant both between the Infant ($M=641$, $s.d.=5.31$) and Move contexts ($M=2.34$, $s.d.=2.54$), $F(1, 198)=75.9$, $p<0.001$, and among the individual animals, whose means ranged from 2.33 to 7.75, $F(8, 198)=7.0$, $p<0.001$. An interaction effect was also found, $F(8, 198)=5.55$, $p<0.001$. Similar analyses of call durations revealed that the sounds were quite brief ($M=109.4$ ms, $s.d.=31.2$), showed no significant difference between the Infant ($M=111.6$ ms, $s.d.=33.2$) and Move ($M=107.1$ ms, $s.d.=29.0$) contexts, $F(1, 198)=1.97$, $p=0.16$, but did show a main effect of Caller, $F(8, 198)=19.4$, $p<0.001$ and a significant interaction effect, $F(8, 198)=2.69$, $p<0.01$.

Sixty-one calls were excluded during the preliminary call-selection process due to compromised recording quality, 28 in the Infant and 33 in the Move contexts, respectively.

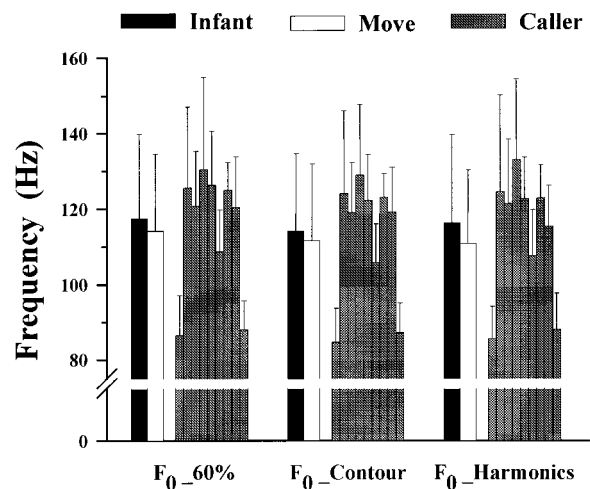


FIG. 3. Means and standard deviations of fundamental frequency values derived from the time waveform ($F_{0_60\%}$), first partial ($F_{0_Contour}$), and intervals between partials ($F_{0_Harmonics}$) are shown for the Infant and Move contexts (left) and for each of the nine adult female baboons (right).

The number of calls excluded did not differ by context, $t(16)=0.19$, $p=0.86$, two-tailed, but did differ by individual caller ($M=6.78$, $s.d.=6.82$), $\chi^2(8)=54.9$, $p<0.001$. The mean SNR value for calls included in the sample was 15.8 dBv (overall range 13.2–17.6 dBv). An ANOVA, using Context and Caller as factors, revealed no significant main effect of the former (Infant: $M=15.2$, $s.d.=6.26$; Move: $M=16.3$, $s.d.=5.94$), $F(1, 197)=1.93$, $p=0.17$. However, both a main effect of Caller, $F(8, 197)=2.05$, $p=0.04$, and an interaction effect, $F(8, 197)=4.98$, $p<0.001$, were found. Follow-up, pairwise Scheffe comparisons of SNR values from individual animals revealed no significant differences. A separate test of mean SNR values from calls produced to subordinate ($M=16.9$, $s.d.=6.54$) versus dominant ($M=14.8$, $s.d.=5.74$) females was not statistically significant, $t(109)=1.68$, $p=0.10$, two-tailed.

Based on the receiver identities noted in the field, the sample of 216 calls produced by the context- and acoustic-quality-related selection procedures was found to include 104 calls that had been given to either adult males, immature animals, or no apparent receiver, 72 calls given to subordinate females, and 40 calls given to dominant females. Analyses involving classification based on the Receiver-Rank variable described above thus included only the latter, 112 grunts.

B. Fundamental frequency measures

Means and standard deviations for each of the three fundamental frequency measures are shown in Fig. 3, by Context and by Caller. A one-way ANOVA (that treated individual calls as subjects, thereby inflating its power) showed that these three sets of fundamental frequency measurements did not differ from one another, $F(2, 639)=0.45$, $p=0.64$. Overall correlations among the three variables were then calculated, based on measurements from individual calls ($n = 212$ calls, with 4 missing values). The resulting Spearman-

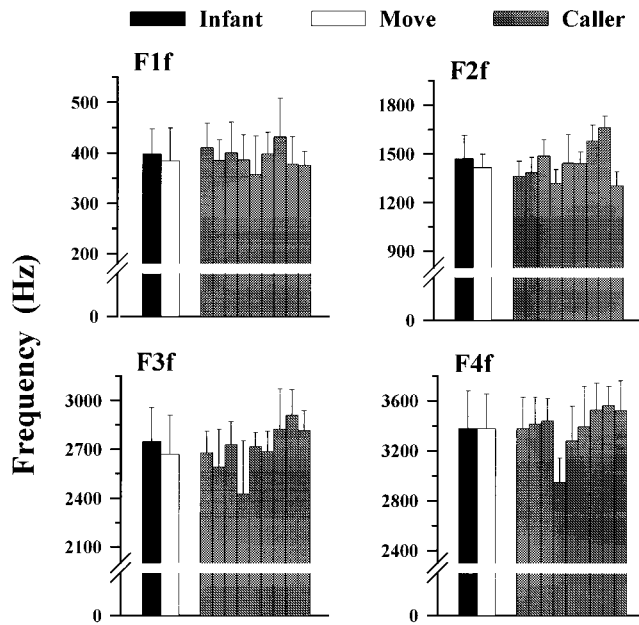


FIG. 4. Means and standard deviations of spectral peak frequencies. Overall values are shown for the Infant and Move contexts (left) and for each of the nine adult female baboons (right). The order of presentation for individual animals corresponds to that of Fig. 3 in each case.

rank correlation matrix revealed coefficients of 0.93, 0.95, and 0.96, comparing $F_{0_60\%}$ with $F_{0_Harmonics}$, $F_{0_Harmonics}$ with $F_{0_Contour}$, and $F_{0_60\%}$ with $F_{0_Contour}$, respectively. The Spearman-rank correlation test is more conservative than the corresponding, parametric Pearson r statistic (according to Hintze, 1995), and each of these correlations was statistically significant at a level of $p < 0.001$. To simplify further analyses, $F_{0_60\%}$ alone was used as a representative measure of fundamental frequency, which will hereafter be referred to as F_0 .

C. Spectral peak frequencies and their relationships to fundamental frequency

Means and standard deviations for spectral peak frequencies are shown in Fig. 4, by Context and by Caller. Multiple regression analysis, in which F_0 was used as a factor and $F1f$ – $F4f$ were entered as dependent variables, produced a statistically significant outcome, $F(4, 179) = 5.50$, $p < 0.001$, which could be traced to individual relationships between F_0 and $F2f$, $t = 3.54$, $p < 0.001$, and between F_0 and $F4f$, $t = -3.08$, $p < 0.01$. However, the adjusted R -squared value of 0.09 observed for the overall relationship indicated that it was weak (this measure serves as a primary index of the equation's goodness-of-fit to the data—a value near 1.0 indicates a strong linear relationship, while a value near 0 indicates little or no association; see Hintze, 1995). Tested individually, F_0 – $F2f$ and F_0 – $F4f$ R -squared values were 0.02 and 0.05, respectively. When each baboon was tested separately using this multiple regression approach, a statistically significant relationship was found between F_0 and the four formant frequencies in only one case, $F(4, 19) = 3.37$, $p = 0.03$. This outcome was associated with

TABLE I. Means and standard deviations of overall formant frequencies (Hz) with resulting estimates of vocal tract length (cm), based on idealized calculations for a straight, uniform tube closed at one end.

Formant	M	s.d.	Estimated vocal tract length (cm)
$F1$	437.7	21.4	19.7
$F2$	1442.1	121.3	18.0
$F3$	2706.3	147.5	15.9
$F4$	3374.6	201.0	17.9
Overall mean estimate			17.9

an adjusted R -squared value of 0.29. The overall mean adjusted R -squared value for the other eight individuals was 0.07.

D. Vocal tract length estimates

Overall means for the first 4 formant frequencies (shown in Table I) were used to estimate the length of an adult female baboon's vocal tract. The estimates are based on modeling the vocal tract as a straight tube, closed at one end. The expected length (l) in cm can be calculated as $l = (2k + 1)c/4F$, where F is the formant frequency, k is the formant number minus 1, and c is 34 400 cm/s, the speed of sound (see Lieberman and Blumstein, 1988). Individual estimates derived from each formant are also shown in the table, where the overall estimated length is 17.9 cm. An ANOVA was used to test the variation among the averaged length estimates derived from each animal's mean formant frequencies, and showed that significant variation did occur, $F(3, 32) = 17.11$, $p < 0.001$. Scheffe tests indicated that this overall effect resulted from the estimates based on $F1$ and $F3$ frequencies. In these cases, the derived lengths were significantly longer and shorter, respectively, than expected (based on the mean overall estimate).

E. Differences by context and caller

1. Analysis of variance

Statistically significant differences were readily found on numerous dependent measures when Context and Caller were used as MANOVA factors, as shown in Table II. In these analyses, each variable was first subjected to z -score transformation to ensure normality, and three groupings of conceptually related acoustic variables were tested separately. As 28 of the calls did not exhibit an identifiable $F4$, all measures related to this formant were excluded from the analyses at this point. A skewing effect resulting from uneven distribution of the 28 missing cases across the Caller and Context variables was thereby avoided. The first grouping of the remaining variables (*Absolute-Spectral*) consisted of the absolute frequency ($F1f$ – $F3f$) and amplitude ($F1a$ – $F3a$) measures for each of the spectral peaks. The second grouping (*Relative-Spectral*) consisted of the relational measurements, including LPC_Tilt values. The third grouping (*Waveform-Related*) consisted of measures related to the time waveform, namely SNR, Dur, and F_0 .

In each case in which a statistically significant context-related difference occurred, the direction of this difference

TABLE II. Overall MANOVA results using Context and Caller as factors in conjunction with three variable groupings—Absolute-Spectral (AS), Relative-Spectral (RS), and Waveform-Related (Wv). Single and double asterisks denote statistical significance at the 0.05 and 0.01 levels, respectively.

Grouping	Factor	df	F value	Statistically significant variables
AS	Context	6,182	7.73**	$F1f^*, F2f^{**}, F2a^{**}, F3f^{**}, F3a^{**}$
	Caller	48,900	8.38**	$F1f^{**}, F2f^{**}, F2a^{**}, F3f^{**}, F3a^{**}$
	Interaction	48,900	1.89*	$F1a^{**}, F2f^{**}, F3f^*$
RS	Context	7,181	7.59**	$F2f-F1f^{**}, F3f-F1f^{**}, F3f-F2f^{**}, F2a-F3a^{**}, LPC_Tilt^{**}$
	Caller	56,980	6.04**	$F2f-F1f^{**}, F3f-F1f^{**}, F3f-F2f^{**}, F2a-F3a^{**}, LPC_Tilt^{**}$
	Interaction	56,980	1.41*	$F2f-F1f^{**}, F3f-F1f^*, F3f-F2f^*, F1a-F2a^{**}, F1a-F3a^{**}$
Wv	Context	1,3	4.84**	$F_{0_60\%}^{**}$
	Caller	8,24	13.01**	$Dur^{**}, F_{0_60\%}^{**}$
	Interaction	8,24	2.71**	$Dur^*, F_{0_60\%}^{**}$

was tallied by subject. In most cases, statistical significance for a given variable did not imply a consistent pattern of underlying mean differences for individual animals. For $F2f$, however, a nonparametric one-proportion test did reveal statistically significant consistency in direction of difference (i.e., eight of nine cases showed a higher mean value in the Infant context $M = 1470.2$, $s.d. = 148.5$, than in the Move context, $M = 1415.3$, $s.d. = 146.1$, $z = 2.33$, $p = 0.02$), while LPC_Tilt closely approached statistically significant consistency (seven of nine cases showed a higher mean value in the Infant context, $M = -6.17$, $s.d. = 131$, than in the Move context, $M = -5.25$, $s.d. = 1.93$, $z = 1.67$, $p = 0.10$).

Receiver-Rank was tested with ANOVA using the same three variable groupings. A significant overall outcome that emerged for Absolute-Spectral measures, $F(6, 99) = 2.81$, $p = 0.01$, was found to be entirely due to the value of $F2f$. Overall, this frequency was significantly lower when the call was produced to a subordinate animal ($M = 1383$, $s.d. = 111.7$) than to a dominant individual ($M = 1468$, $s.d. = 111.3$), $F(1, 104) = 14.2$, $p < 0.001$. However, this outcome did not occur consistently among the various subjects. Relative-Spectral measures also revealed a significant difference by Receiver-Rank, $F(7, 98) = 2.16$, $p = 0.045$, which was due to the variable $F2f-F1f$, $F(1, 104) = 14.2$, $p < 0.001$. As no differences were found in $F1f$ for either grouping, $F2f$ can be presumed to underly the significant effect in $F2f-F1f$. Again, no evidence of consistency in the direction of difference emerged among the various animals. No significant differences were found for Waveform-Related measures, $F(3, 107) = 0.99$, $p = 0.40$.

2. Discrimination function analyses

The overall lack of consistency among individuals in the direction of differences when producing calls in the three social contexts was borne out by discriminant function analyses conducted separately for Context, Caller, and Receiver-Rank. In these analyses (again excluding $F4$ -related measures), classification was based on manual stepwise entry of variables that continued until we had entered all variables whose associated F -value produced a p -value of less than 0.10. Results are illustrated in Table III, including the measures entered in each case. The 8 of 15 variables entered for Caller classification was a significantly higher proportion than the 2 of 15 variables (i.e., $F2f$ and LPC_Tilt) entered for Context classification ($z = -2.32$,

$p = 0.02$, nonparametric two-proportions test), or the 1 of 15 variables (i.e., $F2f$) entered for Receiver-Rank classification ($z = -2.78$, $p < 0.01$). Similarly, error reduction when calls were sorted by Caller was significantly greater than when sorted by Context ($z = 7.23$, $p < 0.001$) or by Receiver-Rank ($z = 4.61$, $p < 0.001$).

3. Relationships between $F2f$ and spectral tilt measures

$F2f$ and LPC_Tilt were found to be significantly correlated, $t = 3.17$, $p < 0.01$, but with an adjusted R -square value of only 0.045. Analogously, testing the relationship of $F2f$ to the relational amplitude variables in multiple regression, an overall significant effect, $F(2, 202) = 5.10$, $p = 0.01$, was associated with an adjusted R -square value of only 0.039. LPC_Tilt was, however, found to be significantly correlated with the various relational amplitude measures when tested with them individually (the latter were themselves significantly intercorrelated). As a result of the strong overall relationship found, $F(2, 202) = 67.5$, p

TABLE III. Variables entered in discriminant function analysis for classification of calls by Context, Caller, and Receiver-Rank, with resulting classification accuracy. Table entries show the order in which each variable was entered and its associated F -to-enter value.

Variable	Context	Caller	Receiver-Rank
Dur		(3) 16.8	
$F_{0_60\%}$		(2) 28.1	
$F1f$		(8) 2.7	
$F1a$		(7) 4.6	
$F2f$	(2) 13.6	(1) 30.2	(1) 14.3
$F2a$			
$F3f$			
$F3a$		(5) 3.1	
$F2f-F1f$			
$F3f-F1f$		(4) 8.1	
$F3f-F2f$			
$F1a-F2a$			
$F1a-F3a$			
$F2a-F3a$			
LPC_Tilt	(1) 16.5	(6) 3.3	
Classification accuracy	Context	Caller	Receiver-Rank
By chance	50.0%	11.1%	50.0%
Statistically	62.8%	65.7%	66.7%
Error reduction	25.6%	58.6%	33.3%

< 0.001 , with an adjusted R -square value of 0.395, LPC_Tilt and the relational amplitude variables were considered to be approximately interchangeable measures of overall spectral tilt.

III. DISCUSSION

A. Baboon vocal production mechanisms

1. Source and filter

As the baboon sound production system anatomically consists of vocal folds and a supralaryngeal vocal tract, the apparent role of vocal-fold vibration was examined by testing the relationship between the evident fundamental frequency of this vibration and energy found elsewhere in the sound spectrum. As primate calls can pose difficulties in automated vocal pitch extraction, for instance due to mismatches between algorithm design and the acoustic properties of particular calls (see Owren and Linker, 1995), three conceptually distinct measures were used for this aspect of the analysis. F_0 _60% values were based on a time-domain method whose combination of short-term autocorrelation and center clipping (i.e., setting low-amplitude samples to zero) is designed to attenuate all but the longest waveform periodicities (see Hess, 1983; O'Shaughnessy, 1987). This approach produced results that were indistinguishable from measurements of the lowest-frequency partial in the frequency spectrum, F_0 _Contour, which in turn were a virtual match to values based on the frequency intervals observed between adjacent partials at higher frequencies, F_0 _Harmonics. This outcome indicates that the primary source energy of grunt calls is periodic vibration, which can be presumed to occur in the vocal folds.

Direct tests of the independence of the source energy and subsequent vocal tract filter were not possible with these data, so an indirect approach of examining correlations between F_0 values and spectral frequency peak values was used instead. No relationships were observed between F_0 and formant peaks $F1$ and $F3$. For $F2$ and $F4$, statistically significant correlations were found, but the degree of relationship was very small in each case. A significant relationship emerged for only one of the nine animals when each was tested individually. While the degree of relationship was more convincing in that one case, the overall dataset demonstrates independence between F_0 values and spectral peak frequencies, leading us to conclude that these measures reflect separable components of the sound-production system.

These data do not, of course, provide evidence concerning the spectrum of the source energy function itself. Such information would be necessary to gauge the relative contributions of frequency-dependent amplitude values and vocal tract resonance effects in shaping the overall amplitude of a given harmonic. In addition, baboons exhibit supra-laryngeal air sacs (see Schön Ybarra, 1995), whose possible role in grunt production cannot be evaluated at this point. Nonetheless, the occurrence of humanlike, formant-related spectral effects is strongly indicated.

These results are consistent with findings of both Zhinkin (1963) and Andrew (1976), who examined spectral

energy peaks in grunts recorded from *hamadryas*, and *hamadryas* and *anubis* baboons, respectively. Testing the peaks using a 32-band spectrometer, Zhinkin found them to generally match the formant frequencies of his own vocal imitations of the sounds. Andrew (1976), who examined the calls of one juvenile male and one infant using a sound spectrograph, also found evidence of a formant-based vocal tract transfer function imposed on a glottally related harmonic structure.

2. Comparisons to human speech production

The mean F_0 of 114.5 Hz (99.9–128.2 Hz interquartile range, based on the F_0 _60% measure) associated with these adult female baboon grunt calls is very similar to the typical spontaneous speaking fundamental frequency of 123 Hz reported for young human adult males (see Baken, 1987). Baboon grunts are often produced in the absence of obvious arousal and are also lower pitched than are calls like *screams* or *fear barks* that are typically associated with such arousal (Seyfarth and Cheney, 1997). This apparent match between baboons and humans in the fundamental frequency of normative vocalizations indicates between-species similarity in the length of the vibrating component of the vocal folds [see Titze (1994) for a review of vocal-fold mechanics]. Adult male baboons typically produce grunts with even lower fundamental frequencies (Seyfarth and Cheney, 1997) and are therefore likely to have longer vocal folds than do either adult female baboons or humans.

Mean frequency values for the first three formants, which were 438, 1442, and 2706 Hz, respectively, can be closely matched to typical formant frequencies reported for human adult males by Peterson and Barney (1952; only three formants are typically reported for vowels). The estimated vocal tract length of 17.9 cm derived from the calls is consequently similar to the frequently cited characterization of the adult human male vocal tract as being 17.5 cm long. Species in the *Papio* genus all exhibit elongated muzzles (e.g., Hill, 1972; Napier and Napier, 1985) with requisitely long mandibles, oral, and nasal cavities (as illustrated in Fig. 1). A chacma baboon skull measured by Elliot (1913) was reported to have a palatal length of 19.0 cm and mandible length of 15.8 cm. Given that the larynx is set just behind and below the oral cavity (e.g., Negus, 1949), the estimated vocal tract lengths derived here are consistent with these measurements.

Figure 5 shows the mean $F1$ and $F2$ values measured in the baboons superimposed on Peterson and Barney's well-known mapping of the American-English vowel space, based on 76 men, women, and children. As is apparent both from the means and from the larger shaded area representing all the data points, the variation in these formant values covers a significant, but centrally located proportion of this space. Overall, this outcome is consistent with Lieberman and colleagues' hypothesis that the typical "vowel space" of primates exhibits a restricted range of formant variation that directly reflects the relatively inflexible nature of their "single-tube" vocal tracts. However, the formant values we measured indicate that the adult female baboon vocal tract is

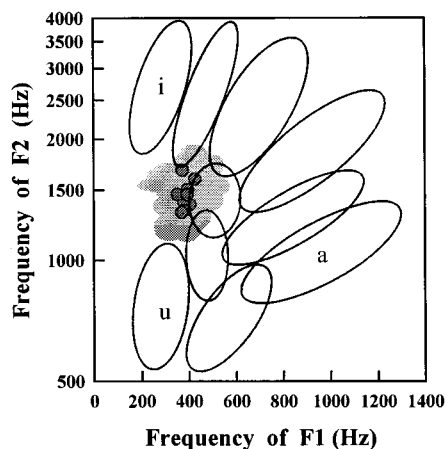


FIG. 5. Mean frequency measurements for F_1 and F_2 from the nine adult female baboons (shaded circles) are shown in the American-English F_2/F_1 vowel-space mapping of Peterson and Barney (1952). The larger shaded area represents the range of variation found among all 216 calls examined.

not entirely uniform over its length and can therefore not be exactly matched by an idealized straight-tube resonator. In addition, further investigation of baboon calls may reveal instances of greater formant variability than was found here, although due to their low pitch and dense harmonic structure, grunts do appear to be the most likely vocalization in the baboon repertoire to exhibit clear, and functionally important formant patterns.

As discussed in greater detail below, the data provided only modest indications of changes in formant-related spectral peak patterns by social context. This outcome is somewhat surprising, in that primates appear capable of a number of articulatory movements that modify the vocal tract transfer function during the production process (see Lieberman, 1975). In fact, the work of Hauser and colleagues (Hauser, 1992; Hauser *et al.*, 1993; Hauser and Schön Ybarra, 1994), indicates that rhesus monkeys routinely do alter the shapes and lengths of their vocal tracts through various articulatory maneuvers, thereby influencing resonant-frequency outcomes. A similar outcome was indirectly indicated by laryngographic measurements involving two Sykes's monkey subjects tested by Brown and Cannito (1995).

While some data specifically relevant to baboons are available from previous work, they shed little additional light on this question of functional modification of vocal tract positioning. Zhinkin (1963), for instance, reported successful x-ray photography of one subject during vocalization and found no evidence of significant tongue movement (specifically, "pharyngeal narrowing"). However, the sound in question was a high-pitched squeal, a call type that is unlikely to be usefully modified by flexible articulator positioning due to having a high fundamental. In contrast, Andrew (1976) specifically reported that one of his young baboon subjects routinely produced changes in formant position through mandible and tongue movements. Unfortunately, the maneuvers (lip-smacking and tongue protrusion) were ones that are not typical of normative call production in adult baboons (Seyfarth and Cheney, 1997).

In sum, then, the data from both the current study and

past work lend little support to the notion that changes in formant position are a critical aspect of baboon vocal communication. However, a firm conclusion must await more comprehensive analyses of various calls in this species, specifically including work with captive animals. The nature of both the source energy and the supralaryngeal vocal tract resonances involved in baboon sound production would be amenable to separate examination using any of a number of the approaches developed in the context of articulatory models of human speech (reviewed by Rubin and Vatikiotis-Bateson, in press). Techniques like laryngography (Brown and Cannito, 1995) and blocking of articulatory control through anesthetic injection (Hauser and Schön Ybarra, 1994) have already been successfully applied to other monkey species.

B. Acoustic variation related to individual identity

1. Functional significance of variation and the importance of formants

Discrimination of kin from nonkin, as well as individuals within both of these categories, has been proposed to be of overriding significance both to the evolution of social behavior in general (Hamilton, 1964; Trivers, 1971) and in the kin-biased behaviors of primates specifically (e.g., Cheney *et al.*, 1986; Smuts *et al.*, 1987). A number of studies of monkeys and apes have shown either individual variability in the acoustic structure of vocalizations, discrimination among individual conspecific callers, or both [reviewed in Snowdon (1986); see also Gouzoules *et al.*, 1986; Gouzoules and Gouzoules, 1990; Owren *et al.*, 1992; Rendall *et al.*, 1996]. While no systematic, vocalization-based evidence is available concerning discrimination or recognition of individuals or kin in baboons, the requisite acoustic cues to support these expected functions appear to be present in grunt calls [see also Cheney *et al.* (1995a, b) for indirect evidence of recognition of individually distinctive grunts, screams, and fear barks in these animals].

Acoustic variation among individuals was found for a variety of the measures used here, of which ten variables were found to be useful predictors of caller identity in discriminant function analysis. While most of the acoustic measures were related to some aspect of the spectral characteristics of these calls, variables Dur and F_0 were also found to be important in sorting calls, at least by individual. However, it should be noted that any given animal is also likely to be able to facultatively alter both of these attributes when grunting, as evidenced by the ability of these baboons to produce other, longer or much higher-pitched, tonal call types (Seyfarth and Cheney, 1997). It is therefore requisitely unlikely that either call duration, vocal pitch, or any combination of the two can serve as the primary cues to individual identity in this vocal repertoire.

In contrast, frequency and amplitude characteristics of the spectral peaks found in the grunts emerged as a predominant source of identity-based classificatory power that can also be expected to be stable both within and among callers. Based on the source-filter principles of vocal production discussed in the Introduction, we hypothesize that these spectral characteristics act as markers of small-scale variation in for-

mant frequencies and amplitudes that in turn result from minor anatomical differences in the structures and cavities of the vocal tracts of individual baboons. Very few spectral peak characteristics appeared to vary among the social contexts examined using our sample. In conjunction with the brief nature of these grunts and an evident lack of within-call formant changes noted during the acoustic analysis process, this evidence suggests to us that vocal tract resonance patterns such as those imprinted on these harmonically rich grunts may provide a particularly stable and dependable source of information marking the identity of the calling animal across a variety of vocalization contexts [discussed by Rendall *et al.* (in review) based on rhesus monkey vocalizations].

2. Perceptual salience of the observed variation in acoustic features

From a perceptual point of view, the acoustic attributes indexed by the measures found to be of greatest interest in statistical analysis also can in general be expected to be salient to listening baboons, although some of the data supporting this view are derived from other primate species. As a group, Old World monkeys appear to be generally similar in overall sensitivity to pure-tone stimuli across the audible frequency range [e.g., Stebbins (1973), but also see Brown and Waser (1984); Owren *et al.* (1988)], including chacma baboons tested by Hienz *et al.* (1982). In comparison to humans, these animals are somewhat less sensitive to frequencies below 500 Hz, significantly more sensitive to frequencies above 8000 Hz, and comparable (albeit slightly less sensitive) in the intermediate range. Thus, the acoustic energy characterized here should be readily perceptible to our subjects.

While macaques and some other species have been found to exhibit six to ten times higher frequency-difference thresholds than humans when both are tested with intermediate-frequency, pure-tone stimuli (see, for instance, Sinnott and Brown, 1993; Sinnott *et al.*, 1985; Sinnott *et al.*, 1987a), the degree of variation in F_0 values between individual animals found here is greater than expected difference threshold values would be. Monkeys have also been found to be quite sensitive to variation in duration, at least with respect to brief sounds (Sinnott *et al.*, 1987b). Finally, based on work by Sommers *et al.* (1992), one can expect that frequency variation within and between the formants demonstrated in the grunts is perceptually salient to the baboons (see Table I). In this study, Japanese macaques tested with single- and multi-formant tone complexes showed formant-frequency discrimination abilities at 500 Hz (e.g., 12.5 Hz) and 1.4 kHz (e.g., 22.8 Hz) that were comparable to those of humans. This finding is consistent with the demonstrations, noted earlier, that both chacma baboons (Hienz and Brady, 1988) and Japanese macaques (Sinnott, 1989) readily discriminate among typical American-English vowel sounds. Both species perform least well when presented with pairs of sounds that human listeners are most likely to confuse, while doing best when the vowel pairs are the ones most readily distinguished by humans. Although direct evidence is not available, one can expect as an associated capability that

baboons are also sensitive to variation in formant amplitude and general characteristics of spectral shape, like tilt.

C. Acoustic variation related to social context

When grunting to an infant or to a mother with an infant, all adult females produced longer call bouts (approximately six grunts per bout) than when grunting while on a move (approximately two grunts per bout). This result suggests that, from the baboons' perspective, these two social contexts were markedly different. However, when we tested for acoustic differences in grunts produced in the two circumstances, there was little evidence of systematic acoustic variation. While ANOVA-based tests revealed statistically significant differences between Infant and Move contexts for a number of variables, concomitant, nonsystematic variation was also likely to be found among individual callers.

In only two cases, namely variables $F2f$ and LPC_Tilt, did the animals in the sample show possibly persuasive consistency in the direction of statistically significant context-related mean differences. These two measures were also the only variables found to be useful in discriminant-function-based call classification. Proportionately, classification accuracy was significantly lower by context and relative dominance rank than by individual caller. Interpreting both LPC_Tilt and the relative amplitude variables broadly, as measures of overall spectral shape, the results indicated a more steeply tilted spectrum when calls were produced in the Infant context and a shallower spectrum when vocalizations occurred in the Move context. Testing Receiver-Rank, $F2f$ was found to be the only important predictor variable in discriminant-function analysis. Due to an overall lack of correlation between tilt-related and $F2f$ values, the two variables may represent functionally independent aspects of social communication in these calls. Overall, the former appears most likely to be related to behavioral context regardless of caller identity or relative rank of the interactants, while the latter is more clearly associated with the identity-related attributes of sender, receiver, or both.

D. Recording-related considerations

As the recordings were made in the field from free-ranging subjects, it is appropriate to consider whether inadvertent but inevitable variation in recording conditions may have affected the results for these often quiet, usually low-arousal calls. On the one hand, as the study animals were part of a baboon group that has been under continuous observation for many years, calls were routinely recorded from close quarters and were therefore of very high quality by the standards of primate fieldwork (cf. Owren and Bernacki, 1988). On the other hand, at least two sources of recording-related variation may have differentially affected the analysis outcomes for various measures. These possibly confounding factors were the distance of the observer from the vocalizing animal and variation in the angle of the microphone relative to the direction of call emission. The former is important due to the greater relative attenuation of high-frequency energy over distance, while the latter is related to the relative directionality of high-frequency energy when broadcast from a

vocal-tract opening [e.g., Brown (1989); see Gerhardt (in press) for a general review of these issues in field recording]. In both cases, then, any circumstantial differences in distance or microphone angle could be expected to create corresponding variability in the relative energy of higher frequencies.

However, with the exception of overall tilt of the spectrum, the variables that were of greatest interest in the analyses (i.e., those related to absolute formant frequency, fundamental frequency, and call duration) are likely to have been quite robust to incidental variation in recording conditions. For instance, as transmission of formant-frequency information represented in a tonal complex depends only on preservation of the relative amplitudes of adjacent harmonics, this aspect of the calls is unlikely to have been affected by global, frequency-dependent effects. Likewise, both fundamental frequency and overall call duration characteristics are arguably unaffected by selective attenuation of high-frequency energy. Spectral tilt, in contrast, directly reflects the relative amplitude of high- and low-frequency energy and may therefore have been more vulnerable to recording-related variation.

IV. CONCLUSIONS

Overall, the data considered here demonstrate a significant acoustic resemblance between grunts produced by chacma baboons and prototypical, centrally located vowel sounds produced by humans. This resemblance is likely to be rooted in a general cross-species similarity in vocal folds and supralaryngeal vocal tract, whose actions are most similar for humans and baboons when the former produce sounds using relaxed, neutral articulator positioning. However, the formant-frequency characteristics of the grunts we measured also indicate that a uniform, straight tube does not provide an entirely adequate model for understanding the evident resonance effects associated with the vocal tracts of adult female baboons.

Differences in grunt acoustics were clearly related to the identity of particular calling animals. Based on statistical analysis of call features, probable asymmetries in the degree of volitional control over various aspects of vocal production, and the demonstrated perceptual salience of spectral sound features (in baboons and Japanese macaques), variability in formant-related attributes is likely to be a predominant source of biologically significant, acoustically based differentiation among grunts. Variation in spectral peaks, shown to be independent of variation in F_0 , was associated both with individual caller and the social context of the vocalization event. In the former case, individual variation in details of vocal tract anatomy appeared to be reflected in a variety of acoustic measures. In the latter case, frequency of the second formant and overall spectral tilt appear to be differentially associated with the Infant and Move contexts, and with the relative rank of the animal to whom the call is directed, respectively.

However, acoustic differences across social contexts were by no means as clear as were acoustic differences among individuals, and the context-based differences were not equally strong for all animals tested. In conjunction with the observation that spectral-tilt-related measures were likely

to be the most susceptible to incidental recording effects, the results indicate that while adult female baboons may have shown differences in the acoustic features of their grunts when they were vocalizing to individuals that were more dominant or more subordinate, they are most likely to have used acoustically distinct grunt types in the Infant and Move contexts. Research is currently underway to test this hypothesis through playback experiments conducted with the study group and to examine how the distance and angle of recording affect the acoustic features of tape-recorded calls.

ACKNOWLEDGMENTS

We thank the Office of the President and the Department of Wildlife and National Parks of the Republic of Botswana for permission to conduct research in the Moremi Game Reserve. We are grateful to W. Hamilton, R. Boyd, Mongabe Kgosiekae, and Mokopi Mokopi for logistical support, to Eric Bokenkamp for selecting, editing, and analyzing grunts, and to Bridget Sly for further assistance in acoustic analysis. Drawings shown in Fig. 1 were generously created and donated by Michael Graham. Interpretation of the results benefited from valuable discussions with Drew Rendall and JoAnne Bachorowski. Two anonymous reviewers and Tecumseh Fitch provided helpful comments on a previous version of this work. Portions of the current results were presented at the November 1995 meeting of the Acoustical Society of America. Recording of vocalizations was supported by grants from the National Geographic Society, the Research Foundation of the University of Pennsylvania, the Institute for Cognitive Science at the University of Pennsylvania, and National Institutes of Health Grant No. HD-29433. Equipment and funding for assistance in vocalization analysis was provided by the College of Liberal Arts and Sciences at the University of Colorado at Denver, by a Division of Sciences Grant at the University of Otago, and the Dean's Development Fund at Reed College. Correspondence concerning this article should be addressed to Michael J. Owren, Department of Psychology, Reed College, 3203 SE Woodstock Blvd., Portland, OR 97202. Electronic mail may be sent via the internet to michael.owren@reed.edu

- Andrew, R. J. (1976). "Use of formants in the grunts of baboons and other nonhuman primates," in *Origins and evolution of language and speech* (Special Issue), edited by S. R. Harnad, H. D. Steklis, and J. Lancaster, Ann. (N.Y.) Acad. Sci. **280**, 673–693.
- Baken, R. J. (1987). *Clinical Measurement of Speech and Voice* (College-Hill, Boston).
- Bolwig, N. (1959). "A study of the behaviour of the chacma baboon, *Papio ursinus*," *Behaviour* **14**, 136–163.
- Brown, C. H. (1989). "The measurement of vocal amplitude and vocal radiation pattern in blue monkeys and grey-cheeked mangabeys," *Bioacoustics* **1**, 253–271.
- Brown, C. H., and Cannito, M. P. (1995). "Modes of vocal variation in Sykes's monkey (*Cercopithecus albogularis*) squeals," *J. Comp. Psych.* **109**, 398–415.
- Brown, C. H., Gomez, R., and Waser, P. M. (1995). "Old World monkey vocalizations: Adaptation to the local habitat?" *Animal Beh.* **50**, 945–961.
- Brown, C. H., and Waser, P. M. (1984). "Hearing and communication in blue monkeys (*Cercopithecus mitis*)," *Animal Beh.* **32**, 66–75.
- Bulger, J., and Hamilton, W. J. (1988). "Inbreeding and reproductive success in a natural chacma baboon, *Papio cynocephalus ursinus*, population," *Animal Beh.* **36**, 574–578.

- Byrne, R. W. (1981). "Distance vocalizations of Guinea baboons (*Papio papio*) in Senegal: An analysis of function," *Behaviour* **78**, 283–313.
- Cheney, D. L., and Seyfarth, R. M. (1982). "How vervets perceive their grunts: Field playback experiments," *Animal Beh.* **30**, 739–751.
- Cheney, D. L., and Seyfarth, R. M. (in press). "Reconciliatory grunts by dominant female baboons influence victims behaviour," *Animal Beh.*
- Cheney, D. L., Seyfarth, R. M., and Palombit, R. M. (1996). "The function and mechanisms underlying baboon 'contact' barks," *Animal Beh.* **52**, 507–518.
- Cheney, D. L., Seyfarth, R. M., and Silk, J. B. (1995a). "The role of grunts in reconciling opponents and facilitating interactions among adult female baboons," *Animal Beh.* **50**, 249–257.
- Cheney, D. L., Seyfarth, R. M., and Silk, J. B. (1995b). "The responses of female baboons (*Papio cynocephalus ursinus*) to anomalous social interactions: Evidence for causal reasoning?" *J. Comp. Psych.* **109**, 134–141.
- Cheney, D. L., Seyfarth, R. M., and Smuts, B. (1986). "Social relationships and social cognition in nonhuman primates," *Science* **234**, 1361–1366.
- Chiba, T., and Kajiyama, J. (1941). *The vowel: Its Nature and Structure* (Tokyo-Kaiseikan, Tokyo).
- Clark, A., and Wrangham, R. (1993). "Acoustic analysis of wild chimpanzee pant hoots: Do Kibale forest chimpanzees have an acoustically distinct food arrival pant hoot?" *Am. J. Prim.* **31**, 99–109.
- Elliot, D. G. (1913). *A Review of the Primates, Vol. 2* (American Museum of Natural History, New York).
- Fant, G. (1960). *The Acoustic Theory of Speech Production* (Mouton, The Hague).
- Fitch, W. T., and Hauser, M. D. (1995). "Vocal production in nonhuman primates: Acoustics, physiology, and functional constraints on 'honest' advertisement," *Am. J. Prim.* **37**, 191–219.
- Gautier, J.-P., and Gautier, A. (1977). "Communication in Old World monkeys," in *How Animals Communicate*, edited by T. A. Sebeok (Indiana U.P., Bloomington, IN), pp. 890–964.
- Geist, F. D. (1933). "Nasal cavity, larynx, mouth and pharynx," in *The Anatomy of the Rhesus Monkey*, edited by C. G. Hartman and W. L. Straus, Jr. (Hafner, New York), pp. 189–209.
- Gerhardt, C. (in press). "Acoustic signals of animals: Recordings, field measurements, analysis, and description," in *Animal Acoustic Communication: Sound Analysis and Research Methods*, edited by S. L. Hopp, M. J. Owren, and C. S. Evans (Springer-Verlag, Heidelberg).
- Gouzoules, H., and Gouzoules, S. (1990). "Matrilineal signatures in the recruitment screams of pigtail macaques, *Macaca nemestrina*," *Behaviour* **115**, 327–347.
- Gouzoules, H., Gouzoules, S., and Marler, P. (1986). "Vocal communication: A vehicle for studying social relationships," in *The Cayo Santiago Macaques: History, Behavior, and Biology*, edited by R. G. Rawlins and M. J. Kessler (SUNY, Albany, New York), pp. 111–129.
- Green, S. (1975). "Variation of vocal pattern with social situation in the Japanese monkey (*Macaca fuscata*): A field study," in *Primate Behavior, Vol. 4*, edited by L. A. Rosenblum (Academic, New York), pp. 1–102.
- Grimm, R. J. (1967). "Catalogue of sounds of the pigtailed macaque (*Macaca nemestrina*)," *J. Zool. Soc. London* **152**, 361–373.
- Hamilton, W. D. (1964). "The genetical evolution of social behavior. I, II," *J. Theor. Biol.* **7**, 1–52.
- Hauser, M. D. (1992). "Articulatory and social factors influence the acoustic structure of rhesus monkey vocalizations: A learned mode of production?" *J. Acoust. Soc. Am.* **91**, 2175–2179.
- Hamilton III, W. J., and Arrowood, P. C. (1978). "Copulatory vocalizations of chacma baboons (*Papio ursinus*), gibbons (*Hylobates hoolek*), and humans," *Science* **200**, 1405–1409.
- Hamilton, W. J., and Bulger, J. (1992). "Facultative expression of behavioral differences between one-male and multimale savanna baboon groups," *Am. J. Prim.* **28**, 61–71.
- Hauser, M. D., and Fowler, C. A. (1991). "Declination in fundamental frequency is not unique to human speech: Evidence from nonhuman primates," *J. Acoust. Soc. Am.* **91**, 363–369.
- Hauser, M. D., and Schön Ybarra, M. (1994). "The role of lip configuration in monkey vocalizations: Experiments using xylocaine as a nerve block," *Brain Lang.* **46**, 232–244.
- Hauser, M. D., Evans, C. S., and Marler, P. (1993). "The role of articulation in the production of rhesus monkey, *Macaca mulatta*, vocalizations," *Animal Beh.* **45**, 423–433.
- Hess, W. (1983). *Pitch Determination of Speech Signals* (Springer-Verlag, Heidelberg).
- Hienz, R. D., and Brady, J. V. (1988). "The acquisition of vowel discriminations by nonhuman primates," *J. Acoust. Soc. Am.* **84**, 186–194.
- Hienz, R. D., Turkkan, J. S., and Harris, A. H. (1982). "Pure tone thresholds in the yellow baboon (*Papio cynocephalus*)," *Hearing Res.* **8**, 71–75.
- Hill, W. C. O. (1972). *Evolutionary Biology of the Primates* (Academic, New York).
- Hill, W. C. O., and Booth, A. H. (1957). "Voice and larynx in African and Asiatic *Colobidae*," *Bombay Nat. Hist. Soc.* **54**, 309–321.
- Hilloowala, R. A. (1975). "Comparative anatomical study of the primate hyoid apparatus," *Am. J. Anat.* **142**, 367–385.
- Hintze, J. (1995). *NCSS 6.0 User's Guide* (Number Cruncher Statistical Systems, Kaysville, UT).
- Hohmann, G. M. (1989). "Vocal communication of wild bonnet macaques (*Macaca radiata*)," *Primates* **30**, 325–345.
- Hohmann, G. M., and Herzog, M. O. (1985). "Vocal communication in lion-tailed macaques (*Macaca silenus*)," *Folia Primatol.* **45**, 148–178.
- Hopp, S. L., Owren, M. J., and Evans, C. S., Eds. (in press). *Animal Acoustic Communication: Sound Analysis and Research Methods* (Springer-Verlag, Heidelberg).
- Kelemen, G. (1938). "Comparative anatomical studies on the junction of larynx and resonant tubes," *Acta Oto-Laryngol.* **26**, 276–283.
- Kelemen, G. (1949). "The anatomical basis of phonation in the chimpanzee," *J. Morphol.* **82**, 229–256.
- Kelemen, G. (1969). "Anatomy of the larynx and the anatomical basis of vocal performance," *Chimpanzee* **1**, 165–186.
- Lieberman, P. (1968). "Primate vocalizations and human linguistic ability," *J. Acoust. Soc. Am.* **44**, 1574–1584.
- Lieberman, P. (1969). "On the acoustic analysis of primate vocalizations," *Behav. Res. Methods Instrum. Comput.* **5**, 169–174.
- Lieberman, P. (1975). *On the Origins of Language: An Introduction to the Evolution of Human Speech* (Macmillan, New York).
- Lieberman, P., and Blumstein, S. E. (1988). *Speech Physiology, Speech Production, and Acoustic Phonetics* (Cambridge U. P., Cambridge, MA).
- Lieberman, P., Crelin, E. S., and Klatt, D. H. (1972). "Phonetic ability and related anatomy of the newborn, adult human, Neanderthal man, and the chimpanzee," *Am. Anthropol.* **74**, 287–307.
- Lieberman, P., Klatt, D. H., and Wilson, W. A. (1969). "Vocal tract limitation on the vowel repertoires of rhesus monkey and other nonhuman primates," *Science* **164**, 1185–1187.
- Masataka, N., and Thierry, B. (1993). "Vocal communication of Tonkean macaques in confined environments," *Primates* **34**, 169–180.
- Milenkovic, P. (1987). "Least squares measures of voice perturbation," *J. Speech Hear. Res.* **30**, 529–538.
- Napier, J. R., and Napier, P. H. (1985). *The Natural History of the Primates* (MIT, Cambridge, MA).
- Negus V. E. (1949). *The Comparative Anatomy and Physiology of the Larynx* (Hafner, New York).
- O'Connell, S. M., and Cowlshaw, G. (1994). "Infanticide avoidance, sperm competition, and mate choice: The function of copulation calls in female baboons," *Animal Beh.* **48**, 687–694.
- O'Shaughnessy, D. (1987). *Speech Communication* (Addison-Wesley, Reading, MA).
- Owren, M. J. (1990). "Acoustic classification of alarm calls by vervets (*Cercopithecus aethiops*) and humans: II. Synthetic calls," *J. Comp. Psych.* **104**, 29–40.
- Owren, M. J., and Bernacki, R. H. (1988). "The acoustic features of alarm calls in vervet monkeys (*Cercopithecus aethiops*)," *J. Acoust. Soc. Am.* **5**, 1927–1935.
- Owren, M. J., and Bernacki, R. H. (in press). "Applying linear predictive coding (LPC) to frequency spectrum analysis of animal acoustic signals," in *Animal Acoustic Communication: Sound Analysis and Research Methods*, edited by S. L. Hopp, M. J. Owren, and C. S. Evans (Springer-Verlag, Heidelberg).
- Owren, M. J., and Casale, T. M. (1994). "Variations in fundamental frequency peak position in Japanese macaque (*Macaca fuscata*) 'coo' calls," *J. Comp. Psych.* **108**, 291–297.
- Owren, M. J., and Linker, C. D. (1995). "Some analysis methods that may be useful to acoustic primatologists," in *Current Topics in Primate Vocal Communication*, edited by E. Zimmermann, J. D. Newman, and U. Jürgens (Plenum, New York), pp. 1–27.
- Owren, M. J., Linker, C. D., and Rowe, M. P. (1993). "Acoustic features of tonal 'grunt' calls in baboons," *Acoustical Society of America biannual meeting*, Denver, CO.
- Owren, M. J., Seyfarth, R. M., and Cheney, D. L. (1995). "Acoustic indices of production mechanisms underlying tonal 'grunt' calls in baboons,"

- Acoustical Society of America biannual meeting, St. Louis, MO.
- Owren, M. J., Dieter, J. A., Seyfarth, R. M., and Cheney, D. L. (1992). "'Food' calls produced by adult female rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques, their normally-raised offspring, and offspring cross-fostered between species," *Behaviour* **120**, 218–231.
- Owren, M. J., Hopp, S. L., Sinnott, J. M., and Petersen, M. R. (1988). "Absolute auditory thresholds in three Old World monkey species (*Cercopithecus aethiops*, *C. neglectus*, *Macaca fuscata*) and humans," *J. Comp. Psych.* **102**, 99–107.
- Peters, E. H. (1983). "Vocal communication in an introduced colony of feral rhesus monkey (*Macaca mulatta*)," Doctoral Dissertation, University of Florida, Gainesville.
- Peterson, G. E., and Barney, H. L. (1952). "Control methods used in a study of the identification of vowels," *J. Acoust. Soc. Am.* **24**, 175–184.
- Rendall, D., Owren, M. J., and Rodman, P. S., "The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalizations," *J. Acoust. Soc. Am.* (in review).
- Rendall, D., Rodman, P. S., and Emond, R. E. (1996). "Vocal recognition of individuals and kin in free-ranging rhesus monkeys," *Animal Beh.* **51**, 1007–1015.
- Richman, B. (1976). "Some vocal distinctive features used by gelada monkeys," *J. Acoust. Soc. Am.* **60**, 718–724.
- Rubin, P., and Vatikiotis-Bateson, E. (in press). "Measuring and modeling speech production," in *Animal Acoustic Communication: Sound Analysis and Research Methods*, edited by S. L. Hopp, M. J. Owren, and C. S. Evans (Springer-Verlag, Heidelberg).
- Schön Ybarra, M. (1995). "A comparative approach to the nonhuman primate vocal tract: Implications for sound production," in *Current Topics in Primate Vocal Communication*, edited by E. Zimmermann, J. D. Newman, and U. Jürgens (Plenum, New York), pp. 185–198.
- Seyfarth, R. M., and Cheney, D. L. (1997). "Vocal behavior in chacma baboons," (unpublished data).
- Seyfarth, R. M., and Cheney, D. L. (1984). "The acoustic features of vervet monkey grunts," *J. Acoust. Soc. Am.* **75**, 1623–1628.
- Seyfarth, R. M., Cheney, D. L., Harcourt, A. H., and Stewart, K. (1994). "The acoustic features of gorilla double-grunts and their relation to behavior," *Am. J. Prim.* **33**, 31–50.
- Sinnott, J. M. (1989). "Detection and discrimination of synthetic English vowels by Old World monkeys (*Cercopithecus*, *Macaca*) and humans," *J. Acoust. Soc. Am.* **86**, 557–565.
- Sinnott, J. M., and Brown, C. H. (1993). "Effects of varying signal and noise levels on pure-tone frequency discrimination in humans and monkeys," *J. Acoust. Soc. Am.* **93**, 1535–1540.
- Sinnott, J. M., Owren, M. J., and Petersen, M. R. (1987a). "Auditory frequency discrimination in primates: Species differences (*Cercopithecus*, *Macaca*, *Homo*) and humans," *J. Comp. Psych.* **101**, 126–131.
- Sinnott, J. M., Owren, M. J., and Petersen, M. R. (1987b). "Auditory duration discrimination in Old World monkeys (*Macaca*, *Cercopithecus*) and humans," *J. Acoust. Soc. Am.* **82**, 465–470.
- Sinnott, J. M., Petersen, M. R., and Hopp, S. L. (1985). "Frequency and intensity discrimination in humans and monkeys," *J. Acoust. Soc. Am.* **78**, 1977–1985.
- Smuts, B. B., Cheney, D. L., Seyfarth, R. M., Wrangham, R. W., and Struhsaker, T. T. (1987). *Primate Societies* (University of Chicago, Chicago).
- Snowdon, C. T. (1986). "Vocal communication," in *Comparative Primate Biology, Volume 2A: Behavior, Conservation, and Ecology*, edited by G. Mitchell and J. Erwin (Liss, New York), pp. 495–530.
- Sommers, M. S., Moody, D. B., and Prosen, C. A. (1992). "Formant frequency discrimination by Japanese macaques (*Macaca fuscata*)," *J. Acoust. Soc. Am.* **91**, 3499–3510.
- Stebbins, W. C. (1973). "Hearing of Old World monkeys (*Cercopitheciinae*)," *Am. J. Phys. Anthropol.* **38**, 357–364.
- Stevens, K. N., and House, A. S. (1955). "Development of a quantitative description of vowel articulation," *J. Acoust. Soc. Am.* **27**, 484–493.
- Titze, I. R. (1994). *Principles of Voice Production* (Prentice-Hall, Englewood Cliffs, NJ).
- Trivers, R. (1971). "The evolution of reciprocal altruism," *Q. Rev. Biol.* **46**, 35–57.
- Waser, P. M. (1982). "The evolution of male loud calls in mangabeys and baboons," in *Primate Communication*, edited by C. T. Snowdon, C. H. Brown, and M. R. Petersen (Cambridge U. P., New York), pp. 117–143.
- Zhinkin, N. I. (1963). "An application of the theory of algorithms to the study of animal speech," in *Acoustic Behaviour of Animals*, edited by R. G. Busnel (Elsevier, Amsterdam), pp. 132–180.
- Zimmermann, E. (1981). "First record of ultrasound in two prosimian species," *Naturwissenschaften* **68**, 531.

A backpropagation network model of the monaural localization information available in the bat echolocation system

Janine M. Wotton and Rick L. Jenison^{a)}

Departments of Neurophysiology and Psychology, University of Wisconsin, Madison, Wisconsin 53706

(Received 29 July 1996; accepted for publication 10 December 1996)

The information echolocating bats receive is a combination of the properties of the sound they emit and the sound they receive at the eardrum. Convolving the emission and the external ear transfer functions produces the full spectral information contained in the echolocation combination. Spatially dependent changes in the magnitude spectra of the emission, external ear transfer functions, and the echolocation combination of *Eptesicus fuscus* could provide localization information to the bat. Principal component analysis was used to reduce the dimensionality of these complex spectral data sets. The first eight principal component weights were normalized, rotated, and used as the input to a backpropagation network model which examined the relative directionality of the emission, ear, and the echolocation combination. The model was able to localize more accurately when provided with the directional information of the echolocation combination compared to either the emission or ear information alone. © 1997 Acoustical Society of America. [S0001-4966(97)00605-X]

PACS numbers: 43.80.Ka, 43.80.Lb, 43.66.Qp, 43.60.Lq [FD]

INTRODUCTION

A. Echolocation combination

An echolocating bat emits a high-frequency sound that is reflected from objects and returns to the bat's ears as an echo. The bat uses the information contained in the echoes to navigate and locate objects in its environment. The spectral information used by the bat is therefore produced by combining the properties of the sound emission and sound reception (Fuzessery *et al.*, 1992; Shimozawa *et al.*, 1974; Schnitzler and Grinnell, 1977). In all the species of bats studied so far, the echolocation combination appears to be more directional than either the ear or emission alone (*Myotis*: Shimozawa *et al.*, 1974; *Pteronotus*: Fuzessery *et al.*, 1992; *Eptesicus*: Wotton *et al.*, 1997). It has therefore been hypothesized that localization performance is improved by combining the ear and emission information compared to using either element alone. We aim to test this hypothesis using the emission, ear, and combination information of the big brown bat, *Eptesicus fuscus*.

There are spatially dependent changes in the spectra of the emission and the external ear and consequently also in the combination echolocation pattern of *Eptesicus* (Hartley and Suthers, 1989; Jen and Chen, 1988; Wotton *et al.*, 1995; Wotton *et al.*, 1997). These changes in spectra provide directional information which the bat could use to localize targets. Backpropagation models have been used successfully to examine whether similar spectral changes could be used by cats to localize sound sources (Neti *et al.*, 1992) and to study whether the information contained in sonar/biosonar signals enables discrimination between different objects (Gorman and Sejnowski, 1988; dolphin: e.g., Roitblat *et al.*, 1989, 1995; Moore *et al.*, 1991; bats: Dror *et al.*, 1995). These

models can be thought of as near optimal observers, they do not necessarily reflect the manner in which the brain processes information but they can indicate if the available information is adequate for the network to learn and perform the required task. Here, we use a backpropagation model as a near optimal observer, to test how well a network model can localize when given either emission, ear, or echolocation combination information. The architecture of the model is not intended to reflect the physiology of the bat's auditory or vocal systems, instead the purpose is to provide a common platform to examine the information contained in each data set. Thus, the relative localization performance of the ear, emission, and the echolocation combination can be directly tested. Comparisons are made of the localization performance of the network under each of these three conditions to determine if the echolocation combination information really does allow the model to localize more accurately.

B. Principal component analysis

Highly variable data such as acoustical transfer functions (Kistler and Wightman, 1992; Middlebrooks and Green, 1992), speech (Li *et al.*, 1969; Plomp *et al.*, 1967; Pols *et al.*, 1969; Zahorian and Rothenberg, 1981), and evoked potentials (e.g., Donchin, 1966; Suter, 1970; Wood and McCarthy, 1984) include a large amount of redundant information which can be reduced using principal component analysis (PCA). PCA is a statistical procedure which reduces the dimensionality of the data while retaining as much of the information bearing variation as possible. PCA uses a small number of vectors and weighting coefficients to describe a complex spectral data set, thus increasing the efficiency of subsequent pattern recognition. For example, Kistler and Wightman (1992) used PCA to model the head related transfer functions (HRTF) of humans. Their data set was decomposed into a set of basis functions (eigenvectors) which had characteristic spectral shapes. Each of the original HRTF

^{a)}Address reprint requests to: Rick Jenison, Department of Psychology, University of Wisconsin, 1202 W. Johnson St., Madison, WI 53706.

spectra could be approximated by a unique weighted sum of these basis functions. The weights therefore determined the proportional contribution of each basis function to each HRTF spectrum.

The axes of the basis functions of the PCA can be rotated to further reduce the complexity of the variables (Harman, 1960). A widely used method of rotation is the Varimax rotation (Kaiser, 1958) which maximizes the positive and negative loadings on some variables and makes other loadings close to zero. If a single component has high loadings for a set of variables and all the remaining variables have zero or negligible loadings on this component then the structure of the data is simplified (Harman, 1960). Rotation of the basis functions can be useful because it tends to separate out independent clusters of variables (Reyment and Joreskog, 1993). Varimax rotation normalizes the factor loadings and groups interdependent variables together (Reyment and Joreskog, 1993).

We use PCA to reduce the dimensionality of the spectral data and then use the rotated weights as the input to a back-propagation model. The structure of the spectral data is therefore both reduced and simplified. The inputs are normalized and the number of inputs needed to capture the main features of the spectra is decreased, which in turn improves the efficiency of the model (Hertz *et al.*, 1991; Sanger, 1989).

I. METHODS

A. Echolocation information

The directional measurements of the emission (Hartley and Suthers, 1989) and external ear transfer functions (Wotton *et al.*, 1995) of *Eptesicus* were taken at 10° intervals from 0° to 70° in azimuth and +60° to -70° in elevation, for a total of 112 spatial locations. The echolocation combination was calculated by convolving the emission and ear spectra at the corresponding spatial locations (Wotton *et al.*, 1997). The frequency range examined was from 20 to 80 kHz, sampled at 1 kHz increments.

The ear, emission, and echolocation combination spectral data were represented in three separate 112 by 61 matrices with a different position in each of the rows ($n=112$) and a different frequency in each column (61 samples in increments of 1 kHz).

B. Principal component analysis

A small set of basis functions was derived from the spectral data set and used to calculate the principal components. The spectral data were represented in a 112 by 61 matrix \mathbf{X} . The singular value decomposition of the matrix \mathbf{X} was calculated to produce three matrices such that,

$$\mathbf{X} = \mathbf{U} * \mathbf{S} * \mathbf{V}'$$

where the columns of \mathbf{V} are the eigenvectors of $\mathbf{X}' * \mathbf{X}$. These orthogonal eigenvectors are the basis functions used in the PCA. The square of the diagonal of \mathbf{S} shows the variance accounted for by each of the eigenvectors. The eigenvectors are ordered by the amount of variance explained, with the first eigenvector accounting for most of the variance and the

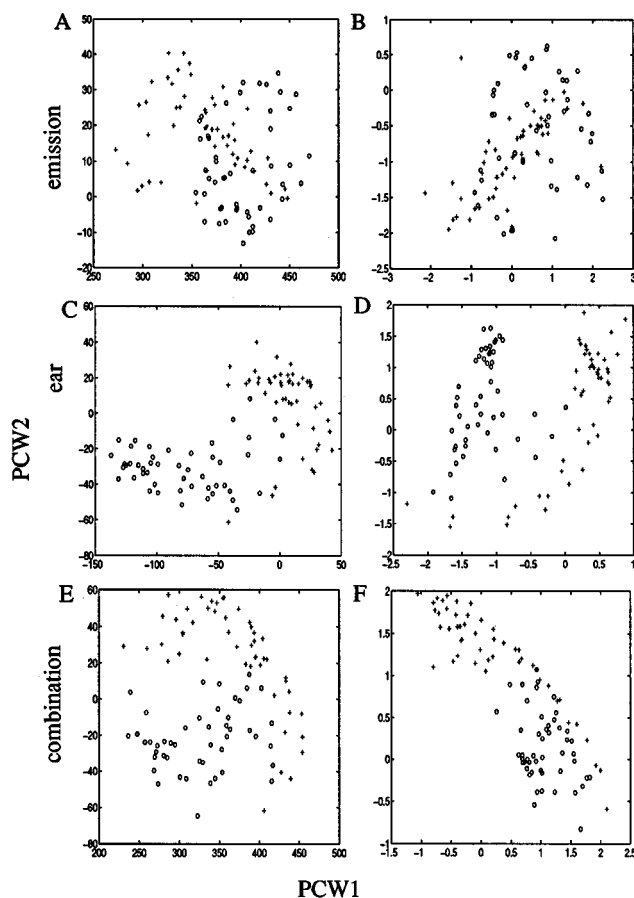


FIG. 1. Comparison of original and rotated principal component weights. PCW1 is plotted versus PCW2 for the emission (A) and (B), the ear (C) and (D), and the echolocation combination (E) and (F). Crosses represent weights of positions at and above the horizon (0° to +60°) and circles represent positions below the horizon (-10° to -70°).

remaining eigenvectors accounting for systematically diminishing amounts of variance.

The weights representing the contribution of each basis function to the spectral data were calculated by

$$\mathbf{W} = \mathbf{X} * \mathbf{V},$$

where each column of the \mathbf{W} matrix contains the weights associated with a single basis function. The first column of \mathbf{W} contains the weights associated with the first basis function, the second column the weights associated with the second basis function, and so on. The weights associated with basis function 1 are designated PCW1, those of basis function 2 are PCW2, and so on.

Although the eigenvectors are mathematically unique, they are only one of a set of factors that will describe the data equally well, with respect to the total variance explained (Harman, 1960; Reyment and Joreskog, 1993). The entire set of eigenvectors can be rotated to provide a different but equally valid solution. Varimax rotation maintains the orthogonality of the eigenvectors and normalizes the principal components. The total variance explained remains constant with rotation but the distribution of variance between components is altered.

The loadings for the first eight principal components were calculated and rotated using the algorithm of Harman (1960) in a routine described by Reyment and Joreskog

(1993). Figure 1 plots PCW1 against PCW2 for the original unrotated case [Fig. 1(A), (C), and (E)] and after rotation [Fig. 1(B), (D), and (F)] for the emission [Fig. 1(A) and (B)], the ear [Fig. 1(C) and (D)] and the echolocation combination [Fig. 1(E) and (F)]. The cross symbols represent weights associated with elevations from 0° to $+60^\circ$ and the circles represent elevations from -10° to -70° . This labeling was chosen *post hoc* to illustrate one aspect of the emergent structure. It is clear that after rotation the weights are separated into two distinct clusters, representing those positions mostly above the horizon and those mostly below, for both the ear and the combination conditions. The rotation accentuates the differences between PCW1 and PCW2 forcing a separation on even this rough category of position. Note, the magnitude of the weights are much more similar after normalization and rotation.

The spectral data can be reconstructed, as matrix \mathbf{Y} , using the basis functions and weights such that,

$$\mathbf{Y} = \mathbf{W} * \mathbf{V}',$$

where $\mathbf{Y} = \mathbf{X}$ if all the weighted basis functions are used. However, the purpose of PCA analysis here is to reduce the number of dimensions used to represent the data; so in practice \mathbf{Y} is calculated using fewer weighted basis functions and is only an approximation of the original data. The number of basis functions required to adequately approximate the data depends on the amount of redundancy and structure in the data.

Deciding how many principal components to use is usually based on how much variance is to be explained (Kistler and Wightman, 1992). The first eight principal components accounting for 98% (ear data) or 99% (emission, combination data) of the variance were used to represent the spectral data. Figure 2 shows how closely these reconstructed spectra (solid line) approximate the original spectra (dashed line) at one location (elevation -10° at azimuth 0°) for the emission [Fig. 2(A)], the ear [Fig. 2(B)] and the echolocation combination [Fig. 2(C)].

C. Network model

Figure 3 shows all levels of the model, the first level is the original spectra and has 61 frequency inputs (not all shown in the figure). These data are then preprocessed, using PCA and Varimax rotation, to reduce the dimensionality and simplify the relationships between the weights. The data are reduced to 8 inputs using PCA and then these inputs are normalized and rotated. These stages of the model are linear and produced the eight inputs to the nonlinear backpropagation model.

The network was a fully connected, three layer model, with an input layer of eight nodes, a hidden layer of five nodes, and an output layer of two nodes. The number of hidden units was chosen to optimize the generalization of the network to novel stimuli. The input to the model was a vector composed of the first eight (rotated) PC weights (PCW1 to PCW8), each input node received one PC weight. Input vectors from 112 different spatial locations were available but only 102 input vectors were used in the supervised training of the network. Ten input vectors were intentionally left

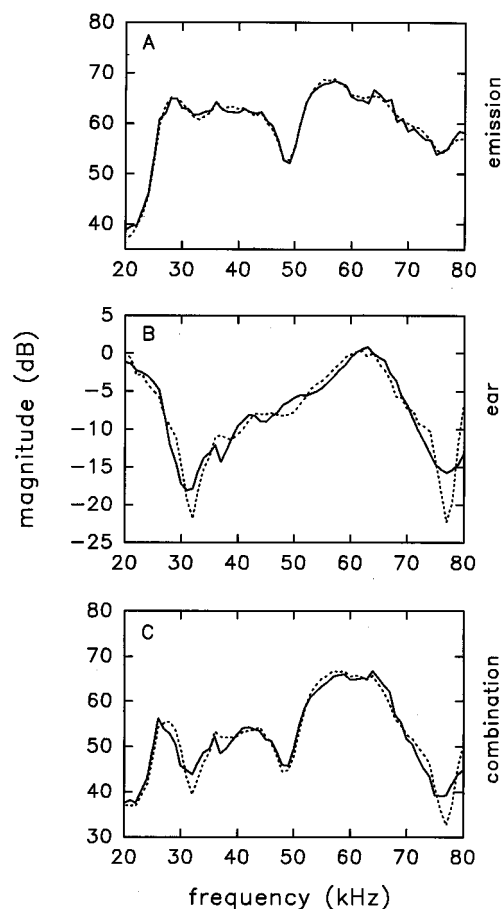


FIG. 2. The original spectra compared to the reconstructed spectra using eight weighted basis functions for the emission (A), the ear (B) and the echolocation combination (C).

out of the training data set. The two output nodes represented azimuth (0° to 70° in 10° steps) and elevation ($+60^\circ$ to -70° in 10° steps).

There are interconnection weights between the nodes at one layer of the network and those at the next upward layer. Initially all weights between the network model nodes have small random values and the output is computed using these values. During the training this output is compared to the desired output and corrections are made to the weights to minimize the error (Hertz *et al.*, 1991; Rumelhart *et al.*, 1986). The weights are adjusted by “backpropagating” the error correction from the output to the input (Rumelhart *et al.*, 1986). The “squashing function” between the input of the network and hidden layer was the hyperbolic tangent and between the hidden layer and the output a linear function was used. Thus, the only nonlinear part of the entire model (see Fig. 3), is the hyperbolic tangent function between the network input and hidden nodes. The learning rate was adaptive and a momentum constant (0.9) was included. The model was given 2000 iterations to adjust the weights and minimize the output error. The weights were then saved and used to test how well the model generalized from the training. The network was trained separately, using the emission, ear, and echolocation combination principal component information as different input sets. The output set was main-

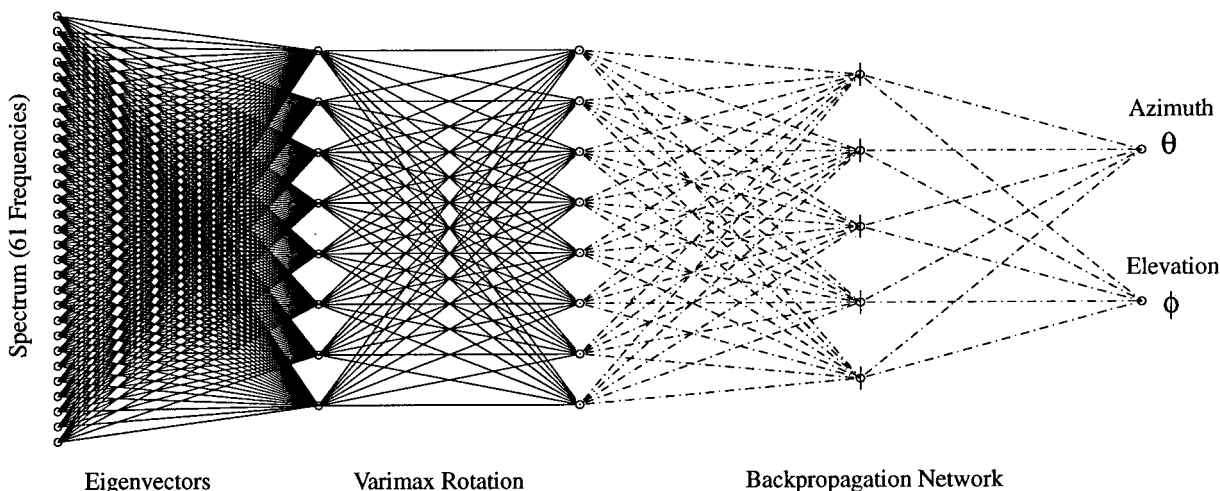


FIG. 3. Schematic of the model. The first level shows the original spectra, then PCA reduces the inputs from 61 to 8. Rotation normalizes the magnitude of the inputs but maintains the same number of inputs. These eight input nodes are used for the network and connected to the five hidden units. The five hidden units are in turn connected to the two output nodes. Plain circle nodes indicate a linear transformation between levels and nodes with circles and a horizontal bar indicate a nonlinear transformation. Solid lines show that the parameters have been solved directly and are fixed. The dotted line indicates that the parameters are adapted by backpropagation.

tained as the same set of spatial locations for all three conditions.

The network model was tested using the 10 novel input vectors. These vectors were chosen to be evenly distributed among all the spatial locations ($n=112$). The output location calculated by the network for each of these 10 inputs was compared to the real location to determine the localization performance of the model. The absolute difference between the network output and the real location was calculated in degrees. The absolute distance for a given azimuth angle varies with elevation due to the shape of the sphere, that is, the distance between two points at the equator (0° elevation) is greater than the distance between two points separated by the same angle near the pole (e.g., $+80^\circ$ elevation). To compensate for this, we weighted the azimuth error as a function of the elevation of the real location to obtain the solid angle in azimuth. If this adjustment is not made then azimuth errors would be inflated at the more extreme elevations. The errors were calculated as:

$$\begin{aligned} \text{elevation} &= (|\text{real elevation} \\ &\quad - \text{output elevation observed}|), \\ \text{azimuth} &= (|\text{real azimuth} \\ &\quad - \text{output azimuth observed}|) * \\ &\quad \times \sin(\text{real elevation}). \end{aligned}$$

The errors were calculated for each pair of real and observed locations and then averaged over the 10 pairs to provide indicators of localization accuracy in azimuth and elevation.

The training and testing were repeated 19 times using different test positions to examine the reliability of the model. A different test set was created by randomly selecting ten locations without replacement from the same original position set ($n=112$). This selection was repeated 19 times to produce 19 different test and training sets (each training set was composed of the remaining 102 input vectors). The

training was repeated for each of these new training sets and the weights saved. The model was then tested on the appropriate novel test set and the localization accuracy was calculated for each of the 19 test sets. Thus, a total of 20 different training sets were saved and then tested for each of the three input conditions: the emission, the ear, and echolocation combination information.

All analyses were conducted separately for the emission data, the external ear transfer function data, and the echolocation combination using Matlab (The Math Works, Inc.).

II. RESULTS

A. The network model

Figure 4 shows the real locations (each circle is a location) of the 102 input vectors in the first training set plotted

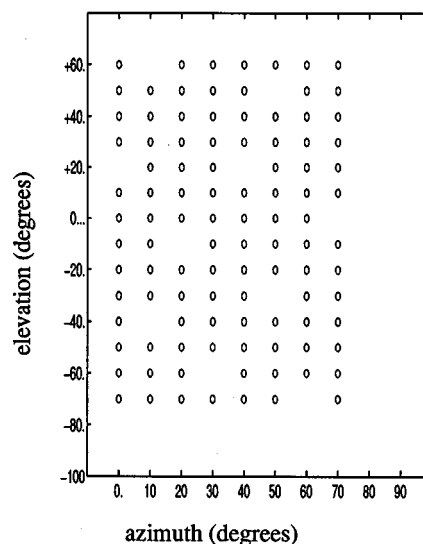


FIG. 4. The real locations of the input vectors used in the first training set. Each circle represents a discrete location and the azimuth range is from 0° to 70° in 10° steps, the elevation range is $+60^\circ$ to -70° in 10° steps.

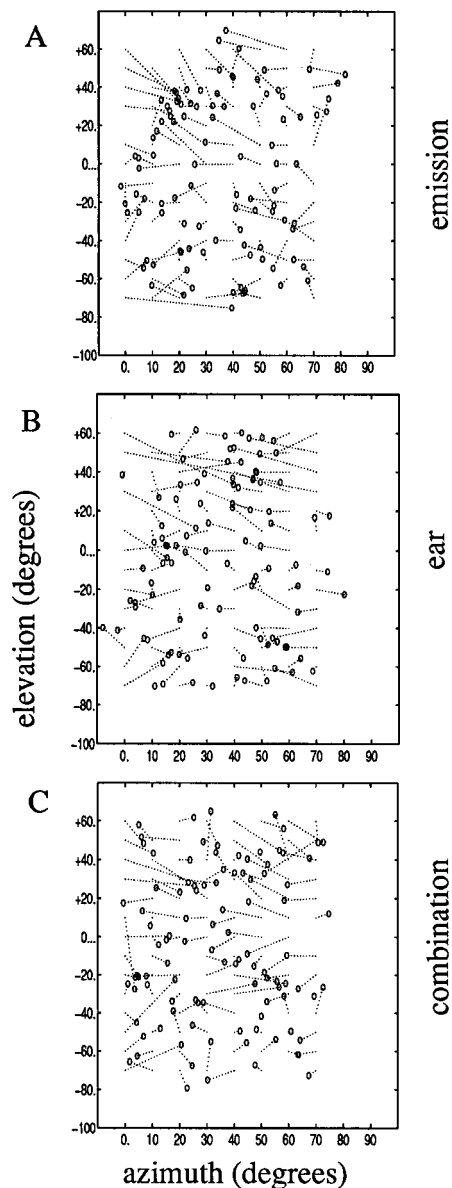


FIG. 5. The locations assigned to the input vectors by the network after training was completed. The circle represents the location assigned by the network and the end point of the line shows the real location for each input vector. The network was trained separately for the emission (A), the ear (B), and the echolocation combination (C) input sets.

as azimuth (0° to 70° in 10° steps) against elevation ($+60^{\circ}$ to -70° in 10° steps). The full set of 112 locations would reveal that each circle is always 10° from adjacent locations in both azimuth and elevation. However, in the training set of 102 locations there are gaps in each column (azimuth) which indicate the location of the 10 input vectors removed for the test set. This first test set was selected to produce an even distribution of locations.

If the network learns to associate the input vectors of the training set with their locations then the output should resemble the distribution of the real locations (shown in Fig. 4). Figure 5 shows the output of the network after training on the emission [Fig. 5(A)], the ear [Fig. 5(B)], and the echolocation combination [Fig. 5(C)]. The circle indicates the output location that the network predicted for a given input

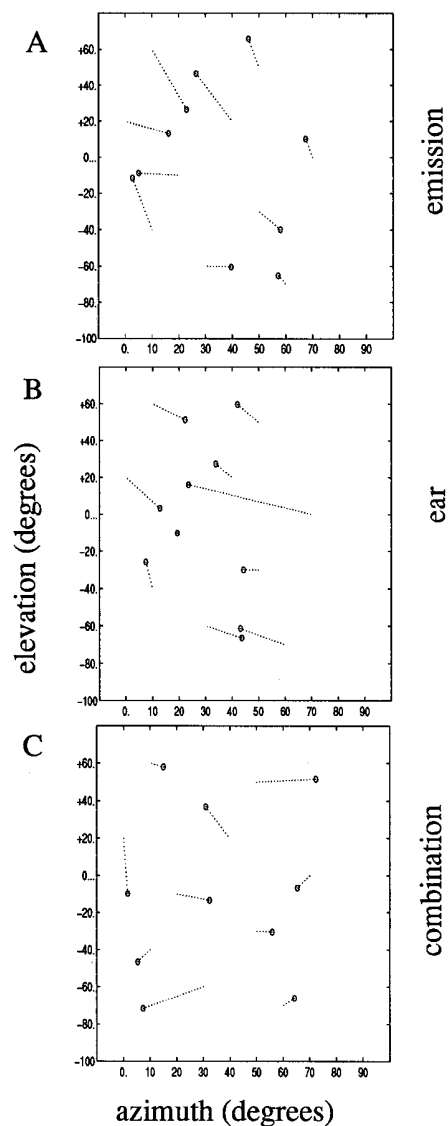


FIG. 6. The locations that the network assigned to the novel test inputs. The circle represents the location assigned by the network and the end point of the line shows the real location for each input vector. The network was tested separately on the saved weights for the emission (A), the ear (B), and the echolocation combination (C) input sets.

vector and the end point of the line shows the real location for that input vector. In all cases the output roughly resembles the distribution of real locations, clearly the output is not a straight line or a tiny clump of locations. The sum squared error at the end of training was very similar for each of these three conditions (5.4 emission, 5.6 ear, 5.3 echolocation combination) but the echolocation combination training output [Fig. 5(C)] appears to be a little more uniformly distributed than either the emission [Fig. 5(A)] or ear [Fig. 5(B)] training output.

B. Testing the network model

If the input vectors change in some systematic way with location, then the network should learn this and be able to generalize the information and localize the novel test inputs. Figure 6 shows the output of the first test set for the emission [Fig. 6(A)], the ear [Fig. 6(B)], and the echolocation combi-

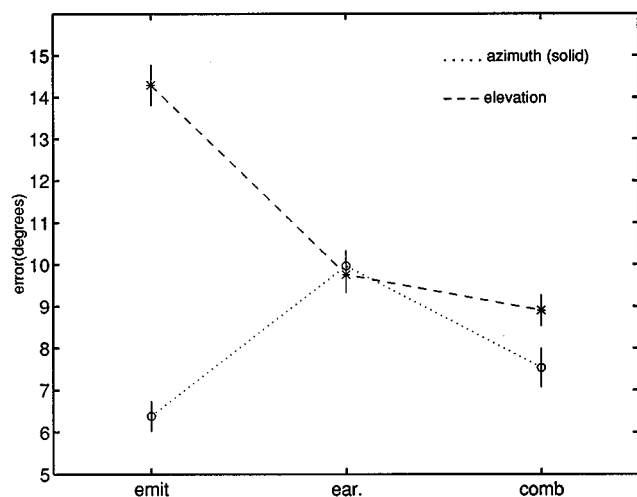


FIG. 7. The mean and standard error of the localization error in azimuth (solid angle) and elevation for each of the three conditions (emission, ear, combination) for the 20 network tests.

nation [Fig. 6(C)]. Again, the circle indicates the location predicted by the network and the end point of the line shows the real location. In all cases the network was able to localize some input vectors quite well and made greater errors on others. Overall, the echolocation combination network made smaller errors in both azimuth and elevation than either the emission or ear networks. For this set of test locations, the average azimuth error was 7.3° for the emission network, 9.6° for the ear, and 6.5° for the echolocation combination. The average elevation error was 13.7° for the emission network, 8.8° for the ear, and 8.2° for the echolocation combination.

The means and standard errors of the outputs from the 20 different test sets for the three conditions (emission, ear, and combination) for each of the two factors (azimuth and elevation) are shown in Fig. 7 and Table I. The azimuth position error is largest for the ear, but small and similar for the emission and echolocation combination, with the emission being slightly better. The mean elevation position error shows the same pattern as seen for the single test set, that is,

TABLE I. The descriptive statistics for the localization error in azimuth (solid angle) and elevation for each of the three conditions (emission, ear, combination) for the 20 network tests.

	Emission azimuth	Ear azimuth	Comb. azimuth	Emission elevation	Ear elevation	Comb. elevation
Mean localization error	6.38°	9.97°	7.53°	14.29°	9.76°	8.90°
Standard error of mean	0.36°	0.37°	0.47°	0.49°	0.44°	0.38°
Minimum localization error	3.24°	7.52°	4.39°	10.08°	5.19°	5.62°
Maximum localization error	8.76°	13.87°	13.25°	17.87°	13.87°	12.45°

TABLE II. The analysis of variance table. There are three conditions (emission, ear, and combination) and two factors (azimuth and elevation). Each cell has 20 entries. (d.f.: degrees of freedom; M.S.: mean square)

	d.f. effect	M.S. effect	d.f. error	M.S. error	F	p
Condition	2	49.61	114	3.57	13.89	<0.001
Azimuth/elevation	1	274.44	114	3.57	76.86	<0.001
Interaction	2	184.91	114	3.57	51.79	<0.001

the emission has the largest errors and the ear and combination are similar, with the combination being slightly better.

To examine the reliability of this result, a multivariate analysis of variance was conducted on the position errors produced by the outputs from each of the 20 different test sets. The results are shown in Table II. There were highly significant main effects of network condition (i.e., emission, ear, or combination), position factor (azimuth/elevation) and an interaction between the network condition and the position factor (azimuth or elevation).

The results of the *post hoc* tests of significance are shown in Table III. The significant condition effect was caused by significant differences in localization accuracy between the emission and the combination and between the ear and the combination. The combination network was more accurate (collapsed across azimuth and elevation) than either the ear or emission networks. There was no significant difference in localization accuracy between the emission and the ear networks.

The significant position factor difference (collapsed across emission, ear, and combination conditions) revealed that the networks performed significantly better in azimuth than in elevation.

The *post hoc* test of the interaction effect reveal several significant differences. For the emission, the elevation error was significantly greater than the azimuth error. However, there were no significant differences between the errors in azimuth and elevation for either the ear or the echolocation combination.

The ear azimuth error was significantly greater than either the emission or combination azimuth errors. There was no significant difference between the azimuth errors of the emission and combination.

The emission elevation error was significantly greater than all other errors, in either elevation or azimuth for all conditions. There was no significant difference between the ear and the combination elevation errors.

III. DISCUSSION

As suggested by the acoustic measurements (Shimozawa *et al.*, 1974; Fuzessery *et al.*, 1992; Wotton *et al.*, 1997) the echolocation combination provides more precise localization information than either the ear or emission alone. The network model's overall localization ability was better using the echolocation combination data. It seems likely that the ear and emission of bats have evolved together to increase the information available for localization. Our model observer does not test all the information which has influenced

TABLE III. The probabilities resulting from the Tukey HSD *post hoc* tests of significance for the main condition effect [Table III(A)], the position factor effect [Table III(B)] and the interaction [Table III(C)].

(A) Tukey HSD tests of main effect of condition

	Emission (10.33°)	Ear (9.87°)	Combination (8.22°)
Emission	...	0.514	<0.001
Ear	0.514	...	<0.001
Combination	<0.001	<0.001	...

(B) Tukey HSD tests of main effect of position

	Elevation (10.98°)
Azimuth (7.96°)	<0.001

(C) Tukey HSD tests of interaction

	Emission azimuth (6.38°)	Ear azimuth (9.97°)	Combin. azimuth (7.53°)	Emission elevation (14.29°)	Ear elevation (9.76°)	Combin. elevation (8.90°)
Em. az.	...	<0.001	0.398	<0.001	<0.001	0.001
Ear az.	<0.001	...	0.001	<0.001	0.999	0.477
Comb az.	0.398	0.001	...	<0.001	0.004	0.201
Em. el.	<0.001	<0.001	<0.001	...	<0.001	<0.001
Ear el.	<0.001	0.999	0.004	<0.001	...	0.703
Comb el.	0.001	0.477	0.201	<0.001	0.703	...

the coevolution of ear and emission because we have not included binaural information. However, the monaural information accounts for a large portion of the overall spectral variation, and in particular, the variation with elevation (Middlebrooks and Green, 1991). Unlike most mammals, bats are able to add spatial information by using a directional emission pattern, and in principle at least, both the bat and our model observer seem to be able to take advantage of this extra information.

The emission network performs significantly better in azimuth than elevation. The emission intensity drops systematically as location moves from the midline and presumably this is a useful azimuth cue. This cue maybe more useful to the network than the bat because at extreme azimuths the intensity of the signal is very weak and maybe undetectable (Hartley and Suthers, 1989; Simmons *et al.*, 1995). The spectral changes for the emission with elevation are much coarser than the changes seen for the ear (Wotton *et al.*, 1997) and seem to provide less localization information.

The ear network did not show a difference in its ability to localize in azimuth or elevation. The information provided to the ear network were the monaural pinna cues which are generally thought to provide largely elevation information (Gardner, 1973; Gardner and Gardner, 1973; Hebrank and Wright, 1974; Middlebrooks, 1992). Vertical localization can be performed well by human subjects under monaural or binaural conditions (Butler and Belendiuk, 1977; Fisher and Freedman, 1968). Human monaural localization vertical acuity is approximately 12° which is only slightly worse than a binaural acuity of 9° (Oldfield and Parker, 1986). However in azimuth, humans show an acuity of about 5° with binaural information but monaural localization is quite unreliable with errors as large as 40° (Oldfield and Parker, 1986).

Interaural intensity differences (IID) are usually available to bats as a cue for azimuth location (e.g., Grinnell, 1963; Fuzessery *et al.*, 1990; Fuzessery and Pollak, 1984). It

appears that bats are also able to localize better in azimuth with binaural information but no estimates of monaural and binaural azimuthal acuity are available. Griffin (1958) reported that when one ear was covered *Myotis lucifugus* and *Eptesicus* showed a pronounced loss of ability to avoid small obstacles, such as vertically strung wires, but that large objects such as walls were still detected. This indicates that the bats suffered a deterioration, but not a total loss, of azimuth acuity with the monaural information. Complete occlusion of one human ear produces a lateral shift in localization judgments (Gardner and Gardner, 1973) and a similar effect may occur for the bats in this experiment.

Bats do not have access to either the emission or ear information alone, the bat always receives the echolocation combination information. In both a discrimination experiment (Simmons *et al.*, 1983) and in a tracking-accuracy study (Masters *et al.*, 1985) *Eptesicus* displays an azimuth acuity of about 1.5°, which is obviously much better than the 7.5° localization acuity achieved by the echolocation combination network. There are several possible reasons for the superior performance of the bats in the discrimination experiment. In psychophysical experiments, discrimination performance and localization performance are tested with quite different paradigms, thus even if both measures of acuity were available for bats they could also be different. The network was trained on a fairly small set of input vectors spaced 10° apart and this may be the best resolution that could realistically be expected under these conditions. The bats had access to binaural information which probably accounts for some of this difference. Neural network models which include binaural information can localize very well in azimuth (e.g., Datum *et al.*, 1996; Neti *et al.*, 1992) and demonstrate an enhancement in performance compared to monaural models (Neti *et al.*, 1992). Although, it is assumed that bats use binaural cues to localize in azimuth, these cues may be restricted to a small range of azimuths (Simmons *et al.*, 1995;

Fuzessery *et al.*, 1992; Grinnell and Schnitzler, 1977; Wotton *et al.*, 1997). The IIDs of *Eptesicus* are quite linear with azimuth over only a limited range of about $\pm 30^\circ$ to 40° (Jen and Chen, 1988). The azimuth angle acuity of the bats was measured near the midline and thus they probably had access to useful binaural information. Generally, the azimuth acuity of mammals is best near the midline and deteriorates as the sound is more lateral (e.g., humans: Makous and Middlebrooks, 1990; cats: Heffner and Heffner, 1988). The echolocation combination network azimuth average of 7.5° is over the entire azimuth range 0° to 70° . It is not known what the average localization acuity over 70° of azimuth space would be for *Eptesicus*.

Clearly, the networks are able to localize in azimuth using monaural spectral information. Although binaural information may improve the bat's performance it is not obvious that this indicates primary reliance on IID information, rather than on binaural fine spectral information, which may be important to cats (Musicant *et al.*, 1990; Rice *et al.*, 1992). Simmons *et al.* (1995) showed that binaural intensity information obtained by combining IIDs and emission intensities results in very small ipsilateral azimuth intensity differences and large contralateral differences. On the ipsilateral side, the IIDs and the emission intensity differences produce opposite effects and therefore the combination intensity differences are smaller than either of these two components. As the monaural intensity ear cues also work in opposition to the emission intensity cues, it might be expected that if intensity difference is the main cue for azimuth, the combination network would perform poorly using only ipsilateral information. The emission network has access to large intensity difference cues and should perform better. However, the emission and echolocation combination network performance do not differ significantly.

The vertical acuity of *Eptesicus* near the horizon in discrimination experiments is about 3° or 4° (Lawrence and Simmons, 1982; Wotton and Simmons, in preparation). At low elevations (between -30° and -50°) the bat's performance deteriorated to about 6° to 8° and at high elevations the bats would not perform the acuity task even at the most extreme acuity tested (21°) (Wotton and Simmons, in preparation). The echolocation combination network performed the elevation localization at about 8.9° which is an average over the range $+60^\circ$ to -70° . The network model appears to be performing close to the vertical acuity range of the bat and presumably if more input vectors were available for training, then the performance of the network may improve.

Close examination of the spectral data by Wotton *et al.* (1997) indicated that there could be an enhancement of the elevation localization cues in the echolocation combination compared to the ear alone. However, the echolocation combination network model did not significantly improve elevation localization compared to the ear network model. The mean error (see Table I and Fig. 7) was smaller in the localization test performance of the echolocation combination but this effect was not significant. In the spectral data the systematic changes of the spectral notches and peaks, thought to be the elevation cues, appear to be present only for locations between about $+30^\circ/-40^\circ$ in elevation and about $\pm 30^\circ$ in

azimuth (Wotton *et al.*, 1997). The network model was both trained and tested over a greater range of locations and it may be that any enhancement of these cues in this smaller region was washed out. The network was trained on 102 positions which is already a fairly small training set and unfortunately restricting training to the small region of these spectral cues would reduce the number of inputs to an impractical 32 positions.

PCA can describe the main features of the spectra of the emission, the external ear, and the combination using only a few dimensions. Kistler and Wightman (1992) found that human observers were able to localize quite well when the HRTFs were synthesized using only five weighted basis functions. It is not known exactly how many principal components are required to produce sufficient information for a bat to localize but as most of the variance (98%–99%) has been accounted for by using the first eight principal components it seems likely that most spectral cues will still be included.

Using the backpropagation network model as a near optimal observer allowed direct comparisons of the localization performance of each condition. It has not been determined that bats utilize the same information as the network models but it is clear that localization information is available in the emission, the ear, and echolocation combination spectra. The echolocation combination appears to combine the best information from both the ear and emission without being affected by the less reliable information.

ACKNOWLEDGMENTS

This research was supported by the Grant NIH Training Program in Hearing (T32 DC00045) to J.M.W. and by NIH Grant No. DC02804 to R.L.J. We would like to thank Dr. Dave Hartley for sending us the data on the sonar emission of *Eptesicus* [D. J. Hartley and R. A. Suthers, *J. Acoust. Soc. Am.* **85**, 1348–1351 (1989)]. Thanks also to the two anonymous reviewers for their helpful comments.

- Butler, R. A., and Belendiuk, K. (1977). "Spectral cues utilized in the localization of sound in the median sagittal plane," *J. Acoust. Soc. Am.* **61**, 1264–1269.
- Datum, M. S., Palmieri, F., and Moiseff, A. (1996). "An artificial neural network for sound localization using binaural cues," *J. Acoust. Soc. Am.* **100**, 372–383.
- Donchin, E. (1966). "A multivariate approach to the analysis of average evoked potentials," *Trans. Biomed. Eng.* **13**, 131–139.
- Dror, I. E., Zagaeski, M., and Moss, C. F. (1995). "Three-dimensional target recognition via sonar: A neural network model, *Neural Networks* **8**, 149–160.
- Fisher, H. G., and Freedman, S. J. (1968). "The role of the pinna in auditory localization," *J. Aud. Res.* **8**, 15–26.
- Fuzessery, Z. M., Hartley, D. J., and Wenstrup, J. J. (1992). "Spatial processing within the mustache bat echolocation system: possible mechanisms for optimization," *J. Comp. Physiol. A* **170**, 57–71.
- Fuzessery, Z. M., and Pollak, G. D. (1984). "Neural mechanisms of sound localization in an echolocating bat," *Science* **225**, 725–728.
- Fuzessery, Z. M., Wenstrup, J. J., and Pollak, G. D. (1990). "Determinants of horizontal sound location selectivity of binaurally excited neurons in an isofrequency region of the mustache bat inferior colliculus," *J. Neurophysiol.* **63**, 1128–1147.
- Gardner, M. B. (1973). "Some monaural and binaural facets of median plane localization," *J. Acoust. Soc. Am.* **54**, 1489–1495.

- Gardner, M. B., and Gardner, R. S. (1973). "Problem of localization in the median plane: Effect of pinnae cavity occlusion," *J. Acoust. Soc. Am.* **53**, 400–408.
- Gorman, R. P., and Sejnowski, T. J. (1988). "Analysis of hidden units in a layered network trained to classify sonar targets," *Neural Networks* **1**, 75–89.
- Griffin, D. R. (1958). *Listening in the Dark* (Yale Univ., New Haven, CT).
- Grinnell, A. D. (1963). "The neurophysiology of audition in bats: directional localization and binaural interaction," *J. Physiol. (London)* **167**, 97–113.
- Grinnell, A. D., and Schnitzler, H.-U. (1977). "Directional sensitivity of echolocation in the horseshoe bat, *Rhinolophus ferrumequinum*, II: Behavioral directionality of hearing," *J. Comp. Physiol. A* **116**, 63–76.
- Harman, H. H. (1960). *Modern Factor Analysis* (Univ. of Chicago, Chicago).
- Hartley, D. J., and Suthers, R. A. (1989). "The sound emission pattern of the echolocating bat, *Eptesicus fuscus*," *J. Acoust. Soc. Am.* **85**, 1348–1351.
- Hebrank, J., and Wright, D. (1974). "Are two ears necessary for localization of sound sources on the median plane," *J. Acoust. Soc. Am.* **56**, 935–938.
- Heffner, R. S., and Heffner, H. E. (1988). "Sound localization acuity in the cat: Effect of azimuth, signal duration, and test procedure," *Hearing Res.* **36**, 221–232.
- Hertz, J., Krogh, A., and Palmer, R. G. (1991). *Introduction to the Theory of Neural Computation* (Addison-Wesley, Redwood City, CA).
- Jen, P. H.-S., and Chen, D. (1988). "Directionality of sound pressure transformation at the pinna of echolocating bats," *Hearing Res.* **34**, 101–118.
- Kaiser, H. F. (1958). "The Varimax criterion for analytic rotation in factor analysis," *Psychometrika* **23**, 187–200.
- Kistler, D. J., and Wightman, F. L. (1992). "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," *J. Acoust. Soc. Am.* **91**, 1637–1647.
- Lawrence, B. D., and Simmons, J. A. (1982). "Echolocation in bats: The external ear and perception of the vertical positions of targets," *Science* **218**, 481–483.
- Li, K.-P., Hughes, G. W., and House, A. S. (1969). "Correlation characteristics and dimensionality of speech spectra," *J. Acoust. Soc. Am.* **46**, 1019–1025.
- Makous, J. C., and Middlebrooks, J. C. (1990). "Two-dimensional sound localization by human listeners," *J. Acoust. Soc. Am.* **87**, 2188–2200.
- Masters, W. M., Moffat, A. J. M., and Simmons, J. A. (1985). "Sonar tracking of horizontally moving targets by the big brown bat, *Eptesicus fuscus*," *Science* **228**, 1331–1333.
- Middlebrooks, J. C. (1992). "Narrow-band sound localization related to external ear acoustics," *J. Acoust. Soc. Am.* **92**, 2607–2624.
- Middlebrooks, J. C., and Green, D. M. (1991). "Sound localization by human listeners," *Annu. Rev. Psychol.* **42**, 135–159.
- Middlebrooks, J. C., and Green, D. M. (1992). "Observations on a principal components analysis of head-related transfer functions," *J. Acoust. Soc. Am.* **92**, 597–599.
- Moore, P. W. B., Roitblat, H. L., Penner, R. H., and Nachtigall, P. E. (1991). "Recognizing successive dolphin echoes with an integrator gateway network," *Neural Networks* **4**, 701–709.
- Musican, A. D., Chan, J. C. K., and Hind, J. E. (1990). "Direction-dependent spectral properties of cat external ear: New data and cross-species comparisons," *J. Acoust. Soc. Am.* **87**, 757–781.
- Neti, C., Young, E. D., and Schneider, M. H. (1992). "Neural network models of sound localization based on directional filtering by the pinna," *J. Acoust. Soc. Am.* **92**, 3140–3156.
- Oldfield, S. R., and Parker, S. P. A. (1986). "Acuity of sound localisation: A topography of auditory space. III Monaural hearing conditions," *Perception* **15**, 67–81.
- Plomp, R., Pols, L. C. W., and van de Geer, J. P. (1967). "Dimensional analysis of vowel spectra," *J. Acoust. Soc. Am.* **41**, 707–712.
- Pols, L. C. W., van der Kamp, L. J. Th., and Plomp, R. (1969). "Perceptual and physical space of vowel sounds," *J. Acoust. Soc. Am.* **46**, 458–467.
- Reyment, R. A., and Joreskog, K. G. (1993). *Applied Factor Analysis in the Natural Sciences* (Cambridge U.P., Cambridge, England).
- Rice, J. J., May, B. J., Spirou, G. A., and Young, E. D. (1992). "Pinna-based spectral cues for sound localization in cat," *Hearing Res.* **58**, 132–152.
- Roitblat, H. L., Au, W. W. L., Nachtigall, P. E., Shizumura, R., and Moons, G. (1995). "Sonar recognition of targets embedded in sediment," *Neural Networks* **8**, 1263–1273.
- Roitblat, H. L., Moore, P. W. B., Nachtigall, P. E., Penner, R. H., and Au, W. W. L. (1989). "Natural echolocation with an artificial neural network," *Int. J. Neural Net.* **1**, 239–248.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). "Learning internal representations by error propagation," in *Parallel Distributed Processing*, edited by D. E. Rumelhart and J. L. McClelland (MIT, Cambridge, MA), pp. 318–364.
- Sanger, T. D. (1989). "Optimal unsupervised learning in feedforward neural networks," MIT Artificial Intelligence Laboratory Technical Report No. 1086 (unpublished).
- Schnitzler, H.-U., and Grinnell, A. D. (1977). "Directional sensitivity of echolocation in the horseshoe bat, *Rhinolophus ferrumequinum*. I. Directionality of sound emission," *J. Comp. Physiol. A* **116**, 51–61.
- Shimozawa, T., Suga, N., Hendler, P., and Schuetz, S. (1974). "Directional sensitivity of echolocation system in bats producing frequency-modulated signals," *J. Exp. Biol.* **60**, 53–69.
- Simmons, J. A., Ferragamo, M. J., Saillant, P. A., Haresign, T., Wotton, J. M., Dear, S. P., and Lee, D. N. (1995). "Auditory dimensions of acoustic images in echolocation," in *Hearing by Bats*, edited by A. N. Popper and R. R. Fay (Springer-Verlag, New York), pp. 146–190.
- Simmons, J. A., Kick, S. A., Lawrence, B. D., Hale, C., Bard, C., and Escudie, B. (1983). "Acuity of horizontal angle discrimination by the echolocating bat, *Eptesicus fuscus*," *J. Comp. Physiol. A* **153**, 321–330.
- Suter, C. M. (1970). "Principal component analysis of average evoked potentials," *Exp. Neurology* **29**, 317–327.
- Wood, C. C., and McCarthy, G. (1984). "Principal component analysis of event-related potentials: Simulation studies demonstrate misallocation of variance across components," *Electroencephalog. Clin. Neurophysiol.* **59**, 249–260.
- Wotton, J. M., Haresign, T., and Simmons, J. A. (1995). "Spatially dependent acoustic cues generated by the external ear of the big brown bat, *Eptesicus fuscus*," *J. Acoust. Soc. Am.* **98**, 1423–1445.
- Wotton, J. M., Jenison, R. L., and Hartley, D. J. (1997). "The combination of echolocation emission and ear reception enhances directional spectral cues of the big brown bat, *Eptesicus fuscus*," *J. Acoust. Soc. Am.* **101**, 1723–1733.
- Wotton, J. M., and Simmons, J. A. (in preparation). "Spectral cues and the perception of the vertical position of targets by the big brown bat, *Eptesicus fuscus*."
- Zahorian, S. A., and Rothenberg, M. (1981). "Principal-components analysis for low-redundancy encoding of speech spectra," *J. Acoust. Soc. Am.* **69**, 832–845.

Acoustic effects of the ATOC signal (75 Hz, 195 dB) on dolphins and whales

Whitlow W. L. Au, Paul E. Nachtigall, and Jeffrey L. Pawloski

Marine Mammal Research Program, Hawaii Institute of Marine Biology, University of Hawaii,
P.O. Box 1106, Kailua, Hawaii 96734

(Received 12 June 1996; accepted for publication 19 January 1997)

The Acoustic Thermometry of Ocean Climate (ATOC) program of Scripps Institution of Oceanography and the Applied Physics Laboratory, University of Washington, will broadcast a low-frequency 75-Hz phase modulated acoustic signal over ocean basins in order to study ocean temperatures on a global scale and examine the effects of global warming. One of the major concerns is the possible effect of the ATOC signal on marine life, especially on dolphins and whales. In order to address this issue, the hearing sensitivity of a false killer whale (*Pseudorca crassidens*) and a Risso's dolphin (*Grampus griseus*) to the ATOC sound was measured behaviorally. A staircase procedure with the signal levels being changed in 1-dB steps was used to measure the animals' threshold to the actual ATOC coded signal. The results indicate that small odontocetes such as the *Pseudorca* and *Grampus* swimming directly above the ATOC source will not hear the signal unless they dive to a depth of approximately 400 m. A sound propagation analysis suggests that the sound-pressure level at ranges greater than 0.5 km will be less than 130 dB for depths down to about 500 m. Several species of baleen whales produce sounds much greater than 170–180 dB. With the ATOC source on the axis of the deep sound channel (greater than 800 m), the ATOC signal will probably have minimal physical and physiological effects on cetaceans. © 1997 Acoustical Society of America. [S0001-4966(97)05905-5]

PACS numbers: 43.80.Nd, 43.80.Lb [FD]

INTRODUCTION

The Heard Island Feasibility Test (Munk *et al.*, 1994) of Scripps Institution of Oceanography and the Applied Physics Laboratory, University of Washington, in January, 1991 has generated considerable controversy. The intensity of ensuing negative publicity and public outcry over the possibility of harming marine mammals and other marine animals (Cohen, 1991) has surprised many underwater acousticians. The follow on Acoustic Thermometry of Ocean Climate (ATOC) program, which will broadcast a low-frequency 75-Hz phase modulated, 195 dB *re*: 1 μ Pa source level acoustic signal over ocean basins to study ocean temperatures on a global scale, has also drawn considerable criticism from concerned individuals. Aside from the intense emotions engendered against ATOC, one of the difficulties surrounding the controversy is the lack of data on the low-frequency hearing sensitivity of whales, making it difficult to rationally address the appropriate issues.

The initial feasibility stage of the ATOC program calls for installing a source at a depth of approximately 900 m on the Pioneer Seamount off Point Sur in California. This source was installed in December 1995 (Mercer *et al.*, 1996) and a study on its effect on marine mammals is underway (Costa *et al.*, 1996). Another source is to be installed in waters north of the island of Kauai in the Hawaiian Archipelago, approximately 7 nautical miles from shore at a depth of approximately 850 m. The California source is located close to the routes of gray and humpback whales. Humpback whales migrate from the waters off Alaska to winter in the waters of the Hawaiian Islands, including the island of Kauai.

In order to obtain some data on the potential effects off the ATOC signal on hearing, communications, and the well being of dolphins and whales, the hearing sensitivity of two pelagic cetaceans was measured using the ATOC signal. The subjects were a false killer whale (*Pseudorca crassidens*) and a Risso's dolphin (*Grampus griseus*). The animals were housed, trained, and tested at the Hawaii Institute of Biology Marine Mammal Research facility in Kaneohe Bay, Oahu, Hawaii.

The ATOC signal is a phase modulated carrier that can be described by the equation

$$s(t) = A \cos(2\pi ft + \phi[2m_i - 1]), \quad (1)$$

where f is equal to 75 Hz, ϕ is equal to 88.2092° , and m_i is a specific sequence of 1's and 0's used to create a phase shift of ϕ when m_i equals 1 and $-\phi$ when m_i equals 0. Two cycles of the 75-Hz carrier correspond to a digit and each bit in the m -sequence determines the phase of a digit. The three signals in Fig. 1 consisting of four two-cycle digits having a duration of 107.7 ms can be used to illustrate how the m -sequence affects the waveform and power spectrum of the signal. A duration of 107.7 ms corresponds closely to the integration time constant of a dolphin's auditory system at 100 Hz (Johnson, 1968). In the top signal of Fig. 1, $m = 1$ for each digit so that there is no phase shift and the signal is a 75-Hz pure tone. In the middle signal, $m = 1$ for the first digit and $m = 0$ for the next three digits so that there is one phase shift of 176.4184° . This single phase shift causes the bandwidth of the signal to increase, as can be seen by its power

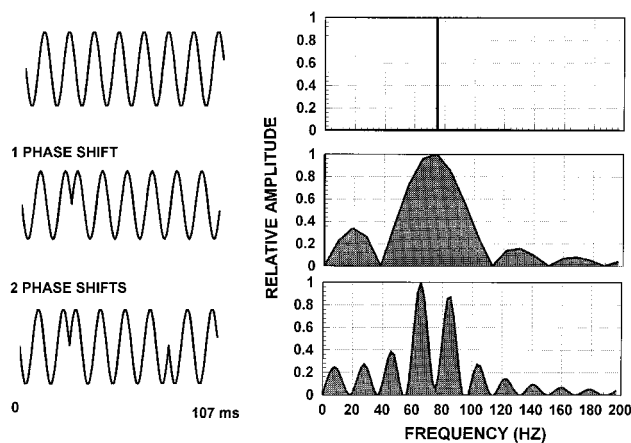


FIG. 1. Examples of the ATOC signal in a 107-ms interval showing the effects of the 88.2092° phase shift.

spectrum. In the bottom signal, $m=1$ for the first digit, $m=0$ for the second and third digit, and then $m=1$ for the fourth digit. Therefore, there are two 176.4184° phase shifts in the bottom signal that cause the spectrum to broaden.

The m -sequence is used in conjunction with cross correlation processing to precisely determine the time of arrival of the signal after it has traveled several thousand miles across the ocean. Changes in the integrated ocean temperature can be determined by examining changes in the arrival time. A correlation type received will also be helpful in extracting the signal out of noise. However, using such a scheme for the signal affects the bandwidth of the signal. Therefore, the hearing threshold of a marine mammal will probably be lower with the ATOC sound than with a 75-Hz pure tone since the broadening of the spectrum may cause energy to spill into critical bands that are adjacent to the one at 75 Hz.

I. METHODS

A go/no-go experimental procedure was used in measuring the animals' sensitivity to the ATOC signal and to other pure-tone signals. A trial began with the subjects swimming into a stationing hoop facing a J-13 transducer two meters from the hoop, as depicted in Fig. 2. When the subject stationing properly, the experimenter activated an underwater

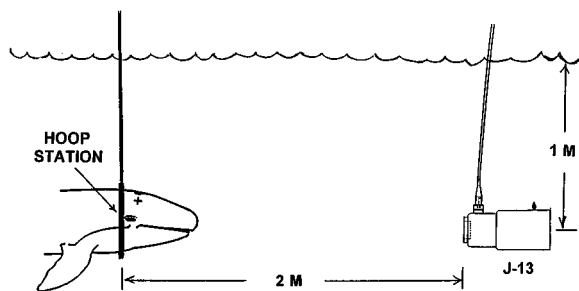


FIG. 2. Experimental geometry showing a *Pseudorca crassidens* facing a J-13 transducer.

light alerting the animal to expect an acoustic signal. The underwater light remained on until the cessation of a trial. For a signal present trial, a signal was projected 1 s after the activation of the light. If the signal was audible to the subject it would back out of the hoop and touch a response paddle (go response). For a signal absent trial, the animal was required to remain in the hoop until the experimenter presented an audible 12-kHz bridge tone to release the animal from station (no-go response). The subject received a food reward for correct responses and did not receive reinforcement for wrong responses. A modified Gellerman series (Gellerman, 1933) was used for the signal presentation schedule in which an equal number of signal present and signal absent trials were conducted per ten trial block with the stipulation that the signal condition not be the same for more than three consecutive trials.

An up-down staircase procedure was used to determine the subject's threshold of hearing. After a ten-trial warm-up period in which a constant supra-threshold level signal was used, the signal level was reduced by 2 dB for every correct signal present response until the animal made its first miss, which constituted a reversal. The signal level was then progressively increased in 1-dB steps until the next correct signal present response was made. Thereafter, the signal level was changed in 1-dB steps depending on the animal's response during signal present trials. A session continued until ten reversals occurred. A threshold was defined when the average signal level per session at the reversal points for two consecutive sessions was within 3 dB.

The simulated ATOC signal was generated using a DATES PC-420 12-bit arbitrary waveform generator board with the 32 736 points of an m -sequence signal stored in on-board memory and clocked out continuously at a rate of 1200 Hz. When the experimenter activated the signal-present switch of a signal shaping circuit, the DATES PC-420 output was amplitude modulated with a 3-s trapezoidal pulse having a 0.3-s rise and fall time to produce the stimulus signal. The signal level was controlled by a rotary attenuator. Pure-tone signals were produced in a similar fashion with the signal gated on and turned off by modulating a continuous signal with the trapezoidal pulse. The signal level at the hoop station was calibrated with an International Transducer Inc. ITC-1032 spherical hydrophone connected in a differential mode to a Princeton Applied Research PAR-113 low-noise amplifier and an oscilloscope.

II. RESULTS AND DISCUSSION

The mean and standard deviation of the *Grampus*' thresholds were 142.2 ± 1.7 dB for the 75-Hz pure-tone signal and 140.8 ± 1.1 dB for the ATOC signal. The *Pseudorca*'s thresholds were 140.7 ± 1.2 dB for the 75-Hz pure-tone signal and 139.0 ± 1.1 dB for the ATOC signal. The difference between the threshold for the 75-Hz and ATOC signals for both animals was significant to the 0.01 level (paired t -test). These thresholds are plotted in Fig. 3 along with other low-frequency pure-tone thresholds for the same two animals obtained by Nachtigall *et al.* (1995) and for another *Pseudorca* measured by Thomas *et al.* (1988). The data of Nachtigall *et al.* (1995) are included here to compare the

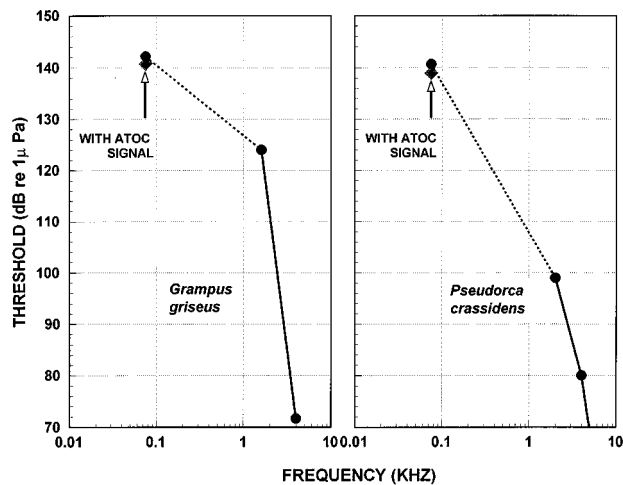


FIG. 3. Hearing threshold results for the ATOC signal and other pure-tone signals. The results at frequencies other than 75 Hz are from Nachtigall *et al.* (1995).

75-Hz threshold with other low-frequency thresholds and to illustrate the fact that the auditory system of these odontocetes is adapted for high ultrasonic frequencies. The pure-tone thresholds are similar to those obtained by Johnson (1967) for *Tursiops truncatus* and for Johnson *et al.* (1989) for *Delphinapterus leucas* at frequencies close to 75 Hz. These odontocetes are relatively insensitive to low-frequency sounds; *Pseudorca* has a maximum sensitivity of approximately 39 dB *re*: 1 μ Pa at 65 kHz (Thomas *et al.*, 1988), which is 100 dB better than at 75 Hz.

The effects of the ATOC signal on these animals can be considered by examining the propagation loss profile of the

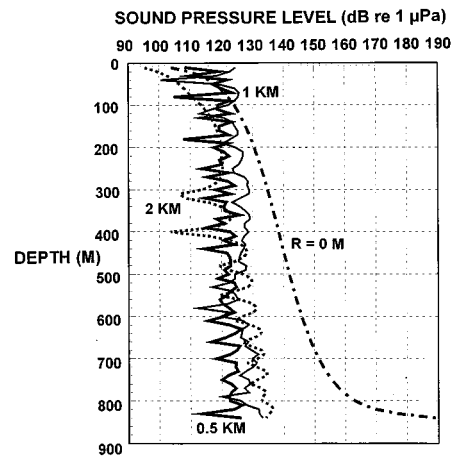


FIG. 5. Sound-pressure level of the ATOC signal as a function of depth for horizontal ranges of 0, 0.5, 1, and 2 km from the source.

ATOC source. The transmission loss profile for the proposed Kauai location was obtained with a computer program that used a finite element parabolic equation (FEPE) solution (Collins and Westwood, 1991) to the two-dimensional wave equation and a vertical sound velocity and a bottom depth profile appropriate to Kauai. The transmission loss profile for the case in which the signal is traveling directly toward shore is shown in Fig. 4, where the black line denotes the bottom contour. This figure illustrates the complexity of the propagation conditions caused by the signal reflectioning off the surface and bottom. The propagation condition reflected in Fig. 4 is probably the most severe in terms of affecting marine mammals since the signal is propagating toward an up-

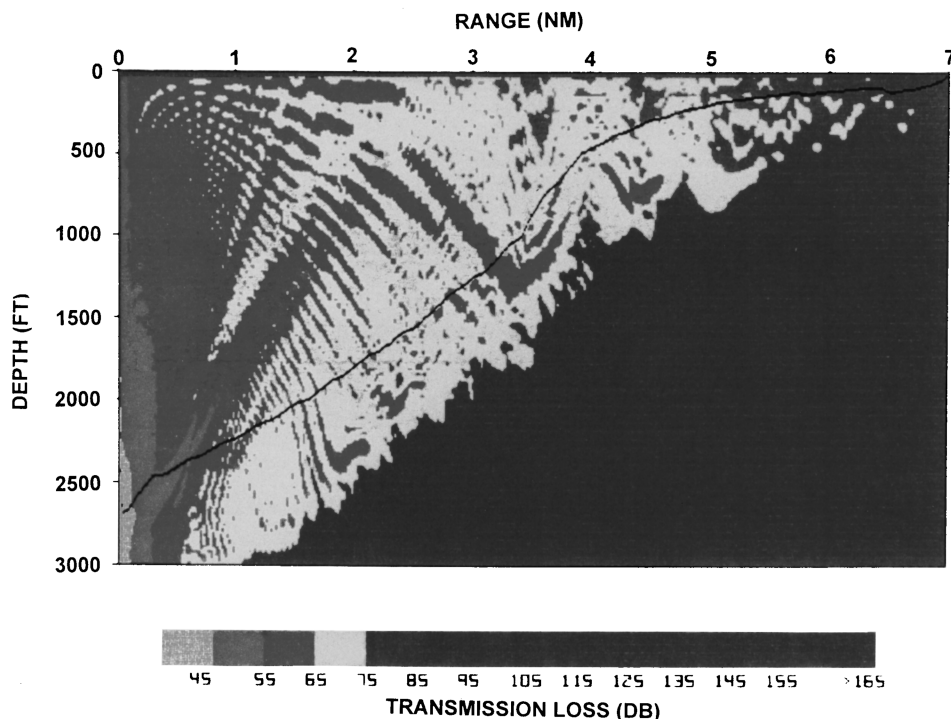


FIG. 4. Propagation loss profile for the ATOC source at the Kauai location obtained with a finite element parabolic equation solution to the two-dimensional wave equation.

TABLE I. Examples of some baleen whale low-frequency sounds (from Richardson *et al.*, 1995).

Species of whale	Signal type	Frequency range (Hz)	Dominant frequencies (Hz)	SPL at 1 m (dB <i>re</i> : 1 μ Pa)	SPL at 50 m (dB <i>re</i> : 1 μ Pa)
Blue	moans	12–390	16–25	188	154
Bowhead	moans	25–900	100–400	128–178	94–144
Fin	moans	14–118	20	160–186	126–152
Gray	moans	20–1200	20–200	185	151
Humpback	moans	20–1800	35–360	175	141
	grunts	25–1900	...	190	156
	horn blast	...	410–420	185	151
	songs	30–8000	120–4000	144–174	110–140
Minke	down sweep	60–1300	...	165	131
	moans/grunts	60–140	60–140	151–175	123–141

ward sloping bottom which can reflect the signal back into the water column.

In order to obtain a more detailed understanding of the sound field close to the location of the source, some of the values used to generate the transmission loss profile were used to calculate the sound-pressure levels as a function of depth and range. The results are shown in Fig. 5, with sound-pressure level depicted as a function of depth at different horizontal ranges from the source. The curve at $R=0$ was obtained by summing the upward moving direct signal with a 180° phase shifted downward moving surface reflected signal, with both components undergoing spherical spreading losses. The $R=0$ curve suggests that small cetaceans such as *Grampus* and *Pseudorca* swimming directly over the ATOC source will not hear the transmitted sound unless the animals dove to a depth of approximately 400 m. If these animals were at a horizontal range greater than 0.5 km, the level of the ATOC signal will be below their hearing threshold at any depth.

The curves of Fig. 5 also indicate that for ranges greater than 0.5 km, the maximum sound-pressure level above a depth of 560 m is approximately 130 dB. As the range increases beyond 2 km, the sound-pressure level will become progressively lower. The in-air equivalence of an underwater sound-pressure level of 130 dB can be determined by equating the intensity (power/area) of an acoustic signal in both media. It would not be meaningful to directly compare the acoustic pressure in both media because the large difference in the density causes the acoustic power or intensity to be very different. Acoustic intensity is defined as

$$I = \frac{p^2}{\rho c}, \quad (2)$$

where ρ is the density and c is the sound velocity of the medium. The product ρc is known as the specific acoustic impedance and is approximately equal to 1.5×10^6 Pa s/m in water and 416 Pa s/m in air. Therefore, an underwater acoustic signal of 130 dB is equivalent to 3.16 Pa with an intensity of 6.7×10^{-6} W/m². Letting the intensity of the underwater acoustic signal be equal to the intensity of an airborne acoustic signal, the resulting airborne acoustic pressure will be 0.053 Pa or 94.4 dB *re*: 1 μ Pa. Airborne sounds are typically expressed in dB *re*: 20 μ Pa and the acoustic pressure using this reference has the notation SPL following the dB value.

Therefore, 94.4 dB *re*: 1 μ Pa in air is equal to 68.4 dB SPL and is equivalent to a 130 dB *re*: 1 μ Pa underwater acoustic signal. An automobile traveling at 35 mph produces a sound of approximately 72 dB SPL 50 ft from it (Kryter, 1985).

An obvious question at this point is the possible effects of the ATOC sound on large baleen whales. One way of addressing this question is to consider sounds that whales themselves make near their conspecifics and to compare these sounds with the ATOC sound field shown in Figs. 4 and 5. It is unfortunate that there are almost no data available on the hearing sensitivity and frequency range of hearing for baleen whales. The frequencies and sound-pressure levels of low-frequency sounds produced by some baleen whales are shown in Table I, with the data extracted from a table in Richardson *et al.* (1995). From Table I it is clear that some whales produce low-frequency sounds that have levels at 50 m that are greater than 130 dB, and levels on the order of 150 dB are not uncommon. If these whales are not physically harmed by conspecifics producing levels of sounds on the order of 150 dB at 50 m, then the ATOC sound at a sound-pressure level of 130 dB or less will probably be relatively harmless to them physically. Furthermore, the bandwidth of the ATOC sound is relatively narrow and would not be as effective in masking as would the broader bandwidth sounds normally produced by some whales themselves.

ACKNOWLEDGMENTS

The authors would like to express their appreciation to Dr. Kurt Metzger of the University of Michigan for supplying the ATOC signal in ASCII format and to Kevin Heaney of the Scripps Oceanographic Institution for providing us with propagation loss calculations. Special thanks to Dr. Robert Gisiner of the Office of Naval Research (Grant No. N0001496WX30364) and Dr. Christopher Clark of the Bioacoustic Research Program, Cornell University for funding the threshold experiment. The comments of two anonymous reviewers were greatly appreciated. This work relates to the Advanced Research Projects Agency Grant No. MDA972-93-1-0003 funded by the Strategic Environmental Research and Development Program (SERDP). The United States Government has a royalty-free license throughout the world in all copyrightable material herein. This is Hawaii Institute of Marine Biology contribution 1020.

- Cohen, J. (1991). "Was underwater 'shot' harmful to the whales?" *Science* **252**, 912–914.
- Collins, M. D., and Westwood, E. K. (1991). "A higher-order energy-conserving parabolic equation for range-dependent ocean depth, sound speed, and density," *J. Acoust. Soc. Am.* **89**, 1068–1075.
- Costa, D. P., Crocker, D. E., Croll, D., Goley, D., Houser, D., LeBoeuf, B. J., Waples, D., and Webb, P. (1996). "Effects of the California ATOC experiment on marine mammals," *J. Acoust. Soc. Am.* **100**, 2581 (A).
- Gellerman, L. W. (1933). "Chance orders of alternating stimuli in visual discrimination," *J. Gen. Psychol.* **42**, 206–208.
- Johnson, C. S. (1967). "Sound detection thresholds in marine mammals," in *Marine BioAcoustics*, edited by W. N. Tavoga (Pergamon, New York), pp. 247–260.
- Johnson, C. S. (1968). "Relation between absolute threshold and duration-of-tone pulses in the bottlenose porpoise," *J. Acoust. Soc. Am.* **43**, 757–763.
- Johnson, C. S., McManus, M. W., and Skaar, D. (1989). "Masked tonal hearing thresholds in the beluga whale," *J. Acoust. Soc. Am.* **85**, 2651–2654.
- Kryter, K. D. (1985). *The Effects of Noise on Man* (Academic, Orlando, FL).
- Munk, W. H., Spindel, R. C., Baggeroer, A., and Birdsall, T. G. (1994). "The Heard Island feasibility test," *J. Acoust. Soc. Am.* **96**, 2330–2342.
- Mercer, J. A., Howe, B. M., Dushaw, B. D., Spindel, R. C., Worcester, P. F., Colosi, J. A., Cornuelle, B. D., Dziechiuch, M. A., Munk, W. H., Birdsall, T., and Metzger, K. (1996). "Acoustic thermometry of ocean climate. A description of the acoustic network," *J. Acoust. Soc. Am.* **100**, 2580 (A).
- Nachtigall, P. E., Au, W. W. L., Pawloski, J. A., and Moore, P. W. B. (1995). "Risso's dolphin (*Grampus griseus*) hearing thresholds in Kaneohe Bay in *Sensory Systems of Aquatic Mammals*, edited by R. A. Kastelein, J. A. Thomas, and P. E. Nachtigall (De Spil, Woerden, The Netherlands), pp. 49–53.
- Richardson, W. J., Green, Jr., C. R., Malme, C. I., and Thomson, D. H. (1995). *Marine Mammals and Noise* (Academic, San Diego).
- Thomas, J., Chun, N., Au, W., and Pugh, K. (year). "Underwater audiogram of a false killer whale (*Pseudorca crassidens*)," *J. Acoust. Soc. Am.* **84**, 936–940.

LETTERS TO THE EDITOR

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

Comments on “Flexural wave reduction using a compliant tube baffle” [J. Acoust. Soc. Am. 99(2), 691–699 (1996)] and “Signal pressure received by a hydrophone placed on a plate backed by a compliant baffle” [J. Acoust. Soc. Am. 89(2), 559–564 (1991)]

Ronald P. Radlinski and Sung H. Ko

Naval Undersea Warfare Center, Newport Division, Newport, Rhode Island 02841

(Received 27 September 1996; accepted for publication 15 October 1996)

The cited formulations for studying the baffling of flexural waves by a compliant tube grating [S. H. Ko, J. Acoust. Soc. Am. 99, 691–699 (1996)] and for determining the acoustic receive sensitivity of a hydrophone placed near a plate backed by a compliant tube grating [S. H. Ko and H. H. Schloemer, J. Acoust. Soc. Am. 89, 559–564 (1991)] are dependent on the validity of the lumped-fluid parameter approximations developed by Junger [M. C. Junger, J. Acoust. Soc. Am. 78, 1010–1012 (1985)] for a grating of closely spaced compliant tubes between two semi-infinite fluid half-spaces. If the fluid between the plate and the grating is great enough to minimize their interaction, the incorporation of the effective layer parameters developed by Junger should be reasonably valid. Unfortunately, the sample calculations in Ko (Ref. 4) and Ko and Schloemer (Ref. 5) have been made for the limiting case of the compliant tube grating in direct contact with the plates where the use of Junger’s approximations is inappropriate. [S0001-4966(97)05803-7]

PACS numbers: 43.20.Gp, 43.20.Jr, 43.20.Tb, 43.20.Ye [JEG]

From the results on acoustical scattering from multiple gratings of compliant tubes published by Radlinski and Simon,^{1,2} Junger³ was able to derive effective fluid parameters for a single layer of closely packed grating of compliant tubes between two semi-infinite fluid half spaces. Ko⁴ recently used the approximations of Junger to study flexural wave baffling by a compliant tube grating and earlier Ko and Schloemer⁵ used the effective parameters to determine, for an incident plane wave, the acoustic signal pressure at a hydrophone placed near a plate backed by a compliant tube grating.

In the examples illustrated in Ref. 4, plates are in direct contact with each side of the compliant tube grating. Since the boundary conditions are no longer those of the expressions derived by Junger, the application would not be expected to be valid. The radiation loading of the plates would shift the effective resonance frequencies of the symmetric compliant modes from that of the tubes in fluid. The periodically placed compliant tubes in direct contact with the plates also would scatter the elastic waves propagated by the plates. In addition, as illustrated in Ref. 1 for two closely spaced gratings of compliant tubes in fluid, as the grating in the fluid approaches the plates, additional higher-frequency noncompliant structural modes of the tubes, which are not in

Junger’s approximations, would be excited from the interaction of the evanescent fields of the grating and of the plates.

Since the calculations in Ref. 5 are performed for hydrophone sensitivity with the compliant tube grating placed in direct contact with the signal conditioning plate, again the use of Junger’s effective fluid layer parameters is not appropriate. However, in the frequency region of the compliant tube resonances, the thickness of the plate has been chosen to shield the hydrophones from being desensitized by the acoustically soft baffle. Thus, except for the shifting of the frequencies of decreased response near the compliant resonances of the tubes, the calculated hydrophone sensitivity may not be strongly dependent on rigorous modeling of the tube grating situated behind the plate. At frequencies below the compliant resonances of the tubes, the plate mass and grating compliance would act more like the lumped elements approximations derived by Junger.

Comparisons in Ref. 5 between theory and experiments give reasonable agreement. This agreement results from the experimental geometry in which the compliant tube grating was encapsulated in a viscoelastic material⁶ and therefore was likely to be outside the evanescent field of the plate. In addition, it has been shown in Ref. 7 that encapsulation of a

compliant tube grating in a viscoelastic material mitigates the near-field excitation of the noncompliant resonant modes found in a fluid medium. Thus, a configuration where the calculations using the lumped elements derived by Junger would be appropriate may be approximated by the experimental arrangement discussed in Ref. 5.

To determine the distances from a plate where Junger's lumped-parameter approximations would be satisfied for the configurations studied in Refs. 4 and 5, calculations using the solution of Junger should be compared with a more rigorous theory. For example, the scattering formulation for normally incident acoustic waves on a compliant tube grating in fluid as discussed in Ref. 1 can easily be extended to include signal conditioning or hull-like plates described by either full elasticity theory or plate theory. Similar modifications can be made to the theory for a compliant tube grating in a viscoelastic material as discussed in Ref. 7. The finite element formulation of Kalinowski⁸ or the combined finite element and analytical development of Hennion and Decarpigny⁹ and Hladky-Hennion¹⁰ provide additional rigorous solutions for investigating either flexural wave baffling or signal conditioning for the cases of the tubes either attached directly to the plate boundaries or with a fluid/viscoelastic layer between the plates and the compliant tube grating. Thus, the formulations in Refs. 1, 7, 8, 9, and 10 can be used to deter-

mine the distances between a plate and a compliant tube grating for which Junger's lumped-parameter approximations might be appropriately used in the formulations developed in Refs. 4 and 5.

¹R. P. Radlinski and M. M. Simon, "Scattering by multiple gratings of compliant tubes," *J. Acoust. Soc. Am.* **72**, 607-614 (1982).

²R. P. Radlinski and M. M. Simon, "Erratum: 'Scattering by multiple gratings of compliant tubes' [*J. Acoust. Soc. Am.* **72**, 607-614 (1982)]," *J. Acoust. Soc. Am.* **74**, 1646 (1983).

³M. C. Junger, "Water-borne sound insertion loss of a planar compliant-tube array," *J. Acoust. Soc. Am.* **78**, 1010-1012 (1985).

⁴S. H. Ko, "Flexural wave reduction using a compliant tube baffle," *J. Acoust. Soc. Am.* **99**, 691-699 (1996).

⁵S. H. Ko and H. H. Schloemer, "Signal pressure received by a hydrophone placed on a plate backed by a compliant baffle," *J. Acoust. Soc. Am.* **89**, 559-564 (1991).

⁶Private communication with Howard Schloemer.

⁷R. P. Radlinski, "Scattering from multiple gratings of compliant tubes in a viscoelastic layer," *J. Acoust. Soc. Am.* **85**, 2301-2310 (1989).

⁸A. J. Kalinowski, "Steady state solutions to the dynamically loaded periodic structures," NASA Conference Publication 2131, Eighth NASTRAN User's Colloquium (1979).

⁹A. C. Hennion and J. N. Decarpigny, "Analysis of the scattering of a plane acoustic wave by a periodic elastic structure using the finite element method: Application to compliant tube gratings," *J. Acoust. Soc. Am.* **87**, 1861-1870 (1990).

¹⁰A. C. Hladky-Hennion, "Application de la methode des elements finis a la modelisation de structures periodiques utilisees en acoustique," thesis, Universite des Sciences et Techniques de Lille, 1990.

Scattering of acoustic waves by intermittent temperature and velocity fluctuations

D. Keith Wilson

U.S. Army Research Laboratory, ATTN: AMSRL-IS-EE, 2800 Powder Mill Road, Adelphi, Maryland 20783

(Received 9 February 1996; revised 22 December 1996; accepted 21 January 1997)

The effect of atmospheric turbulent intermittency on acoustical scattering into shadow zones was discussed in a previous paper [D. K. Wilson, J. C. Wyngaard, and D. I. Havelock, *J. Acoust. Soc. Am.* **99**, 3393–3400 (1996)]. Intermittency was incorporated into the classical scattering theory by defining a local mean scattered intensity, \tilde{I}_0 , which has a log-normal probability density function (pdf), $P_{\tilde{I}_0}$. The pdf P_I for the actual scattered intensity I was then determined from an integral equation involving $P_{\tilde{I}_0}$. However, the discussion in Wilson *et al.* considered only intermittency in the velocity field. In this note, the combined effect of intermittency in both the velocity and temperature fields is considered. It is found that inclusion of temperature intermittency does not cause significant deviations from the log-normal pdf for $P_{\tilde{I}_0}$. Hence the final results for the scattered intensity pdf P_I in Wilson *et al.* remain valid. [S0001-4966(97)05105-9]

PACS numbers: 43.28.Fp, 43.20.Fn [LCS]

In a previous paper by Wilson *et al.*,¹ the effect of turbulent intermittency (the tendency of turbulence to occur in bursts of activity) on scattering into an acoustic shadow zone was discussed. For simplicity, only intermittency in the turbulent velocity field was considered. However, since temperature fluctuations are also known to play an important role in the scattering of acoustic waves by atmospheric turbulence, it is necessary to explore the combined effect of intermittency in the velocity and temperature fields. That is the purpose of this note.

We begin by defining the quantity \tilde{I}_0 as the conditionally sampled mean scattered intensity for particular values of the dissipation rates of turbulent kinetic energy $\tilde{\epsilon}$ and temperature variance $\tilde{\chi}$. The tildes indicate *locally* averaged quantities for the scattering volume. Notationally the definition of \tilde{I}_0 is

$$\tilde{I}_0 = \langle I | \tilde{\epsilon}, \tilde{\chi} \rangle, \quad (1)$$

where I is the scattered intensity, the angle brackets indicate ensemble mean, and the vertical bar indicates conditional sampling. We call \tilde{I}_0 the *local mean scattered intensity*. In the earlier paper we neglected temperature variance, and simply had $\tilde{I}_0 = \langle I | \tilde{\epsilon} \rangle$. Regardless of whether the scattering results from velocity or temperature fluctuations, or both, the following integral equation can still be used to determine the probability density function (pdf) $P_I(I)$ for the unconditionally sampled, scattered intensity I :

$$P_I(I) = \int_0^\infty P_C(I | \tilde{I}_0) P_{\tilde{I}_0}(\tilde{I}_0) d\tilde{I}_0, \quad (2)$$

where $P_{\tilde{I}_0}(\tilde{I}_0)$ is the pdf of \tilde{I}_0 , and $P_C(I | \tilde{I}_0)$ is the probability of measuring I given a particular value of the scattering volume dependent quantity \tilde{I}_0 .

Most results from classical scattering theory are easily adapted to intermittent scattering, simply by reinterpreting

quantities such as the mean scattered intensity and scattering cross section as local properties of the scattering volume. Specifically, since classical scattering theory predicts an exponential pdf for the intensity, we have

$$P_C(I | \tilde{I}_0) = \frac{1}{\tilde{I}_0} \exp\left(-\frac{I}{\tilde{I}_0}\right). \quad (3)$$

Furthermore, the local mean scattered intensity is given by

$$\tilde{I}_0 = I_i \tilde{\sigma} V / R, \quad (4)$$

where I_i is the incident intensity, $\tilde{\sigma}$ is the local scattering cross section, V is the size of the scattering volume, and R is the distance from the receiver to the center of the scattering volume. The scattered intensity depends on both temperature and velocity fluctuations, as well as the angle of the scattering. For inertial subrange scattering, we can write

$$\tilde{I}_0(\theta) = a_T(\theta) \tilde{C}_T^2 + a_V(\theta) \tilde{C}_V^2. \quad (5)$$

In the above, θ is the angle of the scattering, and \tilde{C}_T^2 and \tilde{C}_V^2 are the local structure function parameters for temperature and velocity, respectively. The reader is referred to Ostashev's² Eq. (33) for the specific forms of the proportionality factors a_V and a_T . According to Kolmogorov's refined similarity hypothesis,³ the structure function parameters are related to the dissipation rates by

$$\tilde{C}_V^2 = b_V \tilde{\epsilon}^{2/3}, \quad \tilde{C}_T^2 = b_T \tilde{\epsilon}^{-1/3} \tilde{\chi}, \quad (6)$$

where b_V and b_T are constants,⁴ approximately equal to 2.

Since we know $P_C(I | \tilde{I}_0)$ from the classical theory, the remaining difficulty in determining the pdf for intensity lies in finding an expression for $P_{\tilde{I}_0}(\tilde{I}_0)$. Our starting point is the assumption that $\tilde{\epsilon}$ and $\tilde{\chi}$ have a joint, log-normal distribution, given by⁵

$$P_{\mathbf{X}}(\mathbf{x}) = \frac{1}{2\pi(\det \mathbf{K}_{\mathbf{X}})^{1/2}} \times \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{K}_{\mathbf{X}}^{-1}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})\right], \quad (7)$$

where

$$\mathbf{x} = \begin{bmatrix} \ln \tilde{\epsilon} \\ \ln \tilde{\chi} \end{bmatrix}, \quad \boldsymbol{\mu}_{\mathbf{X}} = \begin{bmatrix} \langle \ln \tilde{\epsilon} \rangle \\ \langle \ln \tilde{\chi} \rangle \end{bmatrix}, \quad (8)$$

$$\mathbf{K}_{\mathbf{X}} = \langle (\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})^T \rangle = \begin{bmatrix} \sigma_{\tilde{\epsilon}}^2 & \sigma_{\tilde{\epsilon}}\sigma_{\tilde{\chi}}\rho \\ \sigma_{\tilde{\epsilon}}\sigma_{\tilde{\chi}}\rho & \sigma_{\tilde{\chi}}^2 \end{bmatrix},$$

$\sigma_{\tilde{\epsilon}}^2$ and $\sigma_{\tilde{\chi}}^2$ are the variances of $\ln \tilde{\epsilon}$ and $\ln \tilde{\chi}$, respectively, and ρ is the correlation coefficient. As discussed in Antonia *et al.*,⁶ $\sigma_{\tilde{\epsilon}}^2 \approx \sigma_{\tilde{\chi}}^2 \approx 0.25 \ln(\ell/r)$, where ℓ is the integral length scale of the turbulence and r is a characteristic dimension of the averaging (scattering) volume, and $\rho \approx 0.5$.

By making the linear transformation

$$\mathbf{y} = \begin{bmatrix} \ln \tilde{C}_V^2 \\ \ln \tilde{C}_T^2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & 0 \\ -\frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} \ln \tilde{\epsilon} \\ \ln \tilde{\chi} \end{bmatrix} \quad (9)$$

we arrive at the joint pdf for the log structure-function parameters, $P_{\mathbf{Y}}(\mathbf{y})$. It has the same form as Eq. (7), but with different mean and covariance matrices:

$$\boldsymbol{\mu}_{\mathbf{Y}} = \begin{bmatrix} \ln b_V + \frac{2}{3}\langle \ln \tilde{\epsilon} \rangle \\ \ln b_T - \frac{1}{3}\langle \ln \tilde{\epsilon} \rangle + \langle \ln \tilde{\chi} \rangle \end{bmatrix} = \begin{bmatrix} \mu_V \\ \mu_T \end{bmatrix},$$

and

$$\mathbf{K}_{\mathbf{Y}} = \begin{bmatrix} \frac{4}{9}\sigma_{\tilde{\epsilon}}^2 & -\frac{2}{9}\sigma_{\tilde{\epsilon}}^2 + \frac{2}{3}\sigma_{\tilde{\epsilon}}\sigma_{\tilde{\chi}}\rho \\ -\frac{2}{9}\sigma_{\tilde{\epsilon}}^2 + \frac{2}{3}\sigma_{\tilde{\epsilon}}\sigma_{\tilde{\chi}}\rho & \frac{1}{9}\sigma_{\tilde{\epsilon}}^2 - \frac{2}{3}\sigma_{\tilde{\epsilon}}\sigma_{\tilde{\chi}}\rho + \sigma_{\tilde{\chi}}^2 \end{bmatrix} \\ = \begin{bmatrix} \sigma_V^2 & \sigma_V\sigma_T\rho' \\ \sigma_V\sigma_T\rho' & \sigma_T^2 \end{bmatrix}. \quad (10)$$

Having found the joint pdf of the log structure function parameters, the problem now becomes one of finding the pdf of the weighted sum (\tilde{I}_0) of two random variables (\tilde{C}_V^2 and \tilde{C}_T^2) whose joint pdf is known. This situation occurs frequently in probability theory and has been well studied. (The integral that needs to be evaluated is given, for example, as Eq. (2.10–13) in Stark and Woods.⁷) Unfortunately the resultant pdf for the sum of two log-normally distributed random variables is unknown in closed form. This problem is discussed in detail by Mityagin and McCulloch,⁸ who prove, in particular, that the sum of two log-normal r.v.'s is not itself log-normal.

Since a closed-form solution is unknown we seek reasonable approximations. The simplest starting point would be to fit the actual pdf for \tilde{I}_0 , $P_{\tilde{I}_0}$, with a log-normal curve. We know this will work well in cases where either the velocity or the temperature fluctuations alone dominate; it is the intermediate cases where the validity of such an approximation must be assessed.

To fit $P_{\tilde{I}_0}$ with a log-normal curve, we need values for the mean $\langle \ln \tilde{I}_0 \rangle$ and variance σ^2 parameters of the distribu-

tion. As shown in Wilson *et al.*,¹ the two are related through the equation $\ln \langle \tilde{I}_0 \rangle = \langle \ln \tilde{I}_0 \rangle + \sigma^2/2$. Since $\langle \tilde{I}_0 \rangle = \langle I \rangle$ is taken as a known quantity, we must actually determine only $\langle \ln \tilde{I}_0 \rangle$.

Formally, $\langle \ln \tilde{I}_0 \rangle$ can be evaluated from the integral

$$\langle \ln \tilde{I}_0 \rangle = \int_{-\infty}^{\infty} \ln(a_V e^{y_1} + a_T e^{y_2}) P_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y}, \quad (11)$$

where $y_1 = \ln \tilde{C}_V^2$ and $y_2 = \ln \tilde{C}_T^2$ [Eq. (9)]. This integral can be approximated using a two-dimensional, steepest descents method. The evaluation is made easier when we first make the transformation

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \sigma_V^2 & \sigma_V^2 \\ -\sigma_T^2 & \sigma_V^2 \end{bmatrix} (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{Y}}). \quad (12)$$

The integral (11) then becomes

$$\langle \ln \tilde{I}_0 \rangle = \int_{-\infty}^{\infty} F(u_1, u_2) P_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}, \quad (13)$$

where

$$P_{\mathbf{U}}(\mathbf{u}) = \frac{1}{2\pi\sigma_V^2\sigma_T^2\sqrt{1-\rho'^2}} \times \exp\left[-\frac{u_1^2(1-\rho') + u_2^2(1+\rho')}{2\sigma_V^2\sigma_T^2(1-\rho'^2)}\right], \quad (14)$$

$$F(u_1, u_2) = \ln\left[a'_V \exp\left(\frac{u_1 - u_2}{\sqrt{2}\sigma_T}\right) + a'_T \exp\left(\frac{u_1 + u_2}{\sqrt{2}\sigma_V}\right)\right], \quad (15)$$

and $a'_V = a_V e^{\mu_V}$, $a'_T = a_T e^{\mu_T}$. We now expand F as a Taylor series about the saddle point $(u_1, u_2) = (0, 0)$:

$$F(u_1, u_2) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} F_{mn} u_1^m u_2^n. \quad (16)$$

Since $P_{\mathbf{U}}(\mathbf{u})$ is an even function, only the even powers in the expansion for F survive the integration. Hence the approximation

$$\langle \ln \tilde{I}_0 \rangle \approx \int_{-\infty}^{\infty} (F_{00} + F_{20}u_1^2 + F_{02}u_2^2) P_{\mathbf{U}}(\mathbf{u}) d\mathbf{u} \\ \approx F_{00} + F_{20}\sigma_V^2\sigma_T^2(1+\rho') + F_{02}\sigma_V^2\sigma_T^2(1-\rho'). \quad (17)$$

The expansion coefficients F_{00} , F_{20} , and F_{02} are easily found by differentiating (15). The result is

$$\langle \ln \tilde{I}_0 \rangle \approx \ln(a'_V + a'_T) \\ + (\sigma_V^2 - 2\rho'\sigma_V\sigma_T + \sigma_T^2) \frac{a'_V a'_T}{2(a'_V + a'_T)^2}. \quad (18)$$

As discussed earlier, experimental data suggest that $\sigma_{\tilde{\epsilon}}^2 \approx \sigma_{\tilde{\chi}}^2$, and $\rho \approx 0.5$. Using Eq. (10), the above equation simplifies to

$$\langle \ln \tilde{I}_0 \rangle \approx \ln(a'_V + a'_T) + \frac{a'_V a'_T}{2(a'_V + a'_T)^2}. \quad (19)$$

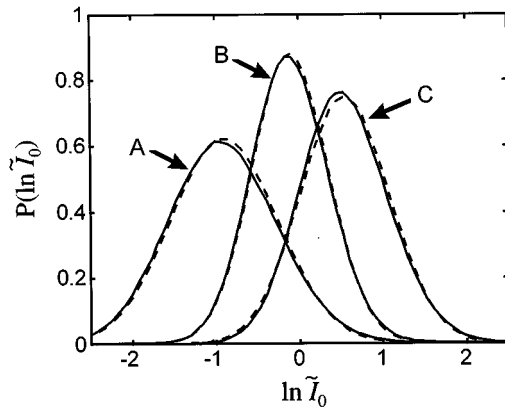


FIG. 1. Comparison of the exact probability density function for $\ln \tilde{I}_0$ (solid line) with the approximation developed in the text (dashed line). In case A the ratio of the integral length scale ℓ of the turbulence to the characteristic dimension of the scattering volume r is 100; in cases B and C the ratio is 10. Cases A and B have equal contributions from velocity and temperature scattering; in case C temperature scattering dominates.

Three test cases of this approximation for $P_{\tilde{I}_0}$ are shown in Fig. 1. For case A, the ratio ℓ/r was set to 100 ($\sigma_\epsilon^2 \approx \sigma_\chi^2 \approx 1.2$), while for B and C this ratio was set to 10 ($\sigma_\epsilon^2 \approx \sigma_\chi^2 \approx 0.56$). (Recall from earlier that ℓ is the integral length scale of the turbulence, and r is a characteristic dimension of the scattering volume; intermittency is important when $\ell > r$.) Hence case A has the largest intermittency effect. In cases A and B the ratio $a_T \langle \tilde{C}_T^2 \rangle / a_V \langle \tilde{C}_V^2 \rangle$ was set to 1, whereas in case C this ratio was 3. Hence cases A and B feature equal contributions from velocity and temperature fluctuations to the mean scattering, whereas C has a greater temperature contribution. The scattering angle θ was fixed at 20° in each case, with 0° being forward scattering and 180° pure backscattering. For the purpose of offsetting the pdf curves, different values for the mean scattered intensity were used in each case ($\langle I \rangle = 0.5$ for A, 1 for B, and 2 for C).

The “exact” pdf curves in Fig. 1 were determined by generating 10^7 random numbers pairs ($x_1 = \ln \tilde{\epsilon}, x_2 = \ln \tilde{\chi}$) that followed a joint-normal distribution. (Such a Monte Carlo method is much easier than the alternative of performing a numerical integration. The pdf curves produced by this method are not entirely smooth, however, due to the finite sample size.) In all three cases, the log-normal approximation curves, with mean chosen according to Eq. (19), are barely distinguishable from the exact pdf’s. The variances $\sigma^2 = 2(\ln \langle I \rangle - \langle \ln \tilde{I}_0 \rangle)$ in the approximate distributions are 0.41, 0.21, and 0.29, for cases A–C, respectively. Agree-

ment is observed to be best for small values of the variance parameter σ^2 . The actual values for a'_V in cases A–C were 0.19, 0.44, and 0.44, respectively. The corresponding values for a'_T were 0.16, 0.40, and 1.20.

Besides the test cases shown in Fig. 1, many others were tried, and in each instance there was still excellent agreement between the exact and approximate results. Although the approximation does become worse with increasing intermittency effect (variance parameter σ^2), for practical values of the ratio ℓ/r the log-normal, saddle-point approximation is quite satisfactory.

In conclusion, the pdf for $P_{\tilde{I}_0}$ for coupled, intermittent temperature and velocity fluctuations is log-normal to a very good degree of approximation. Since a log-normal pdf was used for $P_{\tilde{I}_0}$ in Wilson *et al.*,¹ incorporation of temperature intermittency into the formulation has no practical effect on the main result of that paper, namely the pdf for the actual scattered intensity P_I . The main difference when we include temperature intermittency is that the mean and variance parameters appearing in the log-normal approximation for $P_{\tilde{I}_0}$ have a more complicated functional dependence on the turbulent velocity and temperature statistics, as given by (19).

ACKNOWLEDGMENTS

This work was partially supported by the ONR Marine Boundary Layer ARI, Grant No. N00014-92-J-1688. The author also thanks Walter Bach of the Army Research Office for his support.

- ¹D. K. Wilson, J. C. Wyngaard, and D. I. Havelock, “The effect of turbulent intermittency on scattering into an acoustic shadow zone,” *J. Acoust. Soc. Am.* **99**, 3393–3400 (1996).
- ²V. E. Ostashev, “Sound propagation and scattering in media with random inhomogeneities of sound speed, density and medium velocity,” *Waves Random Media* **4**, 403–428 (1994).
- ³A. N. Kolmogorov, “A refinement of previous hypothesis concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number,” *J. Fluid Mech.* **13**, 82–85 (1962).
- ⁴R. J. Hill, “Acoustic scattering cross section predicted from an accurate model of the spectrum of temperature fluctuations,” *J. Acoust. Soc. Am.* **65**, 1397–1401 (1979).
- ⁵C. W. Van Atta, “Influence of fluctuations in local dissipation rates on turbulent scalar characteristics in the inertial subrange,” *Phys. Fluids* **14**, 1803–1804 (1971).
- ⁶R. A. Antonia, E. J. Hopfinger, Y. Gagne, and F. Anselmetti, “Temperature structure functions in shear flows,” *Phys. Rev. A* **13**, 2704–2707 (1984).
- ⁷H. Stark and J. W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers* (Prentice-Hall, Englewood Cliffs, NJ, 1986).
- ⁸B. Mityagin and H. McCulloch, “Distributional closure of financial portfolio returns,” in *Proceedings of the International Workshop in Analysis and its Applications*, edited by C. V. Stanojevic and O. Hadzic (Novi Sad, Yugoslavia, 1991).

Acoustic scattering by a pair of spheres: Addenda and corrigenda

Guillermo C. Gaunaurd^{a)} and Hanson Huang^{b)}

Naval Surface Warfare Center, White Oak, Silver Spring, Maryland 20903-5640

Hans C. Strifors

National Defense Research Establishment (FOA 6), S-17290 Stockholm, Sweden

(Received 11 August 1996; accepted for publication 15 December 1996)

Our earlier work on acoustic scattering by multiple bodies is being updated by summarizing recent extensions in several new publications. Two figures that were not entirely correct in one of our earlier papers have been replotted and pertinent new physical interpretations have been added. [S0001-4966(97)02805-1]

PACS numbers: 43.30.Ft, 43.20.Fn [JHM]

I. ADDENDA

Our study of sound scattering by multiple bodies started with two impenetrable (i.e., rigid, soft) spheres of arbitrary size, that were insonified by plane acoustic waves¹ in an arbitrary direction. This work was later extended to the case

of two elastic spherical shells.² The shell motions were described by the standard bending theory. Hence, the dynamic response of the pair could only be analyzed in relatively narrow frequency bands, i.e., $0 \leq \Omega \equiv ka \leq 25$. This was followed by a study of the scattering response of a single fluid

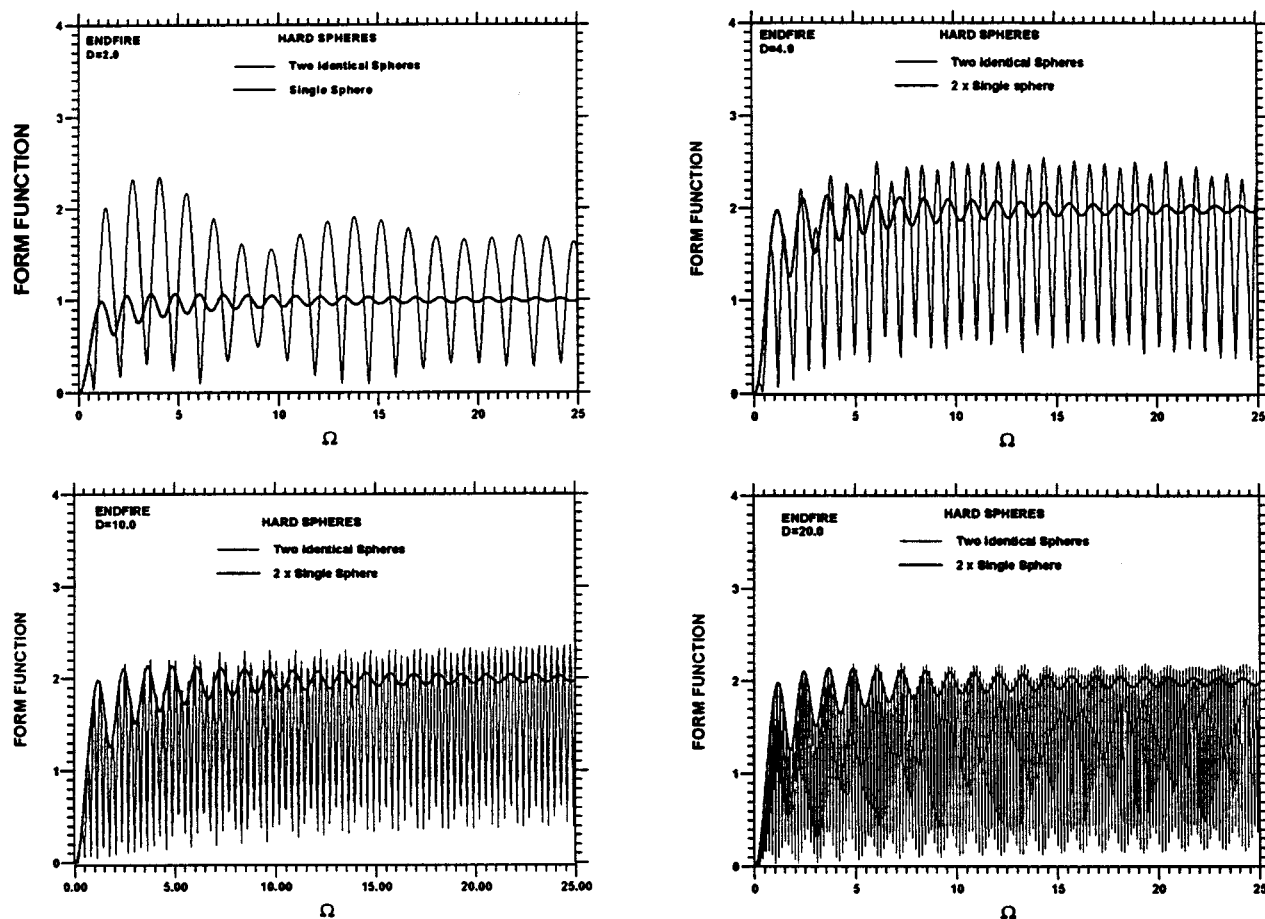


FIG. 1. Exact form functions for two acoustically rigid spheres at end-fire incidence (i.e., $\alpha=0$) vs Ω , for $D=2, 4, 10$, and 20 (thick lines). The thinner curves are for two noninteracting spheres (i.e., the Born approximation case).

^{a)}Carderock Division.

^{b)}Indian Head Division.

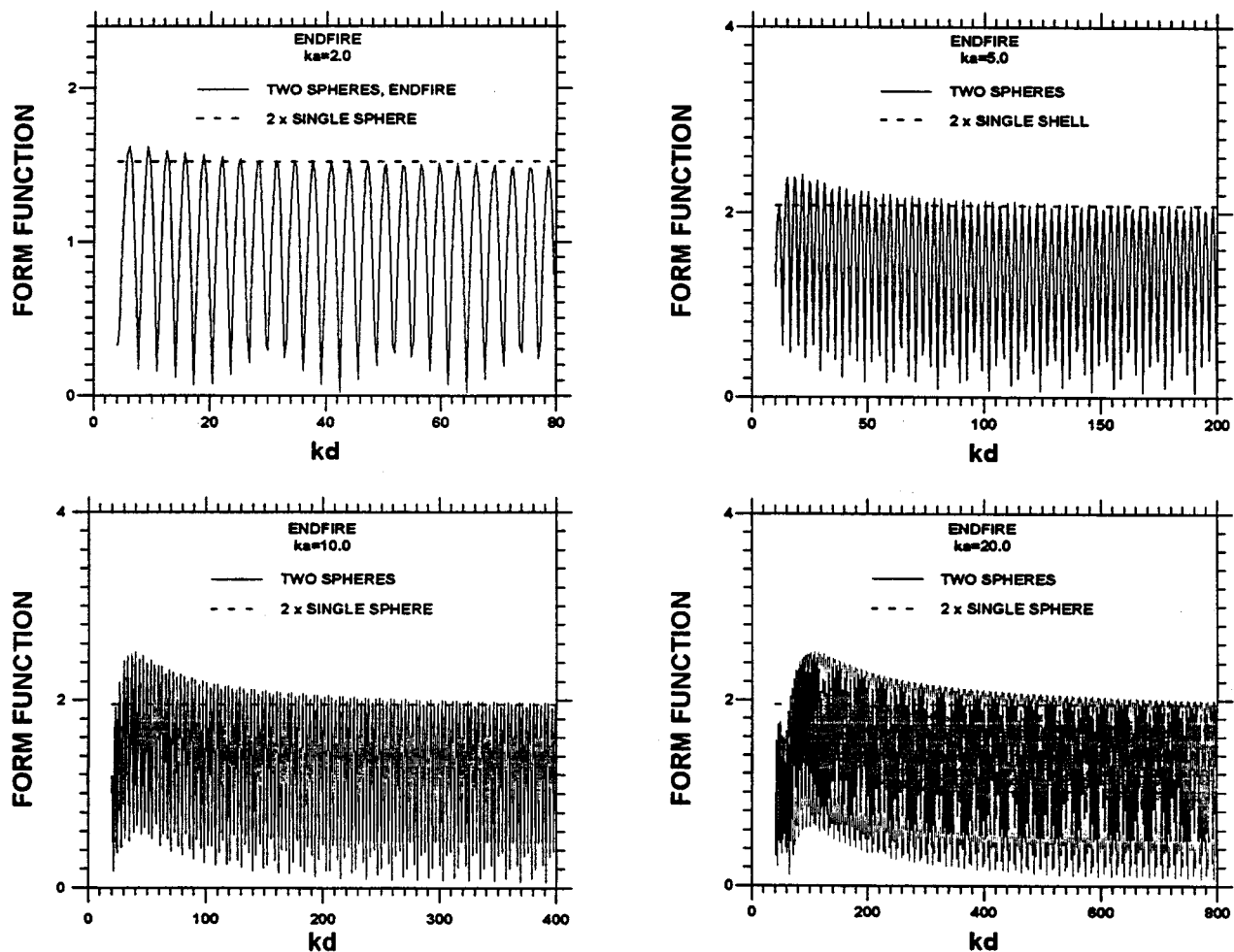


FIG. 2. Exact form functions for two acoustically rigid spheres at end-fire incidence (i.e., $\alpha=0$) vs kd (or ΩD), for $D=2, 4, 10$, and 20 (thick lines). The thinner curves are for two noninteracting spheres at the particular value of Ω used in each plot.

sphere (i.e., an ideal bubble) near the sea surface.³ The method of images facilitated the analysis since then the bubble and its image, introduced by the boundary, formed a pair of scatterers that could be handled by our exact multiple-scattering formulation. This work was later extended to the case of a single-elastic spherical shell near the sea surface.⁴ The method of images and the addition theorem for the spherical wave functions were used in these analyses. If the scattering response of the pair of shells was desired beyond the relatively narrow frequency band that it is possible to analyze with the standard bending shell theory (i.e., such as in Ref. 2), then the shell's motions have to be described by the general three-dimensional theory of elasticity. This more general situation has also been recently published.⁵

II. CORRIGENDA

Taking into account Eqs. (26)–(28) of Ref. 1, the far-field form of Eq. (29) can be rewritten as follows:

$$\phi_{sc}^{A+B} = \frac{e^{ikr}}{kr} \sum_{q=0}^{\infty} \sum_{p=0}^q i^{-(q+1)} \left[b_{pq} P_q^p(\cos \theta) + c_{pq} \frac{r}{r_1} e^{ik(r_1-r)} P_q^p(\cos \theta_1) e^{ikd \cos \alpha} \right] e^{ip\varphi}. \quad (1)$$

In that early notation, $r_1=r'$, $\theta_1=\theta'$, and the factor $\exp(ikd \cos \alpha)$ was included in the definition of the coupling coefficient c_{pq} . We point out that there is a ‘typo’ in the upper limit of the p -sum, in Eqs. (26) and (29) of Ref. 1 which is q , instead of ∞ . Using the normalizations: $R \equiv r/a$, $D \equiv d/a$, $\Omega \equiv ka$, $R' \equiv r'/a = r_1/a$, and the definition of the form function, viz.,

$$|f_{\infty}| = \left| \frac{2r}{a} \phi_{sc}^{A+B} e^{-i\Omega R} \right|, \quad (2)$$

we obtain the final result:

$$|f_{\infty}| \equiv \frac{2}{\Omega} \left| \sum_{q=0}^{\infty} \sum_{p=0}^q i^{-(q+1)} \left[b_{pq} P_q^p(\cos \theta) + c_{pq} \frac{R}{R'} e^{i\Omega(R'-R)} P_q^p(\cos \theta') \right] e^{ip\varphi} \right|, \quad (3)$$

which is the quantity plotted in the figures. However, Figs. 6 and 8 of Ref. 1 require some modification/correction. There is a phase factor incorrectly programmed in Ref. 1. This affects only these end-fire (i.e., $\alpha=0$) plots. Figures 1 and 2 replace and correct the previous two figures, which now display roughly twice as many oscillations as they showed in the displayed band, in Ref. 1.

III. ADDENDA ON INTERPRETATION OF RESULTS

It is now possible to give additional physical interpretation. In Fig. 1, for $D=2$, i.e., touching spheres, there is a slight amplitude amplification at low Ω values. This is a focusing effect. The front sphere focuses the sound energy toward the back sphere, which then reflects it back to the observer amplified, as if a lens had been placed in front of a mirror. The oscillation amplitude of the pair exceed twice those of the single sphere. As Ω increases, still for $D=2$, a shielding effect develops. The back sphere starts to be eclipsed by the front one. The amplitude of the pair is less than twice that of the single. In fact, for sufficiently high Ω , outside the band displayed, the amplitude of the pair decreases toward the value of the single, and the presence of the back sphere will no longer be acoustically noticeable.

At larger separations (i.e., Fig. 1, for $D=20$) the interaction is seen to diminish. At low Ω values, now there is hardly any amplification due to focusing. The eclipsing of the back sphere by the front one goes unnoticed in the displayed band, i.e., $0 \leq \Omega \leq 25$, since much higher Ω values are required to make it clearly visible. The amplitude oscillations in Fig. 1 are contained within an upper and a lower envelope. For the present end-fire situation, these envelopes are given by

$$|f^{\text{env}}| = |f_{\infty}^1| \left(1 \pm \frac{R}{R' + D} \cdot e^{2i\Omega D} \right), \quad (4)$$

where $|f_{\infty}^1|$ is the familiar form function of a single sphere. This result can also be obtained from the asymptotic expression in Eq. (40) of Ref. 1.

The eclipsing effect is not easily seen in Fig. 2 since the condition to be met for it to be noticeable (i.e., Ω large, and D small) is not sufficiently satisfied within the range of the graphs in Fig. 2. However, the same previous observations about the focusing effect are also contained in Fig. 2. The four graphs in Fig. 2 all start at abscissas of value $2ka$ (i.e., touching spheres). As $ka (\equiv \Omega)$ increases from one graph to the next, the amplitude of the pair exhibit progressively larger amplifications in the left portions of the graphs. In brief, for both amplification *and* for eclipsing of the back sphere, it is necessary that the separation, D , be small. But, in addition, for amplification, Ω should be small, while for shielding or eclipsing, it should be large.

¹G. C. Gaunard, H. Huang, and H. C. Strifors, "Acoustic scattering by a pair of spheres," J. Acoust. Soc. Am. **98**, 495–507 (1995).

²H. Huang and G. C. Gaunard, "Acoustic scattering of a plane wave by two spherical elastic shells," J. Acoust. Soc. Am. **98**, 2149–2156 (1995).

³G. C. Gaunard and H. Huang, "Sound scattering by an air bubble near the sea surface," IEEE J. Oceanic Eng. **20**, 285–292 (1995).

⁴H. Huang and G. C. Gaunard, "Scattering of a plane acoustic wave by a spherical elastic shell near the free surface," Int. J. Solids Struct. **34** (5), 591–602 (1997).

⁵H. Huang and G. C. Gaunard, "Acoustic scattering of a plane wave by two spherical elastic shells above the coincidence frequency," J. Acoust. Soc. Am. **98**, 2149–2156 (1996).

Antisymmetric nature of the earliest arrivals of ultrasound propagating in copy paper

Mont A. Johnson^{a)} and Yves H. Berthelot

Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0405

(Received 22 October 1996; accepted for publication 16 January 1997)

Experiments with laser generation and detection of ultrasound in copy paper indicate that earliest arrivals of ultrasound are antisymmetric. A likely explanation is that the out-of-plane interferometer is relatively insensitive to the symmetric compressional pulse because of the low Poisson ratio of copy paper, but rather sensitive to its antisymmetric flexural wake associated with the random heterogeneities and gradients present in copy paper. © 1997 Acoustical Society of America. [S0001-4966(97)05705-6]

PACS numbers: 43.35.Cg, 43.35.Yb, 43.35.Zc [HEB]

INTRODUCTION

Laser ultrasonics is a promising technique for noninvasive, noncontact, on-line monitoring of the physical characteristics of paper during its manufacturing.¹ Even though paper is a fibrous, complex medium, it is usually modeled as an orthotropic plate,^{2,3} i.e., a medium in which the stress-strain relation is fully described by nine elastic constants, C_{ij} 's. When the frequency thickness is sufficiently low that only the two lowest Lamb modes propagate in the orthotropic plate, the fast compressional mode is nondispersive (with a wave speed characterized by either C_{11} and C_{22} depending on the direction of propagation) and the slow flexural mode is characteristically dispersive. In an orthotropic plate, the lowest-order fast mode is symmetric (denoted usually by s_0) and the slow mode is antisymmetric (denoted by a_0). However, in this letter, we report laser interferometric measurements made simultaneously on opposite sides of copy paper which clearly show that the earliest ultrasonic arrivals are antisymmetric. These results point out the limitation of the orthotropic model when applied to paper.

I. EXPERIMENTS

A Q-switched Nd:Yag laser (5.7-mJ, 10-ns Gaussian pulse duration) was focused to a spot (90- μ m diameter) on the copy paper sample (10 cm by 10 cm, 90 μ m thick) and provided the thermoelastic transient source of ultrasound. A 1-W Ar⁺ laser was used for laser Doppler interferometric detection of the ultrasonic pulse, typically on the side opposite to the laser generation. The interferometer was designed to measure the out-of-plane (normal) component of the surface velocity at the detection point which was focused to approximately 10 μ m in diameter. The system was later modified to also measure the in-plane (tangential) component of the surface velocity. Because of some limitations of our optical arrangement, the signal-to-noise ratio was at least 10 dB better with the out-of-plane detection system than with the in-plane one. The Doppler shifted signal was demodulated by means of an FM discriminator, low-pass filtered

with a cutoff frequency of 5 MHz, and displayed on a digital oscilloscope interfaced with a computer. For more details on the experimental apparatus, the reader is referred to Ref. 1. Figure 1 shows three configurations that were investigated. In the first configuration [Fig. 1(a)], the signal was generated on one side of the paper and both the in-plane and the out-of-plane interferometers were used to detect the signals on the opposite side of the paper. For simple plate theory, the ratio of the amplitude of the out-of-plane component to that of the in-plane component, denoted here by $|w|/|u|$, is $\frac{1}{2}\mu(kh)$, where μ is the Poisson ratio, and kh is the wave number-thickness product. As indicated later, reasonable values for our experiment are $\mu=0.15$, $f=1$ MHz, and the wave speed of the earliest arrivals is $c=3.6$ km/s. These values lead to a ratio $|w|/|u|=0.012$. Therefore, we expected the amplitude of the first arrivals to be significantly larger when detected with the in-plane interferometer instead of the out-of-plane system. However, the ratio $|w|/|u|$ was measured to be about 0.4, a very high value indeed considering the extremely low value of the Poisson ratio for copy paper. To force more energy into the symmetric fast mode, we investigated the configuration shown in Fig. 1(b) where the generation laser beam is split into two beams that are focused on opposite sides of the paper. Surprisingly, this arrangement did not significantly increase the amplitude of the fast wave, thus casting some doubts about the supposedly symmetric nature of the first arrivals. This led to an experiment [Fig. 1(c)] in which the laser generation was on one side but where the signal was detected simultaneously on both side of the sample. As indicated in Fig. 2, this experiment clearly revealed the antisymmetric nature of the first arrivals. The two signals shown in Fig. 2(a) were detected on opposite sides of the paper but, for clarity, one of the signals has been inverted so as to illustrate the antisymmetry of the entire waveform. As expected, the slow dispersive wave [clipped in Fig. 2(a), starting at time $t=10$ μ s] is antisymmetric. However, the fast wave (2–8 μ s) is also antisymmetric. (The initial symmetric part between $t=0$ and $t=2$ μ s represents the electromagnetic pickup from the laser discharge.) The waveforms shown in Fig. 2(a) were measured in the fiber direction ("machine direction") and the source receiver distance was 7.0 ± 0.5 mm. Similar results were ob-

^{a)}Present address: Measurex Corporation, One Results Way, Cupertino, CA 95014.

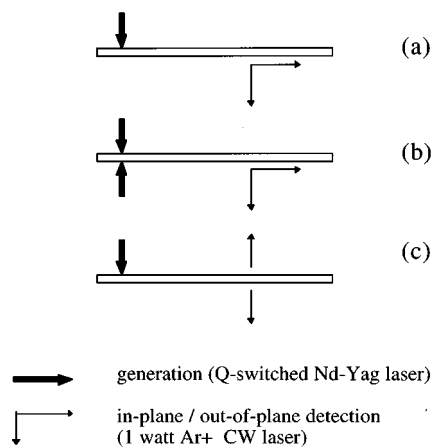


FIG. 1. Schematic of the experimental arrangements.

tained at several points, and also in the cross-fiber direction. Therefore, it appears that the fast mode is antisymmetric, and this result is not very sensitive to slight misalignment of the two opposite probes. To validate the method, the experiment was repeated with a 400- μm -thick aluminum plate. The result is shown in Fig. 2(b). As expected, the fast mode s_0 is symmetric and the slow mode a_0 is antisymmetric.

II. VELOCITY OF THE FAST MODE

In-plane and out-of-plane waveforms were collected for various source/receiver separation distances. The measured signals were time-windowed so as to isolate the fast mode from the slow mode. The windowed fast mode was shown to be nondispersive at least up to 2 MHz. The phase velocity of the fast mode was measured by means of a cross-correlation technique.¹ Table I shows the measured phase velocities obtained from the out-of-plane and the in-plane measurements, in either the machine direction (MD) or the cross direction (CD). Table I also shows the results obtained from independent contact measurements made at the Institute of Paper Science and Technology (IPST). Clearly, the out-of-plane and the in-plane interferometers detect a wave that propagates at the same velocity, within experimental errors. Also, a Fourier analysis of the waveforms reveals that the fast waves detected with in-plane and out-of-plane detectors are both in the same frequency range (peak at 1 MHz with energy spread mostly between 0.3 and 3 MHz). Yet, most surprisingly, the results indicate that the waveform detected with the in-plane interferometer is a pulse (Fig. 3) whereas that detected with the out-of-plane interferometer is more like a wave packet (Fig. 4). This interesting feature illustrates the limitation of orthotropic plate theory for analyzing ultrasound in paper.

III. DISCUSSION

Copy paper is a very anisotropic and heterogeneous medium. It is made up of 20- to 40- μm -thick wood fibers that are 1–2.5 mm in length. The paper thickness is about 90 μm (i.e., 3–5 fibers across the thickness). The bond in the thickness direction is very loose, as indicated by the very low

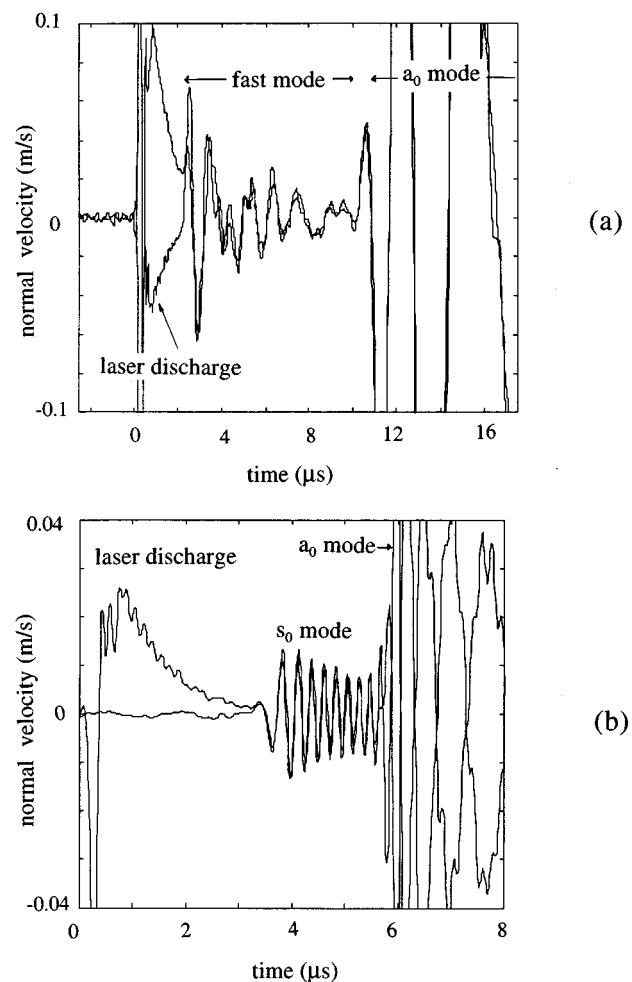


FIG. 2. (a) Signals detected with the experimental arrangement shown in Fig. 1(c). The source/receiver separation is 7 ± 0.5 mm along the machine direction (MD). To illustrate the antisymmetric nature of both the fast and the slow mode, the signal detected from the side opposite the generation point has been inverted. (b) Signals detected with the experimental arrangement shown in Fig. 1(c) when the sample is a 400- μm -thick aluminum plate. The source/receiver separation is 19.0 ± 0.5 mm.

compressional sound speed across the thickness direction (typically 300 m/s, i.e., a factor of 10 below the compressional sound speed in the axial direction). This wave speed corresponds to a thickness-compressional resonance around 1.6 MHz which is in the frequency range of interest for our experiment. Some caution is needed when interpreting the data. Besides the very compliant nature of the paper in the thickness direction, the Poisson ratio of copy paper is very low, typically 0.1–0.2. If paper was an orthotropic material,

TABLE I. Comparison of the phase velocities of the fast mode measured with the out-of-plane interferometer, the in-plane interferometer, and an independent contact method.

	Out-of-plane		In-plane		IPST contact	
	Phase velocity (m/s)	Standard deviation (m/s)	Phase velocity (m/s)	Standard deviation (m/s)	Phase velocity (m/s)	Standard deviation (m/s)
MD	3591	73	3610	97	3550	68
CD	2588	58	2596	74	2578	33

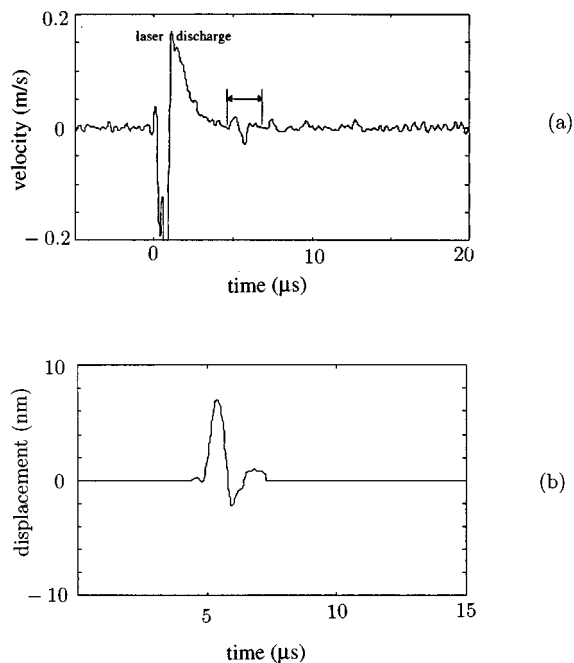


FIG. 3. Signal detected with the in-plane interferometer (source/receiver separation = 13.0 ± 0.5 mm). (a) Velocity; (b) displacement obtained by integrating the windowed signal shown in (a) to isolate the first arrivals.

such a low Poisson ratio would indicate that the fast symmetric wave has a very small out-of-plane component, which contradicts our experimental results. Two mechanisms might be responsible for the antisymmetric nature of the fast mode: strong property gradients in the thickness direction, and random heterogeneities in both the normal and the axial directions.⁴ If the paper has strong property gradients in the thickness direction and negligible random variations in the axial direction, one can show that the fast wave carries a small out-of-plane component which has a significant antisymmetric part superimposed onto the small symmetric part caused by the Poisson effect. If the gradients are strongly antisymmetric across the paper thickness, the antisymmetric part may dominate the symmetric component. If, in addition to strong gradients in the thickness direction, the paper has some significant random heterogeneities in both the thickness and the axial directions, one can show that a longitudinal wave sheds random incoherent antisymmetric flexural waves, leaving an incoherent antisymmetric wake behind the longitudinal pulse.⁴ Therefore, it appears quite plausible that the pulse detected with the in-plane interferometer corresponds to the compressional (symmetric) fast wave but that the wave packet detected with the out-of-plane interferometer corresponds to the antisymmetric flexural wake of the compressional mode, a wake that is caused by the random heterogeneities and gradients within the material. These mechanisms may explain the experimental results presented here, but it will be necessary to quantify the microstructure of the paper more precisely to make quantitative comparisons with the experimental data.

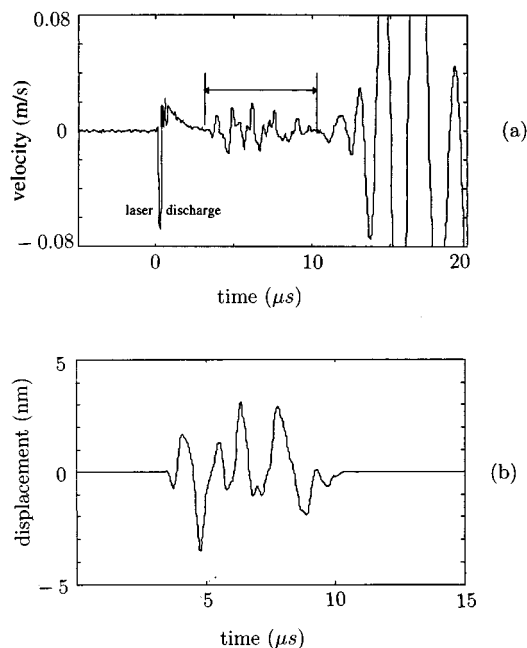


FIG. 4. Signal detected with the out-of-plane interferometer (source/receiver separation = 8.0 ± 0.5 mm). (a) Velocity; (b) displacement obtained by integrating the windowed signal shown in (a) to isolate the first arrivals.

IV. SUMMARY

Lamb waves have been generated in copy paper with a *Q*-switched laser and detected either with an out-of-plane or an in-plane laser Doppler interferometer, either on one side or simultaneously on opposite sides of the paper. It was established that, as expected from orthotropic plate theory: (1) the fast mode is nondispersive; (2) the measured phase velocity of the fast mode agrees whether one measures the waveform with the in-plane or the out-of-plane interferometer; and in addition, (3) the measured phase velocity agrees with independent contact measurements. However, in contradiction with expectations from orthotropic plate theory: (1) the earliest arrivals detected with the out-of-plane interferometer are antisymmetric; and (2) the earliest arrival detected with the in-plane interferometer is a pulse whereas it is a wave packet when detected with the out-of-plane interferometer. These findings suggest that orthotropic plate theory is an oversimplification of wave propagation in paper and that a more complete theory based on the heterogeneities of copy paper and the property gradients across its thickness is needed for quantitative evaluation of the physical properties of paper.

ACKNOWLEDGMENTS

The authors would like to thank Professor Richard Weaver of the University of Illinois at Urbana-Champaign and Dr. Pierre Brodeur of the Institute of Paper Science and Technology in Atlanta for insightful discussions during the course of this research.

- ¹M. A. Johnson and Y. H. Berthelot, "Laser ultrasonics in copy paper," Opt. Eng. **36**(2), 408–416 (1997).
- ²C. C. Habeger, R. W. Mann, and G. A. Baum, Ultrasonics **17**, 57 (1979).
- ³G. A. Baum, Appita **40**, 288 (1987).
- ⁴R. Weaver, Personal communication (1996). [See also J. A. Turner and R. L. Weaver, "Radiative transfer and multiple scattering of diffuse ultrasound in polycrystalline media," J. Acoust. Soc. Am. **96**, 3675–3683 (1994).]

Prediction of sound-pressure level in an occupied enclosure

S. K. Tang, Daniel W. T. Chan, and K. C. Chan

Department of Building Services Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

(Received 7 June 1995; revised 17 January 1997; accepted 23 January 1997)

In a room with inadequate sound absorption, occupants may suffer from low speech intelligibility due to the high noise level resulting from human activities. For effective communication, they tend to raise their voices. In the present study, a method which takes into account the effect of raising voices is derived to predict the variation of A-weighted sound level inside an enclosure in which the number of occupants changes with time. Site measurements are also performed in a canteen of capacity around 250. The model predicts the variations of A-weighted sound levels with a number of occupants with sufficient accuracy for engineering applications. © 1997 Acoustical Society of America. [S0001-4966(97)05605-1]

PACS numbers: 43.55.Hy, 43.55.Br, 43.50.Jh [JDQ]

INTRODUCTION

When people speak freely in an enclosure, the average A-weighted sound level inside the enclosure increases as the number of people increases. However, this growth in noise level gradually interferes with dialogue, and lowers the speech intelligibility. To preserve the latter, the people have to raise their voices. This is a positive feedback, since the people keep on raising their voices, resulting in a rapid increase in the A-weighted sound level inside the enclosure. This “cocktail party effect” was first discussed by Pollack and Pickett.¹ MacLean² proposed a theory to explain this effect but his model involves grouping of occupants, where in reality the group size cannot be determined. The rise in voice level is also not considered in his theory. Legget and Northwood³ conducted noise surveys in cocktail parties, but no suitable method for predicting A-weighted sound level was found. Very often, this “cocktail party effect” is ignored in practice, resulting in overdesign for absorption or failure to achieve an acceptable acoustical environment.

For the design of an acceptable indoor acoustical environment, the variation of A-weighted sound level with the number of people inside the room is required. In this study, a method is derived to predict the A-weighted sound level in an occupied enclosure. It includes a simple model for elevation of the human voice and the effect of sound absorption inside the enclosure. Site measurements were also conducted to test the proposed A-weighted sound level prediction method.

I. THE SOUND LEVEL PREDICTION MODEL

The A-weighted sound-pressure level (SPL) in an enclosure due to a sound source of power W can be estimated by the following well-known formula:

$$\text{SPL} = 10 \log_{10} \left(\frac{W}{W_0} \right) + 10 \log_{10} \left(\frac{Q}{4\pi r^2} + \frac{4}{R} \right), \quad (1)$$

where $W_0 = 10^{-12}$ W, Q is the directivity of the source, R is the room constant of the enclosure, and r is the distance between the observer and the sound source (for instance, see

Ref. 4). The resultant A-weighted sound level from n sound sources at the position of the j th occupant is thus given by

$$\text{SPL}_j = 10 \log_{10} \left[\sum_{i=1}^n \frac{W_i}{W_0} \left(\frac{Q_i}{4\pi r_{ij}^2} + \frac{4}{R} \right) \right], \quad (2)$$

where the suffix i denotes the quantity associated with the i th sound source (that is, the i th occupant), and r_{ij} denotes the distance between the i th and the j th occupants.

In this model, the sound in the enclosure is assumed to be solely due to human voices. Human voice may have a particular directivity pattern such that Q_i in general is not unity. However, in the situation modeled, the orientations of the human faces in the enclosure (and thus the voice directivity) are expected to be random. This randomness makes it reasonable to treat Q_i as unity on average. Other values of Q_i do not appear to be justified.

The room constant R changes with the number of people inside the enclosure. If R_0 , S , and A_p are the room constant of the empty enclosure, the total surface area and the average sound absorption area of each occupant, respectively, the absorption area of the empty enclosure A_0 is given by the expression

$$R_0 = \frac{A_0}{1 - A_0/S} \Rightarrow A_0 = \frac{R_0}{1 + R_0/S}, \quad (3a)$$

and the total absorption area in the enclosure including the presence of n persons is $A_0 + nA_p$. The room constant R when there are n persons present can be approximated by the following formula:

$$R = \frac{A_0 + nA_p}{1 - (A_0 + nA_p)/S} = \frac{R_0 + nA_p(1 + R_0/S)}{1 - (nA_p/S)(1 + R_0/S)}. \quad (3b)$$

Here, A_p is taken to be 0.44 m² in the following numerical computation.⁵

For normal conversational speech without raised voices, the A-weighted sound power level W is estimated from the speech spectrum shown in Fig. 13.3 of Kinsler and Frey⁶ and the A-weighting attenuation curve given in Sharland.⁷ The estimated A-weighted sound level at 1 m from the human mouth is about 62.5 dB. As a speech power level of 73 dB

gives an overall sound-pressure level of 65 dB at 1 m from the human mouth,⁸ W is taken to be 70 dB for all occupants. However, it is not known exactly what A-weighted sound level will trigger the raising of voices, and a theoretical formulation of how people raise their voices according to surrounding noise level does not currently exist. It is proposed in the present study that if the people start raising their voices at an A-weighted sound level of SPL_r dB, they have to produce a sound level that at least keeps the difference between their voice level and the surrounding noise level unchanged during the raising of voices. Therefore, the sound power level generated by the i th occupant W_i is estimated using the relationship

$$10 \log_{10} \left(\frac{W_i}{W_0} \right) - 70 = SPL_i - SPL_r \quad (4)$$

for $SPL_i \geq SPL_r$. The increase in W_i further increases the A-weighted sound level in the enclosure, and thus may result in further increase in W_i as other occupants may raise their voices at the same time. The A-weighted sound level at each occupant position is then calculated using Eq. (2) and iteration continues until the voice power level of each occupant converges (less than 0.1 percent change in two consecutive iterations).

The last parameter required for the prediction of A-weighted sound level is r_{ij} [Eq. (2)]. In reality, r_{ij} cannot be exactly measured due to the randomness of occupancy pattern, unless the capacity of the enclosure is reached. In the present study, the enclosure is modeled as a rectangular mesh and each occupant occupies one small rectangle of this mesh. The predicted A-weighted sound level discussed in the present study is calculated at the center of this mesh, such that

$$r_{ij} = \sqrt{\{[j - (m_c - 1)/2]\Delta_c\}^2 + \{[i - (m_r - 1)/2]\Delta_r\}^2}, \quad (5)$$

where $m_r \times m_c$ is the mesh size (odd numbers), Δ_r and Δ_c the row and column widths in the mesh, respectively.

Since the actual occupant location pattern is not known, the predicted A-weighted sound level does not directly reflect the dynamic variation of A-weighted sound level due to occupancy. It is known that the sound pressure varies with the reciprocal of spatial separation; the highest growth rate of A-weighted sound level at the mesh center occurs when the rectangles closest to the mesh center are occupied first and occupancy occurs gradually outwards from mesh center. That is, calculation commences with the eight rectangles surrounding the mesh center occupied first. Then, the next innermost set of unoccupied rectangles (outside the first eight occupied rectangles) are occupied, and calculation is repeated. This procedure is repeated until all the rectangles are filled up. The results calculated in this manner form the upper bound of the A-weighted sound level. Similarly, the slowest increase in sound level occurs when the rectangles at the perimeter of the mesh are occupied first and occupancy moves toward the mesh center, layer by layer. A lower bound of A-weighted sound level is thus obtained. The actual A-weighted sound level in the enclosure should lie within these two boundaries.

II. SITE MEASUREMENT SETUP

Site measurement results were obtained in the staff canteen of the Hong Kong Polytechnic University, which has a capacity of around 250 and dimensions of 21 m \times 21 m \times 2.8 m high. The average separation of the occupants is about 1 m. A canteen is chosen for the present study because the noise inside it is primarily due to human voices, and there is a rapid change of occupant number within a relatively short period of time. For the measurement, a computer program was developed so that the number of people going into and out of the canteen could be logged into a notebook computer. The corresponding time of people going in and out of the canteen was also recorded by this program using the built-in clock of the notebook computer.

The equivalent sound-pressure level L_{eq} , and the A-weighted level exceedance L_{10} and L_{90} were recorded every 1 min using a calibrated Metrosonic db3100E integrating sound level logger. The built-in clocks in db3100E and the notebook computer were matched before the start of each measurement. The location of the A-weighted sound level measurement was far away from walls and at ear height. Each measurement was done within busy hours and in general lasted for 3 h. Four site measurement trials were carried out in the present study, so as to reduce the impact of random occupancy pattern on the measured A-weighted sound levels.

The room constant of the empty canteen is estimated from the measured reverberation time using Sabine's equation.⁴ The reverberation time was measured using a Brüel & Kjær 2236C precision sound-level meter and Brüel & Kjær 2144 dual channel frequency analyzer in the multi-spectrum mode with averaging time of 1 ms. The sound was generated using a JBL 4604B 20-in. loudspeaker with a white noise signal fed to the loudspeaker through a power amplifier. Four measurements were done with different sound level meter locations, but the difference in the reverberation times obtained in these measurements is insignificant. The average reverberation time for the unoccupied canteen is 0.47 s, giving a room constant R_0 of 680.7 m². The decay curves from the reverberation time measurements also show the presence of flutter echo between parallel walls in the canteen. This canteen is not a very reverberant space. The large sound absorption in the canteen is probably due to the furniture, an open door, measuring 2 m \times 1 m, and the acoustical suspended ceiling having an NRC of 0.95. Other surfaces of the canteen are either painted concrete, glass windows, or concrete floor covered with plastic tiles, which are reflective. Although the canteen may not be very symmetrical acoustically, which makes the application of Eq. (1) questionable, the point is whether the error induced, if there is any, is within engineering tolerance or not. This will be discussed later.

III. EXPERIMENTAL RESULTS AND COMPARISONS WITH PREDICTIONS

Figure 1 is a typical example showing the time variations of occupant number n , and L_{eq} . These variations appear to be very similar to each other. The sound levels increase as the occupant number n increases. However, abrupt

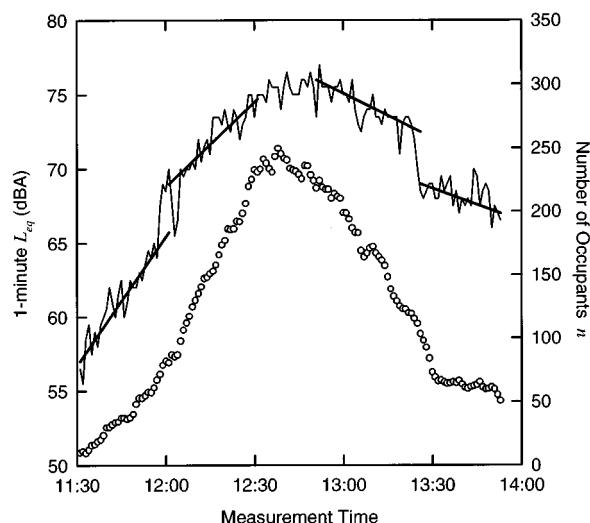


FIG. 1. Time variations of occupant number and A-weighted sound level. Thin continuous line: L_{eq} ; thick straight lines: mean variation lines; \circ : occupant number.

changes of L_{eq} occur at the time 12:00 and 13:23. The gradual variations of occupant number at the times of these abrupt changes in sound level suggest that the first abrupt increase in L_{eq} is due to the raising of human voices, while the other discontinuity results from the restoration of normal voice output. It is noted from Fig. 1 that discontinuities, both in slope and trend, occur at L_{eq} around 65 dB and 73 dB for the increasing and decreasing occupant number cases, respectively. Such a phenomenon basically appears in all trials (not shown here). However, for simplicity, it is more desirable to assume a single L_{eq} at which voice output is raised, because the actual mechanism which determines the raised voice level is still not well understood. It is therefore assumed in the prediction model that the occupants raise their voices when an A-weighted sound level exceeding 69 dB (mean of 65 dB and 73 dB) is heard. This 69 dB A-weighted sound level is also approximately equal to that just after normal voice is restored (Fig. 1), suggesting further the possibility of humans raising their voices at around this noise level. Thus, $SPL_r = 69$ dB.

Although Fig. 1 suggests that L_{eq} depends not only on the number of occupants but may also depend on whether this number is increasing or decreasing, the difference in L_{eq} in these "increasing" and "decreasing" occupant number modes is not significant in view of the large randomness in occupancy pattern (not shown here). Therefore, data from these two modes are not separately considered in the rest of the discussion. Owing to the large degree of scattering in the site measurement data, the variation of noise climate with occupant number is not discussed.

Figure 2 gives a comparison between the site measurement results and the predictions using Eq. (4). A 17×17 mesh is used in the calculations. It is assumed that human beings raise their voices when the A-weighted sound level exceeds 69 dB. Over 90 percent of the site measurement results fall within the predicted A-weighted sound level boundaries. Although sophisticated occupant distribution statistics are not obtained,⁸ the mean value of the two computed

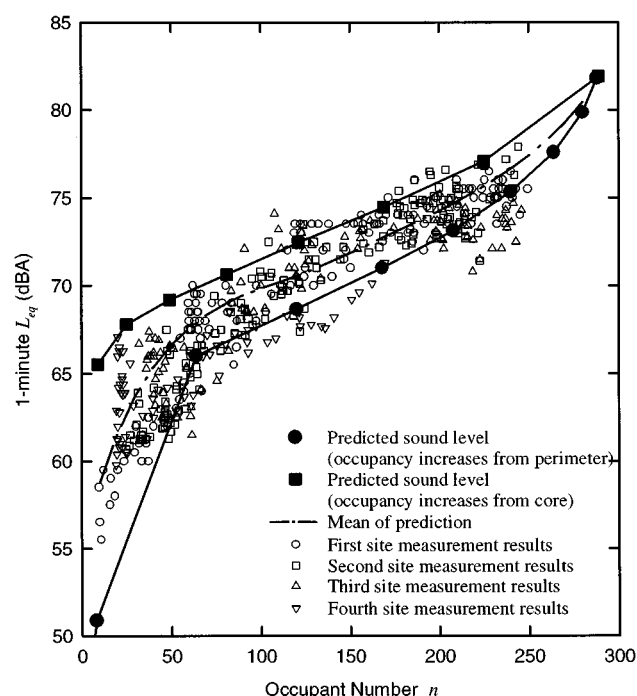


FIG. 2. Comparison of predicted and site measurement results. \bullet : Predictions with occupant number increases from perimeter; \blacksquare : predictions with occupant number increases from core; —: mean values of predictions from the two occupant number increase modes; \circ : first site measurement results; \square : second site measurement results; \triangle : third site measurement results; ∇ : fourth site measurement results.

boundaries seems to give a reasonably good estimation for the variation of A-weighted sound level inside the enclosure. In order to test how well the predicted mean A-weighted sound level variation fits the experimental data, a hypothesis testing using the student t test was performed. The present situation is the same as that in the case of "matched pairs" illustrated in Hoel.⁹ The null hypothesis is that the mean of the differences between the predicted and experimental results vanishes. While the correlation coefficient between the predicted and experimental results is 0.91, the so calculated t value is 1.96, giving a probability ≈ 0.05 and suggesting that there is no significant difference between the two sets of results at the 95 percent confidence level. In view of the relatively high uncertainties in the human voice power level and the critical A-weighted sound level that triggers the raising of voices, the present model gives reasonable estimates of the A-weighted sound levels in the canteen.

A large scattering of data is observed for occupant number $n < 50$. It is probably due to granularity of where people sit, not all of them talking at the same time, and the relatively large difference in individual voice power levels at small n (such that the measured A-weighted sound levels depend substantially on measurement location). When n becomes large, this randomness is smoothed out. It is also noted from Fig. 2 that there is a tendency for lower rate of increase of A-weighted sound level with n for $n > 250$, suggesting that the present model may significantly overestimate the A-weighted sound level at large occupant numbers. This also suggests that there is an upper limit of voice power level,

although this could not be estimated from the present site measurement results.

It is also found that an over 30 percent change in A_p results only in less than 1-dB difference in the predicted A-weighted sound level for $n \leq 200$, showing that the method does not depend very much on A_p in the present situation.

In view of practical use of the proposed sound level prediction method, the base room sound absorption of an enclosure can be estimated once the capacity and function of the enclosure are known. The function of the enclosure determines the maximum allowable A-weighted sound level, while its capacity is used to determine the mesh size in the prediction method whose algorithm is depicted in Sec. II. Optimum room absorption can then be found by iteration. However, further investigation is needed to test the robustness of the prediction model.

IV. CONCLUSIONS

A method incorporating the increase of voice power level is proposed in the present study for the prediction of the A-weighted sound level inside an enclosure in which the number of occupants varies with time. Site measurements were conducted in a canteen to validate the prediction method. Reverberation time measurements are also performed in a canteen to facilitate comparison between predictions and measured results.

Results of the present study tend to indicate that the occupants raise their voices when the A-weighted noise level exceeds 69 dB and suggest that the prediction method, despite its simplicity, gives good estimates for the upper and lower limits of the A-weighted sound level inside an occupied enclosure. Although an exact estimation of the A-weighted sound level cannot be done because of the large

randomness in occupancy pattern, the mean values of the two mentioned A-weighted sound level limits give, allowing for scattering, a satisfactory estimate of the A-weighted sound level.

It should be noted that the present model is limited to one, highly absorptive, space, and the raised voice threshold in this study was chosen empirically and thus may be specific to the present experimental condition. Further investigations into the subject are needed before the present prediction method can be extended to enclosures of other functions.

ACKNOWLEDGMENTS

This study is partially supported by a grant from the Research Committee, the Hong Kong Polytechnic University, Hong Kong. The cooperation of Kenny Ng, the catering manager of the Hong Kong Polytechnic University, is also appreciated.

- ¹I. Pollack and J. M. Pickett, "Cocktail party effects," J. Acoust. Soc. Am. **29**, 1262 (1957).
- ²W. R. MacLean, "On the acoustics of cocktail parties," J. Acoust. Soc. Am. **31**, 79–80 (1959).
- ³R. F. Legget and T. D. Northwood, "Noise surveys of cocktail parties," J. Acoust. Soc. Am. **32**, 16–18 (1960).
- ⁴L. L. Beranek, *Noise and Vibration Control* (McGraw-Hill, New York, 1971).
- ⁵W. C. Sabine, *Collected Papers on Acoustics* (Dover, New York, 1964).
- ⁶L. E. Kinsler and A. R. Frey, *Fundamentals of Acoustics* (Wiley, New York, 1962).
- ⁷I. Sharland, *Woods Practical Guide to Noise Control* (Woods, Colchester, 1979).
- ⁸H. C. Hardy, "Cocktail party acoustics," J. Acoust. Soc. Am. **31**, 535 (1959).
- ⁹P. G. Hoel, *Introduction to Mathematical Statistics* (Wiley, New York, 1984).

The relationship between localization and the Franssen effect

William A. Yost, Dan Mapes-Riordan, and Sandra J. Guzman

Parmly Hearing Institute, Loyola University Chicago, 6525 North Sheridan Road, Chicago, Illinois 60626

(Received 11 July 1996; accepted for publication 19 December 1996)

The relationship between localization and the Franssen effect was studied for noise and tones in a sound-deadened and in a live room. The noise was wideband and the tones were 250, 500, 1000, 1500, 2500, and 4000 Hz. Listeners were asked to determine the location of the stimuli in a localization task and to discriminate the difference between a pair of stimuli used to generate the Franssen illusion and a steady-state tone in a Franssen-effect discrimination task. Poor performance in the Franssen-effect discrimination task is consistent with the stimulus conditions leading to a strong Franssen illusion. Poor performance in both the Franssen effect and localization tasks was obtained for midfrequency tones (near 1500 Hz) and in the live room. Thus, the Franssen illusion is strongest for a live room and for midfrequency tones consistent with the difficulty listeners have in localizing sounds under these conditions. These results are consistent with those of Hartmann and Rakerd [J. Acoust. Soc. Am. **86**, 1366–1373 (1989)] and support their suggestion of a correlation between the Franssen effect and localization in rooms. © 1997 Acoustical Society of America. [S0001-4966(97)01905-X]

PACS numbers: 43.66.Qp, 43.66.Pn [JWH]

INTRODUCTION

The Franssen effect (1962) is generated by producing a sudden onset sound at one loudspeaker (the transient sound) which slowly decays while the same sound is presented with a slowly rising onset for a long period of time at a second loudspeaker (the steady-state sound). The steady-state sound often remains on for many seconds, while the rise and decay times are usually between 50 and 100 ms. Listeners often perceive the source of the sound as coming from the location of the transient sound well after the transient sound has terminated (but while the steady-state sound is still being presented). That is, the steady-state sound is perceptually fused with the transient sound at the spatial location of the transient sound. Hartmann and Rakerd (1989) showed that listeners can have great difficulty discriminating between a single steady-state sound that comes from the location of the transient sound and the illusory sound at the location of the transient sound caused by the two sounds used to generate the Franssen effect.

Hartmann and Rakerd (1989) showed that the Franssen effect exists in reverberant rooms but that it does not exist in an anechoic room. That is, in an anechoic room listeners can determine that the steady-state sound is at the loudspeaker location from which it was presented. They suggested that the Franssen effect exists because in reverberant rooms it is difficult to localize the source of a steady-state sound, while the source of a sudden onset sound is more easily localized (although they did not perform a localization task in their study). Thus, for the Franssen effect the auditory system localizes the source at the location of the transient since its location is more clearly delineated than that of the steady-state sound. In an anechoic space, steady-state sounds are more easily localized, so the steady-state sound in the Franssen effect is localized at its source leading to a breakdown of the Franssen effect or illusion.

In our experiment we studied both the Franssen effect

and localization in a sound-deadened and in a live room in order to test the suggestion by Hartmann and Rakerd (1989) concerning the correlation between the Franssen effect and localization. We also investigated the Franssen effect and localization of different signals (broadband noise and tones of different frequencies), since Hartmann and Rakerd (1989) showed that the type of carrier waveform affected the Franssen effect and since the ability to localize sounds also depends on the waveform (e.g., tonal frequency, see Stevens and Newman, 1936, or Mills, 1958).

I. METHODS

For the Franssen-effect discrimination task, there was a pair of sounds for the Franssen stimuli and a single sound for the non-Franssen stimulus. The transient sound for the Franssen stimuli was presented with a sudden onset and with a 50-ms linear offset (the transient sound was 50 ms in total duration). The steady-state sound for the Franssen stimuli was presented at the same time and level as the transient sound but with a 50-ms linear onset, it remained on for 350 ms, and was turned off with a 100-ms linear offset (for a total duration of 500 ms). The transient and steady-state sound were presented to different loudspeakers. The non-Franssen stimulus was a 400-ms stimulus that was gated on suddenly and was gated off with the 100-ms offset, so that it was 500 ms in duration and equal to the sum of the transient and steady-state sounds used for the Franssen stimuli. The sounds were presented at an average level of 60 dBA measured at the location of the listeners. The overall level of the sounds was randomly roved over a 6-dB range to make it difficult to use any slight loudness differences among the loudspeakers as a cue for detection. A low-level wideband noise (38 dBA) was presented from each of the eight loudspeakers in order to mask amplifier hum (noise was presented to all eight loudspeakers, although only loudspeakers 1–3 and 6–8 were used in the experiments). The stimuli

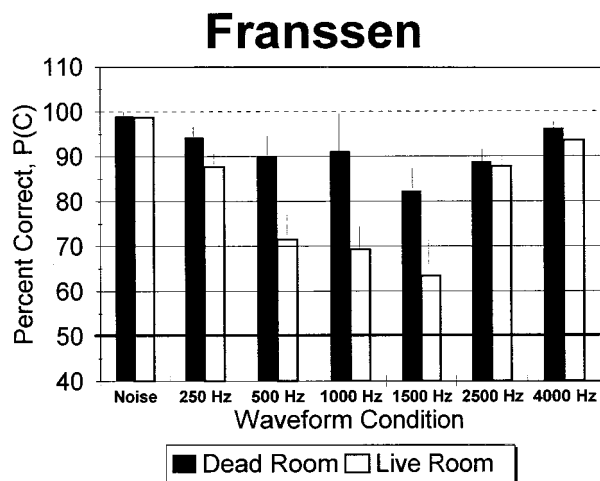


FIG. 1. Percent correct $P(C)$ averaged over the three (live room) or four (dead room) listeners as a function of the waveform type (noise or tones of different frequencies) for detecting the Franssen effect in the dead (filled bars) and live (empty bars) room. The vertical bars for each histogram is one standard deviation. The solid horizontal line at 50% represents chance performance.

were generated on an eight-channel D/A Tucker-Davis Technology (TDT) system at 20 000 samples per second and were low-passed filtered at 5 000 Hz. The sounds were a broadband noise and tones of 250, 500, 1000, 1500, 2500, and 4000 Hz.

The sound-deadened listening room is 3.5 m long, by 2.5 m wide, by 2.1 m high, constructed of sound-deadening office partitions and lined with sound-attenuating foam on all surfaces. Additional sound-attenuating foam was placed in the room at various locations (e.g., directly behind the listener) to equalize (as much as possible) the sound levels at the location of the listener's head that were produced by each of the eight loudspeakers. There were eight loudspeakers (Realistic-Minimus 2.5) in the sound-deadened room that were driven by Crown (Power Line Two) amplifiers. The loudspeakers (numbered 1–8, with 1 on the right, 4 straight ahead, and 8 on the left) were equally spaced around the frontal hemisphere in 25.5° increments (from –90° azimuth on the left to +90° azimuth on the right) at a distance of 1.3 m from the listener, and at a height of 1.2 m. The loudspeaker outputs were matched in dBA output, and the loudspeakers were within 2 dB of each other in the spectral region of 200–5000 Hz, and each loudspeaker was flat to within 2 dB over this frequency region. Only loudspeakers 1–3 and 6–8 were used in the experiments, while all eight loudspeakers were used to measure a few aspects of the room acoustics. The subjects knew that only the six loudspeakers would be used. The live room was constructed by placing 1/4-in. paneling on all of the walls of the listening room covering the walls from floor to ceiling.

To demonstrate that the two rooms were different in their reverberation characteristics we measured the response of the eight loudspeakers to a 122- μ s transient signal at the location of the listener in each room. For each loudspeaker we computed the average time waveform for 10 presentations of the transient which provided eight waveforms (one for each loudspeaker) for each room. We then computed the

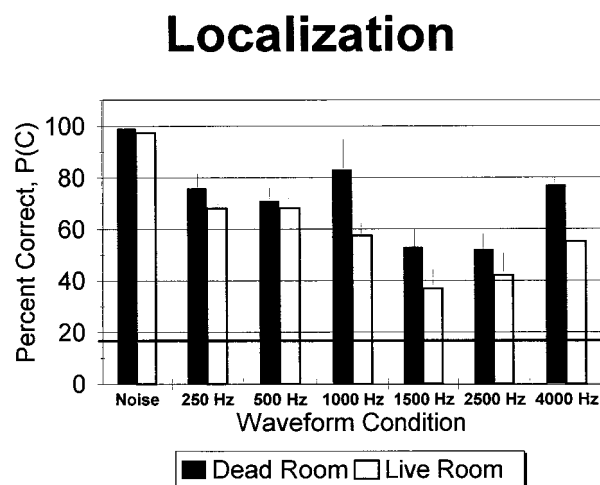


FIG. 2. Percent correct $P(C)$ averaged over the three (live room) or four (dead room) listeners as a function of the waveform type (noise or tones of different frequencies) for localizing the steady-state sound in the dead (filled bars) and live (empty bars) room. The vertical bars for each histogram is one standard deviation. The solid horizontal line at 16.6% represents chance performance.

average rms level (averaged across the eight loudspeakers) for the first 5 ms of the recorded signal and compared this level to the average rms level for the next 25 ms. We then took 20 log of the ratio of these two levels in order to estimate how many decibels down the reflected sound was from the source sound in each room. For the sound-deadened room the last 25 ms of sound was 19.9 dB down from the first 5 ms, and for the live room the difference was 6.7 dB. Thus, there was a 13.2 dB difference in the measurements made between the two rooms.

We also measured the spectrum (from 200 to 5000 Hz) of each of the transient waveforms from each of the loudspeakers for each room. While the spectra of the live room recordings had peaks and valleys that differed by as much as 12 dB (the comb-filter effect of a reverberant room), there were no consistent differences at the frequencies used in this study among the loudspeakers in each of the two rooms. That is, while there were spectral differences between the live and dead room (the comb-filter effect was strongest for the live room), within each room there were no consistent spectral differences across the spectra and among loudspeakers. Therefore, it is unlikely that any differences in performance between conditions for any one room could be attributed to spectral oddities in room acoustics.

The basic Franssen-effect discrimination procedure is based on the procedure used by Hartmann and Rakerd (1989). The basic idea is to ask listeners to discriminate between a single sound (non-Franssen stimulus) presented from one loudspeaker and a pair of sounds (transient and steady state) used to generate the Franssen effect (Franssen stimuli). If the Franssen illusion occurs then listeners should have difficulty discriminating the Franssen stimuli from the non-Franssen stimulus. The transient sound (for a Franssen trial) or single sound (for a non-Franssen trial) was presented on a random basis to loudspeakers 1, 2, 3, 6, 7, or 8 (loudspeakers 4 and 5 were not used). If the transient sound was presented to loudspeaker 1, 2, or 3, the steady state-sound was pre-

Localization

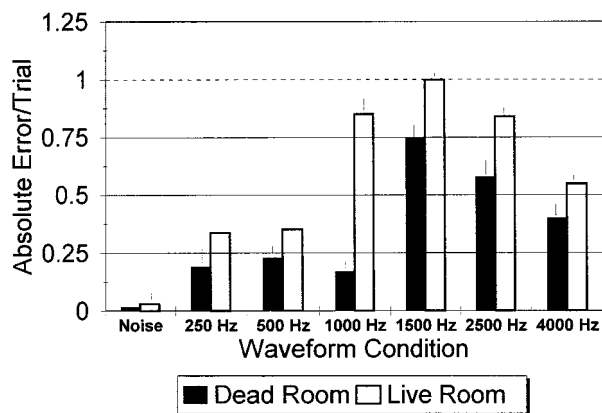


FIG. 3. Absolute error per trial averaged over the three (live room) or four (dead room) listeners as a function of the waveform type (noise or tones of different frequencies) for localizing the steady-state sound in the dead (filled bars) and live (empty bars) room. The vertical bars for each histogram are one standard deviation. See text for further descriptions.

sented on a random basis to loudspeakers 6, 7, or 8; and vice versa. Thus, the transient and steady-state sounds were presented from opposite sides of the listeners and the loudspeaker from which each sound came was determined randomly. Half the trials (determined randomly) were Franssen trials and half non-Franssen trials. In this single-interval procedure, the listener indicated whether the trial was a Franssen or a non-Franssen trial. Trial-by-trial feedback was provided and there was a 350-ms delay after feedback before the next trial was presented. A block was 50 trials and at least 2 blocks were used to estimate percent correct performance, $P(C)$, for any one condition. Four or more blocks were always used if $P(C)$ for the first two blocks was less than 90%. The authors served as listeners in all conditions and an additional undergraduate subject served as a listener in the dead room.

For the localization task, the steady-state sound as used for the Franssen-effect discrimination experiment (i.e., a 500-ms sound with a 50-ms linear rise and 100-ms linear decay time) was the stimulus whose location was to be determined. It was presented randomly from loudspeakers 1, 2, 3, 6, 7, or 8 (the listeners knew that the sounds would come from one of these six loudspeakers). The listener indicated from which loudspeaker they thought the sound came and feedback was provided. All other conditions were the same as those used in the Franssen-effect discrimination experiment.

II. RESULTS

Figure 1 shows the results from the Franssen-effect discrimination experiment as $P(C)$ averaged across the three or four listeners as a function of the stimulus waveform for the sound-deadened and live rooms. Performance near chance (50%) indicates that there is a Franssen illusion. The broadband noise does not produce the illusion in either room. For tones, percent correct, and hence the effectiveness of the illusion, depends on the room and frequency. The effective-

ness of the illusion is greater for the live room and is greatest for frequencies in the 1500-Hz region. The vertical bars indicate one standard deviation.

Figure 2 shows results for the localization task as average $P(C)$ (chance performance would be 1/6 or 16.67%) versus stimulus waveform for the sound-deadened and live rooms. The source of the noise is easily located in both rooms. It is more difficult to locate the source of the tones in the live room and in the midfrequency (around 1500 Hz) region. Figure 3 indicates the error in location judgments. The error is the absolute difference in the loudspeaker number given by the listener and that which presented the stimulus. The total absolute error was computed for each subject and condition and then this total was divided by the number of trials to produce the total absolute error per trial. The average absolute error per trial computed across listeners is shown in Fig. 3. The smallest localization errors occurred for the noise, and larger localization errors were obtained for the live room and for tones in the midfrequency range. For both Figs. 2 and 3 the vertical bars marked one standard deviation.

The good localization performance at 1000 Hz was due to one listener who performed very well at 1000 Hz in the localization task in the dead room. She also performed well in the Franssen task at 1000 Hz, but her performance in the localization task was especially high. We retested her in several other conditions to make sure that she had not learned to use some cue at 1000 Hz that could be used at other frequencies. Her retest performance on these other conditions was within the standard deviation of her original performance. We are not sure why she performed so well at 1000 Hz, especially in the localization task. Except for this one condition and listener the results from all listeners (four for the dead room and three for the live room) were very similar, as the small standard deviations indicate. Hartmann and Rakerd (1989) also found very little intrasubject variability for the Franssen-effect discriminations for those conditions that are similar to the ones used in our study.

For each room a Spearman rank-order correlation was computed, first between $P(C)$ in the localization experiment and $P(C)$ in the Franssen-effect discrimination experiment and then between average absolute error per trial in the localization experiment and $P(C)$ in the Franssen-effect discrimination experiment. For the live room the correlation was 0.86 for percent correct localization and 0.72 for absolute error per trial. For the dead room the two correlations were 0.79 (for percent correct localization) and 0.72 (for absolute error per trial). All of the correlations are significantly different from zero correlation at the 0.05 level of significance.

III. DISCUSSION

The data indicate a significant correlation between the conditions that lead to the Franssen effect and localization. The more difficult it is to localize the sound the more likely it is that the Franssen effect will produce the Franssen illusion. Without the excellent performance of the one listener at 1000 Hz in the localization task, the correlations would be higher. The correlations would have been higher if the dif-

ferences at 4000 Hz seen between the two rooms for the localization task would have been as great for the Franssen discrimination. These results support the suggestion made by Hartmann and Rakerd (1989) that there is a correlation between the ability to localize a sound and the ability for that sound to produce a Franssen illusion, in that the more difficult it is to localize a steady-state sound the more likely it is that there will be a Franssen illusion.

In addition, we have shown that the Franssen illusion is frequency dependent with the strongest illusion occurring in the frequency region around 1500 Hz in both rooms. The frequency dependency of the Franssen effect has not been shown before. This result is also consistent with Hartmann and Rakerd's (1989) suggestion regarding the relationship between the Franssen effect and localization. Since the classic localization experiment by Stevens and Newman (1936) it has been known that tones in the midfrequency region around 1500 Hz are more difficult to localize than lower or higher frequency tones, presumably reflecting the duplex nature of localization based on interaural time and level differences (see Yost and Hafter, 1987). Our results also confirm this finding.

We also found the same trend as Hartmann and Rakerd (1989) in that the Franssen effect works best in live rooms. In the Hartmann and Rakerd (1989) experiment they used several room conditions with different reverberant characteristics. For the most reverberant room conditions and for the stimulus parameters similar to ones used in our study, they showed that listeners were near chance in the Franssen-effect discrimination experiment for a 500-Hz tone. Our listeners were not at chance for any frequency. This may be because our live room was not as reverberant as the ones used by Hartmann and Rakerd (1989) or because of several procedural differences between the two studies.

Although these data support the notion of a relationship between localization and the Franssen effect, they do not explain why tones are more difficult to locate in live rooms and why noise is so easy. As stated above the Duplex Theory of localization does offer an explanation as to why the effects may be frequency dependent (neither interaural time nor level are good localization cues around 1500 Hz, due to the duplex nature of interaural time and level processing). Hart-

mann and Rakerd have a series of papers (Hartmann, 1983; Hartmann and Rakerd, 1989; Rakerd and Hartmann, 1985, 1986) concerning localization in rooms and they should be consulted, since a discussion of localization in rooms is beyond the goal of this paper.

The Franssen illusion can be extremely strong. In many conditions, even when the steady-state tone is on for many seconds, listeners still perceive the source of a single sound at the location of the transient loudspeaker even though no sound has been presented from that loudspeaker for many seconds. In demonstrations of the illusion one can even remove from the room the loudspeaker over which the transient is presented and the audience will continue to indicate that the sound is coming from the approximate location where that loudspeaker had been (assuming, of course, that the steady-state sound is still being presented to the other loudspeaker).

ACKNOWLEDGMENTS

We would like to thank our colleagues at the Parmly Hearing Institute, especially Dr. Toby Dye, Dr. Stan Sheft, Dr. Bill Shofner, and Dr. Greg Sandell for their advice. The work was supported by grants from NIDCD (DC00293) and from the Life Sciences Division of the AFOSR. An audio example of the Franssen effect for noise and a tone can be found at <http://www.parmly.luc.edu/>.

- Franssen, N. V. (1962). *Sterophony* (Phillips Technical Lecture, Eindhoven, The Netherlands), English translation (1964).
- Hartmann, W. M. (1983). "Localization of sound in rooms," *J. Acoust. Soc. Am.* **74**, 1380–1391.
- Hartmann, W. M., and Rakerd, B. (1989). "Localization of sound in rooms IV. The Franssen effect," *J. Acoust. Soc. Am.* **86**, 1366–1373.
- Mills, A. W. (1958). "On the minimum audible angle," *J. Acoust. Soc. Am.* **30**, 237–246.
- Rakerd, B., and Hartmann, W. M. (1985). "Localization of sound in rooms, II: The Effects of a single reflecting surface," *J. Acoust. Soc. Am.* **78**, 524–533.
- Rakerd, B., and Hartmann, W. M. (1986). "Localization of sound in rooms, III. Onset and duration effects," *J. Acoust. Soc. Am.* **80**, 1695–1706.
- Stevens, S. S., and Newman, E. B. (1936). "The localization of actual sources of sound," *Am. J. Psych.* **48**, 297–306.
- Yost, W. A., and Hafter, E. (1987). "Lateralization," in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer-Verlag, New York).

Erratum: “On the spatial coherence and angular spreading of sound forward scattered from the sea surface: Measurements and interpretive model” [J. Acoust. Soc. Am. 100, 748–758 (1996)]

Peter H. Dahl

Applied Physics Laboratory, College of Ocean and Fishery Science, 1013 N.E. 40th Street, Seattle, Washington 98105-6698

(Received 28 September 1995; accepted for publication 22 February 1996)

[S0001-4966(97)02905-6]

PACS numbers: 43.30.Hw, 43.30.Re, 43.30.Zk, 43.30.Gv, 43.10.Vx [JHM]

The following two errata are noted for the above-mentioned paper:

- (1) The phrase above Eq. (8) should read: ... κ is one-half the transverse component...
- (2) Equation (17) should read:

$$\alpha_b = e^{-(\beta_I / \sin \theta_i + \beta_I / \sin \theta_s)}.$$

The published version of Eq. (17) differs by a factor of 2 in the exponent. Note, however, that computer computations presented in the paper utilized the correct form of the equation.

PROGRAM OF

The 133rd Meeting of the Acoustical Society of America

Penn State Conference Center • State College, PA • 16–20 June 1997

1a MON. AM

NOTE: All Journal articles and Letters to the Editor are peer reviewed before publication. Program abstracts, however, are not reviewed before publication, since we are prohibited by time and schedule.

MONDAY MORNING, 16 JUNE 1997

ROOM P, 8:00 TO 11:50 A.M.

Session 1aAO

Acoustical Oceanography and Underwater Acoustics: Acoustical Measurement of Coastal Ocean Processes I

James F. Lynch, Chair

Woods Hole Oceanographic Institution, 203 Bigelow Building, Woods Hole, Massachusetts 02543

Chair's Introduction—8:00

Invited Papers

8:05

1aAO1. Preliminary report on the summer '96 joint China–U.S. internal wave/acoustic wave experiment in the Yellow Sea. Ji-Xun Zhou, Peter H. Rogers, Gary W. Caille (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405), Renhe Zhang, Guoliang Jin, Liangying Lei (Chinese Acad. of Sci., Beijing 100080, PROC), Peter H. Dahl, Robert C. Spindel (Univ. of Washington, Seattle, WA 98105), and Zijun Gan (South China Sea Inst. of Oceanology, Guangzhou 510301, PROC)

Simultaneous observations of internal wave activity and acoustic wave propagation in 70-m water in the Yellow Sea were made in the late summer of 1996. The primary objective of this experiment was to validate the predicated modal coupling and fluctuations induced by shallow-water internal waves. The Yellow Sea provides an ideal environment for such research because it has a very flat and homogeneous bottom and a very strong, sharp thermocline. The environment lends itself to relatively simple models for both the acoustic field and the internal wave field. Propagation was measured over distances up to 55 km in the frequency range of 50 Hz–5 kHz. The receivers were three moored and two suspended hydrophone arrays. Internal wave activity was monitored using thermistor chains, SAR satellite imagery, and high-frequency sonar. Propagation data were taken as a function of range, as a function of time at a fixed range, and as a function of azimuth in the presence of differing levels of internal wave activity. Supporting environmental data obtained during the experiment include ADCP, bottom profile and surface wave-height spectra, and bottom coring. Preliminary data on internal wave field, sound propagation, reverberation, and bottom acoustic parameters will be presented. [Work supported by ONR and the Chinese Academy of Sciences.]

8:25

1aAO2. Intimate '96: Shallow water tomography in the Sea of the Condemned. Emanuel Coelho (Hydrographic Inst., P-1296, Lisbon, Portugal), Sergio Jesus (Univ. of the Algarve, Faro, Portugal), Yann Stephan, Xavier Demoulin (EPSHOM/CMO, F-29275, Brest, France), and Michael B. Porter (Math. Dept., New Jersey Inst. of Technol., Newark, NJ 07102)

As is well-known, the tidal force of the moon and the sun can cause notable changes in the sea level. Besides this so-called barotropic effect, the tidal force also drives internal waves in a daily rhythm. Thus, the internal wave spectrum is often dominated by a single component with perhaps 10 km from crest to crest. This "internal tide" tends to propagate toward shore and has its greatest height near the shelfbreak. As this tide propagates it modulates the surface duct and its acoustic signature is often seen in data. The Intimate '96 experiment (conducted off the coast of Portugal) was specifically designed to acoustically image the internal tide with an

eye toward a more precise understanding of its structure and acoustic impact. A towed source emitted chirps every 8 s for several days and the chirps were received on the SACLANTCEN portable array. The data show a textbook multipath structure with early refracted paths followed by some 30 distinct bottom and surface echoes which shift with the internal tide. The acoustic and oceanographic interpretation of this data will be discussed.

8:45

1aAO3. Observations of cnoidal internal waves and their effect on acoustic propagation in shallow water. David Rubenstein (Sci. Applications Int. Corp., 1710 Goodridge Dr., P.O. Box 1303, McLean, VA 22102, davidr@osg.saic.com)

Packets of cnoidal internal waves were observed during a shallow-water experiment in the Gulf of Mexico. They lasted about 3 h and were always observed at the same time of day, clearly in response to tidal forcing. The cnoidal waves had 2–10 m amplitudes, narrow frequency bandwidths with central frequencies of about 9 cph, wavelengths of about 435 m, and they propagated in the along-slope direction. Data from a set of three thermistor arrays were processed to obtain directional wave spectra. These spectra establish the strong directionality of the cnoidal wave packets. The background “continuum” spectrum is resolved into a small set of discrete, strongly directional components. Observations of low-frequency acoustic propagation along several baselines showed fluctuations induced by cnoidal waves. These effects were simulated using a time-step PE approach. A mode-coupling resonance with the internal wave field results in elevated acoustic variability along a set of discrete spokes, emanating from the acoustic source. While acoustic variability tends to increase with range and with internal wave amplitude, tangential and radial correlation scales do not show a systematic dependence. Other implications of relative directionality between internal wave and acoustic propagation are discussed.

9:05

1aAO4. A review of the shelf edge studies acoustic measurement experiments—SESAME I and II. Robert L. Field (NRL Ocean Acoust. Branch, Stennis Space Center, MS 39529)

The shelf edge studies acoustic measurement experiments (SESAME) were performed in the North West Approaches to the British Isles. The SESAME experiments were a collaborative effort between the Natural Environment Research Council, the Defence Research Agency (both of the United Kingdom), and the Naval Research Laboratory (Stennis Space Center). The objective of the experiments was to determine the effects of internal waves formed at the shelf edge on acoustic propagation. Oceanographic measurements were gathered continuously over an 18-month period and acoustic experiments were performed during the summers of 1995 and 1996. Measurements of temperature, conductivity, and current were made within the water column, sediment cores were taken of the ocean bottom, and SAR images of the sea surface were taken of the experimental area. Range and time-varying acoustic experiments were performed in the along shelf and cross shelf directions over a frequency band of 100–2100 Hz. Both narrow- and broadband (frequency modulated) signals were transmitted. A review of the two experiments and some of the oceanographic and acoustic data collected will be presented. [Work supported by ONR and NRL.]

9:25

1aAO5. The shelfbreak front PRIMER experiment. Glen Gawarkiewicz, Robert Pickart, James F. Lynch (Woods Hole Oceanograph. Inst., Woods Hole, MA 02543), Ching-Sang Chiu, Kevin Smith (Naval Postgrad. School, Monterey, CA 93943), and James H. Miller (Univ. of Rhode Island, Narragansett, RI 02882)

In July and August 1996, a combined physical oceanography and acoustics experiment was performed in the Mid-Atlantic Bight south of Nantucket Shoals. The purpose of this experiment was to examine coastal frontal variability and its impact on acoustic propagation and imaging. Extensive oceanographic measurements of the experimental region were made. These measurements included: (1) SeaSoar hydrography; (2) moored *T*, *S*, and current meter sensors; and (3) satellite infrared imaging of the sea surface temperature. These measurements provided what is believed to be the highest resolution imaging of a shelfbreak frontal region yet obtained. Acoustic measurements in the experimental region included: (1) a three-dimensional tomographic imaging array; and (2) aircraft deployed SUS drop fields. Observations revealed the presence of a strong jet associated with sharp thermal gradients. During this time period, the front was contorted by the presence of a streamer of shelf water moving offshore, which resulted in large thermal gradients in both along the shelf and across the shelf direction. The oceanographic activity produced significant mode coupling in the acoustic signals, leading to large time spreading of the source pulses.

9:45

1aAO6. Acoustic travel time and intensity fluctuations measured in the SWARM95 experiment. Robert H. Headrick, Jr., James F. Lynch (Woods Hole Oceanograph. Inst., Woods Hole, MA 02543), Marshall Orr, Bruce Pasewark, Steve Wolf (Naval Res. Lab., Washington, DC 20375), Mohsen Badiéy (Univ. of Delaware, Newark, DE 19716), Ching-Sang Chiu (Naval Postgrad. School, Monterey, CA 93943), and John Apel (Global Ocean Assoc., Silver Spring, MD 20908)

During the summer of 1995, a multidisciplinary, multilaboratory experiment entitled “SWARM” (shallow-water acoustics in a random medium) was conducted on the continental shelf off New Jersey. The purpose of this experiment was to examine the effects of internal waves, both linear and nonlinear, upon acoustic transmissions in the 10 to 1000-Hz range. Towards this goal, numerous oceanographic sensors were deployed along a fixed across shelf track, together with a fixed acoustic transmission range. Shipborne acoustic sources also allowed out of plane (along shelf) studies. Results will be presented of recent data analyses showing the nature of the measured acoustic travel time and intensity fluctuations, and how they correlate to environmental forcing, particularly by internal waves and the shelfbreak front. [Work supported by ONR.]

10:05–10:20 Break

10:20

1aAO7. Modeling of the out-of-plane pulse propagation in the SWARM95 experiment. X. Tang and M. Badiéy (Marine Studies, Univ. of Delaware, Newark, DE 19716)

The acoustical data collected in the shallow water acoustics in a random medium (SWARM95) experiment on the continental shelf off New Jersey in 1995 have shown that the temporal variations of the time series of the airgun signals (center frequency 75 Hz and bandwidth 50 Hz) are both range and azimuthal dependent. A full-wave broadband PE model is used to simulate such shallow water pulse propagation at different horizontal bearings. The time-dependent shallow water internal wave fluctuations constructed with a broad spectrum of excitation and the solitonlike thermocline displacements based on the *in situ* observations are included in the numerical simulation as well as the bathymetry. The vertical wavefronts and the 2-h time histories of the repeated transmissions from airgun positions above and below the thermocline are calculated and displayed. The simulated data and the experimental data of the receiver arrays at different locations with an approximately 10-km cross-range distance are compared. The results suggest that the azimuthal dependence of the received signals correlate to both the internal wave fields and the bathymetric variations.

10:35

1aAO8. Analysis of internal wave and bottom effects on the acoustic field generated by a towed source during the SWARM95 experiment. Altan Turgut, Bruce H. Pasewark, Stephen N. Wolf, Marshall Orr (Naval Res. Lab., Acoust. Div., Washington, DC 20375), and James F. Lynch (Woods Hole Oceanograph. Inst., Woods Hole, MA)

Thirty-second duration 250- to 375-Hz LFM pulses were transmitted from a towed acoustic transducer to study the effects of internal waves and the sea bottom on an acoustic signal propagating along a 30-km transmission path. The data were taken on the New Jersey continental shelf as a part of the multi-institutional SWARM95 experiment. Environmental data were collected by using CTDs, thermistor strings, a high-frequency acoustic flow visualization system, and a chirp sonar system. Geoacoustic properties of the bottom were inverted from the chirp sonar data with a 1.5-m vertical resolution and 40-m penetration. A deterministic sound speed field in the water column was constructed by imposing the high-frequency visualization measurements on both CTD and thermistor string data. Environmental data indicated that two distinct geological provinces and two major internal wave packets were crossed during the 30-km towed transmission. Using a full-wave broadband PE model, vertical structures of sound field were calculated for comparison to data obtained by a 32-element vertical line receiver array. Preliminary model/data comparisons indicate the dominance of water column effects over the bottom effects. [Work supported by ONR.]

10:50

1aAO9. Broadband propagation and azimuthal dependence through internal waves. M. Badiéy (Marine Studies, Univ. of Delaware, Newark, DE 19716), J. Lynch (Woods Hole Oceanograph. Inst., Woods Hole, MA 02543), M. Orr (Naval Res. Lab., Washington, DC 20375), X. Tang (Univ. of Delaware, Newark, DE 19716), and S. Wolf (Naval Res. Lab., Washington, DC 20375)

During the SWARM 95 experiments the out-of-plane broadband acoustic propagations were used to examine the effect of azimuth as well as the frequency dependence on the propagation of an acoustic pulse through internal waves. A highly repeatable source (a 40-in. airgun with center frequency of 75 Hz and bandwidth of 50 Hz) was used at different water depths. The environmental data were measured at the source and at the two receiver arrays located approximately 25–30 km away perpendicular to the main propagation track. It is shown that the dominant energy of the propagation is distributed between the first and the second modes. The

transfer of energy among the modes as well as the fluctuations of the field are depth dependent as well as time dependent, which can be related to the internal wave characteristics. The coherence of the field in cross range due to the same source signature is also examined.

11:05

1aAO10. Azimuthal and depth-dependent transmission loss caused by shelf internal waves. Marshall H. Orr and Roger M. Oba (Naval Res. Lab., Washington, DC 20375-5320)

The amplitude of an acoustic field propagating through a stratified continental shelf water column in the presence of internal wave packets is shown to be azimuthally and depth dependent when referenced to the propagation vector of the internal wave packet. Internal wave (IW) packets, which propagate across the continental shelf, are known to be regularly generated each tidal cycle at the shelf break edge in stratified water column conditions. They are dispersive in nature and change in physical structure as they propagate away from the generation site. Their wave directional spectrum is highly peaked. Consequently, the wave-number spectrum of the continental shelf sound velocity field is highly directional and will be shown to cause azimuthal dependency in the properties of the acoustic field as a function of range and bearing from the source. Preliminary results of numerical solutions to the parabolic equation approximation of the Helmholtz equation are presented to substantiate the above assertions, in particular for propagation through a uniform internal wave field of a fixed wave number and amplitude. [Work supported by the Office of Naval Research.]

11:20

1aAO11. Wavelet analysis of internal waves in SESAME II data, and the effects on acoustic transmission. Jacob George and Robert L. Field (NRL Ocean Acoust. Branch, Stennis Space Center, MS 39529)

In continental shelf regions, large-amplitude internal tides are believed to spawn smaller wavelength solitonlike waves. During the summers of 1995 and 1996, the SESAME experiments were performed in the North West Approaches to the British Isles. The purpose of these experiments was to measure the properties (length scales, amplitudes, velocities, and the time of occurrence) of these solitonlike waves and determine their effects on acoustic propagation. In this paper wavelet analysis is applied to some of the long-term temperature measurements in order to determine soliton properties. The results of the wavelet analysis and its implications on acoustic transmission will be discussed. [Work supported by ONR and NRL.]

11:35

1aAO12. Acoustic scintillation due to shallow-water internal solitary waves: Theory and experiment. V. R. Djachenko, K. V. Konyaev, D. E. Leikin, K. D. Sabinin, and Yu. I. Tuzhilkin (Inst. of Acoust., 4 Shvernik St., Moscow 117036, Russia)

In August–September 1990, an acoustics experiment was performed in the shallow-water region near Kamchatka. A $1\frac{1}{2}$ -h time series representing amplitude fluctuations of 800-Hz, 160-Hz-bandwidth acoustic pulses was obtained by a hydrophone moored 34 km from the source. During the same period, a time series representing solitonlike vertical displacements of the thermocline near the receiver was recorded using a 37.5-m vertical temperature sensor. The two records were found to display abnormal similarity at time delays of 16.5 and 44.5 min. The effect is shown to be due to modal synchronism that may take place in shallow-water lossy-bottom regions with a small downward-positive gradient of the sound velocity. Theoretical explanation is verified by numerical simulation in which received signals are modeled by Fourier synthesis based on coupled-mode wave-field solutions in a stratified channel with a moving internal solitary wave. On the basis of the theory developed, new possible approaches for acoustic remote sensing are discussed.

Session 1aNSa

Noise and NOISE-CON: Finite Elements and Boundary Elements

Victor W. Sparrow, Chair

Graduate Program in Acoustics, Pennsylvania State University, 157 Hammond Building, University Park, Pennsylvania 16802

Contributed Papers

9:00

1aNSa1. Numerical load generation, vibration analysis, and acoustic signature prediction of a composite combat vehicle. Craig Birkett (Automated Anal. Corp., 2805 S. Industrial, Ste. 100, Ann Arbor, MI 48104, craig@autoa.com), Nickolas Vlahopoulos (Univ. of Michigan, Ann Arbor, MI 48109-2145), Zhen Wu (Ford Motor Co.), and Edward Shalis (U. S. Army Tank—Automotive Command, Warren, MI 48397)

Acoustic signatures of ground combat vehicles are very important to their survivability. In particular, low-frequency tones can be used to identify vehicles beyond the line of site. In the range of 10-200 Hz, the structure-borne acoustic radiation is mainly due to the track force/moment inputs. In this work, full-vehicle models were developed to simulate numerically the radiated acoustic signature as well as interior noise levels at its design stage. The loads applied on the hull of the vehicle were computed from a specialized code previously developed by the U. S. Tank Command. It simulates the interaction of the hull, sprocket, idler, roadwheels, roadarms, torsion bars, and track. The loads generated from this analysis comprise the excitation for the vibration calculations. The boundary element method was used to compute the radiated sound pressure in the semianechoic condition after mapping normal vibration velocities (calculated from a finite-element model) onto the vehicle's surfaces. A description of the entire process of generating the loads, incorporating them in the finite-element analysis, and using the vibration results for the acoustic signature of the army vehicle is presented. [Program partially funded by a U. S. Army SBIR contract.]

[See NOISE-CON Proceedings for full paper.]

9:20

1aNSa2. Sensitivity calculations for structural acoustic EFEM predictions. Fernando Bitsie and Robert Bernhard (Ray W. Herrick Labs., School of Mech. Eng., Purdue Univ., West Lafayette, IN 47907-1077, bernhard@ecn.purdue.edu)

Recent research on the energy finite element method (EFEM) has focused on the development of structural-acoustic coupled capability. This formulation has been developed successfully in terms of radiation efficiency. The approach has been implemented into the finite element method and successfully verified against experimental results for an enclosure with one flexible wall. Calculation of the sensitivity of the response to various parameters of the model is reasonably straightforward in the EFEM formulation. This capability was used for system identification for the experimental verification problem and for understanding the dominant mechanisms which govern the structural acoustic response at high frequency for both acoustical and mechanical excitation of the enclosure problem.

[See NOISE-CON Proceedings for full paper.]

9:40

1aNSa3. Various approaches to estimating natural frequencies of automobile cab cavities from a boundary-element model. Prasanna Kondapalli and Craig Birkett (Automated Anal. Corp., 2805 S. Industrial, Ann Arbor, MI 48104, pkonda@autoa.com)

The boundary-element method has gained wide acceptance as an acoustic analysis tool in the automobile industry. The methodology for extracting eigenvalues of acoustic cavities of complex shapes by the boundary element method is not yet well established. The finite-element method is widely used for eigenvalue analysis, but it is very expensive and time consuming to build a finite-element mesh solely for this purpose. Two ways for estimating the natural frequencies of a cab cavity from its boundary element model are discussed in this paper. One approach is derived from the experimental technique for determining cavity modes. An acoustic (point) source(s) placed inside the cab is used to excite the various cavity modes by doing a frequency sweep in the range of interest. The response is measured at suitable locations and the identified peaks in the frequency response function give an estimate of the cavity modes. By examining the sound pressure level plots on the boundary, the type of mode is identified. The second approach makes use of an automatic mesh generator which automatically meshes the interior cavity by tetra elements from a boundary element model. Then, a finite-element eigenvalue analysis is carried out to determine the cavity modes. Both of the above approaches require significantly less time than constructing a finite-element model. Numerical examples of a rectangular box and a simple car model are considered, and results obtained from a regular finite-element model, tetra finite-element model, and a boundary-element model are compared.

[See NOISE-CON Proceedings for full paper.]

10:00

1aNSa4. In-car noise reduction for a newly developed home elevator. Y-Q. Zhou (Otis Technol. Inst., Nippon Otis Elevator Co., KSP C Wing, 11th Fl., 3-2-1, Sakato, Takatsu-ku, Kawasaki City, 213 Japan, zhou@kawasaki.otis.utc.com)

A serious in-car noise problem was found for our newly developed home elevator, which was designed using low-cost and light-weight material for its cab panels. After a series of real-run measurements, the main source for in-car noise was identified to be the car panel vibration. Numerical analysis using the FEM method has also been performed for understanding the panel dynamic behaviors. The cab panels were redesigned based on measurement and numerical analysis results. The newly designed cab gives a nearly 9-dB in-car noise reduction, which enables that type of home elevator to be sold in the Japanese market as a Nippon Otis product.

[See NOISE-CON Proceedings for full paper.]

1aNSa5. Analysis of periodic acoustic problems with double-sided impedance condition. S. T. Raveendra and B. Konno (Automated Anal. Corp., 2805 S. Industrial, Ann Arbor, MI 48104, ravi@autoa.com)

The modeling of unequal impedance on either side of a vibrating structure is important in many engineering applications. Although the boundary element method is well suited for the analysis of periodic acoustic problems, the modeling of the double-sided impedance condition is not possible in the conventional boundary-element method. In particular, the indirect boundary-element method allows the simultaneous modeling of acoustic domains separated by thin vibrating structures. Thus a facility to model the double-sided impedance condition is very useful in the solution of realistic engineering problems using the indirect boundary-element method. The indirect boundary-element method models the jump of the acoustic pressure and the acoustic velocity between the two sides of the vibrating surface. The impedance relationship that relates the acoustic pressure to the velocity is not reducible to a single relationship between the jump of pressure and velocity. Since the relationship is expressed by two equations, the solution of the equations requires the generation of two sets of integral equations at the impedance surface. The application of this newly developed formulation for the solution of the double-sided impedance condition will be demonstrated by solving a number of example problems. An additional feature of this development is that this formulation can be used to eliminate the irregular frequencies that are associated with the solution of periodic exterior acoustic problems using the boundary element method. The application of this formulation for the elimination of irregular frequencies will also be demonstrated through example problems.

[See NOISE-CON Proceedings for full paper.]

10:40

1aNSa6. Noise prediction of loader by using BEM and SEA. Jae-Woong Choi, Hee-Joon Lee, Shin-Il Kang, and Jong-Min Kang (Dept. of Core Tech. 1, CAIRT, Samsung Heavy Industries, P.O. Box 43, Yusung-gu, Taejeon, Korea, 305-600, jwchoi@halla.shi3.samsung.co.kr)

Though the international regulations about the noise level of earth-moving machinery becomes more and more strict, it is not known that an

analytical tool could be applied to predicting radiation noise. In this study, a systematic procedure is introduced in terms of the noise source modeling, application of the existing analytical tools, and verification. This research selects BEM and SEA by considering the modal density, and does the noise sources modeled in appropriate forms for each method. A half-scale model of real loader and a simple source are used to verify the approach. Comparisons of the simulated and experimental results show good agreements in both the simple and real loader cases. These results suggest that the approach should be applied to the investigation of ways for an effective noise reduction by using a simulated noise prediction.

11:00

1aNSa7. The application of a FEM-BEM model to predict acoustic silencer performance and the BEM model to predict vented earmold dynamics. Chang-Hwan Choi, Larry H. Royster (Dept. of Mech. Eng., North Carolina State Univ., Raleigh, NC 27695-7910, royster@eos.ncsu.edu), and Robert D. Ciskowski (IBM Corp., Waltham, MA 02254)

A directly coupled boundary element and finite element model has been applied to the dynamics of acoustic silencers exhibiting flexible surfaces. In this coupled BEM-FEM formulation, the BEM model used a three-node isoparametric triangular linear continuous element to discretize the acoustic cavity. The BEM model was coupled to a flat shell FEM model. The formulated model was then used to predict the performance of various silencer configurations. In a parallel study, the BEM model was used to study the various parameters affecting the performance of vented earmolds coupled to a headset driver and the ear canal. The preliminary findings from this effort will be presented.

[See NOISE-CON Proceedings for full paper.]

MONDAY MORNING, 16 JUNE 1997

ROOM G, 8:30 TO 11:00 A.M.

Session 1aNSb

Noise and NOISE-CON: Workshop on Power Plant Noise Control I

Robert A. Putnam, Cochair
Westinghouse MC590, 4400 Alafaya Trail, Orlando, Florida 32826

Frank H. Brittain, Cochair
Bechtel Corporation, 50 Beale Street, San Francisco, California 94105

Chair's Introduction—8:30

Invited Papers

8:40

1aNSb1. Issues in modeling outdoor acoustical barriers. Robert A. Putnam (Westinghouse, MC 590, 4400 Alafaya Trail, Orlando, FL 32826, putnam.r.a@wec.com)

In the design of power generation facilities, whether classified as "indoor" or "outdoor," it often happens that the economic evaluation of noise control measures indicates the use of free-standing outdoor noise barriers as part of the overall facility acoustical design. Additionally, the facility buildings and components themselves often function as acoustical barriers. Whenever such barriers are to be modeled, approximations to classical barrier theory are common. The reality of actual facility arrangements, the limitations

regarding placement of acoustical barriers, and the physical or economic constraints on barrier height usually mean that the size and positions of the barriers are not optimal. Furthermore, the presence of nearby reflective surfaces, the finite width of any given barrier, and the compounding effect of successive barriers or surfaces also necessitates adjustments to ideal barrier theory. This paper will discuss a number of particular cases commonly encountered and will recommend specific measures to be used to approximate the reality of actual situations. In addition, a general discussion will be presented which compares the mapping of sound levels beyond acoustical barriers as a function of whether the source is modeled as an idealized point source or as a more realistic distributed source.

[See NOISE-CON Proceedings for full paper.]

9:10

1aNSb2. Issues in predicting community noise for power plants. Frank H. Brittain (Bechtel Corp., 50 Beale St., San Francisco, CA 94105, fhbritta@bechtel.com)

The normal method of predicting community noise levels involves first computing sound power levels from near-field sound-pressure levels and then assuming that the noise radiates from one or several point sources. This approach makes several major leaps of faith. Some of these leaps of faith or issues in predicting community noise levels for power plants are identified. The issues include validity of near-field sound-pressure levels, accuracy of sound power levels computed, radiation from large sources, transmission through building walls, and screening and reflection of noise by power plant equipment and buildings. The validity of representing large sources as one or a series of point sources is reviewed. The role of atmospheric effects and the dangers inherent in typical noise prediction software is briefly reviewed. Several examples and suggestions for making more accurate predictions are given.

[See NOISE-CON Proceedings for full paper.]

Contributed Paper

9:40

1aNSb3. Noise control for a 210-MW cogeneration plant. Charles T. Moritz, Robert D. Bruce (Collaboration in Sci. and Technol., Inc., 15835 Park Ten Pl., Ste. 105, Houston, TX 77084), Reid Bicknell, and Jeff Jacobson (Raytheon Engineers and Constructors, Houston, TX 77042)

In this paper, the authors discuss a case study involving the design of noise control treatments for a 210-MW cogeneration plant. Residential property was located to the north, south, and east of the plant, with industrial zoned property to the west. The local noise regulations limited the A-weighted sound level that could be produced by the plant under all normal operating conditions to 55 dB at residential property lines and 75 dB at the industrial property lines. The nearest residential boundary was

about 100-m east of the turbine buildings. Cost-effective noise control treatments were developed during the design phase of the plant. These treatments included not only sound-level specifications for all major equipment items, special building wall and roof construction, transformer barriers, and silencers for ventilation systems and vents, but also many noisy equipment items were located on the west side of the plant.

[See NOISE-CON Proceedings for full paper.]

10:00–11:00

Panel Discussion

MONDAY MORNING, 16 JUNE 1997

ROOM N, 8:30 TO 11:30 A.M.

Session 1aPA

Physical Acoustics: Quantum Acoustics and Thermoacoustics

Ralph Muehleisen, Chair

Department of Physics, Naval Postgraduate School, Monterey, California 93943

Contributed Papers

8:30

1aPA1. Josephson acoustic radiation in superfluid ^4He : Towards a quantum mechanical pressure standard. Scott Backhaus and Richard E. Packard (Dept. of Phys., Univ. of California, Berkeley, CA 94720, backhaus@physics.berkeley.edu)

Below 2.2 K, ^4He is a superfluid that exhibits nearly dissipationless flow for velocities less than a well defined critical velocity. For flow through submicron apertures, the superfluid begins to dissipate energy at

the critical velocity by the creation and motion of individual quantized vortices across the aperture. Using a recently developed pressure control technique, a constant pressure drop, ΔP , is applied across the aperture. This causes the vortices to cross the aperture at a Josephson frequency given by $f_j = (m_4 \Delta P) / (ph)$ where m_4 is the mass of a helium atom and h is Planck's constant. The periodic vortex crossing causes the superfluid velocity in the aperture to oscillate and emit acoustic radiation at f_j . This sound source will radiate a total power of 10^{-21} W into freespace. Radiation from the vortex motion has been detected using a 1/4 wave resonant detector and a novel cryogenic microphone. This phenomena could lead to

a frequency based pressure standard where the frequency is determined by quantum mechanics. [Work partially supported by the Office of Naval Research and NASA.]

8:45

1aPA2. Thermoviscous effects in steady and oscillating superfluid flow with applications to Helmholtz oscillators. Scott Backhaus (Dept. of Phys., Univ. of California, Berkeley, CA 94720, backhaus@physics.berkeley.edu)

The two fluid nature of superfluid ^4He can lead to interesting behavior when it is used in a Helmholtz oscillator. If the oscillator's constriction is narrow enough, the normal fluid can be locked by its own viscosity leaving only the superfluid component to flow in and out of the oscillator volume. This leads to an accompanying temperature oscillation that can cause dissipation and/or frequency shifts depending on the thermal properties of the walls of the oscillator. If the constriction is somewhat larger, the normal component is able to move and the thermal effects are partially canceled. These same effects lead to dissipation when steady flow is forced through the constriction. Equations have been derived which include superfluid flow, normal fluid flow, compressibility, thermal expansion, and the properties of the oscillator walls. These have been solved for the resistance to steady flow and the resonant frequency and Q of a Helmholtz oscillation. The calculations will be compared with a wide array of recent measurements. [Work partially supported by the Office of Naval Research and NASA.]

9:00

1aPA3. Resonant modes of a constricted annular resonator. Hsiao-Tseng Lin, Yun Choe, Ralph T. Muehleisen, and Anthony A. Atchley (Dept. of Phys., Naval Postgrad. School, Monterey, CA 93943)

One topic of current interest in thermoacoustic research is an annular prime mover [Lin *et al.*, J. Acoust. Soc. Am. **100**, 2846(A) (1996)]. The starting point for this research is an investigation of a constricted annular resonator. A literature search of the field resulted in surprisingly few references. The results of analytic, numerical, and experimental investigations are presented. It was found that introducing a constriction to an annular resonator split each longitudinal duct mode into two modes, one of a higher frequency with a pressure antinode at the constriction, and one at a lower frequency with a velocity antinode near the constriction. While changing the size of the constriction appreciably modified the lower resonance frequency and the position of the velocity antinode, the upper resonance frequency was only moderately modified and the position of the pressure antinode was virtually fixed. [Work supported by the Office of Naval Research and the American Society for Engineering Education.]

9:15

1aPA4. Investigation of an annular thermoacoustic prime mover. Hsiao-Tseng Lin, Ralph T. Muehleisen, and Anthony A. Atchley (Dept. of Phys., Naval Postgrad. School, Monterey, CA 93943)

The study of an annular thermoacoustic prime mover has been presented at a previous meeting [Lin *et al.*, J. Acoust. Soc. Am. **100**, 2846(A) (1996)]. The prime mover described there was unable to reach onset. As a result of improved analysis, the design has been simplified. The resonance frequencies and quality factor of the new prime mover were

measured as functions of the temperature difference imposed across the stack. The experimental work is presented and a numerical model based on an iterative method is proposed. [Work is supported by the Office of Naval Research and the American Society of Engineering Education.]

9:30

1aPA5. Control system analysis of an annular prime mover. Ralph T. Muehleisen, Hsiao-Tseng Lin, and Anthony A. Atchley (Dept. of Phys., Naval Postgrad. School, Monterey, CA 93943)

The thermoacoustic prime mover in an annular resonator is analyzed as a linear feedback control system. Using the techniques of control system analysis, a closed-form solution for the acoustic gain of a constriction with linear temperature gradient in an annular resonator is developed. For a large enough gradient in the constriction, the acoustic gain of the system could become infinite, at which time the system goes into unstable positive feedback creating a thermoacoustic prime mover. An analysis of an existing annular resonator system is made and a comparison to more standard thermoacoustic numerical analysis is presented. [Work supported by the Office of Naval Research and the American Society for Engineering Education.]

9:45–10:00 Break

10:00

1aPA6. Experimental results with a thermoacoustically driven thermoacoustic refrigerator. Thomas J. Hofer, Jay A. Adeff, and Anthony A. Atchley (Dept. of Phys., Naval Postgrad. School, Monterey, CA 93943, hofer@physics.nps.navy.mil)

The construction of a thermoacoustically (heat) driven, thermoacoustic refrigerator (TADTAR) has been completed. The design utilizes a novel resonator topology and should have a high cooling power density. While the design goal calls for a cooling power level of 500 W over a temperature span of 40 °C, the first refrigeration data recently obtained exhibited only small amounts of cooling. However, as the understanding of the engine improves, so does its performance. Variations in resonator tuning and stack location, gas type and pressure, and stack type and geometry are being explored. Low Prandtl number gas mixtures produce particularly strong improvements in onset conditions and amplitude. [Work supported by Office of Naval Research.]

10:15

1aPA7. Heat transfer measurements in a thermoacoustic refrigerator. Martin Wetzel and Cila Herman (Dept. of Mech. Eng., Johns Hopkins Univ., 3400 N. Charles St., Baltimore, MD 21218, herman@titan.me.jhu.edu)

In conventional heat transfer studies, mostly dealing with steady-state flow phenomena, a well-established database is available in the literature. In thermoacoustics, however, the flow oscillates with a zero mean velocity and heat transfer rates in this physical situation cannot be easily predicted. As it was shown previously [M. Wetzel and C. Herman, J. Acoust. Soc. Am. **100**, 2846(A) (1996)], a tool to investigate such high-speed heat transfer phenomena is holographic interferometry combined with high-speed cinematography. Interferometric measurements were conducted for

drive ratios from 1% to 3%. Applying digital image processing to the obtained interferometric fringe pattern, the oscillating temperature field was measured and acoustic temperature fluctuations of the order of 1 K could be resolved. From the temperature fields the important heat transfer parameters such as local heat fluxes, heat transfer coefficients, and Nusselt numbers were determined at the edge of two stack plates. [Work supported by the Office of Naval Research.]

10:30

1aPA8. Measurements of thermoacoustic functions for single uniform pores. Larry A. Wilen (Dept. of Phys. and Astron., Ohio Univ., Athens, OH 45701, wilen@helios.phy.ohiou.edu)

In previous work, a technique to measure the thermoacoustic properties of single pores was described [L. Wilen, J. Acoust. Soc. Am. **101**, 1388–1397 (1997)]. By modulating the volume of a pore and measuring the resultant pressure response, the complex compressibility is measured as a function of frequency. Measurements of different lengths of a given pore allow the end effects to be subtracted out. By working down to very low frequencies, a wide range of values for the ratio of the thermal penetration depth to pore size are accessible. This technique will be briefly reviewed, and recent experimental results for pores of various geometries will be compared to the known theory. Experimental results will also be shown for pores not easily treated by analytic theory. Finally, possible applications to full tests of thermoacoustics in a single pore will be discussed. [Work supported by ONR and NSF.]

10:45

1aPA9. Thermoacoustic streaming on a rigid sphere. Ashok Gopinath (Dept. of Mech. Eng., Code ME/GK Naval Postgrad. School, Monterey, CA 93943)

The fundamental problem of *thermoacoustic streaming* in the presence of a rigid sphere in a strong standing acoustic field has been treated analytically. As a first step the compact sphere assumption is made, and the corresponding governing equations developed in dimensionless form to allow for a consistent treatment of all the energy-exchange mechanisms. It is found that the interaction of the strong acoustic field with the rigid boundaries of the sphere is capable of inducing a second-order time-averaged temperature distribution, in addition to the well known second-order time-averaged fluid motion. The associated temperature gradients

cause localized heating and cooling effects on the surface of the sphere, even in the complete absence of any externally applied temperature potential. The role of a little known second-order thermal expansion coefficient is pointed out. A careful analysis of this conjugate problem coupling the solid–fluid time-averaged transport effects could help in the measurement of this second-order thermodynamics modulus. [Work supported by NASA.]

11:00

1aPA10. Numerical study of the performance of an idealized thermoacoustic stack. Aniruddha S. Worlikar, Omar M. Knio (Dept. of Mech. Eng., Johns Hopkins Univ., Baltimore, MD 21218), and Rupert Klein (Bergische Univ., D-42097 Wuppertal, Germany)

The unsteady, thermally stratified flow in the neighborhood of an idealized stack is simulated using a nonlinear two-dimensional model. The model relies on a vorticity-based formulation of the compressible conservation equations in the low-Mach-number and short-stack limits. The action of the acoustic standing wave is represented in terms of oscillating velocity conditions, and heat exchangers are modeled by maintaining the temperature at the cold and hot ends of the stack at fixed, constant values. Simulations are used to visualize the unsteady velocity and temperature fields, and to analyze the dependence of the “steady-state” cooling load on the drive ratio and the length of the heat exchangers. [Work of ASW and OMK supported by the Office of Naval Research. RK is partially supported by Deutsche Forschungsgemeinschaft (DFG).]

11:15

1aPA11. A time-domain description of the thermoacoustic instability. Sergey Karpov and Andrea Prosperetti (Dept. of Mech. Eng., 122 Latrobe Hall, Johns Hopkins Univ., Baltimore, MD 21218)

In typical conditions the time scale for the development of the thermoacoustic instability in a tube is slow compared with the acoustic period. It is therefore possible to solve the problem by the method of multiple time scales. In this way a time-domain description of the process is obtained. The perturbation results are compared with the exact solution of the linear problem and their range of validity is discussed. It is found that they are accurate only when the spacing of the plates in the stack is sufficiently large. The “short-stack approximation” is also introduced and evaluated with comparable results, provided the effective wave number is chosen appropriately. [Work supported by the Office of Naval Research.]

Session 1aSA

Structural Acoustics and Vibration: Vibration of Structures

J. Gregory McDaniel, Chair

Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Contributed Papers

8:30

1aSA1. Asymmetrical effects and higher harmonics from nonlinear interactions in a driven vibrating wire. Roger J. Hanson, H. Kent Macomber, and Michael P. Kassakatis (Dept. of Phys., Univ. of Northern Iowa, Cedar Falls, IA 50614)

Inclusion of tensional changes and longitudinal motion result in coupled nonlinear equations for motion in, and perpendicular to, the plane of the driving force on a stretched wire. Earlier experimental measurements of nonlinear effects in a sinusoidally driven vibrating wire [Hanson *et al.*, *J. Acoust. Soc. Am.* **96**, 1549–1556 (1994)] will be briefly reviewed. Presented data will show that the onset of motion perpendicular to the plane of the driving force depends very sensitively on asymmetries which produce slight differences in free vibration frequencies in two orthogonal directions. Also reported will be results of amplitude and phase measurements of generated second, third, and fourth harmonics of the driving frequency, both in the plane of the driving force (vertical) and perpendicular to it. These higher harmonics, not significant at frequencies far off-resonance, become very significant near a resonance. With the wire driven at frequencies near its third natural resonance, a very pronounced subharmonic splitting of order three has been observed. Qualitative aspects will be illustrated on video tape.

8:45

1aSA2. Membrane waves on a cylindrical shell with an axial rib. Patricia A. Manning (Dept. of Ocean Eng., MIT, 77 Massachusetts Ave., Cambridge, MA 02139, pmanning@mit.edu)

A test facility has been set up to study the effect of structural modifications on membrane waves in cylindrical shells. Previous research on submerged shells has shown the importance of membrane waves in the mid-frequency, high aspect angle regimes. These shear and compressional waves are important because they are the principle determinants of the shell's radiation and scattering properties. Flexural waves serve mainly as modifiers of the membrane waves because they are poorly coupled to the acoustic medium. The current test structure is a thin plastic shell. The advantages of plastic include higher damping, a lower compressional wave speed, and easy machining. Single modes up to $n=6$ are generated using an axisymmetric source array and the resulting accelerations are measured with an axial line array. The measured accelerations are separated into flexural, compressional and shear waves. Data collected on the unmodified plastic shell show excellent agreement with theory. Membrane and flexural waves in both forward and backward propagation directions are clearly seen. The experimental dispersion curves closely match theoretical predictions. Recent modifications to the shell involve the addition of an axial rib. Data from the modified shell will be presented. [Work supported by ONR.]

9:00

1aSA3. The effects of a rib on the structural intensity in a cylindrical shell. Derek C. Lang and Courtney B. Burroughs (Graduate Program in Acoust., Penn State Univ., State College, PA 16804)

The amount of vibrational energy scattering by a rib in a cylindrical shell will depend on the shape of the rib. This redistribution of energy will affect the coupling between the shell vibration and acoustic radiation. Most models treat the rib attachment to the shell as a line attachment. Here, a rib

with different widths and shapes is considered. Equations for the structural intensity in an infinite cylindrical shell without and with fluid loading are derived. The rib is treated as a change in the shell thickness specified by a rib function with parameters to control height, width and shape. The effects of the rib on the structural intensity in the in-vacuo shell in the farfield of the rib are modeled as a function of frequency and circumferential mode number.

9:15

1aSA4. An experimental study of tire standing waves. Joseph P. Cusumano and John D. Zolock (Dept. of Eng. Sci. and Mech., Penn State Univ., 227 Hammond Bldg., University Park, PA 16802, johnz@psu.edu)

Small-scale test rigs for experimental investigations of tire standing wave instability are described. The rigs are used in experiments in which tire rotational speed is gradually increased past the critical speed at which standing waves appear around the circumference of the tire. Detailed results from one of the rigs are presented. Images grabbed from live video are processed to extract the spatial structure of the standing waves. Frequency and time domain vibrational data and temperature histories are also collected and presented. The spatial frequencies, strains, and logarithmic decrement for the standing wave are computed for four different tire pressures, and the results are shown to be in qualitative agreement with numerical results of previous researchers.

9:30

1aSA5. Spatial scales on fluid-loaded plates and shells. J. Gregory McDaniel (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215), Henno Allik, and Robert Haberman (BBN, New London, CT 06320-6167)

This paper attempts to quantify in a general way the spatial scales that occur on fluid-loaded plates and shells for the purpose of assessing various modeling techniques. Modelers using the SARA finite-element code typically follow a "quarter-wavelength rule," which states that the size of a quadratic element should be less than a quarter of the minimum wavelength in the problem. The minimum wavelength for fluid-loaded plates and shells is often the flexural wavelength. Questions naturally arise regarding the accuracy of this rule near discontinuities. In this paper, a straightforward theoretical justification for the quarter-wavelength rule is presented in terms of a Taylor series expansion of the wave field. This justification is extended to the local evanescent fields generated near discontinuities. The complex wave numbers describing these fields, as well as SARA finite-element data for some canonical problems, are examined to determine the expected accuracy of a given discretized model. The fundamental conclusion is that modeling rules based on the propagating wave fields also give convergent answers for the evanescent fields generated near discontinuities.

1aSA6. The dynamic response of complex structures using hybrid methods. Peter J. Halliday and Karl Grosh (Dept. of Mech. Eng. and Appl. Mech., Univ. of Michigan, Ann Arbor, MI 48109-2125)

Within many complex structures, there are regions which easily lend themselves to closed-form analytic solutions. A hybrid analytic-numeric formulation for the time harmonic structural vibration response of such structures is presented. The seamless inclusion of analytic solutions into the Galerkin finite element method is accomplished by using transition elements which couple the tractions between analytic and numeric solution regions. The main advantage of this formulation is to increase the overall efficiency of a model by eliminating regions of the structure from the computational domain. In this way, the total number of degrees-of-freedom will be reduced, which in turn reduces memory requirements and computational time. This allows for the analysis of higher frequency or more complex problems. One application is the study of systems comprised of structural members which are coupled at nonidealized joints. Using a discretization of the equations of elasticity for a joint, its response may be represented to a desired level of accuracy. This elasticity model is coupled to the reduced beam or shell theory producing an efficient model of the complete structure. The effect of the joint and the level of modeling detail for prediction of the time harmonic response are studied.

10:00–10:15 Break

10:15

1aSA7. Loss factors of a complex composed of a number of coupled harmonic oscillators. G. Maidanik (David Taylor Model Basin, NSWC/CD, 9500 MacArthur Blvd., West Bethesda, MD 20817)

Investigation of the response of a complex composed of a number of coupled harmonic oscillators is conducted. One of the harmonic oscillators is designed as the master, the others as the satellites. The expression for the loss factor of the coupled master harmonic oscillator is sought. Two distinct loss factors are defined; the prevailing loss factor and the effective loss factor. The first is defined in terms of the real part of the inverse of the normalized *in situ* admittance of the master harmonic oscillator. The second is defined in terms of the ratio of the normalized input power onto the master harmonic oscillator to the normalized stored energy in the complex due to that power injection. The relationship and the contrast between these two loss factors are revealed. It is argued that, with the exception of an isolated master harmonic oscillator, the prevailing loss factor is apparent; it is the effective loss factor that is real. Whereas the prevailing loss factor invokes the question, "Where did the energy go?," the effective loss factor renders this question moot.

10:30

1aSA8. Effect of flexibility on rattle noise: Experimental and theoretical analysis. Karen J. L. Fegelman and Karl Grosh (Dept. of Mech. Eng. and Appl. Mech., Univ. of Michigan, Ann Arbor, MI 48109-2125)

Rattle has become an important issue in the automotive and aerospace industries. In order to understand the fundamental mechanics of rattle, a model problem involving a hinged plate rattling against a stiff contact point was formulated. In a previous study, this plate was modeled as a rigid body and the linearized equations of motion were solved explicitly. The calculated tip accelerations were compared to those measured from an experimental test stand and agreed qualitatively. Currently, the model has been expanded to include flexible modes in order to capture higher frequency behavior. In addition, instrumentation has been added to the test stand to measure sound-pressure level (SPL) and contact forces. The data

for accelerations, contact forces, and SPL were compared in the time and frequency domains. Finally, the closed form flexible body solutions were analyzed with regard to stability to predict transitions from periodic to chaotic rattling behavior.

10:45

1aSA9. Vibrations of a two-level underground station. S. A. Kostarev (Tunneling Assoc., Sadovaja-Spasskaja 21, 107217, Moscow, Russia)

A method of vibration calculations of a two-level underground station and its transmission to the ground surface is derived and the algorithm is realized as a part of an integrated package UNSONIC 2.1. A two-level station is constructed with typical sizes significantly exceeding the length of bending waves in the separate constructive elements. The lower level in the considered model is on the ground and the base of the upper floor is suspended on the walls of the station. Forces from the moving trains are localized on four railways of both levels. The package of programs developed in the Acoustics Laboratory of the Tunneling Association (Moscow) permits the calculation of all the vibration levels inside the station and tunnels, and on the ground surface in the adjacent area.

11:00

1aSA10. Structure of the acoustical field generated by an underground tunnel. S. A. Rybak (Tunneling Assoc., Sadovaja-Spasskaja 21, 107217, Moscow, Russia)

The problem considered is about relative magnitudes of the field exited by the modes of cylindrical shell oscillations in soil due to a periodical force on the shell's inner side. The closed form of this problem includes the dynamical equations of the shell, the wave equations of the surrounding elastic medium, and the appropriate boundary conditions. The principal role of the Rayleigh waves in the vibration energy transmission is determined and its value calculated for some examples of layered ground with a free boundary. The effect of the waveguide in the system of elastic layers is estimated.

11:15

1aSA11. Analysis of elastic cavitating wing vibration. Svetlana I. Kovinskaya (Sci. Res. Ctr. for Ecological Safety, Russian Acad. of Sci., Korpusna ya St. 18, St. Petersburg 197042, Russia, svetlana@srce.samson.spb.su) and Eduard L. Amromin (Krylov Shipb. Res. Inst., St. Petersburg, Russia)

A two-dimensional unsteady potential flow outside an elastic partially cavitating wing is analyzed numerically by using the Birnbaum equation on the hydrofoil chord and the Lagrange–Cauchy integral on the cavity. The angle of attack has a small periodic perturbation, and cavity thickness and length also take perturbations. Wing vibration is considered to be vibration of a variable thickness beam with two clamp bolts near the beam center. A monofrequency flow perturbation induces monofrequency flexural vibration of a noncavitating wing, but the vibration of a cavitating wing has multifrequency. Computation of NACA-16009 hydrofoil vibration was made for various free-stream speeds, module of elasticity, fluid, and wing densities, and as a result, three frequency bands of a vibration increase are determined. The low-frequency band is connected with a cavity volume oscillation. There is a considerable influence of cavity length on vibration. The high-frequency band is connected with elastic resonances of the wing. A solitary resonance was found in the middle band. This resonance has a nonlinear dependence on the loss coefficient and does not have a dependence on cavity dimensions and wing elasticity. This resonance appearance is connected with an interaction of lift and media inertia forces.

Session 1aSPa

Signal Processing in Acoustics and Underwater Acoustics: Array Processing and Detection Underwater

Lora Weiss, Chair

Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

Contributed Papers

8:00

1aSPa1. Array element localization for a horizontal hydrophone array in the Arctic Ocean. Stan E. Dosso, Mark R. Fallat (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3Y2, Canada), and Barbara J. Sotirin (NCCOSC RDTE Div., San Diego, CA 92152-6435)

In 1996, an acoustic research array with both horizontal and vertical components was deployed through the polar ice pack north of Ellesmere Island. The horizontal components consist of a 2.4-km horizontal line array (HLA) of 80 hydrophones on the seafloor and a 240-m secondary HLA of 8 hydrophones perpendicular to the main HLA. To accurately localize these hydrophones, a series of recordings were made immediately after deployment. Glass light bulbs were imploded at 50-m depth at eight locations surrounding the array. Source locations were known to within approximately 2 m; however, to invert the travel time data to within their estimated uncertainties, it was necessary to include corrections to the source locations as unknowns. Source instants were not measured, and were also included as unknowns. An iterative inversion algorithm was formulated by expanding the ray-tracing equations and neglecting second-order terms to yield a linear system of equations. The system was solved by minimizing the HLA curvature and source-location corrections subject to fitting the data to a statistically meaningful level. Minimizing HLA curvature produces a minimum-structure solution in which any deviations from a straight array are absolutely required by the data, and are not artifacts of the starting model.

8:15

1aSPa2. Preliminary results on environmentally robust vertical array design for shallow water mode filtering. John R. Buck (Dept. of ECE, Univ. of Massachusetts, N. Dartmouth, MA 02747-2300)

Many shallow water acoustics techniques require accurate estimates of the propagating modes from pressure observations at a vertical array. The noise in these estimates is a function of the conditioning of the vertical modes as sampled by the hydrophone array. However, the sound-speed profile and thus vertical mode shapes are usually not known when the array is designed. In addition, economic considerations often dictate the same array be used in a wide range of environments. Consequently, optimizing the array for a specific environment risks a degradation in performance if the experimental environment differs from the environment assumed by the array design. Conventional wisdom holds that in the absence of precise environmental knowledge, the safest approach is to space the hydrophones evenly throughout the water column. This work presents several simplified mode filtering scenarios demonstrating that nonuniform hydrophone spacings exist which are more robust in the minimax sense than uniform spacing, if it is possible to specify a range of operating environments for the array. This means that their worst case performance is better than the worst case performance of the uniformly spaced array over the range of environments.

8:30

1aSPa3. Mode function extraction from a VLA using singular value decomposition. Tracianne B. Neilsen, Evan K. Westwood (Appl. Res. Lab., The Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029, neilsen@arlut.utexas.edu), and Takeshi Udagawa (Univ. of Texas, Austin, TX 78705)

A method is developed for extracting the depth-dependent mode functions from single-frequency measurements on a vertical line array as a source moves out in range. A matrix of the complex pressure field versus receiver depth and source range is formed, and the cross-spectral density matrix (CSDM) is computed. A singular value decomposition (SVD) is performed on the CSDM to obtain the orthonormal eigenvectors and the eigenvalues. Rearrangement of the normal mode equations reveals that under ideal conditions, the eigenvectors correspond to the mode functions and the eigenvalues are proportional to the modal source excitation and wave number. When two or more eigenvalues are nearly equal, the eigenvectors are not unique, and the correct mode functions are linear combinations of the eigenvectors. Other complications arise when the pressure field is not sampled adequately in depth or range. The procedure is applied to simulated data for a Pekeris waveguide and a realistic geoacoustic profile. Extracted mode functions from sources of opportunity are intended to be used for geoacoustic inversion of the sound-speed profile and bottom properties. [Work supported by the U. S. Army Research Office, AASERT Grant No. DAAH04-95-1-0486, and ONR.]

8:45

1aSPa4. Generic beamforming structure for adaptive schemes implemented in two- and three-dimensional arrays of sensors. Ayman Tawfik, Amar C. Dhanantwari, and Stergios Stergiopoulos (Dept. of Elec. Eng., Tech. Univ. of Nova Scotia, P.O. Box 1000, Halifax, NS B3J 2X4, Canada)

The concept for implementing successfully adaptive schemes in two-dimensional (2-D) and three-dimensional (3-D) arrays of sensors, such as planar, circular, cylindrical, and spherical arrays, is similar to that of line arrays. In particular, the basic step is to minimize the number of degrees of freedom associated with the adaptation process. The present investigation has been centered on the definition of a generic beamforming structure that decomposes the beamforming process of 2- and 3-D sensor arrays into subsets of coherent processes. The approach is to separate the computationally intensive multidimensional beamforming 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 and radar 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 investigation includes also the definition of generic sub-aperture schemes for 2- and 3-D sensor arrays. This approach leads to minimization of the associated convergence period and makes the implementation of adaptive schemes in multidimensional sensor arrays practically feasible. [Work supported by NSERC-STR181039.]

9:00

1aSPa5. Three-dimensional noise-field directionality estimation from towed line-array data. Ronald A. Wagstaff and J. Newcomb (Naval Res. Lab., Code 7176, Stennis Space Center, MS 39529-5004, wagstaff@poorsha.nrlssc.navy.mil)

The ideal measurement tool to measure the three-dimensional arrival structure of the noise field is a high-resolution volumetric array sonar system. Unfortunately, such a system is not generally available. However, towed horizontal line arrays are available and can be used, even though they are far from ideal. The beam patterns of a line array are conically symmetric about the axis of the array. When the axis of a horizontal line array is tilted slightly from the horizontal, the vertical character of the beam cones can be used to discriminate between various vertical arrival angles. By combining this character with measurements made on several different towed array headings, an estimate of the three-dimensional directionality of the noise field can be obtained. An algorithm that uses such a technique on single and multiline towed array data to generate an estimate of the three-dimensional directionality of the noise field is described herein. Some results using measured data from a recent noise measurement exercise using a single-towed line array are presented and discussed. [Work supported by ONR and NRL.]

9:15

1aSPa6. Active sonar coherence measurements of underwater scatterers and channels using a cross correlation process. James Kisenwether and Dennis Ricker (Appl. Res. Lab., Penn State Univ., State College, PA 16804)

The concept of coherence is fundamental and quite important in all fields dealing with fluctuating quantities. A new method for measuring coherence based on the cross correlations of the tones of a multi-component 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 associated with scatterers and channels and represents the signal coherence. This paper will provide a review of scattering function representations and aggregate coherence measures as derived from a correlation matrix. The new coherence measure termed correlation mass will be defined. It will be shown that this measure can be decomposed into correlated components resulting from interaction with a scatterer and a channel. In other words 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.

9:30

1aSPa7. Acoustic tomography at basin scales and clock errors. Anisim A. Silivra (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804), John L. Spiesberger, Anatoly Fabrikant (Penn State Univ., University Park, PA 16802), and Harley Hurlburt (Naval Res. Lab., Stennis Space Center, MS 39522)

A basin scale underwater acoustic tomography simulation is carried out for the northeast Pacific Ocean to determine the accuracy with which time must be kept at the sources when clocks at the receivers are accurate. A sequential Kalman filter is used to estimate sound-speed perturbation and clock errors. Using an acoustic array of 4 fixed sources and 20 drifting receivers, it is found that the percentage of the modeled ocean's sound-speed variance accounted for with tomography is 92% at 400-km resolution, practically regardless of the accuracy of the clocks. Clocks which drift up to hundreds of seconds of error (or even more) for a year do not degrade tomographic reconstruction of the model ocean. It is shown that tomographic reconstructions of the sound-speed field are insensitive to clock errors primarily because of a wide variety of distances between the receivers and each particular source. Every receiver "sees" the same clock

error from the source, regardless of section length, but changes of travel time associated with the sound-speed perturbation depend on the section length. The Kalman filter is thus able to reconstruct clock error and to map the sound-speed field accurately. [Work supported by ARPA, #MDA972-931-0004.]

9:45

1aSPa8. Optimal processing and performance evaluation of passive acoustic systems. Peter L. Greene^{a)} (MIT Lincoln Lab., Group 401, 244 Wood St., Lexington, MA 02173)

The acoustic spectrogram is a fundamental input for many underwater acoustic signal detectors. This presentation describes an effort to place theoretical bounds on the performance of these detectors and to provide insight into the structure of practical processors. The presentation begins with a description of the Neyman-Pearson optimal detector, the likelihood ratio test, and its use for signal detection in colored Gaussian noise scenarios. The presentation continues by developing successively more complex detection scenarios that necessitate the use of background estimation and, consequently, the generalized likelihood ratio test (GLRT). Parameter estimation from several common noise spectral estimation (NSE) techniques is included in the GLRT detection scenarios. In addition, the presentation develops guidelines for matching NSE parameter settings to signal characteristics. Then, the relationships among several relevant signal-to-noise ratios (SNRs) is developed and their correlation with detection performance is considered. Finally, the presentation considers departures from the Gaussian distribution assumption for background and signal processes and the resulting consequences. ^{a)}Ensign, USN

10:00-10:15 Break

10:15

1aSPa9. Passive detection of transient signals in ambient shipping noise. Lisa A. Pflug (Naval Res. Lab., Code 7173, Stennis Space Center, MS 39529-5004), George E. Ioup, and Juliette W. Ioup (Univ. of New Orleans, New Orleans, LA 70148)

A comparison between second- and fourth-order moment detectors for passive detection of low-frequency transient signals in both uncorrelated Gaussian and measured noise is presented. The measured noise, recorded in shallow water near the San Diego harbor, is dominated at low frequencies by ship-generated noise and is thus correlated in time and between sensors. The detectors do not assume noise or signal stationarity and use two channels of data to form the second-order moment, and two or four channels of data to form the fourth-order moment. Detection simulations with test transients of varying kurtosis reveal that in the Gaussian noise, the fourth-order moment detector with two channels of data performs best only for signals with high kurtosis. However, in the measured noise, the fourth-order moment detector performs best for most of the test signals if four channels of data are included. If knowledge of the signal passband is included in the processing and detection is restricted to be within the passband, the fourth-order moment detector with four channels of data generally performs best in both the Gaussian and measured noise. [This work was funded by the Office of Naval Research.]

10:30

1aSPa10. A mode space energy detector for acoustic source depth discrimination. Vincent E. Premus (MIT Lincoln Lab., 244 Wood St., Lexington, MA 02173-9185)

A likelihood ratio test based on a newly defined statistic, the modal scintillation index (MSI), is presented for acoustic source depth discrimination in a shallow-water waveguide. The MSI is defined as the variance in the estimated modal excitation normalized by its expected value over some observation interval. The approach represents, in effect, a mode

space energy detector. In a shallow-water waveguide, the acoustic modes are nearly sinusoidal, and thus all share a zero-crossing and maximum derivative at $z=0$. Consequently, surface source modal excitation levels will be very sensitive to low bandwidth source depth modulation due to surface wave interaction. The hypothesis herein is that a surface source will be characterized by a high scintillation index across all modes, while a source at depth will exhibit a low scintillation index for at least one acoustic mode. The test statistic is self-normalizing, so knowledge of source level and source range is not required to separate the two source classes. Estimation of the modal excitation statistics requires knowledge of the water depth and the sound-speed profile at the array. Stable estimates of the MSI require observation intervals on the order of a few minutes. Performance predictions based on KRAKEN simulations will be presented and both vertical and horizontal array configurations will be considered. It is believed that the proposed technique may have potential as a classifier or clutter management tool, particularly in a littoral environment. [Work sponsored in part by SPAWAR, under Air Force Contract F19628-95-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the U.S. Air Force.]

10:45

1aSpa11. Technique for exploiting the temporal order in acoustic fluctuations to enhance signal processor performance. Ronald A. Wagstaff and Jackson A. Mobbs (Naval Res. Lab., Stennis Space Center, MS 39529-5004, wagstaff@nrlssc.navy.mil)

There are scenarios of interest in which the signals have smaller fluctuation amplitudes than the clutter and noise. When this is the case, it has been shown [R. A. Wagstaff, J. Acoust. Soc. Am. **100**, 2853(A) (1996)] that extraordinary gains can be achieved by signal processors that are designed to discriminate according to fluctuation amplitude. It is also possible to achieve similar gains from processors that are designed to discriminate based on the temporal order dependence of the fluctuations. Since this temporal order-dependent processor is unrelated to other order-independent fluctuation-based processors such as WISPR and AWSUM, the gains can be increased by combining the methods. This order-dependent fluctuation-based signal processor will be described, and the gains including signal-to-noise ratio enhancement, clutter reduction, and unalerted-automatic detection will be discussed. The results will be illustrated using measured data. In addition, results from combining the two different fluctuation-based processing methods, order dependent and order independent, to achieve additional gain will also be presented. [Work supported by ONR and NRL.]

11:00

1aSpa12. Fluctuation-based beamformer for the enhancement of low-fluctuation amplitude signals. Ronald A. Wagstaff (Naval Res. Lab., Stennis Space Center, MS 39529-5004, wagstaff@nrlssc.navy.mil)

A formulation of the cross-spectral density matrix (CSDM, also known as the data covariance matrix) is presented, which is sensitive to the fluctuations that occur naturally in undersea acoustic data. This new formulation, called the AWSUM_k cross-spectral density matrix or A_kCSDM, is patterned after the WISPR and AWSUM filters, both of which are sensitive to amplitude fluctuations. The A_kCSDM utilizes a similar approach to achieving sensitivity to fluctuations, but A_kCSDM goes a step further by including sensitivity to phase fluctuations as well as amplitude fluctuations. The formulation is described and results are presented for beamforming the elements of a towed horizontal line array. The results indicate that in the case of a low-fluctuation amplitude tonal, the signal-to-noise-ratio of the A_kCSDM beamformer is much greater than the conventional CSDM beamformer, and that in some cases the noise and clutter signals

were nearly eliminated. Such results suggest that "noise elimination" is a better description of the A_kCSDM beamformer than is "noise suppression." [Work supported by ONR and NRL.]

11:15

1aSpa13. A simple technique for converting a conventional processor to a fluctuation-based processor and enhancing performance. Ronald A. Wagstaff and S. F. Kooney (Naval Res. Lab., Stennis Space Center, MS 39529-5004, wagstaff@nrlssc.navy.mil)

It has been shown [R. A. Wagstaff, J. Acoust. Soc. Am. **100**, 2853(A) (1996)] that extraordinary gains can be achieved in signal-to-noise ratio enhancement, spatial and spectral resolution can be increased, clutter can be reduced, and submerged sources can be detected by exploiting the natural amplitude fluctuations of the signal and noise. Natural amplitude fluctuations are those fluctuations induced during propagation through the undersea acoustic environment. The gains were achieved by biasing the signal processor, through the WISPR and AWSUM controlling equations, to operate at very low percentiles in the distributions. It is shown in this paper that it is not only possible to achieve additional fluctuation exploitation gains for fluctuation-based processors, such as WISPR and AWSUM, but it is also possible to convert conventional processors to fluctuation-based processors by appropriate modification using a fluctuation sensitive parameter such as the standard deviation. The method is discussed and examples are presented that demonstrate the gains that have been achieved. [Work supported by ONR and NRL.]

11:30

1aSpa14. Sesame I: Assessment of enhancement in the minimum detectable level using fluctuation-sensitive processors. S. F. Kooney (Naval Res. Lab., Stennis Space Center, MS 39529-5004, kooney@miata.nrlssc.navy.mil)

Minimum detectable level (MDL) measurements were made using a vertical array in a shallow-water environment during the Sesame-I experiment held in Aug.-Sept. 1995. A tapered source comb of four frequencies, 150, 154, 158, and 162 Hz, were broadcast at different levels and two depths at a range of 22 nmi. The MDL data were evaluated in terms of signal excess and standard deviation using the average-power processor and fluctuation-sensitive processors. It was found that the fluctuation-sensitive processors provided a signal enhancement of up to 7 dB in cases when the average-power processor's signal excess was zero. In terms of standard deviation, the fluctuation sensitive processors could discern a signal buried deeper in the background noise (where the standard deviation of the noise levels were typically 5 to 6 dB) than could the average power processor. It was determined that fluctuation-sensitive processors could provide an enhancement of the MDL over the average-power processor. The methods of analysis will be discussed and the comparative performance of the algorithms will be presented. [Work supported by ONR and NRL.]

11:45

1aSpa15. Environmentally sensitive adaptive fluctuation-based processing. Ronald A. Wagstaff and Jackson A. Mobbs (Naval Res. Lab., Stennis Space Center, MS 39529-5004, wagstaff@nrlssc.navy.mil)

Signals that propagate from submerged sources to submerged receivers in the ocean generally interact less with the fluctuation generation mechanisms near the sea surface, e.g., internal and sea-surface waves, than signals originating near the surface, e.g., surface ship clutter and sea-surface noise. Therefore, submerged source signals generally have smaller fluctuation

tuations than clutter signals and noise. The differences in amplitudes of fluctuations have been used to devise a class of signal processing algorithms that provides preferential gains for signals having fluctuation amplitudes less than those of clutter and noise. Gains include increases in signal-to-noise ratio, clutter suppression, and unalerted automatic detection. The AWSUM filter is an example of such a processor [R. A. Wagstaff, IEEE J. Oceanic Eng. 22(1) (1997)]. Similar gains can be achieved

by exploiting the order dependence of fluctuation amplitudes. By combining amplitude dependence and order dependence, an algorithm has been devised that is both sensitive to fluctuations and adapts to the input data. The resulting algorithm is designated the AWSUM Environmentally Sensitive Adaptation (ESA). The AWSUM ESA algorithm is described and results from processing real data are presented. [Work supported by ONR and NRL.]

MONDAY MORNING, 16 JUNE 1997

ROOM E, 8:15 TO 10:55 A.M.

Session 1aSPb

Signal Processing in Acoustics, Noise, and NOISE-CON: Signal Processing for Condition Based Maintenance I

Daniel J. McCarthy, Chair
BBN Corporation, 70 Fawcett Street, Cambridge, Massachusetts 02138

Invited Papers

8:15

1aSPb1. Prognostics: What does it mean in condition-based maintenance? G. W. Nickerson (Penn State Univ., Appl. Res. Lab., State College, PA 16802)

The machinery condition monitoring community continues its search for the "Holy Grail" of prognostics. The problem is that there is no common agreement on two questions: "What does the term prognostics mean in this context"? and "What must be done to achieve it"? This paper differentiates "predictive maintenance" (in its generally accepted form) from "condition-based maintenance" (which requires prognostics). A justification for development and implementation of the more advanced and cost-effective condition-based maintenance is presented. A working definition of prognostics for the condition-based maintenance community is proposed. Further, a methodology for setting acceptable false alarm rates for both diagnostics and prognostics is presented to frame the problem for the research community in terms of the application environment, particularly for mission critical equipment.

[See NOISE-CON Proceedings for full paper.]

8:35

1aSPb2. Condition monitoring on engines using response noise estimates and reconstructed excitation. Kwin Abram, Paul Hayes (Noise and Vib. Lab., Cummins Engine Co., Columbus, IN), and Robert Bernhard (Purdue Univ., West Lafayette, IN 47907-1077)

A condition monitoring technique is developed for engines and similar machines whereby only external response measurements are used to determine the condition of the machine. The method assumes the response signal is composed of components due to an event of interest and other events which are not of interest. The response due to the events of no interest are treated as deterministic noise. An identification process is used to identify the noise and the frequency response function between the event of interest and the response. This identification procedure uses operating conditions which span the faults of interest. This model is used in an inverse filter process to determine the excitation during operation. Signal filtering and time gating are used to enhance the procedure. The method is verified for combustion faults on a diesel engine.

[See NOISE-CON Proceedings for full paper.]

1aSPb3. Bearing cage fault detection using bandwidth-weighted demodulation. C. James Li (Dept. of ME AE & M, Rensselaer Polytech. Inst., Troy, NY 12180)

When faults occur in bearing cages, they generate an extremely wideband fault signature modulated by the cage rotating frequency as the defective part of the cage rotates in and out of the load zone. Current bearing fault detection methods, which were developed for race and rolling element defects, are aimed at isolating and demodulating relatively powerful high-frequency, narrow-band ringing vibrations, and therefore have difficulty detecting bearing cage faults. Bandwidth-weighted demodulation (BWD), however, is well suited to cage faults because it is more sensitive to wideband, low-power signal modulations than to narrow-band, high-power signal modulations. A bearing cage fault detection algorithm has been established based on BWD. The effectiveness of BWD compared to other methods is demonstrated in simulation and confirmed using experimental data from a failed bearing.

[See NOISE-CON Proceedings for full paper.]

9:15

1aSPb4. Prognostics for a reciprocating air compressor. Daniel J. McCarthy (BBN Corp., 70 Fawcett St., Cambridge, MA 02138, dmccarthy@bbn.com)

The health of a reciprocating air compressor can affect the time of occurrence of valve impacts within the compressor. A computer simulation reveals useful insight into the nature of the timing changes, which provides a basis for automated prognosis of the machine health. As an example, consider that the opening impact of a leaky discharge valve is delayed by an amount proportional to the size of the leak. The effect that the leak has on the efficiency of the compressor can be determined using the simulation, and a threshold for lost efficiency can be tied to the time of occurrence of the valve impact. Using the computer simulation as an aid, the vibration transients generated by the impacts can be clearly identified in the vibration signal measured on the exterior of the compressor. The time of occurrence of the impacts can then be accurately determined from the measured vibration using minimum-phase processing, a sliding exponential window, and statistical curve fitting.

[See NOISE-CON Proceedings for full paper.]

9:35–9:55 Break

9:55

1aSPb5. Extraction of periodic modulation of rotational speed from machinery vibration signals using a modulation template matching algorithm. Michael B. Van Dyke, Karl Reichard, and Dave Swanson (Penn State Univ., Appl. Res. Lab., University Park, PA 16802)

The periodic modulation of the rotational speed of a machine element, such as a bearing element or a gear, is recognized as a valuable feature in characterizing the physical condition of the machine component. A time-domain template-matching algorithm is introduced as a robust nonlinear frequency modulation estimation technique for extracting this feature from perturbations in the acoustic signal emitted from a damaged machine component. The technique is applied to both synthetic signals as well as real accelerometer and microphone data taken from damaged railroad car roller bearings. The results are compared with those obtained using a standard baseband phase demodulation technique. Although the template-matching algorithm is computationally more expensive, it can be applied directly to a limited-duration, raw input signal, unlike the baseband demodulation technique, which relies on bandpass filtering and synchronous averaging over a lengthy time duration.

[See NOISE-CON Proceedings for full paper.]

Contributed Papers

10:15

1aSPb6. Analysis of engine noise for application in predictive maintenance. G. Douglas Meegan, H. Roice Nelson, Martin L. Barlett, and Gary R. Wilson (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029)

The analysis of radiated sound and vibrational signatures of engines may be useful for monitoring an engine's performance and assessing the need for maintenance. Measurements of the vibrational signatures and exhaust noise produced by normal engines and engines with known faults will be presented along with a corresponding model for engine exhaust noise. The results of power spectral and higher-order spectral processing of the engine noise will be compared and their usefulness as a means of detecting faulty engine operation will be discussed. [Work supported by the Office of Naval Research, Contract No. N00014-96-1-0298-7.]

10:35

1aSPb7. Nonlinear acoustic modulation technique for nondestructive flaw detection. J. P. Kim, E. J. Kim, S. W. Yoon (Acoust. Res. Lab., Dept. of Phys., Sung Kyun Kwan Univ., Suwon 440-746, Rep. of Korea), and A. Sutin (Inst. of Appl. Phys., Nizhny Novgorod 603600, Russia)

A new nonlinear acoustic modulation technique was developed to determine flaw locations in a material as a nondestructive evaluation method. Nonlinear interaction of two primary low-frequency continuous and high-frequency burst-type acoustic waves in the material can generate modulated signals. Such modulated signals show significant difference in the material with flaws. The experimental measurements were performed with steel and aluminum rods. Two PZT ceramic elements attached at the ends

of rods were driven with two primary frequencies by separate function-generators through power amplifiers. Modulated signals detected at a receiving sensor were analyzed by a digital storage oscilloscope with the function of fast Fourier transformation. The second harmonic amplitude

was observed 20 dB higher in a sample material with crack than that without crack. The crack location was also very accurately estimated by this nonlinear acoustic modulation technique at the sum and difference frequencies of the primary waves.

MONDAY AFTERNOON, 16 JUNE 1997

DEANS HALL I AND II, 2:00 TO 6:00 P.M.

EQUIPMENT EXHIBITION

The following companies plan to exhibit at the joint meeting:

01 dB, Inc.
ACO Pacific, Inc.
Bang Campbell Associates
Cambridge Collaborative, Inc.
CEL Instruments, Inc.
Crown International
Digisonix, Inc.
Eckel Industries, Inc.
Etymotic Research
H. L. Blachford, Inc.
Industrial Acoustics Company, Inc.
Lake-DSP
Larson-Davis, Inc.
LMS North America

MBI Products Co.
M&P International
Ono Sokki, Inc.
Overly Manufacturing
Polytech PI, Inc.
Scantek, Inc.
Sonic Perceptions, Inc.
Sony Precision Technology, Inc.
Spectral Solutions, Inc.
TEAC America
The Modal Shop, Inc.
Tucker-Davis Technologies
Vibro-Acoustic Sciences, Inc.

MONDAY AFTERNOON, 16 JUNE 1997

ROOM P, 2:00 TO 4:45 P.M.

Session 1pAO

Acoustical Oceanography and Underwater Acoustics: Acoustical Measurement of Coastal Ocean Processes II

Mohsen Badiey, Chair

College of Marine Studies, University of Delaware, Robinson Hall, Newark, Delaware 19716

Chair's Introduction—2:00

Invited Paper

2:05

1pAO1. Recent shallow water measurements of signal coherence and array performance in several shallow water areas.
William M. Carey (Dept. of Ocean Eng., MIT, Cambridge, MA 02139) and James Reese (RDTE Div. NRAD)

Experimental measurements of signal coherence and array signal gain are reviewed for both deep and shallow water sound channels. The signal gain is related to the horizontal coherence length through relationships from the statistical theory of antennas. Signal gain measurements in the transverse direction are proffered as a practical measure of coherence length for both broadband and narrow-band signals. Using this technique, measurements at frequencies near 400 Hz are presented that show for the deep water cases, lengths on the order of 100 wavelengths can be achieved, while in the downward, refraction conditions of shallow water waveguides with sand-silt bottoms, lengths on the order of 30 wavelengths are realized. Recent measurements of broadband and narrow-band coherence and signal gain are discussed with emphasis on the role of partly coherent noise backgrounds due to coastal shipping, multipath interference effects, and averaging constraints. In particular, the results of an experiment conducted in the Straits of Korea will be presented where array gain and sonar images were obtained in the presence of a strong coastal current.

2:25

1pAO2. Horizontal coherence in shallow water. Ilya Rozenfeld (Dept. of Math. Sci., Rensselaer Polytech. Inst., Troy, NY 12180), Peter Cable (BBN Systems and Technologies, New London, CT 06320), William Carey (MIT, Cambridge, MA 02139), and William L. Siegmann (Rensselaer Polytech. Inst., Troy, NY 12180)

For applications such as beamforming and matched field processing, the spatial behavior of the shallow water acoustic field is relevant. Oceanic and geoacoustic variabilities introduce random fluctuations into the field, which can be characterized by the spatial coherence.

of the acoustic field can be described by its correlation function. In this presentation results from an investigation of the sensitivity of the horizontal correlation function to variations in parameters of shallow water propagation will be discussed. In order to demonstrate key sensitivities, stratified environmental models with one or more sediment layers will be used. Variabilities are assumed to arise from small random fluctuations within the layers and at the layer interfaces. Formulas based on a modal perturbation approach are used. The dependence of the correlation function on such parameters as range and frequency is of particular interest, and in addition effects of mode coupling are estimated. Numerical simulations illustrating the results will be presented. [Work supported by NSF.]

Invited Paper

2:40

1pAO3. Acoustic tomography in the autonomous oceanographic sampling network. Henrik Schmidt, James Bellingham, Pierre Elisseff, and Max Deffenbaugh (MIT, Cambridge, MA 02139, henrik@keel.mit.edu)

Ocean acoustics plays a central role in the new autonomous oceanographic sampling network (AOSN) concept. Thus, as in most other underwater systems, acoustics provides the only reliable means for wireless communication and navigation. Also, the sensor suite applied in the AOSN includes various sonars for object avoidance and imaging, and acoustic Doppler current profilers. By adding a tomographic component, the performance of the AOSN in highly dynamic littoral environments can be significantly improved. Thus, the tomography can provide a low resolution, but also a synoptic image of the oceanography which can eliminate the space-time aliasing inherent to the point measurements made by the AUVs. Using real-time processing the tomographic map is used to adaptively focus the AUV measurements in areas with high gradients or large uncertainty. In addition, the AUVs can enhance the tomographic coverage by equipping them with acoustic sources. The feasibility of this acoustically focused oceanographic sampling (AFOS) concept was investigated in the June 1996 Haro Strait PRIMER experiment. The experiment clearly demonstrated the feasibility of the tomographic components. Robust and efficient algorithms have been developed, capable of performing real-time tomographic inversion within a littoral AOSN. The results of the experiment are reviewed, including frontal mapping performed in more than 80 AUV missions. The experiment also clearly demonstrated the severity of the littoral acoustic environment in terms of ambient noise, reverberation, and variability, significantly affecting the robustness of the acoustic communication and navigation systems crucial to the AOSN. [Work supported by ONR.]

Contributed Papers

3:00

1pAO4. High-frequency reciprocal transmissions across a coastal front in Haro Strait, British Columbia. Pierre Elisseff, Henrik Schmidt (MIT, Cambridge, MA 02139), and Mark Johnson (Woods Hole Oceanograph. Inst., Woods Hole, MA 02543)

The Haro Strait experiment was conducted in Haro Strait, British Columbia (Canada) in order to determine the feasibility of acoustically focused oceanographic sampling (AFOS). In the course of the experiment 15-kHz pseudorandom sequences with a bandwidth of 3 kHz were reciprocally transmitted across a front driven by estuarine and tidal forcing. Temperature profiles were continuously measured at the array locations, and array elements were acoustically tracked. Transmissions were performed over a period of 36 h, providing an exhaustive acoustic characterization of the frontal dynamics. The sensitivity of arrival times to the environmental variability is first investigated. Both current and temperature fields across the front are estimated using travel-time tomography. Resolution issues associated with source-receiver parameters such as geometry, frequency, and bandwidth are discussed in light of the experimental data. Finally, the relevance of the high-frequency acoustic tomographic inversion to AFOS is assessed. [Work supported by ONR.]

3:15

1pAO5. Twinkling of low-frequency tones from a moored source transmitting through and away from the surf zone. Gerald L. D'Spain, Lewis P. Berger, K. Megan McArthur, William S. Hodgkiss, Michael J. Ritter, and William A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0704)

During a recent seismoacoustics experiment, a moored, near-bottom, source just outside the surf zone (approximately 500-m seaward) in 9-m deep water, transmitted 8 tones from 70 to 700 Hz. Recordings were made by several fixed sensor systems, including: 1) a bottom hydrophone/3-component geophone package 15-m distant; 2) two horizontal bottom hydrophone arrays in a nearly perpendicular configuration in 12-m water 1.23 km away and 1.06 km farther offshore; and 3) five 3-component geophones buried at approximate 14-m spacing across the beach at 480–530 m range. These data were used to create 5-min time series of the spectral levels and directionality of the tones every 1/2 h over an 8-h period. Results indicate that within a 5-min period, fluctuations of 10–15 dB occur in the bottom hydrophone array signal levels which, typically, are not correlated from frequency to frequency. Fluctuations in the land geophone data are somewhat smaller, and only the 2 or 3 highest-frequency tones are detectable. Variations in water depth, measured by a package also located just outside the surf zone, are used with an adiabatic normal mode model to predict the fluctuation effects of ocean surface wave activity alone. [Work supported by ONR, Code 32].

3:45

1pAO6. Broadband acoustic imaging of breaking waves. Rex K. Andrew (Dept. of Elec. and Comput. Eng., Univ. of Victoria, P.O. Box 3055, Victoria, BC V8W 3P6, Canada), David M. Farmer (Inst. of Ocean Sci., Sidney, BC V8L 4B2, Canada), and R. Lynn Kirlin (Univ. of Victoria, Victoria, BC V8W 3P6, Canada)

An acoustic array was deployed in the near-surface layer in Saanich Inlet, BC, to image breaking waves using only the naturally occurring acoustical radiation in the band (160 and 2000 Hz) from the breaking region. The 15-element array was configured as a horizontal cross with an 8-m aperture, bottom-moored, and positioned nominally 3 m beneath the surface. A novel broadband scheme combined information at six independent frequencies above about 400 Hz to yield unambiguous resolved images. A parametric image analysis shows that the images align closely with the wind and can be observed moving downwind with a speed roughly equal to the dominant phase speed of the wind waves. Absolute power levels are found to be consistent with previously published results. The data also provide inferences regarding the sound generation mechanism at “collective oscillation” frequencies below about 400 Hz. [Work supported by ONR.]

4:00

1pAO7. A tomographic algorithm applied to imaging bubbly water. Daniel Rouseff and Frank S. Henyey (Appl. Phys. Lab., Univ. of Washington, Box 355640, Seattle, WA 98105, rouseff@apl.washington.edu)

A tomographic algorithm believed appropriate for reconstructing the effective index of the refraction structure of bubbly water is developed. The method is optimum in a least-squares sense if the medium can be interpreted as a realization of a Gaussian random process. In contrast to other tomographic algorithms, it can be applied when the number of ray paths is sparse. The measured input may be acoustic travel time, to reconstruct sound speed, or log intensity, to reconstruct attenuation. The data are multiplied by a precalculated matrix that depends on the geometry and the assumed autocorrelation function of the medium. This gives the weights used in the imaging step. To produce an image, the data from each ray are weighted, then backprojected over a smeared region centered about the original ray path. Superimposing the weighted smeared images produces the final reconstruction. The computationally expensive work can be done prior to processing any data, giving the potential to do nearly real-time imaging. The algorithm is tested in numerical simulations of the March 1997 shallow water bubbles experiment. The locations of the sources and receivers are varied. The effects of environmental mismatch are studied.

1pAO8. Heat flux monitoring in the Fram Strait: Limitations on travel-time measurements. Konstantin A. Naugolnykh (CIRES, Univ. of Colorado/NOAA/Environ. Technol. Lab., 325 Broadway, Boulder, CO 80303), Ola M. Johannessen (Nansen Environ. and Remote Sensing Center, Bergen, Norway), Igor B. Esipov, Oleg B. Ovchinnikov, Dimirii E. Leikin, and Yrii I. Tuzhilkin (N. Andreev Acoust. Inst., Moscow 117036, Russia)

Travel-time measurements form the common basis for all inverse techniques in ocean acoustics. These techniques rely on the assumption that the received signal consists of a train of pulses identical to that produced by the source. However, when few modes survive long-distance propagation, the validity of the travel-time measurement can be compromised by diffraction effects. A primary goal of this paper is to evaluate the restrictions caused by strong mode coupling in irregular sound channels. The results of numerical simulations of pulse propagation through a range-dependent medium with a single sound channel will be presented. The time-domain patterns were obtained by Fourier synthesis based on coupled-mode calculations for various inhomogeneities using a 100-Hz frequency band acoustic signal. The effects of modal dispersion and mode coupling on the accuracy of the pulse-travel-time measurements will be discussed.

4:30

1pAO9. The acoustics of snapping shrimps. Whitlow W. L. Au (Hawaii Inst. of Marine Biology, P.O. Box 1106, Kailua, HI 96734) and Kiara Banks (Florida Inst. of Technol., Melbourne, FL 32901)

With an increased interest in acoustic daylight, the contribution of snapping shrimp noise has drawn considerable attention. These crustaceans are among the major contributors of biological noise in shallow waters of temperate and tropical regions. The shrimp produce sounds by rapidly closing one of their frontal claws, snapping the ends together to generate a loud click. The acoustics of the species *Synalpheus paraneomeris* was studied by measuring the sound produced by individual shrimp housed in a small cage located 1 m from an H-52 hydrophone. Ten clicks each from 40 specimens were digitized at a 1-MHz sample rate. A low-frequency precursor signature (never reported before), which may be associated with the plunger on one of the claws that directs a water jet forward, was observed. The peak-to-peak source level varied linearly with claw size and body length and had values from 183 to 189 dB re: 1 μ Pa. The acoustic power produced by a typical snap was calculated to be about 3 W. A typical spectrum was very broad with only a 20-dB difference between the peak at 2–5 kHz and components out to 200 kHz.

Session 1pEA

Engineering Acoustics: Optical Sensors/Techniques and Transducer Calibration

David A. Brown, Chair

*Department of Electrical Engineering, University of Massachusetts, Dartmouth, 285 Old Westport Road,
Dartmouth, Massachusetts 02747*

Chair's Introduction—1:10

Contributed Papers

1:15

1pEA1. Design and development of a marine acoustic accelerometer.

David Brown and Rama Buddhineni (Dept. of Elec. and Comp. Eng. and Center for Marine Sci. and Tech., Univ. of Massachusetts, Dartmouth, MA 02747)

A push-pull fiber-optic acoustic accelerometer using Mach-Zehnder interferometry to detect surface strains in a centrally clamped, peripherally free disk will be described. The measured sensitivity is approximately 20 rad/g below the resonance of 10 kHz. A total of 36 m of optical fiber wound in a pancake spiral was used for each interferometer leg. The disk is housed in a cylindrical case that allows the device to be near neutrally buoyant in water, thus acting as an acoustic motion sensor. The interferometer was interrogated with a Nd:YAG laser at 1310 nm and a 3×3 fiber coupler. Considering a total disk mass of 19.6 gm, the fractional phase change per unit force is $4.0 \times 10^{-6} \text{ N}^{-1}$. Based on an electro-optic noise floor of $5 \mu\text{rad/Hz}^{1/2}$, the minimum detectable signal is $0.25 \times 10^{-6} \text{ g/Hz}^{1/2}$. [Work supported by the University of Massachusetts, Dartmouth.]

1:30

1pEA2. Forward projection of transient signals obtained from a fiber-optic pressure sensor. Gregory Clement, Ruiming Liu, Stephen Letcher (Dept. of Phys., Univ. of Rhode Island, Kingston, RI 02881), and Peter Stepanishen (Univ. of Rhode Island, Kingston, RI 02881)

An analytical/experimental approach is presented to reconstruct the space-time pressure field in a plane and forward project the resultant space-time pressure field using tomographic and wave-vector frequency-domain methods. Transient pressure signals from an underwater ultrasonic planar transducer are first measured using a fiber-optic pressure sensor which is scanned across a plane at a fixed distance from the transducer. The resulting spatial line integrals in the plane are time-dependent signals that are first used to reconstruct the space-time pressure field in the plane via simply implemented tomographic methods. These signals are then used to forward project the space-time pressure field to arbitrary planes further from the transducer. Numerical results will be presented for transient signals to illustrate both the projection and the detection techniques. The results of the projected fields will be compared at various distances for synthetic signals and experimental data. [Work supported, in part, by the URI Ocean Technology Center and by ONR.]

1:45

1pEA3. Photoacoustic signals from small gas leaks. Serdar H. Yönak and David R. Dowling (Dept. of Mech. Eng. and Appl. Mech., Univ. of Michigan, Ann Arbor, MI 48109-2121)

In the manufacture of heat exchangers or other devices filled with compressed gases or liquids, efficient detection of leaks is a critical process. Photoacoustics has been used to quickly detect these leaks in an industrial environment. Although this technology can detect leaks as small as 10^{-6} standard cc/s, it has not been extended to localization or quanti-

fication of these leaks. This work aims to resolve these issues. An acoustic source strength based on the Beer-Lambert law and gas diffusion from a point source is developed and used to predict photoacoustic signals. The relationship between photoacoustic signals and leak test parameters, specifically leak rate, modulation frequency, and laser power, are presented. An effective strategy for illuminating the suspected leak area is shown, and leak detection using multiple microphones is discussed. A carbon dioxide laser tuned to 10.6 microns is used as the radiation source with sulfur hexafluoride as the tracer gas. Comparisons are made between theoretical predictions and experimental results. [Work sponsored by Ford Motor Company.]

2:00

1pEA4. Acoustic field reconstruction using optical interferometry.

Christophe Mattei^a and Laszlo Adler (Industrial, Welding and System Eng., Ohio State Univ., Columbus, OH 43210-1182, cmatt001@postbox.acs.ohio-state.edu)

Acousto-optical measurements inside transparent media is a well-known noninvasive method for acoustic field probing. The laser interferometric detection technique provides a noncontact, local, and quantitative measurement of pressure or dilatation fields in transparent media [X. Jia, Ch. Mattei, and G. Quentin, *J. Appl. Phys.* **77**, 5528–5537 (1995)]. The optical phase shift of a probe beam passing through an ultrasonic beam is detected using a Mach-Zehnder interferometer and is related to the amplitude and phase of the pressure field. Since the optical phase shift integrates the pressure field over the optical beam pathlength, this technique assumes a constant acoustic field along the laser beam path. In order to obtain a complete mapping of the acoustic field, as required in many applications, e.g., transducer characterization, we propose the use of tomographic techniques associated with the interferometric detection which permits a reconstruction of the acoustic field in the acoustic beam cross section from a set of projection data. The time variation of the pressure field in a cross section as well as the amplitude and phase spectra can be displayed. The mapping of the near pressure field (1 mm) of an ultrasonic transducer is achieved using this technique and other examples of field characterization are presented. ^aPermanent address: G.P.S., Univ. Paris 7, CNRS Ura No. 17, 2 place Jussieu, 75251 Paris Cedex 05, France.

2:15

1pEA5. Holographic analysis of standing sound waves in a resonance tube.

Russell S. Hughes, Daniel A. Russell, and David E. Parker (Sci. and Math. Dept., GMI Eng. & Management Inst., 1700 W. Third Ave., Flint, MI 48504)

Optical holographic interferometry was used to observe standing sound waves in air inside a resonance tube. The resonance tube had a rectangular cross section and was driven by a small loudspeaker at one end. The front face of the resonance tube was constructed of Plexiglas, allowing optical interrogation of the interior of the tube. The object beam of the holographic setup was directed through the Plexiglas and reflected off the back wall.

When driven at resonance, the fluctuation in air density at antinodes altered the index of refraction in the tube, causing interference patterns on the resulting hologram. Both real-time and time-averaged holograms were made of the first 15 longitudinal standing wave modes. The first eight modes agree very well with predictions for a closed-closed tube. This talk will focus on the holographic experiment and will demonstrate the visualization of standing sound waves.

2:30

1pEA6. Ultrasonic enhancement of fiber-optic biosensors. Chonghua Zhou (Dept. of Phys., Univ. of Rhode Island, Kingston, RI 02881), Philip Pivarnik, A. Garth Rand, and Stephen Letcher (Univ. of Rhode Island, Kingston, RI 02881)

An enhanced fluorescent fiber-optic biosensor system using ultrasonic concentration of particles and cells has been developed and applied in the detection of *Salmonella typhimurium*. A biosensor test chamber also serves as an ultrasonic standing-wave cell that allows microspheres or cells to be concentrated in parallel layers or in a column along the axis of the cell. A fiber probe along the axis of the cell delivers laser excitation to fluorescent dye molecules and collects the fluorescent signal. In a "sandwich" assay, the dye molecules are bonded to *Salmonella* antibodies that, in turn, attach themselves to *Salmonella* cells that have been captured by unlabeled antibodies that have been immobilized on the surface of polystyrene microspheres. This entire structure can be manipulated acoustically. The most effective procedure is a uniform stratification of the complexes in a plane-wave field followed by a concentration along the axis of the cell by a Gaussian-profile field. The increase in the fluorescent signal, which can be an order of magnitude, indicates the presence of *Salmonella*. [Work supported by USDA/NRICGP and by the URI Partnership for Sensors and Surface Technology.]

2:45

1pEA7. Vibration isolation and detection using a mesoscopic photomechanical unit. Christopher S. Kwiatkowski (Sentel Technologies, NE 1615 Eastgate Blvd., Box 41-1, Pullman, WA 99163), David Welker (Sentel Technologies, Pullman, WA 99163), and Mark G. Kuzyk (Washington State Univ., Pullman, WA 99164-2814)

Due to the high bandwidth of light, optical fibers are being used extensively in data communication systems, and more recently have found applications in sensors capable of monitoring environmental changes such as stress, displacement, temperature, and chemical concentrations. An active region consisting of a Fabry-Perot cavity inside a nonlinear polymer optical fiber, called a Mesoscopic Photomechanical Unit (MPU), has been studied that has applications in both sensing and vibration isolation. These devices consist of a PMMA core doped with DR1 organic dye. It has been shown that the length of these MPUs has a series of stable states, the stability of which is dependent on the intensity of guided light inside the fiber. Preliminary data show that a vibration sensor incorporating an MPU as its sensing element has the advantage of having an intensity-dependent response. This effect could be used to include all-optical logic within the device, making external electronics unnecessary. [Research supported by NASA and ARO.]

3:00

1pEA8. A laser-based experimental-numerical technique for evaluating the bulk and shear dynamic elastic moduli of viscoelastic materials. T. Shane Stone, R. Lance Willis, and Yves H. Berthelot (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405)

A laser-based numerical/experimental technique suitable for laboratory studies has been designed and tested to simultaneously measure the bulk and shear complex dynamic moduli of a viscoelastic sample excited harmonically in the 0.5–3 kHz range by a shaker. The method consists of measuring the dynamic response (amplitude and phase) of the sample at five surface points with noncontact laser Doppler interferometric probes measuring in-plane (tangential) and out-of-plane (normal) surface veloci-

ties. To determine the bulk and shear moduli as functions of frequency, the measured dynamic surface response is matched to the predictions obtained from a finite element model of the sample in which the two complex elastic moduli are the adjustable parameters. This inversion procedure is based on a classical multi-dimensional direction set method (Powell's method) of optimization theory. The results are based on measurements made in air under standard pressure and temperature conditions. However, work is under way to build a similar system that would allow for measurements at high static pressure and controlled temperature. [Work supported by ONR, Code 334.]

3:15–3:30 Break

3:30

1pEA9. New automated microphone reciprocity calibration system and results obtained. Erling Frederiksen (Calibration Systems Dept., Brüel & Kjær, 2850 Nærum, Denmark, erlingfred@bk.dk) and Joergen I. Christensen (Brüel & Kjær, 2850 Nærum, Denmark)

Microphone calibration by the reciprocity technique is performed by national and other high-level acoustic calibration laboratories. In the past, the electrical transfer function of coupled pairs of microphones was determined by manual measurements. As such measurements were very time-consuming, only a few calibrations were performed at a rather limited number of frequencies. A new automated measurement system has been developed. This has significantly increased the efficiency of transfer function measurements and extended the practical applications of the reciprocity technique. The uncertainty of voltage ratio measurements is less than 0.02 dB ($k=2$) between 40 Hz and 30 kHz, while the repeatability is about two times better. These numbers are so low that they have only a minor influence on the resulting uncertainty of laboratory standard microphone calibrations. Today, pressure response calibrations (according to IEC1094-2) can be performed with an uncertainty of 0.02 to 0.04 dB. The probe for measurement of electrical current through the transmitter and the preamplifier for the receiver microphones are both connected via cables. This makes it possible to place the microphones remotely from the main instrument, as required with environmental testing (pressure, temperature) and free-field calibrations. Calibration results obtained with the new system will be presented.

3:45

1pEA10. Calibration. Primary—Secondary—Field. Ernst Schonthal (PCB Piezotronics, Inc., 3425 Walden Ave., Depew, NY 14043)

The primary calibration of acoustic and vibration transducers are well standardized and confirmed by a round-robin inter-laboratory test. Secondary calibration is also very common with traceability to primary calibration. Examples of setups for microphone and accelerometer secondary calibration are discussed regarding accuracy, frequency response testing, etc. Field check methods using reference sources and other methods like Phantom calibration are discussed regarding advantages and shortcomings.

4:00

1pEA11. Evaluation of instrument standards with virtual sound level meters. Lixue Wu and George S. K. Wong (Acoust. Standards, Inst. for Natl. Measurement Standards, Natl. Res. Council, Ottawa, ON K1A 0R6, Canada, george.wong@nrc.ca)

It is not an easy task to evaluate instrument standards such as ANSI S1.4, IEC 651, IEC 804, and IEC 1672: Sound level meters, which is under development. Traditionally, proposed calibration procedures are tested with existing instruments from various manufacturers. Obtaining quantitative data to justify proposals on tolerances and to measure the combined effects of tolerances on the final instrument readings is extremely difficult. The above tasks can be done with relative ease with a computer emulation of virtual instruments. A sound level meter based on virtual instrumentation using the latest software technology, cost-effective data acquisition,

control hardware, and a computer is presented. For manufacturers of acoustical instruments, a graphic designer interface also makes it possible to evaluate "user-defined" instrument systems with reduction in development time and costs.

4:15

1pEA12. On the estimation of uncertainties for the calibration of standard accelerometers using the fringe counting method. Guillermo Silva-Pineda (Vib. and Acoust. Div., Natl. Ctr. of Metrology in Mexico, CENAM, Mexico CP76900, Mexico)

The uncertainties calculation procedure for the estimation of errors when the fringe counting method is used to get the sensitivity of standard accelerometers is developed. The fringe counting method is used in the frequency range from 50 to 800 Hz and into a dynamic range from 20 to 100 m s⁻². Also, the reference condition is set up at 160 Hz and 50 m s⁻², where the main objective is to get the lowest uncertainty between all the frequency and dynamic ranges mentioned above. This calculation procedure agrees with both the "Guide to the expression of uncertainty in measurement" and the ISO 5347-1 "Primary vibration calibration by laser interferometry." Furthermore, type B uncertainties are received from specific technical information from the equipment used during the calibration procedure and the type A uncertainty is received from periodic measurements performed on the standard accelerometer which have been monitored. Data from primary calibrations have been used to validate the estimation of uncertainties. Moreover, this paper shows the differences on the levels of error sources involved in a standard accelerometer calibration.

4:30

1pEA13. The effects of microphone mismatch on the performance of a spherical acoustic vector-field sensor. John W. Parkins, Jiri Tichy (Graduate Program in Acoust., Penn State Univ., Appl. Sci. Bldg., North Atherton St., State College, PA 16801, jparkins@sabine.acs.psu.edu), and Scott D. Sommerfeldt (Brigham Young Univ., Provo, UT 84602)

Elko previously established the performance improvement for intensity measurements of an acoustic sensor using two microphones embedded on the surface of a hard sphere to that of a sensor comprising two infinitely small microphones separated in space [G. Elko, Noise-Con 91, pp. 525–532]. Elko derived his results for perfectly matched microphones, with regard to phase and sensitivity, in a plane-wave field with a variable angle of incidence. The performance of a spherical sensor and a two-point sensor are compared here for the case of a one-dimensional standing wave field with an arbitrary reflection coefficient and incidence angle, when the measurement microphones have a phase and sensitivity mismatch. Results are generated using sensitivity and phase variations typically found in inex-

pensive electret microphones. The effect of bias errors in potential energy density, kinetic energy density, total energy density, and intensity are reported.

4:45

1pEA14. A study of noise and vibration measurements in an engine dynamometer cell using intensity methods. Satish Padiyar (Ford Motor Co., Engine Manufacturing Development Operations, FEAD NVH; 17000, Southfield Rd., Allen Park, MI 48101)

In automobile engine development, the use of dynamometer cell tests has been established as often-times, a reliable and efficient alternative to in-vehicle tests. Engine noise, vibration, and harshness (NVH) characterization involves studying the relationship between the engine dynamics and NVH parameters of interest. In dynamometer cells that have been designed to be fully or partially anechoic in a given frequency range, noise measurements can be made reliably in the designated frequency range, without concern for the effects of reflections from the cell walls. For NVH measurements in dynamometer cells outside the designated anechoic frequency range, engineers and researchers often resort to expensive anechoic enclosures around the engine alone. This paper will discuss the results of the use of intensity techniques for noise measurements, in an attempt to increase the usable frequency range of an engine dynamometer cell for automobile NVH characterization.

5:00

1pEA15. Sensors for ice detection on aerospace structures. Xiao-Qi Bao, Vasundara V. Varadan, and Vijay K. Varadan (149 Hammond Bldg., Penn State Univ., University Park, PA 16802)

Detecting the onset of ice formation on fixed wing and rotary wing aircraft before and during flight is very important. The experimental results of using a new type of piezoelectric sensor oscillator will be presented. The d15 coefficient of PZT (lead zirconate titanate) is exploited by poling a thin rod (1.3-mm-diam) specimen of PZT along its length direction, and the rod specimens reelectroded diametrically. Two rods were attached to an aluminum plate, and an oscillator circuit was set up using an amplifier and a frequency counter. The plate mode determines the oscillation frequency. Since this is a shear mode, the frequency is unchanged if there is a thin layer of water on the plate. The water is then frozen and a jump in the oscillation frequency is noted at the onset of freezing. Shear waves can couple to ice in solid form, and this changes the dispersion relation, and hence the plate mode frequency. Results are presented showing the sensitivity of the measurement, repeatability, effect of temperature and ice thickness. This is proposed as a reliable, inexpensive sensor arrangement for detecting the onset of ice formation on aircraft.

MONDAY AFTERNOON, 16 JUNE 1997

ROOM D, 1:30 TO 5:00 P.M.

Session 1pNSa

Noise, Architectural Acoustics, and NOISE-CON: Room Noise Criteria

Hsien-sheng (Jason) Pei, Chair

Digital Equipment Corporation, P.O. Box 5327, Cochituate, Massachusetts 01778

Invited Papers

1:30

1pNSa1. Practical considerations in methods of rating room noise. Lewis S. Goodfriend and Martin Alexander (Lewis S. Goodfriend & Assoc., Whippany, NJ 07981)

Since 1955, the NC curves have been used with considerable success to specify and rate room sound quality. Although newer rating schemes have appeared, and in several areas supplanted the NC curves, many practitioners still use and prefer the old scheme. The newer RC and NC-B rating techniques exhibit one aspect which makes them both problematic in practical use. They both attempt to include in one rating number a further "quality" factor of the sound. While it is simple to establish the goal that room noise should

meet a specific RC (or NC or NC-B) value, with a specific curve shape (neither “rumbly” or “hissy”), it may be more difficult to evaluate whether a space meets the criteria. If the goal is to provide a background sound level of RC-40, or NC-B 40, does a room which is RC-35R meet the criteria, or fail it? Is an NC-B 30R better or worse than an NC-B 40? These questions are faced by the practitioner on a regular basis, and is one of the factors which makes the NC scheme more acceptable. When the room noise is “rumbly” or “hissy,” or when it is modulated or cyclic, or contains pure tones, a statement can be made to indicate this. The practitioner must first be able to specify the measurable physical quantities without ambiguity. Where there are sound quality factors present, they may indicate that an otherwise acceptable spectrum is unacceptable. Detailed background material is included along with examples, including those from the literature.

[See NOISE-CON Proceedings for full paper.]

1:50

1pNSa2. The ASHRAE Handbook on noise and vibration: A critical review. Richard J. Peppin (Larson-Davis, Inc., 5012 Macon Rd., Rockville, MD 20852)

Noise of heating, ventilating, and air conditioning (HVAC) systems could have been one of the most common single sources of noise in the developed countries, but it is not, partially because of well-known methods of HVAC noise control promulgated by professional organizations such as ASHRAE, and trade associations such as the Air Conditioning and Refrigeration Institute (ARI) and the Air Movement and Control Association (AMCA). Most people who live in air conditioned spaces are unaware of the work that went into making their environment quiet by many private consultants and by many corporate engineers and acousticians working for the major HVAC companies. It is true there are problem jobs, usually low-noise environments that do not meet expectations, but these jobs are constantly being worked on and the numbers are being reduced by people using the ASHRAE Handbook. But, unfortunately, the Noise and Vibration chapter is far from perfect. This paper reviews a standby in HVAC acoustics, the 1995 ASHRAE Handbook: HVAC Applications, I-P Edition, “Chapter 43 Sound and Vibration Control,” published by the American Society of Heating, Refrigerating, and Air Conditioning Engineers, Inc. The review provides an evaluation based on the perspective of a member of the Technical Committee 2.6 on Sound and Vibration, from which the chapter originated. The paper points out just a few of the several areas in which the data and information contained need to be used with caution.

[See NOISE-CON Proceedings for full paper.]

2:10

1pNSa3. Room noise criteria—A consultant’s experience. Jerry G. Lilly (JGL Acoust., Inc., 6421 Lake Washington Blvd., #209, Kirkland, WA 98033)

There has been much controversy in recent years regarding the various room noise criteria that have been proposed and are in use today. Included in these criteria are the NC curves, the PNC curves, the RC curves, and most recently, the NCB curves. The NR curves, which are commonly used in Europe, must also be mentioned. The recent adoption of ANSI S12.2 has not resolved the controversy—in fact, it has actually heightened the debate because it acknowledges two room noise criteria (RC and NCB) that are clearly at odds with each other in the low-frequency region. This paper examines the limitations of each method (RC and NCB) and provides possible suggestions for modifications that might create a path for eventual resolution of this controversy. This paper also emphasizes the important characteristics of any room criteria method: it must be as simple as possible without sacrificing effectiveness, relatively easy to measure, and credible from the standpoint of human perception. In any event, it is the opinion of the author that any resolution will have to await the completion of the ASHRAE research project which is currently in progress in Australia.

[See NOISE-CON Proceedings for full paper.]

2:30

1pNSa4. Further observations on NCB and RC rumble and hiss criteria for predicting response to sound in rooms. Gregory C. Tocci (Cavanaugh Tocci Assoc., Inc., 327F Boston Post Rd., Sudbury, MA 01776)

At NOISE-CON96 in Seattle, this author presented a paper comparing assessments of 238 actual measured room octave band sound spectra using NC (tangency), NCB, and RC methods. Also determined for these spectra were the presence of excessive rumble and hiss characteristics using the NCB and RC definitions. This paper goes one step further by presenting observations on how well building owner/user subjective judgments correlate with determinations of rumble and hiss using the NCB and RC definitions.

[See NOISE-CON Proceedings for full paper.]

2:50

1pNSa5. Room noise criteria—Time for common sense. Marlund E. Hale and Richard E. Burke (Parsons Eng. Sci., Inc., 100 W. Walnut St., Pasadena, CA 91124)

The acoustical environments of the many types of rooms designed for occupancy are supposed to be established in the design phase of a new project or during a project retro-fit so that the intended use of a room is not compromised by intrusive noise sources. A major tool in accomplishing the desired environment is the application of room noise criteria. However, given the many types of noise sources with a wide range in sound quality, frequency content, and temporal features, it seems impossible to achieve a desired, yet cost effective, acoustical room environment by using venerable noise criteria (NC) curves alone. ANSI S12.2-1995 is an attempt to include the means to account for different characteristics in noise sources by introducing the balanced noise criteria (NCB) and room criteria (RC) methods for evaluating noise. As a result of the introduction of ANSI S12.2-1995, a controversy has arisen over the practicality and even the viability of the new standard(s). The people who must live or work in the designed space will ultimately determine whether an adequate acoustical room environment has been achieved. The economics of the acoustical design often dictates whether

a working solution is ever implemented at all. A practical basic criteria with just a handful of adjustment factors and guidelines to account for quality, frequency, and temporal anomalies, would be better suited to standardizing the design methodology in room acoustics. This paper provides examples of applications of standards in the everyday world of environmental acoustics.

[See NOISE-CON Proceedings for full paper.]

3:10

1pNSa6. Rating noise—The experience of the computer industry. Marc Warshawsky, Robert D. Hellweg, Jr., and Hsien-sheng (Jason) Pei (Digital Equipment Corp., 129 Parker St., Maynard, MA 01754)

With the general consensus of practicing noise control consultants that the two different sets of criterion curves for evaluating room noise contained in ANSI S12.2-1995 [room criterion (RC) and balanced noise criterion (NCB)] should be replaced with a single set of criterion curves, the authors believe that lessons can be learned from the computer industry, especially in accounting for noises with abnormal characteristics. This industry realizes that the noise emissions from their products should be compatible with the acoustical requirements of the environment in which the products are installed. The types of computer products cover a wide range with respect to functionality, size, sound power level, and spectral and temporal characteristics of their emissions. The rooms in which these products are installed also cover a wide range, including many types with significantly different levels of acceptable noise. In addition to using A-weighted descriptors (sound power level to describe product emissions and sound-pressure level to describe room noise levels), the computer industry has recognized the need to take into account and to characterize abnormal features of the noises, such as prominent discrete tones (ECMA-74-1996 or ANSI S1.13), tonal modulation or beating, and impulsive or fluctuating sounds. This paper presents some of these experiences from the computer industry in the context of establishing room noise criterion curves.

[See NOISE-CON Proceedings for full paper.]

3:30

1pNSa7. Specifying design noise levels for occupied spaces and units of equipment. Neil Moiseev (Cerami & Assoc., Inc., 518 Fifth Ave., New York, NY 10036)

Acousticians and mechanical engineers specify design noise levels for occupied spaces and systems components (units of equipment). While it is widely accepted that occupied spaces should have a “balanced” noise spectrum, individual units of mechanical, electrical, plumbing, and other building systems equipment do not. Of the three commonly used systems of rating interior noise—noise criteria, room criteria, and noise criteria–balanced curves—only noise criteria lends itself to specifying a maximum noise level for equipment. How can the need for a “balanced” noise environment for people be reconciled with the fact that radiated noise from a fan-powered variable air volume box is predominately in the 125- and 250-Hz octave bands with little sound energy at 1000 Hz and higher? A discussion of the need for a clear standard that is easily understandable to specifiers, clients, manufacturers, and users is undertaken, with suggestions of how to resolve the differing needs of these interacting communities.

[See NOISE-CON Proceedings for full paper.]

3:50

1pNSa8. Room noise criteria—A new approach at low frequencies. Robert D. Hellweg, Jr. and Hsien-sheng (Jason) Pei (Digital Equipment Corp., 129 Parker St., Maynard, MA 01754)

ANSI S12.2-1995, “Criteria for Evaluating Room Noise,” permits two different sets of criterion curves for evaluating room noise—room criterion (RC) and balanced noise criterion (NCB)—with little guidance as to the selection. These different procedures can have significant impacts on the cost of designs, especially in the low-frequency regions, since the difference in the criteria is more than 10 dB in the 16- and 31-Hz octave bands at lower levels. The authors suggest there are sufficient data to support a modification to ANSI S12.2, resulting in a single evaluation procedure based on the strengths of each of the two existing procedures. The authors suggest establishing room criterion curves for steady noise without abnormal temporal or spectral characteristics, with correction factors for noises with abnormal features. The suggested criteria for steady noise are the NCB curves based on time-average octave band sound pressure levels, and as a starting point, the correction factors for low-frequency tones and fluctuation differences between the NCB and RC curves. Analyses of currently funded research and re-examination of existing cases can determine the exact correction factors and measurement techniques to quantify these abnormal characteristics. A revised ANSI S12.2 could have basic criterion curves for steady, continuous, and broadband room noises, which could be used for design with correction factors to these basic criterion curves for sounds with abnormal characteristics.

[See NOISE-CON Proceedings for full paper.]

4:10

1pNSa9. Room noise criteria—A status report. Hsien-sheng (Jason) Pei (INCE TC/Noise Sources, P.O. Box 5327, Cochituate, MA 01778) and Richard J. Peppin (Larson-Davis, Inc., Rockville, MD 20852)

After more than 20 years of effort, American National Standard “Criteria for Evaluating Room Noise” ANSI S12.2-1995 was released. ANSI S12.2 is not a good standard because it offers two different possibilities to evaluate the acceptability of the sound in an occupied space with the potential for large differences in the cost of conformance and insufficient guidance to a user on how to choose between them. As members of both the ASHRAE TC2.6 and ASA S12/WG18 committees responsible for the preparation of the ANSI S12.2 standard the authors see the need for inter-society partnership to resolve the conflict. Two panel discussion sessions (ASA Boston Chapter meeting on 21 November 1995; and NOISE-CON96 Conference on 30 September 1996); one technical session

(NOISE-CON96 "Room Noise Criteria—The Next 25 Years") and the response from the acoustical and engineering community have served to show that there are enough data to develop a road map for consensus. After the NOISE-CON96 sessions, many attendants have asked for a report. A report entitled "Room Noise Criteria—The Next 25 Years" was generated, which documents the findings of these sessions. This paper is a brief summary of the said report.

[See NOISE-CON Proceedings for full paper.]

4:30–5:00

Panel Discussion

MONDAY AFTERNOON, 16 JUNE 1997

ROOM G, 1:30 TO 4:30 P.M.

Session 1pNSb

Noise and NOISE-CON: Workshop on Power Plant Noise Control II

Frank H. Brittain, Cochair

Bechtel Corporation, 50 Beale Street, San Francisco, California 94105

Robert A. Putnam, Cochair

Westinghouse MC590, 4400 Alafaya Trail, Orlando, Florida 32826

Chair's Introduction—1:30

Invited Papers

1:40

1pNSb1. Issues in design of combustion turbine silencers. Jim R. Cummins, Jr. (Universal Silencer Div., Nelson Industries, Inc., P.O. Box 411, Stoughton, WI 53589)

Many power plants now use combustion turbines as their primary power source. Among the major noise sources on combustion turbines are the intake and exhaust. Each of these sources has unique characteristics but a common feature is an in-line silencer to control the noise. In this paper the issues affecting the design and application of silencers on combustion turbines is considered. The primary issues presented are acoustical and include lack of confidence in sound power levels, uncertainties in the application of predicted silencer performance, and the effects of size, geometry, temperature, and flow. While the primary focus is on the exhaust system, some topics related to the inlet are also presented. Other issues and trade-offs concerning pressure drop performance, mechanical survivability, and thermal performance are also presented. Some suggestions for ways to overcome the issues are also presented.

[See NOISE-CON Proceedings for full paper.]

2:10

1pNSb2. Issues in heat recovery steam generator system noise. George F. Hessler, Jr. (Hessler Assoc., Inc., P.O. Box 77, 6400 Wishbone Terrace, Cabin John, MD 20818)

A heat recovery steam generator (HRSG) is a fundamental component of all combustion turbine-based combined cycle power plants. While its primary purpose is to convert exhaust gas heat to steam, an important secondary function is to reduce noise emissions from the combustion turbine exhaust. This source, at about 155 dB (overall) *re*: 1 pW for a 100-MW turbine, is the highest noise emission source in any combustion turbine plant. Therefore, the residual exhaust noise emissions leaving the HRSG walls and stack exit must be predicted with acceptable accuracy to determine the total plant noise level. The sources involved in this prediction methodology will be discussed. The issues include source power levels, wall and duct transmission loss, and the noise reduction characteristics through the HRSG flow path. Special measurement techniques required to quantify HRSG noise emissions are described. Whereas the HRSG is mainly a passive device that attenuates combustion turbine exhaust noise, two HRSG generated sources, steam venting and supplemental duct firing, will also be discussed.

[See NOISE-CON Proceedings for full paper.]

1pNSb3. Issues in mechanical draft cooling tower noise. Randall W. Jameson (Marley Cooling Tower Co., 7401 W. 129th St., Overland Park, KS 66213)

Sound from cooling towers in power plants can be a significant contributor to the overall noise level at plant boundaries. Because of their size, cooling towers cannot be treated as a single noise source, and special considerations must be made in determining the contribution of the cooling tower to the overall plant noise level. The significant sound sources of a cooling tower will be identified, as well as how the total sound emitted from a cooling tower is affected by operational and design changes in these sources. Data from field sound measurements of cooling towers will be used to illustrate some of these points. Discussion of measures to reduce cooling tower sound levels in a new installation will also be reviewed.

[See NOISE-CON Proceedings for full paper.]

Contributed Paper

3:10

1pNSb4. Prediction of cooling tower noise emissions. W. Brent Ferren and Andrew P. Dicke (Black & Veatch Engineers & Architects, 11401 Lamar, Overland Park, KS 66211, ferrenwb@bv.com)

Cooling towers are commonly significant contributors to the overall noise emitted by power generation facilities. Because power generation facilities are often required to meet strict environmental noise criteria, predicting the potential noise impact of the cooling tower during the design of the facility is important in developing noise mitigation strategies. The prediction technique must accurately simulate the noise radiation pattern

associated with the cooling tower. The main components of noise associated with mechanical draft cooling towers generally include the water falling over the fill media, the operation of the fan units, and the operation of the circulating water pumps. This paper summarizes a study of cooling tower noise modeling techniques frequently used to simulate sound propagation. The noise emitted from a mechanical draft cooling tower is modeled by various directional point source configurations to account for each noise component. Additionally, the noise barrier effect of the cooling tower is included. The noise impact predicted by the cooling tower models are compared to field sound level measurements in order to evaluate the effectiveness of the cooling tower models.

3:30–4:30

Panel Discussion

MONDAY AFTERNOON, 16 JUNE 1997

ROOM N, 1:55 TO 5:00 P.M.

Session 1pPA

Physical Acoustics: Radiation, Propagation, and Scattering

Gregory Kaduchak, Chair

Applied Research Laboratories, University of Texas, P.O. Box 8029, Austin, Texas 78713-8029

Invited Paper

1:55

1pPA1. The impact of Rayleigh's "The Theory of Sound" on the British development of the theory of electromagnetic radiation. Philip L. Marston (Phys. Dept., Washington State Univ., Pullman, WA 99164-2814)

At the time of Maxwell's death in 1879, British physicists had not yet formulated models for the generation of electromagnetic waves. A recent history of British electromagnetic theory of that period [B. J. Hunt, *The Maxwellians* (Cornell U. P., Ithaca, 1991)] notes the influence Rayleigh's text on sound had on the calculation of the power radiated by an oscillating magnetic dipole, which G. F. FitzGerald published in 1883. Examination of a series of FitzGerald's papers on sources of electromagnetic waves for the years 1879, 1880, 1882, and 1883 confirms Hunt's observations. Specifically, FitzGerald credits Rayleigh's text for the inspiration to consider traveling-wave solutions, the key to his eventual success, instead of his earlier standing-wave solution. This anecdote, and others to be noted from that period, illustrate the influence of acoustics in helping to resolve difficulties in cases in which the electromagnetic theory was poorly understood. These may be of interest to today's students and teachers.

2:15

1pPA2. Scattering from elastic axi-symmetric objects in a range-dependent waveguide near ocean surfaces. M. F. Werby (Naval Res. Lab., Stennis Space Center, MS 39529) and N. A. Sidorovskaia (Univ. of New Orleans, LA 70148)

The correct formulation for scattering from objects near an interface is described. The method used to couple the related scattered solution to a range-dependent waveguide is then shown. Calculations for elongated elastic objects such as elastic spheroidal solids and shells are then incorporated into the formulation, and the effect of the imbedding of objects in the waveguide are illustrated, particularly in the vicinity of an ocean surface. [Work sponsored by ONR, the Naval Res. Lab., and the Univ. of New Orleans.]

2:30

1pPA3. Matching borehole modes for slowness estimation using constrained optimization. M. P. Ekstrom (Schlumberger Austin Res., 8311 N. RR620, Austin, TX 78720) and C. J. Randall (SciComp, Inc., Austin, TX 78731)

Propagating modes often appear in borehole acoustical measurements, providing information about the surrounding formation. A dipole source, for example, excites borehole flexural waves, whose low-frequency slowness asymptote is the formation shear slowness. However, because the flexural mode may be highly dispersive, and its spectral amplitude is small in the low-frequency regime, the shear slowness is difficult to estimate. Monopole sources excite Stoneley waves whose slowness is dispersive and a complex function of formation properties, likewise making formation slowness estimation problematical. Theoretical modal dispersion characteristics may be predicted over a prescribed frequency range by finding the poles of a simplified analytical borehole response function [Kurkjian, *Geophysics* **50**, 852–866 (1985)]. This model is used to estimate formation parameters based on nonlinear optimization. A new spectral estimator (frequency-wave number) is first used to estimate the slowness dispersion for a measured array dataset [M. P. Ekstrom and C. J. Randall, *J. Acoust. Soc. Am.* **98**, 2867 (1995)]. The parameters of an analytical model are then iteratively adjusted via a nonlinear optimization algorithm to minimize the error between model and estimated slownesses. Optimizer performance depends on the error function construct, constraints imposed, and variational knowledge of the forward model. Results of using monopole and dipole field data sets will be shown.

2:45

1pPA4. Pulsed beam propagation in lossless dispersive media. I. Background. Leopold B. Felsen and Timor Melamed (Dept. of Aerosp. & Mech. Eng., Boston Univ., 110 Cummings St., Boston, MA 02215)

Pulsed beams (PB) are highly collimated wave objects which can simulate actual acoustic fields radiated by time-domain (TD) compact transducers and, by superposition, can analytically synthesize arbitrary pulsed source distributions. While much progress has been made in tracking PB wave packets through nondispersive environments, the TD phenomenologies characterizing PB interaction with dispersive media are as yet inadequately understood. To initiate this study, we review known TD asymptotic techniques for pulsed noncollimated wave objects with *real* frequency and wave number constituents. The results can be interpreted in terms of space-time ray theory and instantaneous dispersion relations, with interesting graphical representations (in the configuration–spectrum phase space) that highlight the relevant wave physics. Of special interest are inputs which focus the pulsed signal at a specified space-time point inside the medium. The extension to pulsed beams requires consideration of *complex* spectra, as performed in part II.

3:00

1pPA5. Pulsed beam propagation in lossless dispersive media. II. Implementation. Timor Melamed and Leopold B. Felsen (Dept. of Aerosp. & Mech. Eng., Boston Univ., 110 Cummings St., Boston, MA 02215)

The space-time asymptotic analysis in lossless dispersive media for noncollimated pulsed wave objects with *real* spectra, as performed in Part I is here extended to the *complex* frequency and wavenumber spectra that characterize pulsed beam (PB) wave packets. To address the problem systematically, all spectral components in the formal asymptotic wave number frequency synthesis are expanded to second order away from the beam axis, which represents a real space-time ray with real spectra. Off-axis observation points in the paraxial region are reached by complex space-time rays with complex spectra. The complex off-axis spectra lead to corrections of the real on-axis spectra, in addition to introducing an imaginary (evanescent) decay component. Efforts are in progress to assess the wave physics that accompany these results, and to determine the modifications, due to dispersion, of the highly resolved PB in a nondispersive environment. Synthesizing PB profiles for focusing at a given space–time point in the dispersive medium are also considered.

3:15–3:30 Break

3:30

1pPA6. Sonic band structure and surface localized modes in a density-modulated system: An experimental and theoretical study. Christopher Carr and Roger Yu (Dept. of Phys., Central Washington Univ., Ellensburg, WA 98926)

A result of the Bloch theory is that the eigenstates of an electron in a strictly periodic potential are extended and the corresponding energy levels form clustered energy bands separated by gaps. In the study reported here, a *finite* air-filled tube with arbitrary mass density modulation which permits the inclusion of disorder in the study of the acoustic frequency bands is theoretically considered. Since the theoretical model is finite, direct comparisons can be made with the experiment. The baffled tube is excited with a sound source at one end. The source can deliver short pulses as well as harmonic oscillations. Both normal mode analysis (NMA) and pulse analysis (PA) have been used. The spectra obtained from both methods are compared with our theoretical prediction. For the tube with evenly spaced baffles, three passing frequency bands have been found. More interestingly, some localized resonant modes are discovered whose frequencies lie in the gaps. The wave functions of these modes are confined within a few sections of divided tube at both ends.

3:45

1pPA7. Propagation and loss in layered viscoelastic composites. Wallace R. A. George and Michael Schoenberg (Schlumberger-Doll Res., Old Quarry Rd., Ridgefield, CT 06877-4108, george@ridgefield.sdr.slb.com)

The acoustic response of a source in an anisotropic viscoelastic medium has not only a directionally dependent speed of propagation, but also a directionally dependent rate of attenuation. The long wavelength properties of a layered composite comprised of two alternating constituent layers, both of which are isotropic, but with one elastic and the other viscoelastic (modeled as a “standard solid”), are investigated. Such a

structure behaves as a homogeneous transversely isotropic viscoelastic medium. The plane-wave phase slowness and quality factor as a function of direction are shown for three cases. The first two are (a) when the slower medium is lossy and (b) when the faster medium is lossy. In these cases, the curves for loss versus angle are highly directional, but the actual patterns depend dramatically upon the choice of (a) or (b). The third case examines the behavior when the velocities of the constituents are approximately the same, but the loss is such that which constituent is fast and which is slow is dependent on frequency.

4:00

1pPA8. Speckle statistics of scatterer distributions described by marked regularity models. Robert M. Cramblitt and Kevin J. Parker (Dept. of Elec. Eng., Univ. of Rochester, Rochester, NY 14627-0231, cramblitt@ee.rochester.edu)

Speckle images characterized by a non-Rayleigh amplitude probability density function (pdf) are formed when the assumptions underlying fully developed speckle are violated, such as when the number of scatterers in a resolution cell is small, or scatterers are organized with some periodicity, leading to nonuniform scatterer phasing. Various random-walk models have been developed to explain how non-Rayleigh cases such as the K , Rician, and homodyned- K pdf's arise. Marked regularity models provide an alternative to specialized random-walk models, because they describe the physical spatial placement of point scatterers in one dimension in a parametric manner. Spatial distributions ranging from clustered to random to periodic can be obtained by adjusting the model parameters governing the scatterer's density and "regularity." It is shown that the amplitude statistics of the coherently detected speckle pattern of a collection of scatterers described by the model may be either Rayleigh or non-Rayleigh, and that it is possible to generate the special cases of Rician, K , and homodyned- K pdf's through various settings of the regularity parameters. The model's second-order properties are known and can be varied, eliminating the need to incorporate nonstochastic variations into the random-walk models.

4:15

1pPA9. Experimental observation of sound propagation through thermal turbulence near a boundary. J. Wasier, Ph. Blanc-Benon, and D. Juvé (Ecole Centrale de Lyon, LMFA UMR CNRS 5509, BP 163 69131 Ecully Cedex, France, acous@ec-lyon.fr)

The statistical properties of an ultrasonic wave that traverses a medium exhibiting temperature fluctuations are investigated under well-controlled laboratory conditions. A heated grid in air is placed horizontally in a large anechoic room and the mixing of the free convection plumes above the grid generates a homogeneous isotropic random thermal field. The spectrum of refractive index fluctuations is accurately described by a modified von Karman model which takes into account the entire spectrum of turbulence. Experimental data are obtained by varying both the frequency of the

acoustic source and the distance of propagation. Measurements will be presented for a rigid boundary and for a finite impedance boundary. In this paper, attention is concentrated on the mean sound pressure levels and the intensity fluctuations. Experimental values will be compared with results of numerical simulations based on a wide angle parabolic code. [P. Chevret *et al.*, J. Acoust. Soc. Am. **100**, 3587–3599 (1996)]. [Work supported by DGA-DRET under Contract No. 94-083.]

4:30

1pPA10. Modal impulse response and wave vector–time domain methods to evaluate transient radiation from axisymmetric planar radiators. Jie Sun (Box 1846, Dept. of Geologic. Sci., Brown Univ., Providence, RI 02912) and Peter R. Stepanishen (Univ. of Rhode Island, Narragansett, RI 02882)

The modal impulse response method and the wave vector–time domain method are two of the existing methods to evaluate the transient acoustic field from an axisymmetric planar radiator. Representations of the transient acoustic field in the space-time domain and the wave vector–time domain are linked through the Fourier/Hankel transform relationships, and the wave vector–time domain method is used to derive the modal impulse response for a space-time separable normal velocity distribution of an axisymmetric source. The classical uniformly shaded piston is first addressed as a benchmark case, and a Bessel shaded aperture is then addressed as a result of its importance in addressing more general source distributions. Numerical results are presented for both cases to illustrate the procedures in using the modal impulse response and the wave vector–time domain methods. [Work supported by URI/Ocean Technology Center.]

4:45

1pPA11. Space-time dynamic focusing from axisymmetric planar radiators with transient excitations: Localized waves. Jie Sun (Box 1846, Dept. of Geologic. Sci., Brown Univ., Providence, RI 02912) and Peter R. Stepanishen (Univ. of Rhode Island, Narragansett, RI 02882)

A geometric space-time approach to evaluate the on-axis response is developed for axisymmetric planar radiators with pulsed transient excitations. The approach is based on combining the Rayleigh time-dependent surface integral with a simple geometric focusing method. The acoustic field at the pulse center is simply related to the source signal or aperture distribution along a focusing line in the space-time domain. This space-time dynamic focusing approach is used to explain the behavior of localized wave solutions to the free-space scalar wave equation. As examples, two different types of localized wave fields are investigated. The characteristics of the space-time source distributions are connected to the on-axis acoustic field and energy distribution pattern, using the space-time dynamic focusing approach. Numerical results are presented to show both the source signals and the resulting acoustic fields for the localized waves. [Work supported by URI/Ocean Technology Center.]

Session 1pSA**Structural Acoustics and Vibration and Musical Acoustics: Time Domain Modeling of Plate and Shell Vibration**

Jean-Louis Guyader, Cochair
LVA, INSA, Bat 303, Villeurbanne 69621 France

Earl G. Williams, Cochair
Structural Acoustics, Code 7137, Naval Research Laboratory, Washington, DC 20375-5350

Invited Papers**1:15**

1pSA1. Time-domain modeling of damping in plates. Christophe Lambourg and Antoine J. Chaigne (Dept. Signal, ENST, 46 rue Barrault, 75634 Paris Cedex 13, France, lambourg@sig.enst.fr)

A time-domain formulation of flexural vibrations in damped rectangular orthotropic plates has been developed, in order to investigate transient and nonlinear excitation of plates. The model includes three basic mechanisms of damping: thermoelasticity, viscoelasticity, and radiation. As a consequence, the four rigidity factors are modified by perturbation terms, each perturbation term corresponding to one specific damping mechanism. The thermoelastic term is derived from the coupling between the thermoelastic stress-strain relationships and the heat diffusion equation. The second perturbation term is obtained from the standard generalized partial differential equation formulation of viscoelasticity. The third term is obtained through a Padé approximant of the damping factor which governs the radiation of an infinite plate coupled with a light fluid. The flexural equations are solved numerically in the particular case of a sphere impacting the plate with the help of a finite difference scheme of second order in time and fourth order in space. A simulated sound pressure is then obtained by a simple discrete form of the Rayleigh integral. The simulated tones show the relative relevance of the three damping mechanisms, for four different materials: aluminium, glass, carbon fibers, and wood.

1:40

1pSA2. Impact of plates: Some questions on numerical precision and scaling. Alison J. McMillan (64 Hillport Ave., Newcastle under Lyme, Staffordshire ST5 8QR, UK)

The deflection of an impacted plate may be treated as a time-dependent forced vibration at the impact point. The position of the impactor head can be determined from its initial velocity and deceleration on impact. The difference between the impactor position and the plate deflection is the relative approach of the contacting surfaces (Hertz' law). This is a Volterra integral equation of the second kind which must be solved numerically. Two structures of different sizes but of the same shape will have proportional modal properties. For similar geometries, it can be shown that there are three basic variables; the fundamental frequency, the mass ratio of impactor and plate, and a contact parameter. It is of interest to determine the extent to which the distribution of the higher modes (or equivalently, the shape of the plate) governs the impact force and duration. For example, is it valid to approximate a circular plate by a rectangular one, or to use thin plate bending theory rather than thick? The benefit is the reduced precision required in the modal analysis: With a numerical modal analyzer such as finite-element method, reliable impact force levels may be determined with a low precision modal model.

2:05

1pSA3. Time-dependent radiation and scattering from elastic plates. David Feit and Gerard P. Carroll (Carderock Div., Naval Surface Warfare Ctr., 9500 MacArthur Blvd., W. Bethesda, MD 20817-5700)

The far-field time-dependent radiation pattern for an elastic plate excited by a transient point or line load is calculated. Comparisons are made between results calculated by both Timoshenko-Mindlin and classical plate theories. Such results have been obtained by others, but by viewing the results as a function of time and in a polar angle plane, one is better able to visualize and describe how the directional beam patterns are generated. The supersonic waves generated by the transient excitation give rise to a radiated field appearing as a precursor (with respect to the time of arrival of the sound emanating from the drive point) in a direction perpendicular to a Mach cone. The cone angle is determined from the ratio of the sound speed in the fluid to that of the flexural wave traveling with a velocity characteristic of the frequency at which the transient force spectrum has a peak. In those cases where an inhomogeneity, such as a dynamic system is attached to the plate, the time-dependent scattering pattern, in response to an ensonifying pulse, allows one to image and estimate the parameters of the scatterer.

2:30-2:45 Break

2:45

1pSA4. Convolution integral methods of evaluating the transient response of fluid-loaded plates and shells with attached structures.

Peter R. Stepanishen (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett, RI 02882-1197, stepanishen@mistral.oce.uri.edu)

A general time-domain approach is presented to address the transient vibratory response of fluid-loaded plates or shells of arbitrary shape with attached substructures when subjected to specified mechanical and/or acoustical excitations. The approach is based on utilizing an *in vacuo* eigenvector expansion with time-dependent coefficients to describe the velocity field of the basic structure. Fluid loading and the effect of the substructures on the basic structure are described via the use of convolution integrals involving the modal velocity coefficients and impulse responses. A universal set of coupled convolution integral equations for the time-dependent modal velocity coefficients of the basic structure are developed using a time-domain method of analysis. Special cases, which include a finite plate and a general shell of revolution, are addressed. A reduced set of the coupled convolution integral equations for the time-dependent modal velocity coefficients is obtained for the plate, and sets of similar equations are also obtained for the shell. Numerical results are presented and discussed for some simple beam, plate, and shell problems involving transient excitations.

3:00

1pSA5. High-frequency backscattering enhancements for thick truncated cylindrical shells in water at oblique incidence.

Scot F. Morse, Philip L. Marston (Phys. Dept., Washington State Univ., Pullman, WA 99164-2814), and Gregory Kaduchak (Univ. of Texas, Austin, TX 78713-8029)

Various authors have shown that the backscattering by a finite thin cylindrical shell at oblique incidence is enhanced when resonance conditions are met for the propagation of a leaky wave on the shell [see, e.g., M. L. Rumerman, J. Acoust. Soc. Am. **93**, 55–65 (1993)]. To explore the backscattering enhancements relevant to high-frequency detection, impulse response measurements were carried out for slender thick cylindrical shells using a broad bandwidth PVDF sheet source, for tilt angles ranging from broadside to end-on incidence. These measurements reveal large backscattering enhancements beyond the shear wave cutoff angle, which extend to near end-on incidence in the coincidence frequency region. A simple approximation for the phase velocities of several surface waves is used to identify the underlying mechanisms. A broad enhancement feature is shown to result from the axial propagation of the supersonic a_0 leaky Lamb wave. The associated meridional ray radiates a backward directed wavefront having a vanishing Gaussian curvature. Also observable are what appear to be resonance loci associated with the helical propagation of the subsonic a_{0-} . [Work sponsored in part by the Office of Naval Research and by the ARL:UT Independent Research and Development Program.]

3:15

1pSA6. Transient axisymmetric response of a finite cylindrical shell with internal fluid loading. Part 1. Theory and numerical results.

Scott E. Hassan (Naval Undersea Warfare Ctr., Bldg. 1246, Newport, RI 02841, hassan@code83.npt.nuwc.navy.mil) and Peter R. Stepanishen (Univ. of Rhode Island, Kingston, RI 02881)

A general approach to evaluate the transient response of finite length shells with internal fluid loading is addressed. As a special case the axisymmetric response of a cylindrical shell with a base plate at one end and internal fluid loading is investigated. The velocity field of the shell is expressed as an *in-vacuo* eigenvector expansion and the pressure field in the fluid is expressed using a Green's function for the mixed boundary value problem. The unknown time-dependent modal velocity amplitudes are formulated as a set of coupled convolution integral equations where the time-dependent fluid loading on the shell and base plate is included using

an impulse response approach. *In-vacuo* and fluid-loaded results will be presented for the special case of a transient point force excitation applied to the base plate. The influence of fluid loading on the shell and base plate will be discussed in detail.

3:30

1pSA7. Transient axisymmetric response of a finite cylindrical shell with internal fluid loading. Part 2. Comparison of theory and experiment.

Scott E. Hassan (Naval Undersea Warfare Ctr., Bldg. 1246, Newport, RI 02841, hassan@code83.npt.nuwc.navy.mil) and Peter R. Stepanishen (Univ. of Rhode Island, Kingston, RI 02881)

An experimental investigation addressing the transient axisymmetric response of a finite length cylindrical shell with a base plate and internal fluid loading is presented. The configuration consists of a steel cylindrical shell with a circular base plate welded to one end. The shell and base plate contain an internal volume of water with a free surface. Data are presented for the special case of an axisymmetric transient excitation due to a force impulse hammer acting on the base plate. The *in-vacuo* and fluid-loaded response of the shell and base plate measured with an accelerometer will be presented and compared with numerical results. Transient field point pressure measurements taken with a hydrophone will also be presented and compared with numerical results. The agreement between experiment and numerical results is noted to be excellent.

3:45

1pSA8. Implementation of time-domain fuzzy structure representations.

Paul E. Barbone (Dept. of Aersp. & Mech. Eng., Boston Univ., 110 Cummington Ave., Boston, MA 02215)

It is shown that complex dynamical vibratory subsystems can be represented in a numerical simulation by an equivalent boundary condition. This boundary condition is usually written in the form of a time convolution with a known kernel. It is shown that the kernel of this boundary condition can be replaced by a simple approximate kernel when the substructure has a high modal density (a "fuzzy" approximation). Depending on the properties of the kernel, the time convolution may be replaced by an equivalent condition that is local in time, but depends on high order time derivatives. The issues related to implementing the boundary condition in both convolution and differential form are discussed. In particular, the effects of the boundary condition on the stability and accuracy of numerical time integration are considered. [Work supported by ONR.]

4:00

1pSA9. Closed-loop finite-element modeling of active/passive structural vibration damping in the time domain.

Young-Hun Lim, Vasundara V. Varadan, and Vijay K. Varadan (149 Hammond Bldg., Penn State Univ., University Park, PA)

Finite-element modeling is used for the study of structures with piezoelectric actuators and sensors augmented by viscoelastic passive dampers. The structure is excited by a transient force; the response of the structure as determined by the computed voltage response of the sensor is used to construct a feedback loop that excites the actuator. A constant gain P-D (proportional and derivative) controller is used. The controller changes the effective stiffness of the structure and the derivative feedback also introduces the desired damping. The increased stiffness also lowers the vibration amplitude. Active vibration damping is compared and contrasted with passive damping provided by a viscoelastic dampers. A BMG model is used for the viscoelastic phase. A hybrid approach using modal superposition in conjunction with time integration is used for solving the matrix equations in space and time in an efficient manner. A parametric set of results is presented showing the contribution of active and passive damping for the transient response of a clamped plate structure.

1pSA10. Nonstationary vibrations of an elastic nonlinear cylindrical shell in external contact with an elastic isotropic medium. A. Rossikhin Yuriy and Marina V. Shitikova (Dept. of Theoret. Mech., Voronezh State Acad. of Construction and Architecture, Ul. Kirova 3-75, Voronezh 394018, Russia)

Nonlinear nonstationary vibrations of an infinite elastic isotropic cylindrical shell of constant thickness in welded or smooth contact with a continuous elastic isotropic medium are considered. Nonstationary vibra-

tions are excited by snap-action loads, resulting in two types of cylindrical shock waves propagating in the medium. Behind the wavefronts, the solution for the desired functions is constructed along the rays in terms of a power series (ray series), whose coefficients are the discontinuities in various orders of partial derivatives of the functions to be found with respect to time. A variable value is the time needed for a disturbance to propagate along the ray from the point under consideration up to the wavefront. In so doing, the power of the variable value corresponds to the order of partial time derivative of the desired function.

MONDAY AFTERNOON, 16 JUNE 1997

ROOM E, 2:40 TO 5:20 P.M.

Session 1pSP

Signal Processing in Acoustics, Noise, and NOISE-CON: Signal Processing for Condition Based Maintenance II

Gene Parker, Chair

Barron Associates, Inc., 3046A Berkmar Dr., Charlottesville, Virginia 22901

Invited Papers

2:40

1pSP1. 24-bit digital accelerometers. Richard Lally (Oceana Sensor Technologies, 1632 Corporate Landing Pkwy., Virginia Beach, VA 23454)

Rapid advances in high quality sensors and integrated circuits have enabled the affordable realization of very high resolution digital output devices. New analog to digital converters and piezoelectric accelerometers are being combined to deliver unprecedented broadband resolution and noise immunity. Standardization of the interface between digital sensors and microcontrollers provides for cost-effective measurement system development. Applications such as condition based maintenance and structural acoustics are poised to expand rapidly with the advent of this technology.

[See NOISE-CON Proceedings for full paper.]

3:00

1pSP2. Statistical change detection using nonlinear models. Gene Parker (Barron Assoc., Inc., 3046A Berkmar Dr., Charlottesville, VA 22901)

An important objective in monitoring mechanical and other systems is the detection of changes in the underlying dynamics governing system behavior. Statistical change detection (SCD), or change-point detection, can play a crucial role in the early detection of small changes in mechanical systems. The primary challenge is to detect changes manifest in very small deviations in the statistical behavior of the observations. Statistical change detection algorithms act as novelty detectors, obviating the need to collect examples of all possible fault signature patterns for the purpose of training pattern classifiers. With respect to SCD, the use of nonlinear models can provide useful benefits. In particular, detector performance may be enhanced through the use of more appropriate model structures, such as when processing nonlinear and/or non-Gaussian data. Second-order statistics are, in general, adequate when data being fitted are Gaussian. When data are non-Gaussian, or represent the output of a nonlinear process, higher-order statistics (HOS) (e.g., bispectrum, trispectrum, etc.) may be more useful in characterizing a process. An attractive aspect of the SCD approach is its ready extensibility to include higher-order statistical moments, where necessary, to exploit additional information in the data. Since HOS are polynomial functions of the observed data, polynomial neural networks are well suited for the implementation of signal detectors based on HOS. In this paper, the performance benefits arising from the use of nonlinear models in SCD algorithms are considered.

[See NOISE-CON Proceedings for full paper.]

3:20

1pSP3. A hybrid physical model—neural network approach to machine diagnostics. Doug Yoon (Logicon, Inc., Los Angeles, CA)

Neural network based diagnostic algorithms have been an effective approach to machine diagnostics. A feature extraction routine calculates parameters indicating the relative changes in signal phase along different paths to multiple accelerometers as might be due to localized damage along one path. These parameters were calculated for accelerometer data from the Westland seeded fault helicopter gearbox experiment. Perfect fault discrimination was achieved using a conventional feed forward neural network classifier on only a small training set, with no special treatment for the confounding influence of varying torque levels. A hybrid neural network was designed, which implemented a simple springs and masses mechanical model of a more complex structure, where sensor locations

served as the mass points. This approach is able to localize structural damage using only a small training set on data generated by a finite elements model, and exhibits excellent interpolation and extrapolation. Extending the hybrid neural network approach, a method is proposed for solving for the external forces applied to surface-mounted sensors, from vibrating machine components. The derived forces will then serve as inputs to the signal phase change feature extraction routine. It is hoped this method will train well on small data sets and reduce ambiguities due to signal multipath effects.

[See NOISE-CON Proceedings for full paper.]

3:40

1pSP4. Passive low-frequency acoustical method of nondestructive testing of the energetic systems. Leonid M. Gelman (Dept. of Nondestructive Testing, Natl. Tech. Univ. of Ukraine, 37, Peremogy pr., Kiev, 252056, Ukraine)

The new preventive passive low-frequency acoustical method of nondestructive testing of the energetic bladed systems is considered. The proposed method is based on reception, processing, and estimation of parameters of the energetic bladed system's own acoustic noise in a low-frequency range (0.002–50 kHz) with the usage information about rotation frequencies of rotors. The method includes the following procedures: (1) detection of the narrow-band acoustic signals on the background of interference; (2) identification of the detected acoustic signals; and (3) estimation of testing date and decision acceptance of the state of the energetic bladed system stages. The method allows for acceptance of the state at the appropriate stage of the energetic system in maintenance when the stage is in a normal state, or when the stage is in a state before damage. Experimental results are presented. The proposed method represents one of the approaches to efficient energy use and conservation.

4:00–4:20 Break

Contributed Papers

4:20

1pSP5. Nonstationary chaos in acoustic emission from lathes. Satish T. S. Bukkapatnam (Dept. of Industrial and Manufacturing Eng., Penn State Univ., University Park, PA 16802), Akhlesh Lakhtakia, and Soundar R. T. Kumara (Penn State Univ., University Park, PA 16802)

Acoustic emission (AE) signals have tremendous promise for monitoring machining processes, but current understanding of AE generation is very limited. In addition, existing techniques to analyze these signals are inadequate. Presented here is a new approach to analyze AE signals based on thorough characterization of the dynamics underlying the signals. The signals were collected from extensive experimentation on five different lathes, performed under various cutting conditions, with different tool-work combinations and sampling rates. First, a battery of statistical tests was applied on the collected AE signals. Tests revealed that AE in turning is chaotic with fractal dimensions ranging between three and six. But owing to nonstationarity, chaos was not clearly revealed in the lag plots. Next, the transients were quantified, based on a compact representation scheme that was developed—called *suboptimal wavelet packet representation*—that captured all salient features of the AE signals. This representation scheme was used in developing a neural network-based estimator for on-line monitoring of tool wear.

4:40

1pSP6. A method of minimizing interference in Wigner–Ville distribution and its application in acoustics and vibration signals. Youn-Kyu Park and Yang-Hann Kim (Ctr. for Noise and Vib. Control (NOVIC), Dept. of Mech. Eng., KAIST, 373-1 Science Town, Taejeon-shi, Korea)

One of the major advantages of expressing signal of interest in terms of Wigner–Ville distribution is that one can see how energy of signal varies with regard to time and frequency. On the other hand, a major drawback of this method is that it also displays “signal interference,” therefore one has to have *a priori* knowledge about the signal. Otherwise, one cannot distinguish true and false information from the distribution. This paper addresses the idea to reduce such signal interferences. The idea simply comes from a hypothesis that there could be a domain on which the interference could be more realizable than the time-frequency axis, which is the

Wigner–Ville domain. In fact, there can be four different quadrants. The ambiguity function of the Wigner–Ville distribution is examined, which can be regarded as the mirror image of the distribution. What was found is rather useful—the interferences which appear in the Wigner–Ville distribution tend to be located at the center of the ambiguity function domain. Several simple signals were examined. Based on this theoretical analysis, attempts were made to develop a kind of window that can effectively eliminate the interferences, which is, in fact, a “rotating window” in ambiguity function domain. Several numerical simulations reasonably confirm the proposed idea. Various acoustic and vibration signals are now under investigation to demonstrate the ability of the method.

[See NOISE-CON Proceedings for full paper.]

5:00

1pSP7. Exploitation of vibration and noise signals cyclostationarity in condition based maintenance. Karel Vokurka (Dept. of Phys., Tech. Univ. of Liberec, Halkova 6, CZ-461 17 Liberec, Czech Republic, karel.vokurka@vslib.cz)

A traditional approach to the analysis of vibration and noise in condition based maintenance of machinery is to assume stationarity of the measured signals. However, in the case of machinery working cyclically (e.g., gearboxes, engines, turbines) the generated vibration and noise are basically nonstationary (cyclostationary). This feature, in the traditional approach partially suppressed by continuous time averaging, can be exploited to give more detailed information on the condition of the monitored machinery. To achieve this, periodic time averaging is used. In this paper, the second-order statistical characteristics of the cyclostationary signals are considered. These are double correlation functions, double autospectral densities, and instantaneous autospectra. The disadvantage of these statistical characteristics is their greater complexity. However, they yield a more detailed description of measured signals. Another advantage is a greater immunity to interfering signals. Algorithms for computation of the characteristics will be presented. Theoretical conclusions will be demonstrated on a concrete example of vibration data measured on a passenger car gearbox. [Work supported by the Grant Agency of the Czech Republic.]

[See NOISE-CON Proceedings for full paper.]

Session 1pUW

Underwater Acoustics and Signal Processing in Acoustics: Matched-Field Processing

Gerald L. D'Spain, Chair

Marine Physical Laboratory, Scripps Institute of Oceanography, 291 Rosecrans Street, San Diego, California 92106

Contributed Papers

1:45

1pUW1. Partially adaptive matched-field processing. Arthur B. Baggeroer (Depts. of Ocean and Elec. Eng., MIT, Cambridge, MA 02139)

Adaptive array processing using a quiescent, or conventional, beamformer with adaptive noise cancellation have been implemented for signals based upon plane-wave representations. They have an estimator/subtractor structure and can be interpreted as generalized sidelobe cancellors. One of the useful features of this implementation is that the dimensionality of the adaption space can be reduced from the number of array sensors. While this reduction leads to a partially adaptive processor, it can lead to better performance in applications which are degrees of freedom deficient, or where the number of snapshots is less than the number of sensors, such as large arrays with short-term stationary fields or active systems. Sidelobes in the range/depth ambiguity plane and a sparse number of snapshots are both problems in current matched-field processing. The partially adaptive array leads to an implementation which has both the useful properties of the conventional processor and adaptive sidelobe cancellation in a reduced dimension space.

2:00

1pUW2. Matched-field source tracking with SWELLEX-1 data. Stacy L. Tantum, Loren W. Nolte, and Jeffrey L. Krolik (Dept. of Elec. Eng., Duke Univ., Box 90291, Durham, NC 27708-0291)

The optimal approach to tracking a moving source in an uncertain environment is one which incorporates *a priori* knowledge about both the continuity of the propagation environment and the nature of the source movement. The optimum uncertain field tracking algorithm (OUFTA) [S. L. Tantum and L. W. Nolte, "Tracking and localizing a moving source in an uncertain shallow water environment," J. Acoust. Soc. Am. (submitted)] extends the optimum uncertain field processor (OUFP) [A. M. Richardson and L. W. Nolte, J. Acoust. Soc. Am. **89**, 2280–2284 (1991)] to include modeling of the source motion as a Markov process. This differs from the suboptimal approach which performs a series of independent source localizations and then combines the results to estimate the path taken by the source. The performance of the OUFTA and the suboptimal approach is evaluated using the SWELLEX-1 data collected off the San Diego coast. Results using the SWELLEX-1 data are also presented which show how this optimal philosophy can be used to guide the modification of some popular beamformers, such as the minimum variance adaptive beamformer with environmental perturbation constraints (MV-EPC) [J. L. Krolik, J. Acoust. Soc. Am. **92**, 1408–1419 (1992)], to incorporate source motion. [Work supported by ONR.]

2:15

1pUW3. "Through the sensor" environmental estimation. George B. Smith (Naval Res. Lab., Code 7173, Stennis Space Center, MS 39529-5004, gsmith@nrlssc.navy.mil)

Classification of an echo at tactical ranges in a shallow-water environment is made difficult by dynamic shallow-water multipath. Stable deconvolution (or other) techniques can adapt the processing to the environment

if reasonable and dynamic estimates of the environmental impulse response can be obtained. Technology that obtains real-time environmental impulse responses directly from the data record being analyzed acoustically is called "through the sensor environmental estimation." This work develops a new technique to be added to the already existing arsenal of "through the sensor" techniques and is meant to complement the already-existing technology. This technique uses the ambient noise field and the fact that spatial diversity of the noise sources results in uncorrelated modes. As a consequence of this, the depth modal eigenfunctions for the waveguide can be obtained directly from an eigenvector decomposition of the data spatial covariance matrix, and the wave numbers can be obtained from beamforming these eigenfunctions. The estimated environment is tested here with passive matched-field simulations with limited success and with active matched-filter and biosonar simulations. [This work was supported by the Office of Naval Research. Technical management was provided by the Naval Research Laboratory, Stennis Space Center, Mississippi.]

2:30

1pUW4. Source localization at very long ranges in the presence of large-scale spatial variability. Lisa C. Gresham and Loren W. Nolte (Dept. of Elec. Eng., Box 90291, Duke Univ., Durham, NC 27708-0291, lwn@ee.duke.edu)

Large-scale spatial variability in the ocean environment results in uncertain sound-speed variations on the order of 10–15 m/s, a potentially significant factor in many ocean acoustic applications. The effect of this variability on source localization performance at a range of one megameter is investigated. The optimum uncertain field processor (OUFP) [A. M. Richardson and L. W. Nolte, J. Acoust. Soc. Am. **89**, 2280–2284 (1991)], in its narrow-band and wideband forms, is introduced as an alternate and superior processor to the traditional Bartlett processor for this application. The Bartlett processor, which assumes the environmental parameters to be the average values, essentially ignores the large-scale spatial fluctuations. Optimally incorporating *a priori* information about the environmental and source parameters, as does the OUFP, makes this processor robust to uncertainty in these parameters and provides an upper bound on localization performance. Measures such as *a posteriori* distributions of source location, the probability of correct localization, and the standard deviation of range estimation error reveal that the spatial variability has a significant degrading effect on localization performance, the magnitude of which is dependent on the source frequency. Results are discussed for 25- and 75-Hz sources as well as their coherent and random combinations. [Research supported by ONR (Ocean Acoustics).]

2:45

1pUW5. Three-dimensional matched-field tracking in shallow water using SWELLEX-96 Data. Gregory J. Orris, Laurie T. Fialkowski, Michael Nicholas, and John S. Perkins (Naval Res. Lab., Code 7145, Washington, DC 20375)

A three-dimensional version of matched-field tracking (MFT), a generalization of matched-field processing (MFP), is used to locate and track a towed source using a vertical array during SWELLEX-96—an experiment conducted off the San Diego coast in May 1996. Archival and *in situ*

environmental data are used as input into a forward propagating adiabatic normal mode (ANM) model and a reciprocal propagating parabolic equation (PE) model to create the three-dimensional replica fields used to compute ambiguity volumes (the three-dimensional generalization of the ubiquitous two-dimensional ambiguity surface). Interaction with variable bathymetry caused by a downward refracting sound-speed profile and shallow water is not sufficient to provide uniqueness of the data and replica fields, thus creating many ambiguities for any given data within the fiducial volume. These ambiguities can be minimized via MFT, which uses source motion to reduce the number and relative magnitude of false ambiguities by averaging the ambiguity functions corresponding to trial source tracks defined by an initial position and velocity vector. As increasing numbers of ambiguity functions are averaged, steady movement within a fluctuating background can be detected.

3:00

1pUW6. Matched-field processing in a wedgelike ocean. Paul A. Baxley (Propagation Div., Acoust. Branch, NCCOSC RDTE Div., Code D881, 53560 Hull St., San Diego, CA 92152-5001)

The neglect of the three-dimensional nature of acoustic propagation in a wedgelike shallow-water ocean can cause significant localization errors in underwater acoustic array processors. This study examines in detail the effect of such neglect on the performance of matched-field processors in wedgelike environments, along with an analysis of the performance improvement gained by including known three-dimensional effects into the processor. Results were obtained via computer simulations of signals received on a vertical line array (VLA) from a source moving in the cross-slope direction relative to the ASA benchmark penetrable wedge. Typical effects observed when the three-dimensional nature of the propagation was ignored include an increase in the range error as the source moves away from the VLA (giving the impression of an accelerating source), distinct jumps to incorrect depths as modes are stripped, and loss of signals upon entering a horizontal shadow zone (all modes stripped). Incorporation of the known three-dimensional propagation into the processor not only eliminates the localization errors but provides azimuthal discrimination. Recent progress has also been made in the computer simulation and data analysis of a real wedge-shaped environment off San Clemente Island near San Diego. [Work supported by ONR/NRaD.]

3:15–3:30 Break

3:30

1pUW7. Regularized modal decomposition for matched-mode processing. Stan E. Dosso and Nicole E. Collison (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3Y2, Canada)

Long-range acoustic propagation in the ocean is often well modeled by a discrete set of propagating normal modes. When this is the case, the acoustic field measured at an array of sensors can be decomposed into its modal components providing the basis for matched-mode processing (MMP) methods. Modal decomposition represents a discrete linear inverse problem. For vertical arrays which do not adequately sample the water column or for horizontal arrays, the inverse problem is ill-conditioned and modal decomposition can lead to poor results for noisy measurements. Regularization is a technique for stabilizing inverse problems based on trading off minimizing the data residuals (misfit) with minimizing some function of the solution. In the case of modal decomposition for MMP, an appropriate choice for the regularizing function is the deviation of the modal excitations of the solution from those of the replica field at each trial source location. This results in the best possible match between the measured and replica mode excitations given the noise and array characteristics and can lead to substantial improvements in MMP.

3:45

1pUW8. Detection of low-level broadband signals using adaptive matched-field processing. Newell O. Booth, Phil Schey (Ocean and Atmospheric Sci. Div., NCCOSC RDT&E DIV Code D881, San Diego, CA 92152-5000), and William S. Hodgkiss (Scripps Inst. of Oceanogr., San Diego, CA 92152-6400)

The detection performance of low spectral level, broadband signals from a submerged acoustic source in the presence of strong surface interference is described and illustrated with examples obtained in the May 1996 shallow water evaluation cell experiment (SWellEx-96). The experiment was carried out west of Point Loma in 200-m water of complex bathymetry. A vertical line array (VLA) was deployed next to an identical line array tilted at 45 deg. The detection performance of each of the arrays with linear and adaptive matched-field processing is measured and compared. [Work sponsored by Office of Naval Research, Code 321(US).]

4:00

1pUW9. An intercomparison of low-level signal detectability in a shallow water environment using horizontal, vertical, and tilted vertical arrays. W. S. Hodgkiss, J. J. Murray (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0701), N. O. Booth, and P. W. Schey (NCCOSC RDTE DIV, San Diego, CA 92152-5001)

An intercomparison is presented of multi-tone, low-level signal detectability in a shallow water environment using various array configurations and processing approaches. The data were collected west of Point Loma in 200-m-deep water during SWellEx-96 which was carried out in May 1996. The multi-tone transmissions covered the 50-400 Hz band and consisted of 13 sets of tonals each having a pilot tone and 4 lower-level tonals. These were broadcast from a source towed at 60-m depth. Two 120-m aperture, 64-element arrays were deployed from the R/P FLIP—a vertical array and a tilted vertical array (tilted westerly at 45 deg). In addition, a 240-m aperture, 32-element horizontal array (slightly bowed) was deployed on the seafloor approximately 2-km south of FLIP and oriented SW to NE. Broadband adaptive spatial processing has been carried out on the data; matched-field processing in the case of the two vertical arrays and both matched-field processing and plane-wave beamforming in the case of the horizontal array. The low-level signal detectability results for the various array configurations and processing approaches will be intercompared. [Work supported by ONR, Code 321(US).]

4:15

1pUW10. Analytical and experimental comparison of azimuth/range/depth bias errors in MFP source localization using vertical and tilted arrays. Gerald L. D'Spain, William S. Hodgkiss, James J. Murray (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0704), Phil W. Schey, and Newell O. Booth (NCCOSC RDT&E DIV, San Diego, CA 92152-5001)

During a recent shallow-water (200-m) experiment west of San Diego, two 120-m aperture hydrophone arrays were deployed from the R/P FLIP. One array was oriented in the vertical direction, while the second was deployed at a 45-deg tilt. A set of 13 tones, from 49 to 388 Hz, were transmitted from a 60-m-deep towed source traveling cross slope over a 6-km track. At CPA, the source was at endfire to the horizontal aperture of the tilted array. The data from both arrays were used to perform matched-field processing (MFP) for source localization in range and depth using replica vectors generated from a range-independent environmental model.

In addition, the data from the tilted array were used to simultaneously localize the source in azimuth. The range and depth MFP biases of the two arrays agree quite well with one another and with analytical predictions from previous work. Here, those results are extended to successfully predict the offset bias in the MFP azimuth results, which reaches a maximum of 20 deg at endfire. [Work supported by ONR, Code 321(US).]

4:30

1pUW11. Matched-field processing of multi-tone acoustic signals in shallow water with a horizontal line array. Phil Schey, Newell O. Booth, Gary Dorrance, Al Aburto (Naval Command, Control and Ocean Surveillance Ctr, RDT&E DIV, San Diego, CA 92152-5000), and William S. Hodgkiss (Scripps Inst. of Oceanogr., San Diego, CA 92152-6400)

Matched-field processing results are presented which compare the resolution (in range, depth, and azimuth) and peak-to-sidelobe ratios obtained using a bottom-mounted horizontal line array and a tilted vertical line array. The arrays were deployed in the May 1996 shallow water evaluation cell experiment (SWellEx-96) carried out west of Point Loma in 200-m water of complex bathymetry. Results are presented for a towed source transmitting signals between 49 and 388 Hz at ranges between 0.7 and 6 km and at bearings between broadside and endfire. [Work sponsored by Office of Naval Research, Code 321(US).]

4:45

1pUW12. Matched-field source localization in a random ocean. George R. Minassian (Underwater Acoust. Lab., N. N. Andreyev Acoust. Inst., Shvernik St. 4, 117 036 Moscow, Russia)

This paper is devoted to the statistical approach to matched-field processing (MFP). In Sec. I, a statistical model of the Green's function of a random oceanic waveguide is described using the mean field and the second moment. In Sec. 2, using Bayesian approach, optimum target detection and localization algorithms matched to a deterministic and a stochastic ocean are derived. First, the optimum decisive statistic (log-likelihood ratio) is derived for a deterministic ocean. In the general case, ocean parameter fluctuations make the Green's function random and so we need to average the log-likelihood ratio. The MFP algorithms for a random ocean waveguide developed on the basis of the Bayesian approach are less precise than the algorithms for a deterministic oceanic waveguide model, but they are more robust. In Sec. 3 a comparative analysis of the proposed MFP methods and the known suboptimum MFP techniques is carried out. The ambiguity surfaces for the proposed algorithms, Bartlett, Minimum Variance (MV), MV with neighboring location constraints, and MV with sound-speed perturbation constraints methods are presented. The proposed MFP method outperformed other algorithms under random ocean conditions, giving the least bias error and the lowest sidelobes.

MONDAY EVENING, 16 JUNE 1997

ROOM S, 7:00 TO 9:00 P.M.

Note: Separate registration is required to attend this Tutorial Lecture. See page A28 for details.

Session 1eID

Interdisciplinary: Tutorial Lecture on Medical Ultrasonic Imaging

Kenneth J. Plotkin, Chair

Wyle Laboratories, 2001 Jefferson Davis Highway, Suite 701, Arlington, Virginia 22202

Chair's Introduction—7:00

7:05

1eID1. Medical ultrasonic imaging. James G. Miller (Dept. of Phys., Washington Univ., St. Louis, MO 63130)

Diagnostic ultrasound has become a standard tool in the delivery of healthcare. Although some work was carried out at least as early as the 1940s, medical imaging with ultrasound grew from a modest novelty in the early 1970s to become the widely used and still rapidly growing diagnostic modality of today. In this tutorial lecture, the use of ultrasonic imaging will be illustrated with the aid of video tapes and large-screen projection of representative diagnostic studies. Some of the physics and engineering underlying the generation of clinical images will be reviewed at an intuitive, nonmathematical level. Some of the advantages and limitations of ultrasound in comparison with alternative diagnostic modalities will also be highlighted. Specific applications used to diagnose heart disease will provide opportunities to illustrate ultrasonic imaging of the beating heart. Real-time two-dimensional ultrasonic images of the heart with superimposed color encoded visualizations of Doppler-based estimates of blood flow in the cardiac chambers, which have become the standard for the diagnosis and management of most forms of heart disease, will be shown.

EQUIPMENT EXHIBITION**See page 3030 for list of exhibitors**

TUESDAY MORNING, 17 JUNE 1997

ROOM I, 8:00 TO 11:45 A.M.

Session 2aAA**Architectural Acoustics, Noise, and NOISE-CON: Speech Intelligibility in Rooms**

John S. Bradley, Cochair

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

L. Gerald Marshall, Cochair

*KMK Associates, 59 South Greeley Avenue, Chappaqua, New York 10514***Chair's Introduction—8:00*****Invited Papers*****8:05****2aAA1. Comparisons of measures of speech intelligibility in rooms.** J. S. Bradley (IRC, Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

The speech transmission index, useful-to-detrimental ratios, and percent articulation losses are three quite different types of measures of speech intelligibility in rooms. They each combine a measure of the speech-to-noise ratio and a room acoustics measure to better relate to speech intelligibility. Variations of all three types of measures were calculated from 91 impulse responses obtained from a wide range of acoustical conditions, and for different speech-to-noise ratios. Several broadband forms of these measures are shown to be reasonably well related to each other. The results are used to derive conversion equations between the three types of measures.

8:25**2aAA2. Diagnostic assessment of speech intelligibility in auditoria with the STI.** Tammo Houtgast (TNO-TM, Kampweg 5, P.O. Box 23, 3769 ZG, Soesterberg, The Netherlands) and Herman J. M. Steeneken (TNO-TM, Soesterberg, The Netherlands)

The quality of speech transmission in an auditorium equipped with a public address system is determined by the characteristics of both the electro-acoustic system and the acoustics of the auditorium. Potentially, various types of distortions and interferences in the sound transmission from talker to listener may have a negative effect on speech intelligibility. Assessment of the intelligibility by means of subjective tests with speakers and listeners offers a representative quality measure of the total chain, without however providing diagnostic information on the underlying type of distortion. Objective methods, like the speech transmission index (STI) [H. J. M. Steeneken and T. Houtgast, *J. Acoust. Soc. Am.* **67**, 318–326 (1980)] may not only predict overall speech intelligibility, but may also provide valuable diagnostic information. However, in order to properly represent the effect of all potential types of distortion in a single index, and to obtain relevant diagnostic information, the test signal and the measuring procedure should be chosen very careful. In this context the STI and other methods, like the use of impulse response measurements (MLSSA), will be discussed. Examples will be presented of results obtained in various environments (e.g., railway stations, highway tunnels, and aircraft cabins).

[See NOISE-CON Proceedings for full paper.]

8:45

2aAA3. Investigation on relationships between STI and speech intelligibility in listening tests on Japanese language for sound system design of auditoria. Motoo Komoda (Nagata Acoust., Inc., Minami-Shinjuku-Hoshino Bldg., 8F, 5-23-13, Sendagaya, Shibuya-ku, Tokyo, 151 Japan, komoda@nagata.co.jp)

In order to clarify the applicability of STI as a parameter in sound system design, a series of investigations on speech intelligibility were carried out. This report summarizes the results of the investigation on the following items: (1) STI round-robin tests with different measuring systems; (2) relationships between measured STI and the intelligibility in listening tests on Japanese language; (3) STI measurements in various types of auditoria in Japan.

9:05

2aAA4. Review of speech intelligibility indicators: Their relationship and applications. P. W. Barnett (AMS Acoust., 43 Chase Side, Southgate, London N14 5BP, England)

This paper will review selected speech intelligibility indicators including word scores, $\%AL_{cons}$, and speech transmission index. The implied relationship between indicators will also be deduced based on work and experiments by others. These relationships will be presented on a new scale known as the "Common Intelligibility Scale," which is currently under consideration for incorporation into measurement standards in the United Kingdom. Finally, the use and applicability of various indicators will be discussed, including sensitivity, ease of prediction, and measurement.

9:25

2aAA5. Speech intelligibility in rooms. Johan van der Werff (Peutz & Associates B.V., Postbus 66, 6585 ZH Mook, The Netherlands, mook@peutz.nl)

Speech is a method of communication between two or more minds. The actual forming and understanding of the message are in the physiological domain. The transmission of the speech by sound waves is in the physical domain. Only the latter can be controlled by means of acoustical engineering. The relevant processes in the physiological domain can be contained in the term "proficiency of speaker and listener." The relevant processes in the physical domain can be contained in the term "information loss." This paper will deal with: (a) the underlying factors of information loss and assessing them by measurement and calculation as accurately as possible; (b) the means and limitations of quantifying the information loss: AL_{cons} , STI, etc.; (c) for the calculation of the information loss, an adaption of the "classic" Peutz equation, suitable for data generated by contemporary computer programs, will be given; (d) the accuracy of the data, with some examples like comparison of results from word lists, STI measurement and AL_{cons} calculation; (e) results in extreme situations like a large, closed sports stadium with a 9-reverberation time and a tunnel with a length of 1 k; (f) design considerations for sound systems regarding the proficiency of speaker and listener.

[See NOISE-CON Proceedings for full paper.]

9:45–9:55 Break

9:55

2aAA6. Some comparison of C_{50} with other speech-intelligibility predictors. L. Gerald Marshall (KMK Assoc., 59 S. Greeley Ave., Chappaqua, NY 10514, mark2acous@aol.com)

STI (including the abbreviated version, RASTI), $\%Alcons$ and the 50-ms early-to-late sound energy ratio (C_{50}) are well known as predictors of speech intelligibility. Among these, C_{50} is the least used both as a design equation and for room-acoustic and sound-system measurements, and its performance as a predictor of intelligibility is probably the least well validated. To better understand the performance of C_{50} , data from miscellaneous sources are examined with respect to: (1) correlations between the predictors; and (2) their ability to predict actual intelligibility from measurements and from calculations using statistical-behavior formulas. Relationships are shown between calculated values, between calculated and measured values, and between measured values for the three predictors.

10:15

2aAA7. Predicted and actual intelligibility-rating values produced by a unique church speech reinforcement system. Dan Clayton (720 Fort Washington Ave., Apt. 1R, New York, NY 10040) and L. Gerald Marshall (KMK Assoc.)

A new speech-reinforcement system was included in the 1996–97 reconstruction of historic St. James's Church in Richmond, VA, following a major fire in 1994. Intelligibility of the spoken word was of primary importance to the congregation, yet they also wished to improve the already lively sanctuary acoustics for enhancement of the choral and organ music so integral to traditional Anglican worship. Considerations of superior electro-acoustic performance and sensitivity to appearance in this architecturally significant and

reverberant room overcame the usual client preoccupation with cost, and led to the selection of a system incorporating digitally controlled, steerable-pattern, vertical-line-source loudspeakers. Evaluations of system performance during design were based on both "drawing-board" calculations and computer modeling with a sound system design program, and involved three speech-intelligibility predictors: C_{50} , STI (RASTI), and %Alcons. Comparisons of the predicted intelligibility ratings are made with post-completion measured ratings. Results of effective-directivity (Q_{ds}) and frequency-response measurements are included.

10:35

2aAA8. Factors influencing perceived listening quality in university classrooms. Murray R. Hodgson and Susan M. Kennedy (Occupational Hygiene Program, Univ. of British Columbia, 3rd Fl., 2206 East Mall, Vancouver, BC V6T 1Z3, Canada, hodgson@mech.ubc.ca)

The speech transmission index (STI) can be used to describe acoustical quality in classrooms. The objective of the work reported here was to establish the relationship between STI and user subjective perception of the listening environment by studying thirty university classrooms. STI measurements were made in these rooms when unoccupied. A questionnaire was developed to create a summary measure of perceived listening quality (PLQ) and to enquire about user-related factors that could affect PLQ. The questionnaire was administered to over 5700 students in 107 classes in the 30 rooms. Preliminary results suggest that the main factors contributing to increased PLQ are the perceived instructor speech level, the perceived quality of the classroom air and lighting, the proportion of time that the instructor speaks toward the audience, and the "unoccupied" STI. Use of a speech-reinforcement system reduced PLQ.

[See NOISE-CON Proceedings for full paper.]

10:55

2aAA9. Limitations of speech intelligibility assessment methods. Peter Mapp (Peter Mapp Assoc., Acoust. Consultants, Colchester CO3 4JZ, UK)

Although there are a number of electronic/acoustically based objective intelligibility rating methods (as opposed to subject based word score tests), each suffers from a number of limitations. The essentially nonlinear behavior of most sound systems can and does affect the accuracy of electronically based assessment methods, yet such effects receive little or no mention in the literature. The paper sets out to redress the balance and highlights, and discusses many of the practical limitations of carrying out such measurements. In particular, %Alcons, RASTI, and STI methods are reviewed. The paper describes how system nonlinearities, distortion, compression, digital signal processing, and bandwidth restrictions, as well as late acoustic reflections, can all affect and possibly corrupt the final result. The choice of test stimulus and method of analysis are also shown to affect the apparent result.

[See NOISE-CON Proceedings for full paper.]

Contributed Papers

11:15

2aAA10. Concepts and applications of directivity controlled loudspeaker arrays. Gerald W. J. van Beuningen and Geert H. J. de Vries (Koxkampseweg 10, 5301 KK Zaltbommel, The Netherlands, gerrit.duran@tip.nl)

One way of improving speech intelligibility in acoustically difficult environments is the application of highly directional loudspeaker arrays in order to increase the ratio of direct to reverberant sound. Separate processing of the signals applied to the individual transducers results in a nearly frequency-independent directional behavior of the array. The directional characteristics (expressed in terms of beamwidth and beam-angle) may be changed on-the-fly by digitally implementing the required transfer functions. Basic directional aspects of one-dimensional controlled arrays will be outlined and general constraints regarding the required hardware will be derived. Simulations will indicate the advantages which might be achieved by the application of digitally-controlled loudspeaker arrays over more traditional methods of speech reinforcement. A hardware setup will be described and the results of some real-life implementations will be discussed and compared to the simulations.

11:30

2aAA11. Relationships between speech intelligibility and objective acoustical parameters for architectural features in Catholic churches. António P. Carvalho, António E. Morgado (Acoust. Lab., Dept. of Civil Eng., College of Eng., Univ. of Porto, R. Bragas, P-4099 Porto Codex, Portugal, carvalho@fe.up.pt), and Luís Henrique (Polytech. Inst. of Porto, P-4000 Porto, Portugal)

This study reports on subjective and objective acoustical field measurements made in a major survey of 36 Roman Catholic churches in Portugal built in the last 14 centuries. Monaural acoustical measurements (RT, EDT, C80, D50, TS, L, and RASTI) were taken at several source/receiver locations in each church. A group of college students was asked to judge the intelligibility of speech by evaluating live speech at similar locations in each room. This paper complements those presented at the 1996 Indianapolis and Honolulu ASA Meetings and concentrates exclusively on the relationships of the speech intelligibility averaged values with the objective room acoustics measures and with some architectural features of the churches. The averaged results by church are graphed and analyzed by comparisons. Correlation analyses and statistical modeling identified some relationships among the measures. For instance, squared cor-

relation coefficients (R^2) of 0.67 were found for the relationships: SPEECH-RT AND SPEECH-TS. Between SPEECH and RASTI only a maximum R^2 of 0.50 was found. Regarding the churches' architectural features, the maximum R^2 found was 0.52 between SPEECH and NAVE

HEIGHT. A general linear model including several architectural features increased the R^2 to 0.72. [Work supported by FEUP and ESMAE/IPP-Portugal.]

[See NOISE-CON Proceedings for full paper.]

TUESDAY MORNING, 17 JUNE 1997

ROOM O, 8:55 TO 11:45 A.M.

Session 2aEA

Engineering Acoustics and Noise: Sensor Self Noise

Gerald C. Lauchle, Cochair

Applied Research Laboratory, Graduate Program in Acoustics, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

Thomas B. Gabrielson, Cochair

Applied Research Laboratory, Pennsylvania State University, 204M New ARL Building, P.O. Box 30, State College, Pennsylvania 16804

Chair's Introduction—8:55

Invited Papers

9:00

2aEA1. Self-noise: The relationship between a transducer and its electronics. Thomas B. Gabrielson (Appl. Res. Lab., Penn State Univ., 204M New ARL Bldg., P.O. Box 30, State College, PA 16804, tb3@psu.edu)

The dominant source of self-noise in a sensor system is often assumed to be the first-stage preamplifier. While this can be a poor assumption especially for new-technology sensors, the relationship between the transducer and its electronics is a critical one to consider in the design process. A poor choice for the electronics can negate any potential gains in noise performance from the transducer. This paper explores the inter-relationships between the transducer and its electronics for the important varieties of receiving transducers whether those transducers are simple, self-generating transducers, or complex, feedback-controlled, ac-biased transducers. While this paper will approach the problem from fundamentals, the common figures-of-merit will also be examined and their limitations noted. A number of design guidelines and comparison charts will be presented to facilitate identification of the dominant noise source early in the design process and to permit rough estimation of the achievable noise performance of a particular sensor configuration.

9:25

2aEA2. Sources of background noise in membrane-based microphones. Allan J. Zuckerwar (NASA Langley Res. Ctr., MS 236, Hampton, VA 23681-0001)

The background noise is one of the most important specifications of a microphone because it determines the minimum resolvable sound pressure. Two types of membrane-based microphones will be considered: the condenser and the fiber optic lever (intensity-modulated) microphones. The noise sources in the condenser microphone originate in the preamplifier and the membrane (generated in the acoustic resistance of the air layer). The primary sources in the fiber optic lever microphone are shot noise in the photodetector and electrical Johnson noise in the output amplifier. In both cases the noise sources are modeled, equivalent electrical circuits analyzed, and the results compared to measurements. A brief discussion of $1/f$ noise, both electrical and mechanical, will be included.

9:50

2aEA3. Noise in fiber-optic sensors and sensor systems. A. Dandridge (Naval Res. Lab., Code 5670.1, Washington, DC 20375, adandrid@ccs.nrl.navy.mil)

Over the past decade there has been considerable interest in using fiber-optic sensors for a variety of applications. One area which has attracted substantial attention is the use of these sensors for underwater acoustic arrays. In this paper the fundamental noise performance of these sensors will be discussed. The primary focus will be the noise associated with the transducer itself; however, attention will be paid to the noise associated with the interrogation and multiplexing of these sensors. Transducer noise is usually determined by the thermal noise associated with the structure to which the sensing fiber is attached and the intrinsic thermal noise of the fiber itself. The noise associated with the interrogation of these sensors is usually associated with the characteristics of the optical source. In many systems noise generated by frequency, amplitude, and polarization modulation of the light as it passes through the telemetry cable will also limit the systems noise performance. Consequently, ultimate system noise is a complex trade-off of many parameters and can lead to different solutions for particular applications.

10:30

2aEA4. Self-noise reduction and control in seaborne acoustic systems. James F. McEachern (Technol. and Adv. Concepts Div., Avionics Dept., Naval Air Warfare Ctr. Aircraft Div., Code 4.5T, Bldg. 2185, Unit 6, 22347 Cedar Point Rd., Patuxent River, MD 20670-1161)

Noise sources and noise reduction techniques in cable suspended SONAR systems (i.e., sonobuoys) are discussed. The sonobuoy has evolved from a hydrophone on a simple cable to a sophisticated suspension system as a result of the requirement to provide acoustic performance over extended bands of the acoustic spectrum in the hostile ocean environment. An overview of the progressive identification and solution of self-noise problems in sonobuoys illustrates the layered nature of noise sources. Sonobuoy self-noise specifications evolved from those of traditional audio systems to the specification of hydromechanical performance requirements to limit vibration-induced self noise. The specification of self-noise performance further evolved to produce specifications that relate flow- and vibration-induced self noise directly to the sonobuoy acoustic performance. Examples of noise attenuation, energy decorrelation, and translation of noise energy out of the band of interest are presented. The possible mechanisms of wave- and flow-induced hydrophone noise are discussed, including the concept of vibration-induced head changes as a source of self noise on acoustic pressure hydrophones. A brief review of the progress in the reduction of flow noise on gradient hydrophones is included.

10:55

2aEA5. Flow-induced self-noise on a spherical acoustic velocity sensor. Gerald C. Lauchle (Penn State Univ., Graduate Program in Acoust., P.O. Box 30, State College, PA 16804) and Andrew R. Jones (General Electric Co., Louisville, KY 40225)

An experimental effort to characterize the broadband flow-induced self-noise on a spherically shaped acoustic velocity sensor operating underwater is described. The transducer is a small geophone encased in a near-neutrally buoyant sphere, 76.2 mm in diameter. The research described in this paper quantifies the flow-induced self-noise signal as a function of the sphere diameter Reynolds number ($2.15 < Re < 34\,290$) and it is found that the broadband (10- to 100-Hz) flow-induced signal may be separated into two regimes characterized by: (1) Stokes flow which scales on a viscous time scale, and (2) turbulent flow which scales on an inertial time scale. During Stokes flow and at higher sub-critical Reynolds numbers where the shed vortices are laminar, virtually no broadband noise is produced. The detected, but weak, flow noise appears to scale on the magnitude of the viscous forces in this range. Above the transition Reynolds numbers, turbulent flow structures shed by the bluff body generate the dominant flow noise. This noise scales on the inertial forces of the flow. [Work supported by NAWC/AD/WAR and ONR.]

11:20

2aEA6. Flow-induced self noise in hydrophones resting on the sea floor. M. Strasberg (David Taylor Model Basin, NSWC, 9500 MacArthur Blvd., W. Bethesda, MD 20817-5700)

Measurements of underwater sound at very low frequencies made with a hydrophone resting on the sea bottom may be contaminated by flow noise. Although the flow velocity may be negligible at the very bottom, the turbulent boundary layer of natural currents flowing above the bottom induces nonacoustic local pressure fluctuations on the bottom which may cause self-noise interference. Estimates of these pressure fluctuations are presented based on pressures measured on the ground in an open field in the presence of wind, using dimensional analysis to extrapolate from air to water. The estimated values are compared with underwater ambient noise levels measured with bottomed hydrophones to determine how large the current must be to cause concern with self-noise interference at very low frequencies. Methods for reducing the interference are discussed.

TUESDAY MORNING, 17 JUNE 1997

ROOM H, 8:30 TO 10:10 A.M.

Session 2aEDa

Education in Acoustics and NOISE-CON: Education in Acoustics and Noise Control

Jiri Tichy, Chair

*Graduate Program in Acoustics, Pennsylvania State University, 217 Applied Science Building,
University Park, Pennsylvania 16802*

Invited Paper

8:30

2aEDa1. Education in acoustics and noise control. Tor Kihlman (Dept. of Appl. Acoust., Chalmers Univ. of Technol., S-412 96, Gothenburg, Sweden)

The tasks of designing quieter products, a quieter society, and acoustically good environments demand skilled engineers. It is essential to develop good curricula for the education of the engineers, also considering the difficulties that many students have in developing a deep understanding in the acoustic discipline. However, it is also necessary to develop strategies within the university in order to attract some of the best students. Such strategies may partly be based on judgments and demands from industry and society. Results of some inquiries to industry will be reported. They show that the universities' own judgments and experiences from the past also need to be taken into account, otherwise the tasks will be difficult to solve.

8:50

2aEDa2. A computer-based certificate program in acoustics (noise control engineering). Alan D. Stuart (Penn State Univ., Graduate Program in Acoust., College of Eng., 101 Hammond Bldg., University Park, PA 16802)

The Pennsylvania State University, funded by a grant from the Alfred P. Sloan Foundation, is designing and developing a five-course Certificate Program in Acoustics. Scheduled to be offered in the Fall of 1997, this five-course program will allow geographically dispersed engineers working in, or seeking a career in, noise control engineering to participate in a variety of collaborative and individual learning activities that integrate basic engineering principles with real world examples and applications. Subject material is presented in multiple-media modules that blend CD-ROM-based instruction with text materials and computer conferencing. Each module is enhanced by a series of MatLab-based exercises and/or LabView-based experiments that provide instruction as well as serve as a resource that students can use in solving design problems. This computer-based certificate program makes appropriate use of technology to reach and serve students who are geographically distant from the faculty—and also from fellow students—and bring them together in new and interesting ways to create a unique kind of professional learning community. This paper will present an overview of this program.

[See NOISE-CON Proceedings for full paper.]

9:10

2aEDa3. CD-ROM development for a certificate program in acoustics. Victor W. Sparrow (Penn State Univ., Graduate Program in Acoust., 157 Hammond Bldg., University Park, PA 16802) and Vickie S. Williams (Penn State Univ., University Park, PA 16802)

The Pennsylvania State University is currently developing a computer-based certificate program in acoustics aimed at practicing noise control engineers. This presentation will describe the development of a CD-ROM to deliver instruction in the program. Supplemented by a study guide, reading assignments, and homework, the purpose of the CD-ROM is to take the place of, and improve upon, traditional lectures in a course in acoustics. On each CD-ROM are placed Hypertext Markup Language (HTML) files which are viewable in a provided World Wide Web browser. These HTML files contain text, graphics, sounds, animations, and equations, the latter in the form of Graphics Interchange Format (GIF) files. The students move at their own pace from page to page through the HTML files as they would view overheads in a lecture presentation, but the hypertext, sound, and animation links allow for a more thorough treatment of the material. Having the HTML files already on the CD-ROM eliminates the communications delays inherent if all the files were only on the Internet. [Work supported by the Sloan Foundation. The authors appreciate the

advice and contributions of all of the members of the Cohort-based Asynchronous Learning in Acoustics development team.]

[See NOISE-CON Proceedings for full paper.]

9:30

2aEDa4. A virtual sound level meter for use in acoustics distance education. Dean E. Capone (Penn State Univ., Appl. Res. Lab., P.O. Box 30, State College, PA 16804) and Gerald C. Lauchle (Penn State Univ., State College, PA 16804)

A virtual sound level meter has been designed using the LabVIEW® programming language for use in a CD-ROM-based distance education course, developed and taught by instructors from the Penn State Graduate Program in Acoustics. The purpose of the virtual instrument is to reinforce concepts presented in the lectures by providing students hands-on experience in applying the principles. The meter was designed using ANSI standard S1.4, Specification for Sound Level Meters. All of the functionality prescribed by standard S1.4 is incorporated in the virtual sound level meter. As a signal input the user can specify the frequency and magnitude of a number of deterministic signals, such as a sine wave. Uniform or Gaussian random noise of a user selected magnitude can then be added to the deterministic signal. The student can apply various weightings, third-octave filters, and integration times to the signals. The raw time series, filtered time series, and calculated SPL are displayed on the front panel of the virtual instrument allowing the student to visualize the original signal and the results of the signal conditioning.

9:50

2aEDa5. Development and testing of courseware in acoustic phonetics. Michel T. T. Jackson (Sensimetrics Corp., 26 Landsdowne St., Cambridge, MA 02139, jackson@sens.com)

Sensimetrics Corporation has been engaged in developing and testing interactive, multimedia educational software in acoustic phonetics for more than three years. The course software contains a nonmathematical presentation of major topics in acoustic phonetics, a lab in which students can record and analyze their own speech, and a library in which students can look up technical terms, hear and see spectrograms of IPA examples, and find readings related to the topics covered in the course software. Topics covered in the software include reading spectrograms, basic vowel acoustics, articulatory-acoustic relations in vowels, speaker differences in vowels, acoustic characteristics of constants, aerodynamic effects in consonant production, perception of acoustic cues, identification and discrimination experiments, native and nonnative perception of consonants, and perceptual normalization of vowels. Faculty volunteers at various educational institutions used components of the course and solicited anonymous, written feedback from over 200 graduate and undergraduate students in Communication Disorders/Speech and Hearing Science during the '94-'95 and '95-'96 academic years. The volunteers also provided written comments on the course software. The results indicate that the course software is a useful adjunct to traditional teaching methods. [Work supported by NIH Grant No. 9 R44 MH 51970-02.]

Session 2aEDb**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!

2a TUE. AM

Session 2aMU**Musical Acoustics: Lip Reed and Brass Instruments**

Robert Pyle, Cochair

11 Holworthy Place, Cambridge, Massachusetts 02138

Peter L. Hoekje, Cochair

*Physics Department 0150, University of Northern Iowa, Cedar Falls, Iowa 50614***Chair's Introduction—8:40*****Invited Papers*****8:45**

2aMU1. Impulse responses and effective lengths for brass instruments. R. Dean Ayers (Dept. of Phys. and Astronomy, California State Univ., 1250 Bellflower Blvd., Long Beach, CA 90840, rdayers@csulb.edu)

A brass instrument has three major parts: an input segment, a cylindrical or conical central segment, and a bell. (The latter two are often treated together.) It is conceptually useful to isolate the input segment from the rest of the air column and study the acoustical behavior of that complete segment rather than its components. Features in time- and frequency-domain plots can be correlated very easily with each other and with the physical structure. Combining the input segment's time-domain behavior with that for the rest of the air column yields some novel impulse responses for use with any model of the lip reed [R. D. Ayers, J. Acoust. Soc. Am. **100**, 1190–1198 (1996)]. In the frequency domain, effective lengths for the input segment and the rest of the air column can be combined to give a useful indicator for the degree of harmonic alignment among the input impedance peaks [R. D. Ayers, J. Acoust. Soc. Am. **98**, 81–87 (1995)]. Specific examples in both domains will be discussed.

2aMU2. Application of a novel algorithm for calculating a brass instrument's impulse response to sound simulation. Seiji Adachi (ATR Human Information Processing Res. Labs., 2-2 Hikaridai Seika, Kyoto 619-02 Japan, adachi@hip.atr.co.jp) and R. Dean Ayers (California State Univ., Long Beach, CA 90840)

A novel algorithm for calculating a brass instrument's acoustic pressure response to an impulsive volume flow at the input end [R. D. Ayers, J. Acoust. Soc. Am. **100**, 1190–1198 (1996)] is applied to sound simulation. The algorithm employs several impulse responses that have no multiple reflections and that decay very rapidly. These are convolved with the mouthpiece pressure and volume flow in the past to give the pressure at the present. The impulse responses calculated from the measured dimensions of an actual B♭ trumpet can be considered for simulation purposes as being confined in a period of 13 ms, which is four times shorter than the nonzero period of the reflection function used in the conventional algorithm. Due to the shorter convolution time, the new algorithm, incorporated with a nonlinear driver model of lip vibration and airflow, can run about 4.5 times faster than the method using the reflection function. Simulations with the two algorithms indicate the same behavior for the transition among the air-column resonance modes. Individual waveforms synthesized with the new algorithm are in very good agreement with their counterparts synthesized with the conventional algorithm.

2aMU3. Phase relations in the lip reed flow controller. P. L. Hoekje (Dept. of Phys., Univ. of Northern Iowa, Cedar Falls, IA 50614-0150)

Recent research into the mechanism of the lip reed flow controller provide stroboscopic confirmation [Copley and Strong, J. Acoust. Soc. Am. **99**, 1219–1226 (1996)] and theoretical support [Adachi and Sato, J. Acoust. Soc. Am. **99**, 1200–1209 (1996)] of the two-dimensional model of lip motion. In this model the lips undergo a rotary motion in which the phase difference between the (upper) lip's upward and outward components may shift in a manner that favors self-sustained oscillation near the impedance peaks of the air column. In the present work, the pressure difference p across the lips and the flow u into the instrument are estimated from measurements of pressures p_u and p_d upstream and downstream of the lips and the input impedance Z_d of the instrument. The measured relations between p and u are consistent with the predictions of the two-dimensional model, and in particular illustrate how the brass player is able to vary the blown frequency both above and below the impedance maximum of a given mode. [Work supported by the Carver Foundation and a UNI Graduate College Project Grant.]

2aMU4. How brass instruments are built: Art, craft, perhaps even science. Robert W. Pyle, Jr. (11 Holworthy Pl., Cambridge, MA 02138)

Most brass players believe that the playing qualities of their instruments are affected by the method of construction and the materials used. Composition, thickness, and hardness of the metals used are all cited as important. Phrases such as "hand-hammered bell" are often found in advertising literature. This paper will give an overview of the tools, techniques, and materials used by brass-instrument makers today and in centuries gone by. Traditional techniques are still used along with the latest metalworking methods, so that a present-day instrument may contain some parts produced by computer-controlled machine tools and others made using processes familiar to 16th-century artisans. Players' "lore" will be compared with available experimental data. It is likely that more questions will be raised than answered.

Contributed Papers

2aMU5. Internal sound field of vibrating trombone bell. Andrew Morrison and P. L. Hoekje (Dept. of Phys., Univ. of Northern Iowa, Cedar Falls, IA 50614-0150)

Many of the vibration patterns of a free trombone bell show rotational symmetry, with equal numbers of oppositely moving anti-node regions of equal size. Thus it is predicted that the acoustic sources due to these patterns will cancel, particularly at low frequencies, resulting in little or no net coupling between the body vibrations and the air. This has implications for both the radiated field and the acoustic response of the instrument as seen by the player. An isolated trombone bell was driven at certain of its structural resonance frequencies by attaching a magnet chip and exciting with a solenoid carrying an oscillating current. The resulting acoustic field was probed with a miniature microphone throughout the bell. For each structural mode, the acoustic field pattern was also compared with the corresponding structural vibration pattern imaged with holographic interferometry. The results show acoustic short circuiting between adjacent anti-nodes of the structure, as expected, contributing little bell-to-air coupling. However, there are also significant asymmetries in the acoustic field that do not result in cancellation, and therefore can produce increased signals in the radiated field, as well as at remote regions in the bore. [Research supported by a UNI College of Natural Sciences SOAR grant.]

2aMU6. An efficient numerical method for solving the horn equation. Robert W. Pyle, Jr. (11 Holworthy Pl., Cambridge, MA 02138)

For lossless horns, there are many contours for which exact solutions of the so-called "Webster" horn equation are known. For quantitative work on brass instruments, however, it may be necessary to solve the equation for arbitrary contours, perhaps known only numerically from measurements of existing instruments. The present paper describes a numerical technique that is efficient and accurate. The effects of viscous and thermal damping at the walls of the horn are easily incorporated. Like most numerical methods for solving differential equations, this breaks the length of the horn into a number of smaller pieces. However, the pieces are not quasi-infinitesimal, but are initially as long as a quarter wavelength. The number of pieces is then increased and the error is estimated and reduced in a series of consecutive approximations using a method called "the deferred approach to the limit." Where an exact solution is known (e.g., lossless exponential or catenoidal horn), results accurate to about six or seven decimal places are easily obtainable when the calculations are carried out with a resolution of about 15 significant figures.

Session 2aNS

Noise: Environmental Noise, Annoyance and Related Topics

Lawrence S. Finegold, Chair

Armstrong Laboratory, AL/OEBN, 2610 Seventh Street, Wright Patterson AFB, Ohio 45433-7901

Contributed Papers

9:30

2aNS1. Guidelines for reporting community noise-reaction studies.

James M. Fields (10407 Royal Rd., Silver Spring, MD 20903, jfields@capaccess.org)

Difficulties in comparing studies that use inconsistent reporting and measurement procedures have long been recognized as major impediments to the accumulation of knowledge about community response to noise. Gaps in the information presented in publications have been especially serious problems for combined social and acoustical surveys of residents' responses to environmental noise. In an attempt to alleviate this problem an ICBC team, Community Response to Noise, recommends guidelines for reporting on combined social/acoustical surveys of residents' reactions to noise. The guidelines are the result of a 3-year development process that involved consultation with more than 70 experts from 12 countries. If implemented, the reporting guidelines would contribute to the development of cumulative, scientific findings from combined social and acoustical community noise surveys. The guidelines have been developed for three types of publications for information on 18 topics that come under the general headings of overall survey design, survey sample description, data collection methods, acoustical parameters, dose/response reporting, and explanatory variable reporting.

9:45

2aNS2. Environmental noises from different sources and overall annoyance ratings: An interdisciplinary evaluation concept. Brigitte Schulte-Forkamp (Dept. of Phys., Oldenburg Univ., P.O. Box 2503, D-26111 Oldenburg, Germany)

In different studies of noise annoyance ratings, the overall annoyance of a multiple-noise-source situation is often rated to be less annoying than the most annoying single source. There have been various explanations for this phenomena, one of which is that people are not able to rate multiple-noise-source situations, because overall noise annoyance from more than one source presented a task too complicated to be achieved. Literature studies show the need for different methodical procedures and acoustical parameter sets to identify the particular effects of annoyance. In a socio- and psychoacoustical field and laboratory study on this issue, different combinations of synergetic environmental traffic noise are evaluated. Frames of reference for judging the annoyance of a single noise source are compared to the overall annoyance in a multisource environment. The meaning of social features like status, mode of living, courses of daily and weekly activity, strategies of noise perception and noise management, and degrees of the inhabitants' capacity to deal with traffic and noise interference is analyzed. Further, psycho- and socioacoustic indicator models on the basis of these and future comparative case studies are discussed.

10:00

2aNS3. Acoustic source location using a neural network. Tarun Bhatt (Dept. of Mech. Eng., Tennessee Tech. Univ., Cookeville, TN 38505, tb7101@tntech.edu), Esther T. Ososanya, and Corinne M. Darvennes (Tennessee Tech. Univ., Cookeville, TN 38505)

This paper presents an artificial neural network (ANN) approach to locate an acoustic source and determine its distance from a microphone. First, two sets of experiments were performed to locate a rattling bolt on a plate using the unwrapped phase method, by comparing the acceleration

signal from the vibrating plate to the acoustic signal received by an array of four microphones. The first test used a wide frequency bandwidth random signal and the second test used a short frequency bandwidth signal. Then the neural network was trained using this experimental data. The effectiveness of the ANN at locating a source using four microphones was evaluated, in light of experimental problems associated with frequency resolution and effects of reflecting surfaces. The ANN was also used to locate the same source using only three microphones. The expectation is that fewer microphones will be needed if an ANN is used and still provide the same level of accuracy on source location. This would result in better accuracy, or substantial savings for acoustic source location evaluation.

10:15

2aNS4. Using FUZZY logic to clean unattended blast noise monitor data. Paul D. Schomer and Hans J. van Slooten (U. S. Army Construction Eng. Res. Lab., P.O. Box 9005, Champaign, IL 61826-9005)

Unattended noise monitors are a common tool used to manage the noise generated by large distributed sources such as airports and military installations. In most instances, the desire is to measure just the noise from the source under study or to separate out the noise from the source under study from other noises. In practice, this can be a difficult task. In the case of blast noise from military weapons (e.g., artillery, tanks, bombs), it is necessary to separate this particular noise from other low-frequency, short-duration sounds which can be generated by helicopters, quarries, and road construction, etc., or from wind-induced pseudonoise. The use of FUZZY logic to perform this separation and to designate those measured noise events that come from weapons fire at the military installation is described.

10:30

2aNS5. Airport noise monitoring system accuracy limitations. Alan Dent and Ron Brown (Wyle Labs., Inc., Res. Dept., 128 Maryland St., El Segundo, CA 90245)

In recent years, there have been discussions about airport noise monitoring systems and their ability to more accurately measure aircraft noise. Most of these systems are built to meet the type 1 requirements of ANSI S1.4-1983, Specification for Sound Level Meters. With advancements in microphone components and design, along with improved testing methods, the tolerances should be reviewed. Many airports use the measured noise data in support of noise mitigation programs and enforcement of noise abatement procedures. However, participation in some programs considers measured data more accurate than the microphone's tolerance. This could affect the spending of millions of dollars. Also, the noise abatement requirements at some airports invoke measurement accuracy to within 0.1 dB. These goals signal the need for new consideration. The microphone measurement error can be considered a combination of electronic, mechanical, and acoustic errors. The accuracy and linearity of electronic circuits is better than in the past and there have also been improvements in temperature, humidity, and static pressure effects. Field calibration devices and acoustic limitations must also be considered. Considering all of these factors, the system tolerances may deserve to be reduced.

2a TUE. AM

2aNS6. On the evolution of the capability to experience annoyance: Behavioral-ecological considerations of the effects of acoustical noise. Karl Th. Kalveram (Dept. of Cybernetical Psych., Univ. of Duesseldorf, 40225 Duesseldorf, Germany)

“Traditional noise research” is based upon the psychophysics of auditory perception and neglects the ecological meaning of physical noise measurements. In contrast, “ecological noise research” emphasizes that sounds have a psychological function, regarding annoyance caused by acoustical signals not compatible with these psychological functions. In the present paper, this functional approach is extended, also implying the biological function: Assuming a harmful environmental variable affecting the individual’s (Darwinian) fitness, there is a possibility that a neural

detector will evolve (the input of which is the sensory stimulation correlated with this variable, while the output is motivating the actions capable of diminishing that sensory input) thereby possibly interrupting ongoing behavior. Event-related integration of sensory data is an essential part of this detector, causing the experienced annoyance to be composed of components representing number, average intensity, and average duration of the disturbing events. Because the energy-equivalent sound level L_{eq} is composed in this manner, annoyance, measured by the L_{eq} , can indeed be considered a possible-loss-of-fitness signal (PLOF signal). Therefore, “traditional” and “ecological” noise research are not as contradictory as suggested above. However, the old problem of how to find the conditions influencing the annoyance rating can now be replaced by the question: Which conditions shaped the PLOF-detector in the evolution of men?

TUESDAY MORNING, 17 JUNE 1997

ROOM N, 8:30 TO 11:30 A.M.

Session 2aPA

Physical Acoustics: Drops, Filaments and Bubbles

D. Felipe Gaitan, Chair

National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677

Contributed Papers

8:30

2aPA1. Theory on streaming generated by a bubble in an ultrasound field. Junru Wu (Dept. of Phys., Univ. of Vermont, Burlington, VT 05405, jwu@zoo.uvm.edu) and Gonghuan Du (Inst. of Acoust., Nanjing Univ., PROC)

Theory has been developed to calculate microstreaming velocity inside and outside an isolated air bubble in a liquid produced by a bubble-scattered sound field, taking account of two predominant modes of the bubble’s motion: a monopole (pulsation) and a dipole (translational harmonic vibrations). It has been demonstrated that streaming velocity inside a bubble is much greater than that outside the bubble. It has also been shown that the microstreaming is most pronounced for a bubble undergoing volume resonance, which is in agreement with Elder’s experimental observations [S. A. Elder, J. Acoust. Soc. Am. **31**, 54–64 (1959)]. If the viscosity of the gas medium in a bubble is neglected in theory, the results obtained agree qualitatively with Davidson and Riley’s [B. J. Davidson and N. Riley, J. Sound Vib. **15**, 217–233 (1971)].

8:45

2aPA2. Acousto-electric generation of charged clusters and drop arrays. Robert Apfel, Yuren Tian, and Yibing Zheng (Yale Univ., New Haven, CT 06520)

Novel, compact instrumentation for studying the behavior of drop arrays and sprays, and of clusters of drops, permits fundamental research into the behavior of reacting and nonreacting fluid species. The new capability is made possible by simultaneous levitation and charging of “seed” droplets (10–30 μ m in diameter) which come together in 2-D clusters. These clusters are interesting in their own right because of their crystalline and quasi-crystalline forms, which depend on the acoustic and electric field parameters. By varying the electric and acoustic field intensities, one can cause the clusters to condense into larger drops (e.g., 50–300 μ m) which, because of their charge, form uniformly spaced 2-D arrays of monodispersed drops. One or more layers of these 2-D arrays can form in the acoustic standing wave. Such a configuration permits a wide range of fundamental studies of drop evaporation, combustion, and nucleation. The

drops can be single or multicomponent. Therefore, fundamental materials studies can also be performed. Instrument characteristics and initial studies with the new apparatus will be reviewed. [Work supported by JPL 958722.]

9:00

2aPA3. Drop dynamics in microgravity. Joseph C. Jankovsky, Jeffrey A. Ketterling, Robert E. Apfel (Dept. of Mech. Eng., Yale Univ., New Haven, CT 06520), and R. Glynn Holt (Boston Univ., Boston, MA 02215)

The dynamics of liquid drops in microgravity were observed in the Drop Physics Module flown aboard the second United States Microgravity Laboratory (USML-2) in 1995. Water drops with varying concentrations of the surfactants Triton-X 100 and Bovine Serum Albumin were suspended in air and deformed by an acoustic standing wave. When the deforming force was removed, the drops oscillated freely about a spherical equilibrium shape, as opposed to the oblate equilibrium shape of a drop levitated in 1-g. Drop diameters between 1 and 3 cm were observed. Frequencies and decays for the dynamics have been extracted from the data. Results are compared to theoretical models and experiments in 1-g. Results for frequency shifts due to large amplitude oscillations are also presented. [Work supported by NASA through JPL, contract 958722.]

9:15

2aPA4. Oscillations of a water drop with surfactant Triton X-100 in zero gravity. Xiaohui Chen, Yuren Tian, Joe Jankovsky, Robert Apfel (Dept. of Mech. Eng., Yale Univ., 9 Hillhouse Ave., New Haven, CT 06520-8286, xhchen@minerva.cis.yale.edu), and William Tao Shi (Thomas Jefferson Univ., Philadelphia, PA 19107)

Large amplitude nonlinear oscillations of an axially symmetric water drop of 2.51-cm diameter, initial aspect ratio 3.4, with surfactant Triton X-100 of 0.5 mM, in zero gravity are studied by a boundary element method. Included in the analysis are surface-shear and dilatational viscos-

ity, under the assumption that the shear viscosity of the bulk phase is small. Numerical simulations of the drop oscillations are in good agreement with the experimental results of drop oscillations measured in space during the second United States Microgravity Laboratory, USML-2. The evolution of the drop oscillations for both experiment and simulation is given. The simulation provides predictions for the values of both surface dilatational viscosity and shear viscosity of 0.20 sp and 0.10 sp, respectively. In addition, with the simulated data, the damping constants, frequencies, and decomposed oscillation modes are computed. [Work supported by NASA through JPL, Contract No. 958722.]

9:30

2aPA5. Improved active stabilization of capillary bridges using acoustic radiation pressure and dynamics of a breaking bridge. Mark J. Marr-Lyon, David B. Thiessen, and Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814, o9938156@unicorn.it.wsu.edu)

Liquid bridges between two solid surfaces have applications in low gravity, such as the solidification of floating zones. Long bridges naturally become unstable to a symmetric mode by bulging near one end while the opposite end thins. For a cylindrical bridge of radius R and length L , the slenderness $S=L/2R$ has a natural (Rayleigh) limit of π beyond which the bridge breaks. Using acoustic radiation pressure in simulated low gravity as described previously [J. Acoust. Soc. Am. **99**, 2540 (1996)], stable bridges significantly beyond the Rayleigh limit can be produced. Improvements in the acoustics, optics, and electronics have enabled stabilization of bridges with S as large as 4.3. To be useful in low gravity, this technique will have to be modified to work on liquid bridges in air. Currently, acoustic resonators in air that have the required property that the sound amplitude can be spatially redistributed rapidly are being investigated. The laboratory experiments with simulated low gravity have made it possible to observe the dynamics of a breaking super-long bridge. [Work supported by NASA.]

9:45–10:00 Break

10:00

2aPA6. Computer analysis of the resonant scattering of sound by a bubbly fluid encapsulated by a submerged spherical shell. Murray S. Korman (Dept. of Phys., U. S. Naval Acad., Annapolis, MD 21402), Scott R. Swain (College of William and Mary, Williamsburg, VA 23185), and Lawrence A. Crum (Univ. of Washington, Seattle, WA 98105)

The theoretical results for the resonant backscattering form factor f ($\theta=\pi, k$) versus wave number k for scattering from a submerged elastic shell filled with a bubbly fluid are presented using MATHEMATICA 2.2. The elastic shell is a castable urethane known as Smooth-On 775 which has a density of $\rho=1.03$ g/cm³ and longitudinal and shear speeds of $c_d\sim 1450$ m/s, $c_s\sim 70$ m/s, respectively. The bubbly fluid is taken to be air bubbles of uniform radius (50 μ m) and spacing, suspended in a host liquid of castor oil, instead of water, due to its greater viscosity. Resonant scattering is compared over a range of volume void fractions $0<\beta<0.01$ for the frequency range 0–20 kHz. Shell inner radius b and thickness t are varied. Results are compared with an “ideal” spherical cloud of bubbly water. Comparisons are virtually identical for shells where $b=5.08$ cm and $t=0.635$ cm. The scattering theory by Ayres *et al.* [Int. J. Solids Struct. **23**, 937–946 (1987)] is used, along with the theoretical model by K. W. Commander and A. Prosperetti [J. Acoust. Soc. Am. **85**, 732–746 (1989)] for the complex sound speed and density in a bubbly medium.

10:15

2aPA7. Acoustical resonator frequency shift due to the migration of particles and bubbles. Christopher S. Kwiatkowski and Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814)

A change of the resonant frequency of an acoustic chamber can arise due to the presence and migration of objects inside the chamber. In the investigations of collinear four-wave mixing mediated by a suspension, it was noticed that for sufficiently high particle concentrations, the resonant frequency of the chamber shifted as the particles formed periodically spaced bands, due to the induced radiation pressure from two counter-propagating pump waves. This frequency shift, which also affects the Q factor of the cavity, is another manifestation of the interaction of sound with sound. A transfer-matrix approach was developed which uses a continuum approximation for the particle suspension, and is in good agreement with previous theory and experiment. By using sufficiently gassy water, pump pressures of large enough amplitude led to the measurement of the response of the system to cavitation bubbles. By proper tuning, the frequency shift of higher azimuthal modes of the cylindrical cavity were probed, with experimental results showing both stable and unstable modes. [Work was supported by the Office of Naval Research.]

10:30

2aPA8. On acoustic cavitation of slightly subcritical bubbles. Ali Nadim (Dept. of Aeronaut. and Mech. Eng., Boston Univ., Boston, MA 02215) and Tasso J. Kaper (Boston Univ., Boston, MA 02215)

An analytic expression is obtained for the threshold acoustic pressure amplitude which causes gas bubbles that are slightly smaller than the critical size to undergo cavitation in a liquid. The derivation is based on the Rayleigh–Plesset equation which describes nonlinear bubble oscillations. In the limit where the steady part of the pressure field in the vicinity of the bubble is decreased to a value less than the vapor pressure inside the bubble, it is well known that there exists a critical bubble radius above which bubbles spontaneously cavitate. For bubbles which are below this critical size and would otherwise be stable, imposition of a time-harmonic acoustic pressure of the right frequency causes the bubbles to become unstable and undergo cavitation. The threshold value of this acoustic pressure is found by employing a nonlinear analysis which reduces the Rayleigh–Plesset equation to a damped and forced Mathieu equation for the case of such slightly subcritical bubbles.

10:45

2aPA9. Comparisons of the calculated and measured acoustic pressure amplitude from single-bubble sonoluminescence. Thomas J. Matula, Larry A. Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105), and William C. Moss (Lawrence Livermore Natl. Lab., Livermore, CA 94550)

The acoustic emissions from single cavitation bubbles have been measured with a state-of-the-art needle hydrophone. The measurements correspond to a stable cavitation bubble driven below and above the sonoluminescence threshold. Previous measurements above this threshold indicated that the lower-bound peak pressure at a distance of 1 mm from the bubble was nearly 2 bar [I. M. Hallaj *et al.*, J. Acoust. Soc. Am. **100**, 2717(A) (1996)]. The measured pressure amplitude below the sonoluminescence threshold is much weaker. Comparisons of the measured amplitudes with calculations are considered. In one case the acoustic emission mechanism is considered as being due to internal shock waves diverging into the fluid. In a second case the acceleration of the bubble wall is considered as the mechanism for the acoustic pulse, assuming a spatially uniform gas pressure in the bubble. Both methods are consistent with the measured values. However, neither method agrees well with the measured pulse amplitudes of the afterbounces. The addition of damping mechanisms not currently

considered in the theories should provide better agreement. [Work supported by NSF. WCM's work was performed under the auspices of the U. S. Department of Energy at Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.]

11:00

2aPA10. Calculated pulse widths and spectra of sonoluminescing nitrogen and argon bubbles. William C. Moss, Douglas B. Clarke, and David A. Young (Lawrence Livermore Natl. Lab., L-200, 7000 East Ave., Livermore, CA 94551, wmoss@llnl.gov)

By modeling the hot compressed gas in a sonoluminescing bubble as a thermally conducting plasma it is shown that the measured picosecond pulse widths are due to electron conduction and the rapidly changing opacity of the plasma at the onset of ionization. The model shows that these mechanisms are also responsible for the absence of an "afterglow" subsequent to the sonoluminescent flash, as the hot bubble expands and cools. The calculated spectra for sonoluminescing nitrogen and argon bubbles suggest that a sonoluminescing air bubble probably contains only argon, in agreement with a recent theoretical analysis. [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.]

11:15

2aPA11. Mass flux resonance and the threshold for light emission in single bubble sonoluminescence. R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02115) and D. Felipe Gaitan (Univ. of Mississippi, University, MS 38677)

The acoustic pressure, equilibrium, and time-varying radius of single air bubbles acoustically levitated in a 20-kHz acoustic standing wave have been measured in the region of parameter space where light emission from bubbles is observed. A resonancelike response of the bubble oscillation is observed along the path in parameter space to which the bubble equilibrium radius is constrained by an unknown mass flux mechanism. A partial explanation of the anomalous stability observed previously [R. G. Holt and D. F. Gaitan, Phys. Rev. Lett. **77**, 3791 (1996)] is presented. The bubble response is correlated with the threshold for light emission and light intensity as a function of increasing acoustic pressure. Finally, the measured bubble response is shown to agree with the Rayleigh-Plesset equation, even though the theory cannot predict the equilibrium path in parameter space. [Work supported by NASA.]

TUESDAY MORNING, 17 JUNE 1997

ROOM S, 8:30 TO 11:30 A.M.

Session 2aPP

Psychological and Physiological Acoustics: Physiological and Psychological Aspects of Amplitude Modulation Processing I

Neal F. Viemeister, Chair

Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, Minnesota 55455

Chair's Introduction—8:30

Invited Papers

8:40

2aPP1. Detection and discrimination of amplitude modulation. Stanley Sheft and William A. Yost (Parmly Hearing Inst., Loyola Univ., 6525 N. Sheridan Rd., Chicago, IL 60626, ssheft@wpo.it.luc.edu)

Amplitude-modulated stimuli have been used in psychophysical experiments to address a variety of issues including intensity discrimination, frequency and temporal resolution, and roughness perception. Thresholds for detecting amplitude modulation (AM) and discriminating a change in AM rate or depth are the basis for much of this work. Early work emphasized the role of the peripheral excitation pattern in AM detection. More recent studies have often concentrated on the temporal aspects of modulation coding. With sinusoidal AM, the change in detection threshold as a function of modulation rate is commonly interpreted as a temporal modulation transfer function (TMTF) and has been used to assess temporal resolution. Stimulus parameters that affect the TMTF include carrier spectrum, level, and gating characteristics. The effects of these parameters and implications for AM processing will be reviewed. This current work measures the ability to discriminate a change in either modulator phase or envelope regularity as a function of AM rate. Irregular AM was generated with frequency modulation of a sinusoidal envelope modulator. Results from these experiments will be discussed in terms of auditory processing of modulation as characterized by the TMTF. [Work supported by NIH.]

9:05

2aPP2. Encoding of amplitude modulation in the peripheral auditory system. Philip X. Joris (Div. of Neurophysiol., Univ. of Leuven, Campus Gasthuisberg, B-3000 Leuven, Belgium, philip.joris@med.kuleuven.ac.be) and Tom C. T. Yin (Univ. of Wisconsin, Madison, WI 53706)

Auditory-nerve fibers synchronize to the envelope of SAM stimuli over a wide range of modulation parameters. Synchronization is found at all characteristic frequencies and shows moderate gain differences between fibers of different spontaneous rate. Modulation transfer functions are uniformly low pass, with an upper frequency limit that systematically increases with characteristic frequency. As expected, envelope phase locking in bushy cells of the anteroventral cochlear nucleus is navelike, but marked transformations occur in other cell classes. Phase-locking gain is very high in onset cells in posteroventral cochlear nucleus. High gains are also seen in the dorsal cochlear nucleus; however, small changes in stimulus parameters often result in large changes in gain and phase. In general, it is unclear whether high gains are functionally significant, or rather a mere reflection of the transient response properties of cells. There

is evidence for the latter in one class of high-gain cell. Effects of modulation on average rate are small in the auditory nerve, but can be large in cochlear nucleus and, with binaural stimulation, in the olivary complex. The lateral superior olive and its bushy cell afferents provide an example of a circuit in which temporal envelope information is converted into average rate changes. [Work supported by NIDCD.]

9:30

2aPP3. The psychophysics of complex envelope processing. Sid P. Bacon (Psychoacoust. Lab., Dept. of Speech and Hearing Sci., Arizona State Univ., Tempe, AZ 85287-1908, spb@asu.edu)

During the past decade, psychophysical research on the processing of amplitude modulation (AM) has expanded to relatively complex situations where subjects are asked to make detection or discrimination judgments regarding some aspect of AM in the presence of competing AM. This AM-masking paradigm has been referred to as modulation masking when the signal and masker carriers are overlapping noise bands, and as modulation detection (or discrimination) interference (MDI) when the two carriers are sinusoids. An important finding in the MDI paradigm is that this AM masking occurs even when the sinusoidal carriers are 2 or more oct apart, and hence under conditions where peripheral interactions within a single auditory filter are unlikely. It also occurs when the signal and masker carriers are presented to opposite ears, further evidence that at least some of the AM masking is occurring at a relatively central site in the auditory system. Finally, a common finding in these masking experiments is a rather broad tuning for the detection (or discrimination) of AM. Whether this tuning should be taken as evidence for AM channels in the auditory system will be discussed, as will recent research on subjects with cochlear hearing loss. [Work supported by NIDCD.]

9:55–10:10 Break

10:10

2aPP4. Coding of amplitude modulation in the central auditory system. C. E. Schreiner, P. Bedenbaugh, S. Cheung, B. Bonham, S. Taha, and S. Nagarajan (W. M. Keck Ctr. for Integrative Neurosci., Med. School, UCSF, Box 0732, 513 Parnassus Ave., San Francisco, CA 94143-0732, chris@phy.ucsf.edu)

The coding of temporal information, as reflected in the response to amplitude-modulated signals, changes dramatically between the auditory midbrain and the auditory cortex. Most neurons in the inferior colliculus are tuned to specific AM rates between ten and several hundred Hertz. This tuning is expressed either in their capacity to phaselock their response to the AM rate or by their overall response magnitude. In the primary auditory cortex, this capacity is largely reduced to AM rates below 20 Hz. Other cortical fields show even lower AM limiting rates and usually only one field (the anterior field in the cat, and the posterior field in New World monkeys) shows a slightly higher following capacity. The consequences of this overall reduction in repetition coding capacity for the representation of temporal properties of complex signals, such as fundamental frequency or details of the temporal envelope, are profound. A second property of the temporal coding in the midbrain, its systematic distribution along the isofrequency-axis of the nucleus, is also diminished or absent at the cortical level. Potential origins of this transformation in temporal coding and consequences for the perception of temporal information will be discussed. [Work supported by NIH R01-02660.]

10:35

2aPP5. Modeling auditory processing of amplitude modulation. Torsten Dau (AG Medizinische Physik, Fachbereich Physik, Univ. Oldenburg, D-26111 Oldenburg, Germany, torsten@medi.physik.uni-oldenburg.de)

A popular model of the processing of amplitude modulation (AM) [Viemeister, J. Acoust. Soc. Am. **66**, 1364–1380 (1979)] consists of a bandpass filter, a rectifying nonlinearity, a low-pass filter, and a decision device. The low-pass filtering (or integration) stage was intended to simulate the temporal resolution limit by attenuating rapid changes in the envelope of the signal. Fleischer [Acustica **47**, 155–163 (1981)] developed a model describing the influence of the inherent fluctuations within a noise carrier on the detectability of added modulation. The modulation spectrum was weighted by a certain factor which also essentially represented a low-pass characteristic. In a recent modeling approach [Dau *et al.*, submitted to J. Acoust. Soc. Am. (1997)], a modulation filterbank replaced the low-pass filter, to analyze the envelope fluctuations of the stimuli in each peripheral auditory filter. The inclusion of a modulation filterbank was motivated by results from several studies on modulation masking which provided evidence for frequency selectivity in the modulation-frequency domain. In addition, the role of peripheral filtering for the processing of amplitude modulation, the envelope statistics, and the spectral distribution of modulation power in the modulation spectrum are discussed.

11:00–11:30

Open Discussion

2a TUE. AM

Session 2aSA

Structural Acoustics and Vibration: Vibration of Plates and Beams

Courtney B. Burroughs, Cochair

Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

J. Stuart Bolton, Cochair

School of Mechanical Engineering, 1077 Ray W. Herrick Laboratories, Purdue University, West Lafayette, Indiana 47907-1077

Contributed Papers

7:45

2aSA1. Effects of transducer mass on resonance frequency and quality factor of free-free bars. Qiushuang Guo and David Brown (Dept. of Elec. and Comput. Eng., and Ctr. for Marine Sci. and Technol., Univ. of Massachusetts Dartmouth, N. Dartmouth, MA 02747, dbrown@umassd.edu)

The use of resonant free-free bars to determine the elastic properties of materials has been presented previously [Garrett, "Resonant acoustic determination of elastic moduli," J. Acoust. Soc. Am. **88**, 210-221; Brown, "Characterization of viscoelastic materials using the resonant modes of a 'free-free' bar," J. Acoust. Soc. Am. **94**, 2294(A) (1993)]. The preferred excitation and detection method is based on electrodynamic transduction in which a coil of wire is attached to each bar end, which in turn is subjected to a strong magnetic field. The effect of the added transducer mass on the observed resonant frequency and quality factors of longitudinally driven bars used to determine Young's modulus and damping will be reported. [Work supported by the University of Massachusetts Dartmouth.]

8:00

2aSA2. Vibro-acoustic response reduction of a fluid-loaded beam using unequally spaced absorbers and concentrated masses. Chih-Chun Cheng (Dept. of Mech. Eng., Natl. Chung Cheng Univ., 160, San-Hsing, Ming-Hsiung, Chia-Yi 621, Taiwan, R.O.C., imeccc@ccunix.ccu.edu.tw)

A systematic method based on optimization techniques has been developed to determine locations of absorbers or concentrated masses required to reduce the vibration level, surface sound-pressure level, and total acoustic radiation power from a fluid-loaded structure. The method is demonstrated using a simply supported, fluid-loaded beam driven by a harmonic point force. In the optimization procedure, the unequally spaced absorbers or concentrated mass locations are the design variables and the designated location of vibration level, surface pressure, and the radiated acoustic power serve as the objective functions, respectively. Results show that the displacement response, surface sound pressure, and total radiated acoustic power can be reduced or controlled due to the presence of the absorbers or concentrated masses in an optimal arrangement. Although the reduction in vibroacoustic response achieved through concentrated masses is not as efficient as that of absorbers (as expected) satisfactory results in local vibration and surface sound pressure confinement are achieved using the concentrated masses. Note that the structure demonstrated here is simple, but it is believed that this methodology can be simply applied to other complex fluid-loaded structures. [Work supported by NSC, Taiwan.]

8:15

2aSA3. Quantitative verification of laser-modulated phase-stepping digital shearography for monofrequency vibration measurement. Benjamin A. Bard (Graduate Program in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804), Shudong Wu (Penn State Univ., State College, PA 16804), and Bernhard R. Tittmann (Penn State Univ., University Park, PA 16802)

An offspring of holographic and shearing interferometry, digital shearography is a full-field, optical technique. Its full-field nature allows examination of approximately a one square foot area, requiring no scanning. In its usual implementation for vibration studies, time-averaged speckle correlation fringes are formed. While sufficient for observing mode shapes, these fringe patterns carry several disadvantages when quantitative interpretation is attempted. Digital shearography, combined with stroboscopic laser illumination and optical phase stepping, overcomes these limitations. The differential displacement field during monofrequency vibration is calculated independently at each point, and is no longer dependent upon fringe interpretation. The system, a new combination of several well-established techniques, has not been validated in either the optics or acoustics literature. The goal of the current work is to verify the technique via side-by-side measurements made with shearography and laser vibrometry. A simple cantilever beam is vibrated at several resonance and antiresonance frequencies and examined by both methods. Analysis techniques for comparing velocity output of the vibrometer to differential displacement output from shearography will be discussed. Agreement between the two systems is excellent, demonstrating the potential usefulness of this new method. [Work supported by the Penn State University Applied Research Lab.]

8:30

2aSA4. Instability of an elastically supported Timoshenko beam under a traveling inertia load. Seroj Mackertich (Penn State Univ.—Harrisburg, Middletown, PA 17057-4898)

The instability of an elastically supported infinite Timoshenko beam under constant velocity traveling inertia load has been investigated. The regions of dynamic instability are determined for different values of the elastic foundation constant. Since the occurrence of this dynamic instability reduces the axial buckling load of the beam, the result is important for the study of buckling of a continuous beam.

8:45

2aSA5. Frequency response of a cantilever beam attached to a plate with varying boundary conditions. Robert C. Bouws and Richard G. DeJong (Eng. Dept., Calvin College, 3201 Burton, Grand Rapids, MI 49546)

The frequency response function at the free end of a cantilever beam attached to an irregularly shaped flat plate has been measured experimentally. Variations in the drive point acceleration due to changes in the boundary conditions of the plate are presented. The ensemble of frequency

response functions shows some deterministic features (due to the modes of the beam) and some statistical features (due to the varying boundary conditions of the plate). A numerical model using a beam impedance element and a simply supported rectangular plate with the same area and thickness as the test plate gives results that are similar to the ensemble of test results.

9:00

2aSA6. Spatial statistics of plate vibration modes with irregular boundary conditions. Michael J. Bakker and Richard G. DeJong (Eng. Dept., Calvin College, 3201 Burton, Grand Rapids, MI 49546)

The spatial statistics of several resonant bending modes in an irregularly shaped flat plate have been measured experimentally and calculated in a finite element model. The variations due to changes in the boundary conditions have also been included. The spatial statistics for one boundary condition are essentially the same as the statistics obtained over random variations in the boundary conditions. This is consistent with an ergodic model for the vibration pattern in an irregular plate in that one realization of the plate boundary condition represents the same statistics as the ensemble over varying boundary conditions.

9:15

2aSA7. Signal response of elastically coated plates. Richard F. Keltie (North Carolina State Univ., Raleigh, NC 27695)

The signal response of a two-dimensional elastic coating attached to a submerged plate has been examined. The coating response was generated by an incident acoustic plane wave impinging on the surface. The main points of interest were the normal and tangential velocity components of the coating response as functions of frequency, signal arrival angle, location through the coating thickness, and coating parameters. Primary attention was focused on developing physical understanding of the processes affecting the coating response, and the mechanisms by which this response could be tailored to achieve desired results. The waveguide behavior of the coating was found to be useful in explaining the signal response. In particular, it was found that the dilatational resonances were useful in explaining the normal velocity response, while the shear resonances were found to be associated with the tangential velocity maxima.

9:30–9:45 Break

9:45

2aSA8. Two-dimensional wave-number descriptions of flexural vibration in ribbed plates. Richard N. Brown (Sci. and Technologies Div., BBN Corp., Inc., 70 Fawcett St., Cambridge, MA 02138)

Regularly ribbed plate constructions are common in marine, aircraft, and other structures, and their vibrational response has been a subject of many theoretical studies. Most studies have been concerned with propagating waves in one-dimensional structures. In order to understand two-dimensional structural vibration, radiation, and near fields due to point-force excitation, it is helpful to study the distribution of surface velocity in two-dimensional wave-number space. This presentation is intended to clarify the physics of the complex response of such a periodic structure. An exact solution for the surface velocity response of a fluid-loaded ribbed plate subjected to a point-force load is solved for the distribution in wave number of the surface velocity. The result is graphically presented as a “movie” of color images of the velocity distribution where time is proportional to frequency. Overlaid indicators delineating the acoustic response region, Bloch-wave dispersion contours, and rib response filtering actions help to clarify the complex response patterns and their interactions.

10:00

2aSA9. Fluid-borne and Lamb-type waves on elastic plates in contact with two different fluids. X. L. Bao (Dept. of Mech. Eng., Auburn Univ., Auburn, AL 36849), H. Franklin (LAUE, URA, CNRS 1373, Univ. of Le Havre, Pl. Robert Schuman, 76610 Le Havre, France), O. Poncelet (Univ. Bordeaux I, 33405 Talence, France), P. K. Raju (Auburn Univ., Auburn, AL 36849, pkraju@eng.auburn.edu), and H. Uberall (Catholic Univ. of America, Washington, DC 30064)

The subject of propagating waves on fluid-loaded plates and shells, and their acoustic excitation, has become of great current interest. Besides the existence of the Lamb-type plate or shell wave modes, a water-borne Scholte-Stoney wave has been found on plates or shells specifically with one-sided water loading. Another specific case is that with two-sided water (or same-fluid) loading; here a symmetrical Scholte-Stoney wave was found in addition to an antisymmetric one. In this paper, more generally, the plate loaded with two different fluids is discussed via the corresponding dispersion curves of plate and fluid-borne waves. The existence of two Scholte-Stoney waves has been confirmed for this case.

10:15

2aSA10. A new trigonometric functions set for predicting plate bending, using the hierarchical finite element method. Olivier Beslin and Jean Nicolas (GAUS, Univ. of Sherbrooke, Sherbrooke, QC J1K 2R1, Canada, olivier.beslin@gme.usherb.ca)

This communication proposes a new hierarchical functions set to predict flexural motion of platelike structures in the medium-frequency range. This functions set is built from trigonometric functions instead of polynomials as classically encountered in the literature. It is shown that such a trigonometric set presents all the advantages of a classical hierarchical polynomials set and additional ones which are of interest if very high-order functions are desired to be used. It is stated that this new trigonometric set can be used at very high orders, up to 2048, without taking care of computer roundoff errors, while the polynomials set fails at order 46 because of the limited numerical dynamic of computers. This trigonometric set can be easily implemented on a computer. It does not require quadruple precision precomputed arrays. Only a very low number (which do not depend on the function order) of basic operations is needed when calling such functions. Moreover, it is shown that this trigonometric set presents a better convergence rate than polynomials when predicting high-order natural flexural modes of rectangular plates with any boundary conditions.

10:30

2aSA11. Convergence analysis of p -version FEM on vibration problem. Andre F. Cote, Francois Charron, and Noureddine Atalla (GAUS, Mech. Dept., Univ. of Sherbrooke, 2500 University Blvd., Sherbrooke, QC J1K 2R1, Canada, andre.cote@gme.usherb.ca)

The p -version FEM has proven to be very efficient for static problem. An analysis of the possibilities of this method on vibration problem is presented. The method is first explained and compared with the classic h -version FEM. Then, convergence indicators are defined and a convergence analysis is performed. To follow the performance on diverse situations, three types of plates are used as test problems: a rectangular plate, an L plate, and a stiffened plate. It is shown that the p -version outperforms the classic h -version by an average factor of 5 (the number of converged eigenvalues, for the same number of DOF, is five times greater). [Work supported by NSERC and FCAR.]

10:45

2aSA12. Vibroacoustic behavior with total or partial layer of constrained viscoelastic. Jean Nicolas, Olivier Foin, Noureddine Atalla, and Bertrand Mercier (GAUS, Mech. Eng. Dept., Univ. of Sherbrooke, Sherbrooke, PQ J1K 2R1, Canada, jean.nicolas@gme.usherb.ca)

The actual efficiency of constrained viscoelastic layers is sometimes misunderstood. To study in depth this configuration, a theoretical approach has been developed by generalizing Ungar's formulation [E. E. Ungar, J. Acoust. Soc. Am. **34**, 1082–1089 (1962)]. The shear effect, “in-plane”

2a TUE. AM

motion, and flexion are taken into account into the viscoelastic layer while only flexion is considered for the metallic layer. By expressing “in-plane” and shear displacements in terms of transverse displacements, a linear system is obtained. Radiation impedances are calculated via Nelisse’s model [H. Nelisse *et al.*, J. Sound Vib. **198**(4), 485–506 (1996)]. The theoretical results have been systematically validated experimentally with a dedicated setup. Basic configurations of interest are: basic plate, basic plate and constrained viscoelastic layer with total or partial covering under or not under the excitation forces. A case of nonmetallic (fiber resin) material is also studied, leading to very interesting results. Finally, an optimization study brings new meaningful design possibilities.

11:00

2aSA13. Experimental validation of a finite element model predicting the vibro-acoustic behavior of double-plate structures with Biot poroelastic materials. Raymond Panneton, Noureddine Atalla (GAUS, Mech. Eng. Dept., Univ. of Sherbrooke, Sherbrooke, PQ J1K 2R1, Canada, raymond. panneton@gme.usherb.ca), and Nicolas Dauchez (Univ. du Maine, Le Mans, France)

Recently, a tri-dimensional finite-element model predicting the vibro-acoustic behavior of multilayer systems containing Biot poroelastic materials have been developed [R. Panneton and N. Atalla, J. Acoust. Soc. Am. **100**, 346–354 (1996)]. The objective of the current paper is to validate the finite-element model with experimental measurements. Three experimental validations are proposed: (1) measurements versus FEM predictions of the surface impedance of a poroelastic material using an impedance tube; (2) measurements versus FEM predictions of the vibrations of a plate bonded to a poroelastic layer; and (3) measurements versus FEM predictions of the vibration transmission loss through a double plate system containing a poroelastic material. Results of the FEM predictions agree very well with the measurement results. [Work supported by NSERC, FCAR, and Bombardier Inc. Canadair.]

11:15

2aSA14. Vibrations of an orthotropic plate with an internal flat circular slot. H. S. Paul, M. Venkata Subramanyam (47/20 First Main Rd., Gandhi Nagar, Chennai-600 020, India), and M. Venkatesan (Anna Univ., Chennai-600 025, India)

Vibrations of free and clamped transversely isotropic plates with an internal flat circular slot are studied. According to the physical conditions, the normal stresses become uniform within the circular slot and displacements, and the shearing stresses are zero outside the circular slot. These are mixed boundary value problems which give rise to dual integral equations. Following Noble, the dual integral equations are reduced to Fredholm integral equations. Kernels of the integral equations are evaluated by the contour integration technique. The stress intensity factors are obtained for both cases and corresponding isotropic results are deduced. Numerical analysis is performed in the case of Beryl material. The relationship between the dimensionless frequency and the stress intensity factor is presented. The results thus obtained are discussed.

11:30

2aSA15. Vibration analysis of ship beams. Fuat Kara (Dept. of Naval Architecture, Istanbul Tech. Univ., Ayazağa Kampüsü, 80626, Istanbul, Turkey), G. Aşkar Altay (Boğaziçi Univ., Bebek, 80815, Istanbul, Turkey), and M. Cengiz Dökmeci (Istanbul Tech. Univ.—Teknik Üniversite, P.K. 9, Gümüşsuyu, 80191, Istanbul, Turkey)

This paper is essentially addressed to the vibrations of ships which were extensively studied both experimentally and analytically by modeling a ship as a beam of uniform cross section and predicting its vibration characteristics by Euler’s theory of elastic beams. To begin with, the free vibration characteristics of a ship beam are predicted from Timoshenko’s theory of elastic beam of variable cross section by use of Galerkin’s method. The results are evaluated by use of those obtained recently from Euler’s theory [e.g., S. Abrate, J. Sound. Vib. **185** (4), 703–716 (1995)]. Next, a numerical algorithm that is based on the method of moments [e.g., G. A. Altay and M. C. Dökmeci, Z. angew. Math. Mech. **77**, 1–17 (1997) and references therein] is developed so as to predict the vibration characteristics of a ship beam at both low and high frequencies. The equations derived include the effects of warping and shear deformations of cross section and, in fact, they incorporate as many higher order effects as deemed necessary in any case of interest. They govern all the extensional, flexural, and torsional as well as coupled motions of ship beam. Also, the results are compared with those of earlier ones. [Work supported in part by TÜBA.]

11:45

2aSA16. Moment method formulation for piezoceramic plate vibrations. M. Cengiz Dökmeci (Istanbul Tech. Univ.—Teknik Üniversite, P.K. 9, Gümüşsuyu, 80191, Istanbul, Turkey) and G. Aşkar Altay (Boğaziçi Univ., Bebek 80815, Istanbul, Turkey)

Presented herein is a numerical algorithm based on the method of moments that has been used recently in structural electroelastic elements [e.g., M. C. Dökmeci and G. A. Altay, Z. angew. Math. Mech. **77**, 1–17 (1997) and references therein] in order to predict the dynamic response of a piezoceramic plate. The piezoceramic plate of uniform thickness is coated completely with very thin perfectly conducting electrodes on and under an electric potential applied to its faces. In the first part of the paper, the two-dimensional equations for the piezoceramic plate are derived in both variational and differential forms by use of the differential variational principle [M. C. Dökmeci and G. A. Altay, Int. J. Eng. Sci. **34**(7), 769–782 (1996)] together with a series expansion of the field variables. In the second part, the method of moments [M. C. Dökmeci, *Proceedings of the 1992 Ultrasonics Symposium* (IEEE, New York, 1992), pp. 1025–1028] is described for the dynamic response. The significant mechanical and electrical effects are all taken into account. The equations derived accommodate all the types of motions as well as coupled motions of the piezoceramic plate. In the third part, attention is focused on certain cases involving special geometry, motions, and polarization directions of the piezoceramic plate. [Work supported in part by TÜBA.]

Session 2aUW

Underwater Acoustics and Acoustical Oceanography: Bottom Interaction

Kevin D. LePage, Chair

Bolt, Beranek and Newman, 70 Fawcett Street, Room 15-208, Cambridge, Massachusetts 02138

Contributed Papers

8:00

2aUW1. Extracting modal wave numbers from data collected in range-dependent environments. Ronald T. Kessel (School of Earth and Ocean Sci., Univ. of Victoria, P. O. Box 3055, Victoria, BC V8W 3P6, Canada)

Properties of the seafloor can be determined from the horizontal wave numbers of the normal modes of sound propagation. The wave numbers can be extracted from the harmonic sound field measured in range-independent media using a single hydrophone, by applying a Hankel transform to compute the wave-number spectrum from the field when either the source or hydrophone (not both) are towed [S. D. Rajan *et al.*, J. Acoust. Soc. Am. **82**, 998–1017 (1987), for example]. In range-dependent environments, the wave numbers vary owing to changes in bathymetry and geoacoustics, and this smears the wave-number spectrum, making the extraction of individual modes uncertain, if not impossible. This is particularly true for high-order modes which, penetrating more deeply into the sediment layers, provide more information about the sediments than low-order modes. A wave-number demodulation technique is derived from an adiabatic mode model whereby range-dependent wave numbers can be extracted from a considerably smeared spectrum. This not only extends the wave-number-based measure of seafloor properties to weakly range-dependent environments, but it provides a method for localizing a moving source in range-dependent shallow waters, simply by matching wave-number tracks in time with bathymetry maps.

8:15

2aUW2. Scattering from rough fluid–fluid interfaces using the finite-difference time-domain method. Frank D. Hastings, John B. Schneider, Shira L. Broschat (School of Elec. Eng. & Comp. Sci., Washington State Univ., Pullman, WA 99164-2752), and Eric I. Thorsos (Univ. of Washington, Seattle, WA 98105)

The finite-difference time-domain (FDTD) method is a numerical technique that can be used to solve a wide variety of problems via time-domain simulations. Results have been presented for scattering from pressure-release surfaces [F. Hastings *et al.*, IEEE Trans. Antennas Propagat. **43**, 1183–1191 (1995)] based on the FDTD method. This paper describes the use of the FDTD technique to study scattering from rough fluid–fluid interfaces. Precise results are obtained by employing several enhancements to the FDTD algorithm. Most importantly, a grid conforming to the rough interface is used to achieve accurate results at a reasonable computational cost. This is in contrast to a traditional stair-stepped model of the interface which gives spurious overprediction of the backscattered field. Numerical results are presented for the scattering strength obtained using a Monte Carlo technique and are compared with results obtained using a Monte Carlo integral equation approach. For both methods, the incident field is a tapered plane wave, and 50 surface realizations are used to obtain the average scattering strength. One advantage of the FDTD method is that it allows the addition of heterogeneities at little extra computational cost. [Work supported by ONR.]

8:30

2aUW3. Coupled-mode modeling of acoustic scattering from axisymmetric, three-dimensional objects. John A. Fawcett (SACLANT Undersea Res. Ctr., 19138 La Spezia, Italy)

A coupled-mode method for computing the wavefield scattered by three-dimensional, axisymmetric objects is presented. Only the compressional properties of the object and the medium are modeled. The method provides a unified approach to solving three-dimensional scattering problems for axisymmetric objects in free-space, in a waveguide, or partially or fully buried in a basement. In the cases of an infinite space, the computational domain is truncated and an artificial attenuation profile added near the boundaries to attenuate artificial reflections. Galerkin's method is used to construct the mode set in this case. Numerical examples of scattering from cylindrical disks surrounded by a homogeneous space and from disks partially buried in an underlying basement are presented. A benchmark case of scattering from an infinitely long cylinder, for which analytic solutions can be constructed, is also discussed.

8:45

2aUW4. Scattering from objects near the ocean bottom at very high frequency and geometrical diffraction theory. M. F. Werby (Naval Res. Lab., Code 7181, Stennis Space Center, MS 39529) and N. A. Sidorovskaia (Univ. of New Orleans, New Orleans, LA 70148)

For extremely high frequencies, the geometrical theory of diffraction is probably a sufficient, if not necessary, method to carry out predictions of scattering from ellipsoids or objects of more complicated structures for ka values on the order of a few hundred to tens of thousands. It is possible to extend the Helmholtz integral expression to the problem of ellipsoids near an ocean bottom by making use of a Green's function with a reflection coefficient that characterizes the bottom. In carrying out the integral expressions, saddle point methods were experimented with to determine the reflected signal, and the Maggi transformation and principles from differential geometry to reduce the diffraction integral over the shadow side of the object to line an integral at the illuminated shadow region of demarcation. The latter two devices are necessary due to time constraints in evaluating surface integrals for ultra-high frequencies. Examples are presented. [Work sponsored by ONR, the Naval Res. Lab., and the Univ. of New Orleans.]

9:00

2aUW5. Estimating in-plane bistatic scattering strength from a monostatic experiment. Paul C. Hines, D. Vance Crowe, and Dale D. Ellis (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada)

Acoustic backscatter data were collected from the ocean bottom at four sites on the SohM Abyssal Plain. Using a simple array consisting of a free-flooding ring projector, omnidirectional in azimuth, and an omnidirectional hydrophone, data were collected over the frequency range of 800 to 1700 Hz. To obtain sufficient signal-to-noise ratio without impairing spatial resolution, 200-Hz-wide linear FM pulses were employed. The array was deployed 500 m above the seabed and bottom backscatter data

2a TUE. AM

were obtained at grazing angles down to 4° before the onset of the first surface return. For data arriving after the first-bottom first-surface interaction, the geometry approximates an in-plane bistatic experiment. Therefore, the in-plane bistatic data and the monostatic data can be compared to examine the validity of using the separable and half-angle approximations used in some bistatic scattering models [J. Acoust. Soc. Am. **89** (1991)]. In this paper the bistatic geometry is explained and the data are interpreted in light of the bistatic arrivals. Then, an algorithm is presented for extracting the in-plane bistatic scattering strength from the data. Finally, bistatic and monostatic data are compared using the separable and half-angle approximations.

9:15

2aUW6. Mechanisms that cause changes to bottom reverberation fluctuation statistics. Linda S. Deliman and Diana F. McCammon (ARL/PSU, P.O. Box 30, State College, PA 16804-0030, lsd2@psu.edu)

An investigation into the underlying mechanisms that cause bottom reverberation fluctuation statistics to deviate from the typically assumed Rayleigh distribution reveals that sediment "patchiness" is a major contributor. The investigation involved processing both shallow water data and simulated data. The shallow-water data, collected during a SACLANT Centre experiment, include low-frequency measurements made over chalky sediments with varying degrees of roughness using different pulse types (cw and FM) and pulse lengths (2.2 and 8 s). The processing produced both Rayleigh and non-Rayleigh statistics. The results indicate that sediment type, pulse type, and pulse length are not a factor *per se* in the deviations of the statistics from Rayleigh. However, sediment "patchiness," defined as an ensouffled area containing patches which contribute largely different backscattering strengths within the area, does cause the statistics to deviate from Rayleigh. The simulated data reinforce these conclusions and provide insight into the spacial and temporal extent of "patchiness" required to cause a deviation in the statistics.

9:30

2aUW7. A generalized reverberation model. Kevin LePage (BBN Systems and Technologies, Cambridge, MA 02138)

A model for bottom and subbottom scattering is proposed which estimates the short time average of the received level as a function of the processing bandwidth and the channel characteristics. The model also estimates the degrees of freedom available in the received level as a function of time. The model describes how the backscatter from individual modes decorrelates as a function of time and processing bandwidth. For narrow bandwidths, modes interrogate the same bottom patches at large time, giving structure to the expected received level. Narrow-band signals also interrogate more roughness correlation lengths, yielding lower expected fluctuations about the mean. Results from the model are compared to experimental observations and conclusions are drawn regarding the validity of the proposed approach.

9:45

2aUW8. An investigation of seabed acoustic penetration and scattering via broadband echo soundings. Frank A. Boyle, Nicholas P. Chotiros (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029), Nicholas G. Pace, Oddbjorn Bergem, and E. Pouliquen (SACLANT Undersea Res. Ctr., 19138 La Spezia, Italy)

Recent experiments were conducted in the Mediterranean Sea with the intention of developing a better understanding of the physics of acoustic penetration into and scattering from the seabed. Two shallow water sites were studied, one near Viareggio, Italy and the other off the coast of Sardinia. One site consisted of a fine mixture of silt and sand, with a large amount of trapped gas, while the other included a layer of marine vegetation that produced a bright acoustic return ahead of the bottom echo. Echo soundings were collected with a broad beam impulsive sound source op-

erating over a band between 1 and 15 kHz. The data acquisition system was configured to record the acoustic pulse before and after encountering the seabed, from which a calibrated measure of the seabed's reflective character was obtained. The measured reflection coefficient was compared with model predictions, based on a poroelastic theory for the seabed. In general the measured reflection coefficients exceeded the predictions significantly. A likely hypothesis is that trapped gas bubbles act to enhance the reflection. A description of various signal processing techniques that were employed will be included, along with a discussion of relationships between the processed acoustic data and known structural features. [Work supported by ONR.]

10:00–10:15 Break

10:15

2aUW9. *In situ* measurements in marine sediment. P. Philip Thomson and John I. Dunlop (School of Phys., UNSW, Sydney 2052, Australia)

Investigations on the acoustic properties of marine sediments have been carried out on wide range of core samples from the Australian continental shelves over a wide range of frequencies. *In situ* measurements in marine sediments are providing better data for characterizing the propagation of acoustic signals in shallow water environments. An *in situ* acoustic probe has been designed and used to measure the *p*-wave velocity at 50 kHz using pulse transit timer. The transducer probe is inserted about 100 mm into the sediment with minimum distortion, the transmitter-receiver distance being about 150 mm. The dilatational velocity measurements were found to be of acceptable precision between those of *in situ* and laboratory measurement. Four sets of three matched piezoelectric transducers each have been mounted in colinear frames to enable measurements of sound speed ratio and attenuation constant at higher frequencies in the laboratory. The experimental and theoretical results were in good agreement and well fitted to mathematical models based on Biot-Stoll theory.

10:30

2aUW10. Sound-speed profile estimation using empirical orthogonal functions for stochastic propagation over poro-elastic sediments. Michael Jaye, William L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180), and Mohsen Badiy (Univ. of Delaware, Newark, DE 19716)

The efficacy of empirical orthogonal functions (EOFs) in modeling ocean fluctuations for acoustic propagation predictions is well known, and recent work has demonstrated their advantages for sediment sound-speed variabilities. In this study, using EOFs for the direct estimation of sediment sound speed, density, and attenuation profiles is described, when relatively extensive data sets such as those from the Atlantic Generating Station (AGS) site are available. First, interpolations are generated based on minimizing mean-square errors. The resulting fields are appraised and compared thoroughly with those obtained from universal kriging, a method which is commonly employed by geostatisticians. Not only do the EOF interpolations have advantages in numerical efficiency, but also their construction is modified here to incorporate lateral terrain orientations as occur at the AGS site. Uncertainties in the geoacoustic fields are then modeled by stochastic variations in the EOF coefficients. The field representations are used for environmental input to propagation modeling of poro-elastic sediments in the AGS region. Statistics of transmission loss fluctuations are discussed in terms of geoacoustic variabilities. [Work supported by ONR.]

10:45

2aUW11. Radiation impedance of a circular vibrating plate on the surface of sand. Masao Kimura (Dept. of Ocean Eng., Tokai Univ., 3-20-1 Orido, Shimizu, Shizuoka, 424 Japan)

Radiation impedance of a circular vibrating plate on the surface of sand varies with the viscoelastic properties of sand. Therefore, it is possible to predict the viscoelastic properties using the characteristics of the radiation impedance. In this study, the characteristics of the radiation impedance of

a circular vibrating plate on the surface of viscoelastic media were obtained by numerical analysis. The results show that the frequency at which the imaginary part of the radiation impedance becomes zero, is proportional to the shear wave velocity of viscoelastic media. The density of the media can be obtained using the values of the real part of the radiation impedance and the predicted shear wave velocity. Next, the radiation impedance for dry and water saturated beach sand were measured, and the measured values were compared with the calculated values using the equivalent Lamé's constant of sand.

11:00

2aUW12. Numerical simulation of shallow-water reverberation.

George R. Minassian (Underwater Acoust. Lab., N. N. Andreyev Acoust. Inst., Shvernika St. 4, 117 036 Moscow, Russia)

A computer model of monostatic shallow-water reverberation is described. The model includes two types of reverberation: surface and bottom. In this paper, hydrophysical and acoustical characteristics of a shallow-water environment needed for reverberation modeling are described. The Pierson-Moskowitz and Leikin-Rosenberg spectra of wind seas are used to compute surface-scattering coefficients in the base of a two-scale model. The Lysanov-Ivakin models of bottom roughness are applied to compute bottom-scattering coefficients. A numerical procedure for computing the reverberation angular power spectrum as a function of time is described. A procedure for generating a random quasistationary reverberation time series at the array output is given. The computer simulation results are presented.

11:15

2aUW13. Averaging a finely layered poroelastic Biot medium.

Andrey V. Bakulin (St. Petersburg State Univ., Geological Faculty, Dept. of Geophys., St. Petersburg, 199034, Univ. embankment 7/9, Russia) and Lev A. Molotkov (St. Petersburg Branch of Steklov Mathematical Inst., St. Petersburg, 191011, Russia)

Finely layered porous structures are of practical interest in geophysics. Thin layers in such structures are described by an isotropic Biot model. Backus averaging on stratified poroelastic medium has been generalized and it has been found that effective long-wave equivalent medium in this case is a Biot model with tensor densities, which is called a generalized Biot model. This means that total and fluid densities are different along and across lamination. Only if the fluid density is the same is the effective model an ordinary transversally isotropic Biot model. Total density is a tensor even in the case when the layers differ only by fluid properties. Solid and fluid layers are included in consideration as partial cases for a Biot model. Wavefronts (group velocities) are found and demonstrated for some typical cases. When one of the layers is elastic, it may be considered as a porous Biot layer with infinite tortuosity. In this case the effective model has a triangular wavefront of second longitudinal waves which has zero velocity across lamination. When one of the layers is fluid, then the effective model is a generalized transversally isotropic Biot model, which is quite similar to the effective model of a medium of alternating solid and fluid layers.

2a TUE. AM

TUESDAY AFTERNOON, 17 JUNE 1997

12:00 NOON TO 5:00 P.M., DEANS HALL I AND II

EQUIPMENT EXHIBITION

See page 3030 for list of exhibitors

Session 2pAAa**Architectural Acoustics, Noise, and NOISE-CON: Classroom Acoustics**

David Lubman, Chair

David Lubman & Associates, 14301 Middletown Lane, Westminster, California 92693

This session will be followed by a workshop (2pAAb) on standards for the acoustical performance of classrooms.

Invited Papers**1:30**

2pAAa1. America's need for standards and guidelines to ensure satisfactory classroom acoustics. David Lubman (David Lubman & Assoc., 14301 Middletown Ln., Westminster, CA 92693, dlubman@ix.netcom.com)

Architectural guidelines and standards for classroom acoustics are needed for the general population, as well as for special groupings such as hearing- and learning-disabled students, non-native listeners and talkers, and teachers with mild-to-severe hearing handicaps. Without specific guidelines, economic considerations or ignorance may tempt architects and school boards to use "standard commercial practices" for classroom construction such as "Noise Criteria" (NC) classifications. Proposed NC-40 ratings are unacceptable for many people in special groupings. An NC-40 classroom is sub-optimum for serious learning situations, even for the general population of children and adults. Needed acoustical guidelines can ensure adequately low noise and reverberation so that the speech-to-noise ratio and speech-to-reverberation ratio allow satisfactory communication and learning. Acoustical guidelines are also needed in connection with the growing use of classroom amplification to ensure that amplified speech is not used inappropriately, such as in reverberation-limited conditions or where inadequate sound isolation between adjacent classrooms produces distracting crosstalk. Guidelines can be developed with the cooperation of acousticians, architects, audiologists, school building officials, and government. Some of the costs for higher acoustical standards will be returned indirectly. Quiet classrooms may be a good investment for America.

1:50

2pAAa2. Impact of hearing loss on children in typical school environments. Peggy B. Nelson (Div. of Otolaryngol., Univ. of Maryland School of Medicine, 16 S. Eutaw St., Ste. 500, Baltimore, MD 21201)

Every school can expect to have children with hearing loss. Every teacher in the early grades can expect that every day there is a child in that classroom who has trouble hearing. Ear infections (otitis media, or OM) are the most frequent medical diagnosis for children. The National Center for Health Care Statistics estimates that (per 100 children per year) there are 70 OM cases in children under the age of 5 years, and 14 cases in children ages 5–17 years. Additionally, two have other permanent hearing losses. Hearing loss causes significant educational problems. Academic delays occur in children with mild hearing loss as early as first grade. Traditional classrooms are noisy. In many classrooms, the child in the middle of the room functions at 0-dB signal-to-noise or worse. Noise and reverberation combine to produce adverse effects on speech understanding by hard-of-hearing children. As S/N and reverberation times get worse, the performance gap between children with and without hearing loss increases. Children with hearing loss require 10–12 dB more favorable S/N than others to achieve the same word recognition scores. Hearing loss and poor classroom acoustics adversely impact learning for significant numbers of children.

2:10

2pAAa3. Speech interference by noise in classrooms and remedial solutions. Michel Picard (Ecole d'orthophonie et d'Audiol., Univ. of Montreal, Montreal, PQ H3C 3J7, Canada)

A review of the effects of ambient noise and reverberation on speech intelligibility in classrooms has been completed, initiated by the long-standing lack of agreement on permissible noise levels for good speech communication in educational facilities. Thus an overwhelming body of evidence has been collected to suggest that noise levels in particular are usually far in excess of any reasonable prescription aimed at producing optimal conditions for understanding speech in classrooms, assuming representative conditions of occupation by children. Quite surprisingly, poor classroom acoustics seem to be the prevailing condition for both normal-hearing and hearing-impaired students. In particular, reported ambient A-weighted noise levels are approximately 5–35 dB above values currently agreed upon for optimal understanding by normal-hearing children, and 17–32 dB too high for hearing-impaired children. Accordingly, strategies are discussed to decrease classroom noise levels and to improve classroom acoustics for speech communication purposes, taking into account the current state of economic constraints plaguing all school administrations.

2pAAa4. Pilot studies of speech communication in elementary school classrooms. Carl C. Crandell (Dept. of Commun. Processes and Disord., Univ. of Florida, 461 Dauer Hall, Gainesville, FL 32611), Martin A. Gold, M. Joyce Hasell, Christopher R. Herr, Hee Won Lee, Mitchell Lehde, and Gary W. Siebein (Univ. of Florida, Gainesville, FL 32611-5702)

Numerous investigations [Crandell, Smaldino, and Flexer (1995); Nabelek and Nabelek (1994); Olsen (1998)] have demonstrated that the acoustical environment in a classroom is an important variable contributing to the academic and psychosocial achievement of children. Additional studies have shown that children do not obtain adultlike abilities to perceive speech in noise or reverberation until approximately 15 years of age [Crandell (1992); Elliott (1979); Neuman and Horchberg (1985)]. Despite these data, there is very little literature concerning appropriate classroom acoustics for young children. Acoustical measures of reverberation time, early reverberation time, early-to-late energy ratios, loudness (or relative strength), Alcons, speech transmission index, background noise levels, and signal-to-noise ratios were made across a range of elementary school classrooms to document the acoustical conditions. Speech recognition of children in the rooms was assessed by nonsense syllables and monosyllabic words. Relations among the speech recognition measures, acoustical measurements, and the architectural design features of the classrooms were examined. Observations of teachers' methods and children's behavior were also conducted to analyze the relationships between the acoustical tests and the actual communication processes that occur in daily teaching. Data obtained will be discussed in view of developing appropriate classroom acoustics for elementary school children.

[See NOISE-CON Proceedings for full paper.]

Contributed Papers

2:50

2pAAa5. Acoustical design parameters for hearing impaired classrooms: A case study. Darron A. Chin-Quee (Rowan Williams Davies & Irwin, Inc., 650 Woodlawn Rd. West, Guelph, ON N1K 1B8, Canada)

In 1993, Toronto's York University completed construction of a classroom designed for teaching future teachers of the hearing impaired. Student teachers using the facility have various degrees of hearing, ranging from the deaf to the normal hearing. The classroom was equipped for multimedia presentation with assistive listening devices and designed with room acoustics consistent with extremely high speech intelligibility, particularly in the context of the hearing impaired. Considerations incorporated into the design included: maximum 250 and 500 Hz RT60 values of 0.4 s; signal to noise ratios of 20 dB with normal vocal effort; early reflections within 20 ms; and maximum background noise levels of NC-15 to NC-20. After 3 years of use, a number of acoustical parameters, such as RT60, STI, %Alcons, background noise, and sound isolation, were recently measured and compared to the subjective assessment of the users and the design criteria. The findings suggest that the room functions as intended with a high degree of user satisfaction. As such, it is hoped that the design criteria and approaches used can be of assistance in future designs for hearing impaired classrooms.

3:05

2pAAa6. Measurement of typical speech and background noise levels in university classrooms during lectures. Murray R. Hodgson and Rod Rempel (Occupational Hygiene Program, Univ. of British Columbia, 3rd Fl., 2206 East Mall, Vancouver, BC V6T 1Z3, Canada, hodgson@mech.ubc.ca)

A measurement system has been developed for determining typical long-term speech and background noise levels in university classrooms during lectures. A particular objective was to determine typical levels of student activity noise. A total of 18 lectures were recorded at up to four positions in 14 classrooms. These recordings were digitized and filtered using digital octave-band and A-weighting filters. Software processed the

resulting pressure time histories as follows: The signals were squared; short-term mean-square pressures were calculated; sound-pressure levels were calculated; sound-pressure-level frequency distributions were determined and plotted; and the resulting distributions were fitted by one, two, or three normal-distribution curves to identify peaks. Three curves generally gave the best fit. The levels of the peaks so identified were associated with speech, student activity, and background noise levels. The corresponding values were determined for each test case; the results are summarized and conclusions are drawn.

3:20

2pAAa7. Can noise levels at school gymnasias cause a hearing loss: A case study of a physical education teacher. Tao Jiang (Nova Scotia Hearing and Speech Clinic, 5599 Fenwick St., Halifax, NS B3H 1R2, Canada)

There is mounting evidence that noise is widespread at schools. Poor acoustic conditions affect speech and academic performance. Noise can be detrimental to teachers. Recent research has found noise at schools can become a hazard to the hearing. Unacceptable noises at gymnasias have been documented. Nevertheless, cases are rare that teachers are at risk of noise-induced hearing loss (NIHL). Presented is a case study of a physical education (PE) teacher diagnosed with NIHL. The audiologic evaluation indicates a bilateral sensorineural hearing loss characterized by a typical noise notch. To rule out presbycusis, the ISO 1999 standard is utilized. Consequently, medical and audiological evaluations have confirmed his NIHL. To investigate its cause, a noise study was conducted in the gymnasium. The noise levels ranged between 72 and 119 dBA during classes. Dosimetry recorded $L_{Aeq\ 8h}$ 90.8 dB(A) with the peak at 125 dB. The calculated dose exceeded the safety limit by 300%. Furthermore, to determine *in situ* noise levels of blowing whistles used by PE teachers, a real-ear analysis was performed using a probe microphone. The whistle noise reached 130 dB at the tympanic membrane. In conclusion, the teacher has developed a NIHL caused by hazardous noises associated with his teaching work.

[See NOISE-CON Proceedings for full paper.]

3:35

2pAAa8. Factors affecting children's speech communication in classrooms. Sigfrid D. Soli and Jean A. Sullivan (Dept. of Human Commun. Sci. and Devices, House Ear Inst., 2100 W. 3rd St., Los Angeles, CA 90057)

A child's ability to hear and communicate with speech in the classroom is influenced by several factors, including talker characteristics (level, clarity, dialect), the acoustical environment (noise level, reverberance), and listener characteristics (hearing status, cognitive development, language level, exposure to English). This presentation will review and assess the relative contribution of each factor to the communication requirements of the classroom. An emphasis will be placed on the large individual differences in listener characteristics that are likely to exist in contemporary classrooms, and their impact on the requirements of the acoustic environment for effective speech communication. Recent findings from this laboratory and from other researchers that document the range of speech communication abilities in both difficult and easy listening conditions among school-age children will be reviewed. Finally, an initial proposal for acceptable levels of background noise and reverberance in classrooms based on these findings will be offered. The rationale for this proposal is that noise and reverberance levels must be low enough to allow children with a wide range of listener characteristics to communicate with speech easily and effectively in the classroom.

[See NOISE-CON Proceedings for full paper.]

TUESDAY AFTERNOON, 17 JUNE 1997

ROOM P, 4:05 TO 5:30 P.M.

Session 2pAAb

Architectural Acoustics, Noise and Speech Communication: Workshop on the ASA's Role in Promoting and Implementing National Standards for the Acoustical Performance of Classrooms

Michael T. Nixon, Co-moderator
2810 Urbandale Lane, Plymouth, Minnesota 55447

Dana S. Hougland, Co-moderator
Omnivest Technologies, LLC, 1401 Seventeenth Street, Suite 450, Denver, Colorado 80209

A workshop will be held by the Subcommittee on Classroom Acoustics following the invited papers. All Meeting attendees and outside interested parties are invited to participate. The goal of this session is to develop a consensus for action on the appropriate standards, proposals for additional research, and the viable avenues for implementation. Small working groups will be organized within the workshop to develop specific proposals.

Session 2pEA

Engineering Acoustics: Loudspeakers: Past, Present and Future

John L. Butler, Cochair

Image Acoustics, Inc., 97 Elm Street, Cohasset, Massachusetts 02025

Alexander L. Butler, Cochair

Image Acoustics, Inc., 97 Elm Street, Cohasset, Massachusetts 02025

Chair's Introduction—1:30

Invited Papers

1:35

2pEA1. A short history of the dynamic loudspeaker, with comments. Edgar Villchur (Foundation for Hearing Aid Res., P.O. Box 306, Woodstock, NY 12498)

The history of the dynamic speaker, a speaker in which the signal is applied to a voice coil that is part of the moving system, begins with the Rice and Kellogg paper [Trans. AIEE **44**, 461–475 (1925)] and the “Radiola” loudspeaker of 1926. Innovations since then include the ported or bass reflex enclosure of Thuras, patented in 1932, and the acoustic suspension system. (The author received a patent for the latter in 1956, but a court later ruled that it had been anticipated by prior art.) Present design practice often departs, for good or ill—in the author’s opinion, mostly for ill—from design approaches of the past. For example, some current designs are aimed at creating directional sound rather than wide dispersion; some designs are aimed at controlling phase; and there is an explosion of design meant to influence loudspeaker performance through accessories such as speaker cables and speaker stands.

2:00

2pEA2. Radial magnetic topology for dynamic loudspeakers. Ted R. Telesky (AuraSound Professional Products Div., Aura Systems, Inc., 2335 Alaska Ave., El Segundo, CA 90245, ttelesky@aurasystems.com) and Neil A. Shaw (Menlo Sci. Acoust., Inc., Topanga, CA 90290)

Magnetic topologies for loudspeakers have remained essentially unchanged since the conversion from electrodynamic to dynamic magnetic structures. By utilizing the magnetic technology developed for long-stroke linear actuators, a neodymium magnet circuit was developed by dynamic loudspeakers that uses a radially charged magnetic topology. Benefits of this topology include large linear travel, low distortion, and high-power handling. The topology is also inherently shielded. The technology has been successfully demonstrated in dynamic cone-type direct radiator transducers that range in size from 2 to 18 in. in diameter.

2:25

2pEA3. Considerations in loudspeaker system directivity and power response. Kenton G. Forsythe, David Gunness, and Jeff Rocha (Eastern Acoust. Works, 1 Main St., Whitinsville, MA 01588, kenton@eaw.com)

The influence of power response of a loudspeaker system on sound quality and intelligibility is an often overlooked aspect of a speaker system’s performance. This paper will examine the importance of achieving uniformly predictable power response and directivity and the factors leading to a successful design. Traditional axial response, beamwidth, and directivity plots fail to convey a complete information set. Data sets from traditional and recent designs will be compared to demonstrate the important factors and the means of identifying them in proposed products. This will be examined for a range of loudspeaker system sizes and applications.

2:50

2pEA4. The DTC loudspeaker. Gabriel Weinreich (Randall Lab. of Phys., Univ. of Michigan, Ann Arbor, MI 48109-1120, weinreich@umich.edu)

The property of “directional tone color,” especially characteristic of a violin [G. Weinreich, J. Acoust. Soc. Am. **101** (to be published)], consists of a directional radiation pattern which is a very strong function both of angle and of frequency. This can produce the illusion that each note (in fact, each harmonic of each note) is coming from a different direction, confusing the common psychoacoustic cues and endowing the sound with a specially striking spatial sense. A loudspeaker has been constructed in which a single driver drives four pipes of different lengths, with their mouths forming a three-dimensional pattern of roughly 20-cm spacing. This produces four sources whose phases vary rapidly with frequency, and which interfere with each other so as to produce the type of pattern desired. The result (which will be demonstrated), when driven by a monophonic signal, appears to “fill space” more strikingly than many stereo systems. The paper will also discuss other ways in which the effect might be produced (for example, by using a number of drivers with varying signal delays), various applications (in particular, for pipe organ music), and implications for the way in which stereo music is perceived. [Work supported by NSF.]

3:30

2pEA5. The creation of audible sound from ultrasonic energy: A fundamental paradigm shift. Elwood G. Norris (American Technol. Corp., 12725 Stowe Dr., Poway, CA 92064, woody@atcsd.com)

The author presents HSS™ Technology, the creation of audible sound waves derived as difference information from ultrasonic energy. It will be demonstrated that the original experiments of German physicist, Hermann von Helmholtz, regarding the creation of difference frequencies by sounding two tones loudly in air, can be dramatically enhanced and applied to audio reproduction using ultrasonic frequencies. It will be shown that by emitting a complex ultrasonic wave into air from a single transducer, it is possible to produce difference frequency sound waves. Further, it will be shown that this audible “difference” sound is propagated as an independent wave once it is created in air. The technique eliminates the traditional need for multiple transducers, crossovers, and enclosures. The ability to “project or shine” sound in a particular direction is shown to be an inherent benefit of this technique. The reproduction of broadband music using this methodology will be demonstrated. An overview of the underlying science, physics, and mathematics supporting this technology will be presented.

Contributed Papers

3:55

2pEA6. The rebirth of a ribbon loudspeaker. Alexander L. Butler and John L. Butler (Image Acoust., Inc., 97 Elm St., Cohasset, MA 02025)

Although the ribbon driver has been around for many decades, it is still quite unknown to many people. A particular ribbon driver, which acts as a tweeter/midrange, was developed by Image Acoustics, Inc. The driver evolved from a multiple ribbon with substrate to a single corrugated ribbon design with a transformer. Throughout the development and testing the goal was a flat response, broad bandwidth (650 Hz to 20 kHz), wide beamwidth, and high sensitivity. This was not easy to achieve. A low-loss and low-inductive electric system as well as a strong magnetic field were required. The reduction of the inductance played a large part in achieving a flat response in the 15- to 20-kHz band. The approach and techniques used to achieve the desired results will be presented.

4:10

2pEA7. PANEL: A novel piezoelectric air acoustic transducer. V. D. Kugel, Sanjay Chandran, and L. E. Cross (Intercollege Mater. Res. Lab., Penn State Univ., University Park, PA 16802, vxk7@psu.edu)

The PANEL (Piezoelectric Acoustic Noise ELimator) transducer was developed for air acoustic noise cancellation in the frequency range of 200–1000 Hz. It has a thin panel structure, occupies less volume, and is lighter than an equivalent electromagnetic transducer. This device consists of a U-shaped configuration of piezoelectric bimorph and unimorph actuators bridged by a triangular amplifying diaphragm, which acts as an effective source of the sound radiation into air. The role of the diaphragm is twofold: it amplifies by a factor of 5–20 the tip displacement of the piezoelectric actuators and it also increases the sound emitting area. Amplitude of the vibrations and usable frequency range of the transducer depend on the dimensions and material of the actuators and diaphragm. The overall dimensions of the device fabricated were 52×50×24 mm. Soft PZT ceramics were used. The diaphragm was made of a carbon fiber reinforced composite. The diaphragm vibrations reach 500 μm rms well below the fundamental resonant frequency of bending vibrations which lies between 200 and 500 Hz. Above resonance, the amplitude of vibrations gradually decreases to 25 μm rms at 1000 Hz. The experimental results indicate that the transducer developed is promising as a sound transmitter for acoustic noise canceling devices.

4:25

2pEA8. The Cooper Union “super monitor” speaker system. Seho Kim (Cooper Union), Daniel R. Raichel (Cooper Union, Acoust. Res. Ctr., 51 Astor Pl., New York, NY 10003 and CUNY Grad. Ctr., raiche@cooper.edu), and John Torres (Cooper Union)

A pair of three-way transmission-line speaker systems was designed, constructed, and tested for the purpose of providing a high-accuracy reference monitor capable of reproducing the full audio range with extremely low harmonic distortion and fast transient response with minimal coloration. Each of the speaker cabinets is a mirror image of the other; and edges are made rounded to minimize diffraction effects. The use of six mid-range drivers enclosed separately from the woofer in each of the extremely stiff cabinets ensure linear operation even at fairly high levels of power input. The tweeters, also enclosed separately, are arranged to provide wide dispersion. While a three-way crossover network is incorporated, separate input jacks allow for bypass in order to accommodate tri-amplified inputs. The transmission line provides a perfectly resistive termination to the woofers, functions as a low-pass filter and virtually eliminates cabinet resonance. Damping material varies in density, tapering off from 0.5 lb/ft³, where the last 30 cm of the transmission line is unfilled. Directivity through 180° proved to be extremely even in the test ranges of 101 to 13 212 Hz.

4:40

2pEA9. Horn loudspeakers for the home: Past, present, and future. Bruce C. Edgar (Box 1515, Redondo Beach, CA 90278)

At the start of recorded sound, acoustical horns were an integral part of early home audio. As electrical reproducers were developed, horns were still required because of electromechanical driver and amplifier limitations. In 1928, the wider bandwidth of the Rice–Kellogg direct radiator loudspeaker and availability of higher-power amplifiers all but removed the horn loudspeaker from home audio systems. After 1928, horn loudspeakers were only found in theater PA systems, until the introduction in the late '40s of the Klipschorn, which spawned a revival of horn loudspeakers in the 1950s. The introduction of high-power solid-state amplifiers and small bookshelf speakers in the 1960s removed horns from home audio systems. In the early '90s, the reintroduction of low-power class A tube amplifiers spurred a new interest in new horn designs. The new horn designs utilize the Tractrix horn contour for high frequencies and hyperbolic-exponential horn contours for the bass frequencies. Also, the new horn design efforts have removed colorations in the sound which plagued many of the vintage horn systems. Future horn systems will feature lighter diaphragms, better magnets, and (maybe) room temperature superconducting voice coils for higher efficiencies.

2pEA10. Multipole expansion of loudspeaker as sound source.

Kaarina Melkas (Nokia Res. Ctr., P.O. Box 100, FIN-33721 Tampere, Finland, kaarina.melkas@research.nokia.com)

Higher-order vibration modes of a loudspeaker diaphragm are represented as multipole sources of a sound field. As each normal mode corresponds to a multipole series, the radiation induced by a vibration pattern containing many normal modes is the image of a linear mapping of the

participation factors to the coefficients of a multipole series. In terms of linear algebra, the participation factors and the set of corresponding coefficients are written as vectors. The generated sound field is then computed as a superposition from the multipole sources with strengths given by the coefficients in the series. Correction factors are introduced to count for the finite extent and other geometry related effects of the source. The higher-order normal modes form separate loads for the driving circuit. This method allows one to apply lumped parameter models at frequencies where the loudspeaker diaphragm does not move as a piston anymore.

TUESDAY AFTERNOON, 17 JUNE 1997

ROOM E, 1:30 TO 5:45 P.M.

Session 2pMU

Musical Acoustics: Acoustics of Bells

Thomas D. Rossing, Chair

Physics Department, Northern Illinois University, DeKalb, Illinois 60115

This session will be followed by a concert of handbell music.

Invited Papers

1:30

2pMU1. Acoustics of Eastern and Western bells, old and new. Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115)

Bells have been a part of nearly every culture in history. The acoustical properties of many different types of bells are compared, including church bells, carillon bells, handbells, ancient Chinese two-tone bells, and temple bells from several countries. Special attention is paid to modes of vibration and sound radiation.

2:00

2pMU2. Tower bell music through the centuries. Margo Halsted (Carillonist, Univ. of Michigan, 900 Burton Tower, Ann Arbor, MI 48109-1270)

From a single bell thousands of years ago to a modern carillon of up to seven octaves, bell sounds and bell music have fascinated, informed, regulated, and entertained the people within their sphere. The talk will concentrate on the Western traditions of marking the hours, signaling, banishing evil spirits, enriching market place activity, and filling the air with the distinctive and sometimes complicated sounds of a carillon concert. Slides and recordings will illustrate the instruments and the music.

2:30

2pMU3. Partial tones and tuning of American-made bells. Richard M. Watson (Meeks, Watson & Co., 10402 West Fork Rd., P.O. Drawer 700, Georgetown, OH 45121)

A study over the past 20 years has involved recording the relationship of the first five important partials in the sound of a large number of American-made bells by several founders, together with major physical dimensions of the same bells. This has allowed the correlation of the partial configurations with what might be termed "family profile" characteristics of these several major founders. It has been found that in such bells the deviation of the prime from the theoretically ideal octave relationship with the nominal is generally in the range $+1/2$ to $+2$ semitones; while, at the same time, the deviation of the hum from the theoretically ideal 2-oct relationship with the nominal has been found generally to be -1 to -3 semitones. Such deviations had formerly been considered uncorrectable by standard tuning methods (machining). However, from these data, and from practical experiments, it has proved possible to develop a general approach to tuning such bells, allowing principally the correction of the important octave partials (hum, prime, and nominal). This approach has made practical the tuning of several peals and chimes, one of which has been added to, with similarly scaled bells, to form a musically successful carillon.

3:00

2pMU4. Experimentally observed acoustic characteristics of the King Songdok bell. Yang-Hann Kim, Doo-Byung Yoon (Ctr. for Noise and Vib. Control, Dept. of Mech. Eng., KAIST, 373-1 Science Town, Taejeon-shi, Korea), and Young-Ok Jhin (Kyung-Hee Univ., Soowon, Korea)

The bells cast during the Silla dynasty were hung over various types of hollows or cavities in the ground. The sound pipe, called the eumtong, which opens through the top, appears in most of the bells. Another characteristic of these bells is that they ring with a very long beating frequency, sometimes for more than 3 s. The King Songdok bell ("Emille" bell) cast during the Silla dynasty (771 A.D.) is the most famous one among these. Its height is 3.66 m, the diameter of bell mouth is 2.23 m, and its mass is about 20 000 kg. A very extensive experiment was designed and conducted. To visualize how the bell radiates sound, a cylindrical acoustic holograph was attempted. The total measurement points for the holograph are 32×12 ; 32 microphones at every 30 degree. The acceleration signals were also measured extensively at many points, which were 11×12 . The acoustic characteristics of the inside cavity of the bell were also analyzed. This paper explains the role of the acoustic cavity in the ground, the sound pipe, and the source of long beating. Sound visualization based on the holographic measurement will demonstrate how its sound radiates in space.

3:30–3:40 Break

3:40

2pMU5. The "exercise" of changeringing: Acoustics, athletics, and algebra. Ronald D. Edge (Phys. Dept., Univ. of South Carolina, Columbia, SC 29208)

Ringling of changes on tower bells (the "exercise") involves acoustics, in the tuning of the bells, sound insulation, and propagation from the tower. Muscular power is required to ring, though even young girls can ring the light (treble) bells. At the same time husky teenagers ring the heavy tenors—all can join in. In England, 16th-century ringers developed the techniques to sound their bells (tuned diatonically) sequentially, using a rope and wheel to swing them through almost a full circle, mouth vertical. The period lengthens toward the balance point, and by using the rope to change the energy, the order in which the bells sound may be changed. Each bell is tuned for the strike note, the hum note (about an octave below), and partials a third (the tierce) and a fifth (quint) above. Methods such as "plain bob major," "superlative surprise maximus," or "grandsire caters" organize the "changes" (possible permutations of the bells rung). Mathematics is required to compose peals (5000 different changes) which are true (no repeats). Ringling in America died out after the Revolutionary War, but in the last 30 years, at least 25 new towers have been built here, and several pre-Revolutionary rings refurbished.

4:10

2pMU6. Bass handbells of aluminum. Thomas D. Rossing, Edward R. Mansell, Deepak Gangadharan (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115), and Jacob H. Malta (Malmark, Inc., Plumsteadville, PA 18949)

In order to obtain a higher radiation efficiency and thereby enhance the sound of bass handbells, Malmark has created a new bell design using aluminum rather than the traditional bronze. These bells are larger in diameter, and they have lower coincidence frequencies, both of which lead to more efficient radiation of bass notes. Using holographic interferometry, vibrational modes of a new aluminum bell are compared with those of a bronze bell having the same fundamental frequency. Sound radiation studies of these large bells reveal a prominent sixth-harmonic partial that is not present in smaller handbells.

4:40

2pMU7. The art of handbell ringing. Kathleen Ebling-Thorne (Westminster Choir College of Rider Univ., Lawrenceville, NJ)

The history and art of handbell ringing will be discussed. Members of the Westminster Concert Bell Choir will demonstrate some of the techniques used in modern handbell choirs, such as martellato, plucking, thumb damping, and four-in-hand ringing, as well as techniques for special effects, such as echo and gyro.

5:10–5:15 Break

5:15–5:45

Handbell Concert

Westminster Concert Bell Choir

Westminster Choir School of Rider University

Directed by Kathleen Ebling-Thorne

Session 2pNSa

Noise: Sound Propagation and Related Topics

Gilles A. Daigle, Chair

National Research Council, Institute of Microstructural Science, Ottawa, Ontario K1A 0R6, Canada

Contributed Papers

2:00

2pNSa1. Estimation of far-field sound-pressure levels based on near-field sound-pressure level data for various industrial equipment. Andrew Dicke and W. Brent Ferren (Black & Veatch, 11401 Lamar, Overland Park, KS 66211)

In order to estimate the noise radiated by an industrial facility, it is essential to accurately determine the far-field noise emissions of each piece of equipment. However, many times the only available equipment to determine acoustic information is near-field noise data. Therefore, accurate conversion of near-field to far-field noise data is necessary to determine equipment noise emissions at distant locations. This paper examines the process for converting measured near-field sound levels to far-field sound levels in actual equipment installations. Near-field data for various types of equipment are converted to sound power level and then extrapolated to a far-field sound level. The extrapolated sound levels are then compared to direct measurement data to verify the accuracy of near- to far-field calculations. The sensitivity of measurement results to environmental factors such as background noise and reflective planes is also addressed. The measurements and sound power level calculations are conducted in accordance with applicable ISO noise standards.

2:15

2pNSa2. Theoretical study of diffraction by a jagged-edge noise barrier. Penelope Menounou and David T. Blackstock (Appl. Res. Labs., and Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78713-8029)

An approximate analytical method for calculating the diffraction field behind a jagged-edge noise barrier, idealized as a thin rigid half-plane having an irregular edge, is presented. Experimental studies have shown that for apertures, disks, and plane barriers, making the edge irregular can significantly decrease the diffracted signal [Bailey *et al.*, J. Acoust. Soc. Am. **92**, 2359(A) (1992); Ho *et al.*, J. Acoust. Soc. Am. **101**, article in press (1997)]. Straight-edge diffraction from a rigid half-plane has a known exact solution for plane, cylindrical, and spherical incident waves. It is observed that in all three cases the shadow zone signal, in the far field, is equivalent to radiation from an infinite line source, but modified by a certain directivity function. In the present work the jagged edge is modeled as a crooked line source having the same directivity as a straight edge. The radiation is calculated by treating the crooked line as an infinite set of point sources, summing their contributions, and adding the appropriate directivity. The method is simple to apply. Results are compared with those found by other methods. [Work supported by Texas Advanced Technology Program.]

2:30

2pNSa3. Numerical simulation of jagged-edge noise barriers. Won-Suk Ohm, David T. Blackstock, and Ilene J. Busch-Vishniac (Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78712-1063)

The performance of a jagged-edge noise barrier (thin, infinitely long) is investigated numerically. The source is a spherical N wave. The theoretical approach is to treat the incident field in the plane above the barrier as a plane piston radiator (Kirchhoff approximation). Numerical calculation

gives the radiated field (pressure time waveform) at any point in the shadow zone. A discrete Fourier transform then yields the spectrum. The numerical result for a straight barrier compares favorably with the asymptotic result of U. J. Kurze [J. Acoust. Soc. Am. **55**, 504–518 (1974)]. Because of the intrinsic limitation imposed by the Kirchhoff approximation, however, the numerical method does not give accurate results when the listener is deep in the shadow zone. For the jagged-edge barrier, the numerical simulation confirms the experimental findings of Ho *et al.* [J. Acoust. Soc. Am. **101** (to be published 1997)] in that making the barrier edge jagged (1) increases the insertion loss at high frequencies, but (2) degrades the performance somewhat at low frequencies. Suggestions for improving the insertion loss formula of Ho *et al.* for a jagged-edge barrier are also made. [Work supported by the Texas Advanced Technology Program.]

2:45

2pNSa4. Continued investigation of noise reduction by a random-edge barrier. Eric Rosenberg and Ilene J. Busch-Vishniac (Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78712)

Recent experimental work [Ho *et al.*, J. Acoust. Soc. Am. (to be published)] indicates that replacing the straight-edge top of a noise barrier with a random edge of the same average height can result in improved insertion loss, except at low frequencies. It is thought that the jaggedness reduces the coherence of the sound diffracted off the edge so that the signal amplitude is diminished. These experimental results were obtained in a laboratory using a spark source and defining the maximum peak-to-peak sound pressure at the receiver as the metric. However, because the jagged edge has the effect of increasing the time of duration of sound observed from an impulsive source, it is not clear from the original results that the random-edge barrier reduces the sound energy received. Here the results of a new set of experiments using a similar test setup, but with more rigorous signal analysis are reported. This permits determination of whether the introduction of the randomness reduces the energy received. Additionally, it allows one to determine more precisely the conditions under which the jagged-edge barrier performs worse than a straight-edge barrier.

3:00

2pNSa5. A review of parallel traffic noise barrier literature. Nicole L. Martel and Ilene J. Busch-Vishniac (Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78712)

It is widely accepted that a single noise barrier beside a highway will provide a degree of noise reduction for neighboring homes. An interesting issue arises when residential areas line both sides of the highway, requiring two parallel barriers, one on each side of the highway. An abundance of research has been carried out to examine this situation. The research suggests that this configuration allows for multiple reflections to occur between the two barriers, unfortunately resulting in less noise reduction for both residential areas. Researchers attacking this problem have uncovered many relevant issues. This paper will summarize these issues and will class them into those in which there is consensus and those in which there is no consensus. Researchers agree that multiple reflections exist in a parallel wall situation, that these reflections cause an insertion loss degradation, that the insertion loss degradation can be controlled, and that making

2p TUE. PM

barriers absorptive or tilting them is effective in reducing the degradation. They do not agree on the extent of insertion loss degradation, the significance of environmental variables such as ground cover, or the limits to improvements from absorptive or tilting barriers.

3:15

2pNSa6. Comparison of traffic noise levels from various bridge/soundwall configurations. Michael Greene (Woodward-Clyde Int., 2020 E. First St., Ste. 400, Santa Ana, CA 92705)

Noise levels at receptor locations of varying distances and heights above a lake in Geneva, Switzerland were calculated. The traffic noise levels were computed for two bridge heights and six receiver heights. In addition, traffic noise levels for six distances (15-, 30-, 125-, 250-, 500-, and 1000-m) from the receptor to the near edge of the roadway were computed for each of the two bridge heights and six receiver heights. Four separate soundwall design scenarios were analyzed: (1) no soundwalls (base case); (2) two edge-of-shoulder plus one center divider reflective soundwalls, each 1 m in height; (3) two edge-of-shoulder reflective soundwalls, each 2 m in height; and (4) two edge-of-shoulder absorptive soundwalls, each 2 m in height. Comparison of the results shows that for all receptor distances and heights, noise levels would be lower for the 30-m-high bridge than for the 10-m-high bridge. This effect tends to diminish with receiver distance, however. The results also show that there are calculable and audible noise effect differences among the various antinoise barriers considered. For all receptor distances and heights, the soundwall design providing the highest degree of mitigation would be soundwall design alternative (4).

3:30

2pNSa7. On the use of pavement surfaces to attenuate traffic noise. Paul T. Calamia (Dept. of Elec. and Comput. Eng., Univ. of Texas, Austin, TX 78712-1084, calamia@ece.utexas.edu), Ilene J. Busch-Vishniac, Tracy A. Turen, and Michael T. McNeerney (Univ. of Texas, Austin, TX 78712)

Efforts to limit the propagation of traffic noise tend to focus on the use of noise barriers and their associated insertion loss. Bearing in mind that the source which contributes most to roadside automobile noise at highway speeds is tire/pavement interaction, an investigation into the use of "quiet" pavements to reduce tire/pavement interaction noise, and thus attenuate traffic noise at the source, has been undertaken. Using the trailer method, onboard and roadside measurements have been made of tire/pavement interaction noise on various pavements currently in use in the state of Texas, as well as in the Western Cape and Gauteng provinces, South Africa. Preliminary results obtained with a test-vehicle speed of 60 mph show that reductions in roadside noise levels from 4 to 15 dB can be achieved with "quiet" pavements. [Work supported by the Texas Department of Transportation.]

3:45

2pNSa8. A measurement method of an oblique-incidence absorption coefficient. Masaru Okoshi, Norio Shiroko (Chiba Inst. of Technol., 2-17-1 Tudanuma, Narashino, 275 Japan), Ken'iti Kido, Hideo Suzuki (Ontek R&D Co., Ltd., Yokohama, 226 Japan), and Takahiko Ono (Ono Sokki Co., Ltd., Yokohama, 226 Japan)

There are two well-known methods for the evaluation of absorption of a material. One is the acoustic tube method (normal incidence) and the other is the reverberation room method (random incidence). In a real situation, however, oblique-incidence absorption coefficients are required. For

example, it is necessary to have knowledge of absorption coefficients for various angles in order to evaluate the effect of the road surface on traffic noise. In the previous paper [H. Suzuki *et al.*, ASA-ASJ Third Joint Meeting, Honolulu, Hawaii, J. Acoust. Soc. Am. **100**, 2588(A) (1996)], a method of using the sound intensity technique was proposed. An absorbing material was placed on the floor of a semi-anechoic room. A line source with 16 loudspeakers with 100 mm diameter was used in order to try to generate a uniform sound field in the direction normal to the incidence-reflection plane. Measured results showed some irregular changes as a function of frequency. One possible reason for this irregularity was that the multi-loudspeaker source had a more complex directivity near the main axis than a single loudspeaker source. Results obtained by use of the single loudspeaker source are presented in this paper.

4:00

2pNSa9. Acoustic stop bands for periodic arrays of metallic rods. M. S. Kushwaha (Universidad Autonoma de Puebla, Apdo. Post. J-48, Puebla 72570, Mexico) and P. Halevi (Instituto Nacional de Astrofisica, Optica y Electronica, Apdo. Post. 51, Puebla 72000, Mexico)

Extensive band structure computations are performed for two-dimensional periodic arrays of rigid cylinders in air, with the Bloch vector being perpendicular to the cylinders. It is found that for any value of the period (of the square lattice), there is no acoustic gap for filling fraction $f < 30\%$; the magnitude of the gap, for $f > 30\%$, is inversely proportional to the period. Moreover, if the filling fraction exceeds 40%, a second gap, higher in frequency, opens up. Within these gaps (or stop bands) sound and vibrations are forbidden, irrespective of the direction of propagation. It is noted that for cylinders 2.9 cm in diameter and period of 10 cm (i.e., $f = 0.066$), there is no acoustic gap for frequencies below 6.4 kHz. However, the density of states reveals prominent minima at 1.7 and 2.4 kHz. These frequencies do agree with those of the first two attenuation maxima observed in a recent experiment [Nature **378**, 241 (1995)] and are indeed related to diffraction from the [100] and [110] planes. Thus, even with idealization, Sempere's sculpture (investigated in the above reference) exhibits pseudogaps—not full gaps. A proposed application of the above-mentioned band structure calculations is presented in the following abstract.

4:15

2pNSa10. Ultra-wideband filter for noise control. M. S. Kushwaha (Univ. Autonoma de Puebla, Apdo. Post. J-48, Puebla 72570, Mexico) and P. Halevi (Inst. Nacional de Astrofisica, Optica y Electronica, Apdo. Post. 51, Puebla 72000, Mexico)

It is proposed that arrays of periodically positioned metallic bars could serve as a low-tech, low-cost, and effective device for filtering out environmental or industrial noise. The proposed scheme is composed of five sets of periodic arrays in tandem, each characterized by a certain value of the period (of square lattice) a and cylinders of radius r ; the filling fraction is $f = (r^2/a^2)$. The purpose of the tandem structure is to achieve an ultra-wideband filter of noise. The tandem structure is designed such that, for certain values of the period, the stop bands of the respective sets of structure superimpose over one another. It is found that, for $f = 0.55$, the tandem structure should shield from noise in the entire range between 2.0 and 11.0 kHz. This is attributed to the formation of a "super stop band" due to overlapping of the five individual stop bands. An ultra-wideband filter for environmental or industrial noise could thus be fabricated according to the specific requirements. This idea is based on the band structure calculations reported in the preceding abstract.

Session 2pNSb

Noise and NOISE-CON: Quality Control, Noise Measurements, and the Pursuit of Quiet

Robert J. Comparin, Chair

Sound and Vibration Group, Copeland Corporation, 1675 West Campbell Road, Sidney, Ohio 45365

Chair's Introduction—1:15

Contributed Papers

1:20

2pNSb1. Implementation of a "buy quiet" policy for equipment purchases at the NASA Lewis Research Center. Beth A. Cooper (NASA Lewis Res. Ctr., MS 6-4, 21000 Brookpark Rd., Cleveland, OH 44135, beth.a.cooper@lerc.nasa.gov) and David A. Nelson (Hoover & Keith, Inc., Houston, TX 77082)

At the NASA Lewis Research Center, more than 1100 civil servant and contractor employees are included in a hearing conservation program that incorporates the requirements of the Occupational Safety and Health Administration standard on occupational noise exposure, as well as NASA's policy, which prohibits unprotected exposure to A-weighted noise levels above 85 dB. In many work areas, multiple, complex, or distributed noise sources require emphasis on hearing protection and personnel enclosures as the primary means of noise exposure management. Gradual, yet cost-effective, reductions in area noise levels are expected over time as equipment is replaced by selecting low-noise equipment whenever possible. To achieve maximum results from each equipment purchase, a "Guide to Specifying Equipment Noise Emission Levels" assists engineers and designers with identifying appropriate equipment noise emission requirements and incorporating those requirements into comprehensive specifications. The guide covers noise emission from a broad variety of fixed and portable equipment expected to produce A-weighted noise levels approaching 80 dB and higher. Developed to support hearing conservation program goals, the guide also addresses relevant community noise and speech communication issues.

[See NOISE-Con Proceedings for full paper.]

1:40

2pNSb2. Development of a sound audit program for a high-volume product. Robert J. Comparin (Copeland Corp., 1675 W. Campbell Rd., P.O. Box 669, Sidney, OH 45365)

Once sound has been identified as an important product feature, it becomes necessary to ensure that the sound levels of the product remain within the desired specification. A method for developing a sound audit program for a high-volume product is presented. The audit program is based on the use of statistical process control techniques. The sound audit program is used to monitor sound performance, improve quality control, and as an aid in the design of new products and new manufacturing processes.

[See NOISE-CON Proceedings for full paper.]

2:00

2pNSb3. Effect of blade profile on sheet metal shear noise. Ahmad Bahrami and Hugh M. Williamson^{a)} (Acoust. and Vib. Unit, University College, Univ. of New South Wales, Australian Defence Force Acad., Canberra, ACT, 2600, Australia)

Research has been carried out on the effect of different blade profiles on the radiated noise of an experimental high-speed sheet metal shear. The specimens being cut were roll formed sheet steel products of an approximately sinusoidal profile produced by BHP Building Products, Australia.

For this product, five different shear blade profiles of varying blade angles were designed and produced. In the tests, three parameters, i.e., blade angle, speed of cutting and blades alignment, could be varied independently. A pressure transducer was used as a measure of cutting force, allowing an investigation of the relationship between radiated noise and the time history of the cutting force. Results show that the blade angle in shear cutting is of great significance for the reduction of the radiated noise. Speed of cutting was found to have a significant but lesser effect. Blade misalignment gave some advantage when the blade angle was less than 0.5 degrees. ^{a)}Currently at the Institute for Research in Construction, National Research Council, Ottawa, Ontario, Canada.

[See NOISE-CON Proceedings for full paper.]

2:20

2pNSb4. Comparative testing of acoustic properties of cinder blocks packed with recycled materials for noise barrier applications. Alek Jamgocian, Daniel R. Raichel, and Constantine Yapijakis (The Cooper Union, Acoust. Res. Ctr., New York, NY 10003)

In the search for more effective noise attenuation methods, a number of candidate materials for road noise barriers and partition walls were tested at the Cooper Union Acoustics Research Center in a specially constructed enclosure divided into two sections by a wall constructed of these materials. The transmission losses of cinder blocks filled with various recycled materials constitute the acoustical property of principal interest. Three tests were executed for each type of material: (1) transmission loss measurements; (2) absorption coefficient measurements; and (3) diffraction effects. The intended applications of the designed materials are as noise barriers for urban highways and subway stations, and as partition walls for factories, classrooms, libraries, etc. Recycled materials experimented with would normally end up in land fills or incinerators.

[See NOISE-CON Proceedings for full paper.]

2:40

2pNSb5. Determining sound power using sound intensity in a reverberation room and other spaces. Robert Hickling (Sonometrics, Inc., 8306 Huntington Rd., Huntington Woods, MI 48070, sonomet@aol.com), Peng Lee, and Wei Wei (Univ. of Mississippi, University, MS 38677)

Tests were conducted with two sound-power reference sources (Bruel and Kjaer model 4204) in three different acoustical environments: half of a reverberation room (overall room dimensions 6.7×5.5×4.1 m), a large semi-anechoic room, and an indoor work space. Sound power (narrow band 136–6400 Hz) was determined using a semicircular array of eight sound-intensity probes rotated in steps of 20 deg (72 measurement positions) over a hemispherical integration surface, radius 1.52 m, with the source at the center. The floor base completing the surface enclosing the

source was assumed to be rigid with zero normal intensity. Despite prevailing concerns about sound-intensity measurement in diffuse fields and reverberant spaces, it was found that there was quite close agreement between sound power determined in the reverberation room and sound power determined in the other two environments. Sound power determined in the semi-anechoic room and work space was found to be highly repeat-

able, within 2%. For both sources the sound power was about 10% less than the manufacturer's calibrated value. With the source outside the integration surface, sound power determined in the semi-anechoic room was two orders of magnitude less than sound power with the source inside.

[See NOISE-CON Proceedings for full paper.]

TUESDAY AFTERNOON, 17 JUNE 1997

ROOM G, 3:15 TO 5:20 P.M.

Session 2pNSc

Noise and NOISE-CON: Sound Quality

B. John Feng, Chair

Ford Research Laboratories, P.O. Box 2053, Dearborn, Michigan 48121-2053

Chair's Introduction—3:15

Contributed Papers

3:20

2pNSc1. A psychoacoustic annoyance concept for application in sound quality. U. Widmann (Mueller-BBM, Planegg, Munich, Germany)

Industry engineers have to deal with the customer raising demands for acoustic comfort. At present, in the case of the judgment of sounds, manufacturers and customers go different ways. In the development labs noise is analyzed mainly by means of physical parameters. However, the customer judges a noise by means of human sense of hearing. These views can result in different assessments for the same sound. Therefore, objective quantities describing the different attributes of a sound in line with the human hearing system are necessary. In this study a psychoacoustic attempt is introduced aiming at a procedure for an objectively measurable psychoacoustic annoyance. The first results of a pilot study with synthetic noises with different loudness, sharpness, fluctuation strength, and roughness are presented, showing which psychoacoustic hearing sensations are related to the psychoacoustic annoyance evaluation of subjects. It is shown how the psychoacoustic annoyance evaluation changes if variations of more than one hearing sensation occur at the same time. A model summarizing the results for psychoacoustic annoyance is presented. General guidelines for application of the results for the development of consumer product sound is also presented.

[See NOISE-CON Proceedings for full paper.]

3:40

2pNSc2. Effect of definition and jury testing experience on perceived generality and importance of selected sound quality attributes. Kim Abouchakra and Tomasz Letowski (AMSRL-HR-SD, Bldg. 520, Aberdeen Proving Ground, MD 21005-5425, kaboucha@arl.army.mil)

People make judgments of sound every day. Sound judgments, like judgments of any other entity, can vary in their generality (general-specific) and the importance of the judged characteristic (primary-secondary). The development of sound quality criteria systems and jury testing protocols requires knowledge of both the generality and importance of specific sound assessment criteria. Smith and Letowski [J. Acoust. Soc. Am. **89**, 1938 (1991)] reported intuitive generality and importance ratings of selected sound quality attributes measured on naive listeners and proposed a set of sound quality attribute definitions. The object of the present study was: (1) to determine whether listeners maintain their intuitive generality and importance ratings for selected sound quality attributes follow-

ing exposure to the proposed definitions, and (2) to assess the effect of jury testing experience on the preservation of intuitive generality and importance ratings. The results support the notion that while some sound attributes have strong intuitive meaning (e.g., clarity), the others need clear definitions and experienced listeners (e.g., roughness) in order to be used successfully in sound quality studies. Some of the definitions tested in this study agreed well with the listeners' intuitive meaning of the defined attribute whereas others did not.

[See NOISE-CON Proceedings for full paper.]

4:00

2pNSc3. Influence of modulated complex tones on power window regulator noise quality. Richard D. Shadden and Teik C. Lim (Ctr. for Automotive Res., Ohio State Univ., 930 Kinnear Rd., Columbus, OH 43212-1443, shadden.1@osu.edu)

The influence of amplitude and frequency-modulated complex tones on the perceived quality of power window regulator noise is examined in this paper. Modulation characteristics of sounds generated by electro-mechanical actuators in automotive components, such as power window regulators, play a major role in affecting the perceived overall quality of a vehicle system. Annoying noise with a high degree of modulation tends to create the impression of an inferior product with inherent design flaws and manufacturing problems. Current subjective quality rating methods based on simple jury evaluation schemes and idealized metrics cannot be used to quantify modulation effects accurately. This is because modulated tones are temporal phenomena that are closely related to the rotational velocities of motors that are not typically accounted for in existing time-averaged sound quality metrics. A new time-varying based index is developed that incorporates operational velocity, noise spectrum, and binaural effects to quantify sound quality from acoustic signals measured using a binaural head system and rotational velocity data of the window regulator mechanism. Numerical simulations based on the proposed algorithm are subsequently used to predict the noise quality of various systems.

[See NOISE-CON Proceedings for full paper.]

2pNSc4. Sound quality evaluation of vehicle interior power train noise. J.-W. Kim (Noise and Vib. Team, Central Res. Ctr., Hyundai Motor Co., San1-1, Mabook-Ri, Kusung-Myun, Yongin-Shi, Kyungki-Do, 449-910, Korea, phin7429@chollian.dacom.co.kr)

In this paper, the sound quality of interior noise of three 1500-cc passenger cars driving at constant speed is evaluated. Paired comparison for three different cars and engine rpm was applied for subjective assessment with respect to important factors, i.e., booming, roughness, sharpness, powerfulness, and loudness. Based on the results, loudness is the most important factor on the perceived annoyance. Finally, an objective measure to evaluate: Each of five factors was conducted and each of these measures had good correlation and subjective evaluation.

[See NOISE-CON Proceedings for full paper.]

2pNSc5. A computer-aided paired-comparison algorithm using a 5-point comparison scale and its application to performing subjective ratings. Hsiao-an Hsieh (15031 S. Commerce Dr., Rm. G308, Dearborn, MI 48120)

The paired-comparison method (PCM) is a common tool used for sound quality studies in the automobile industry. In using PCM, evaluators are asked to compare two noise samples and indicate their preference for every possible sample-pair combination. A preference ranking of all samples is then generated. This paper describes the development of a PCM based on a five-point comparison scale. The primary motivations for this work were: (1) the need of predicting preference ranking for an individual or small group of evaluators, and doing so with reasonable repeatability and accuracy; and (2) predicting an individual's subjective acceptability ratings for all the samples without additional testing. Using a computer-optimization scheme on numerous simulated events, the author developed a weighting system for this comparison scale. A computer tool was sub-

sequently written which incorporated this weighting system and measured the relative preference-distances among all noise samples based on an individual's PCM inputs. The preference ranking was readily available, and the predicted subjective ratings were obtained by mapping these measured distances on a linear scale of 1 to N (1=totally unacceptable, N =perfectly acceptable). Some successes of using this method on real-life noise samples are presented and discussed.

[See NOISE-CON Proceedings for full paper.]

2pNSc6. Modeling judgments on sound quality by neural networks. Holger U. Prante (Tech. Univ. of Berlin, Inst. of Tech. Acoust., Einsteinufer 25, 10555 Berlin, Germany, prn2133@mailszrz.zrz.tu-berlin.de)

The current noise control methods are not sufficient for covering the ratings of different kinds of noise. They are mainly based on experiments with stationary and artificial test signals. New judgment models have to be developed to account for the time fluctuations of environmental sounds. During the last decade algorithms emerged in connectionism that not only find optimal mappings within stationary data (for example the backpropagation algorithm) but also within time series. One architecture for this task is called finite impulse response (FIR) neural network. Due to the architecture the networks include preceding time steps for the calculation of the current output. This contribution presents FIR neural networks to model sound quality judgments on environmental sounds. A listening test has been conducted for collecting training samples for the network and for testing its generalization abilities. The sound patterns and the subjects answers are pre-processed using factor analysis. The analysis of the subjects answers yielded 2 factors, which can be identified to represent the sound attributes annoying and powerful. The FIR neural networks are trained to estimate both attributes. The results of training and generalization show high correlation between the sound attributes and predicted judgments of the neural networks.

TUESDAY AFTERNOON, 17 JUNE 1997

ROOM N, 1:30 TO 4:45 P.M.

Session 2pPA

Physical Acoustics: Nonlinear Acoustics

Bruce C. Denardo, Chair

National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677

Contributed Papers

2pPA1. Second-order solutions for arbitrary piston radiators. Peter J. Westervelt (Dept. of Phys., Brown Univ., Box 1843, Providence, RI 02912)

Eckart's equation in the presence of the volume source density q is $\square^2[\pi + 2(\Psi^2)_{,00}] = \Gamma V_{,00} - \rho_0 c_0 q_{,0}$. The interaction terms $[(\rho_0 - \rho)c_0 q_{,0} - q c_0 \rho_{,0}]$ are absent, a consequence of $\square^2 \rho_s = -\rho_0 c_0 q_{,0}$. If the approximate first-order solution for a radiator with directivity $D_1(\omega, \theta, \phi)$, $p_1 = P_0 r_0 |r'|^{-1} D_1[\cos(\omega t - kr + \Psi)] \exp[-r\alpha(\omega)]$ in which $r' = r_0 + ir$, r_0 = Rayleigh distance, $\Psi = \arctan(rr_0^{-1})$, $P_0 = \rho_0 c_0 u_0$ with piston velocity u_0 , is introduced into Eckart's equation the second-order pressure is obtained: $p_2 = \Gamma(4\rho_0 c_0^2)^{-1} P_0^2 D_1^2 k r_0^2 |r'|^{-1} [(ln|r'|r_0^{-1})^2 + \Psi^2]^{1/2} [\cos(2\omega t - 2kr + \Psi - \chi)] \exp[-2r\alpha(\omega)]$ in which $\chi = \arctan[\Psi(ln|r'|r_0^{-1})^{-1}]$. Matching solutions at $r=r_0$ yields the term generating fingers: $p_f = P^* r_0^{-1} D_2[\cos(2\omega t - 2kr + 1/4\pi - \delta)] \exp[(r_0 - r)\alpha(2\omega) - 2r_0\alpha(\omega)]$ in which $\delta = \arctan[1/4\pi(ln\sqrt{2})^{-1}]$, $D_2 = D(2\omega, \theta, \phi)$, and $P^* = \Gamma(4\rho_0 c_0^2)^{-1} P_0^2 k r_0 2^{-1/2} [(ln\sqrt{2})^2 + (1/4\pi)^2]^{1/2}$.

For rectangular radiators the zero's of D_1 coincide with the zero's of D_2 therefore fingers do not occur in the case. The total pressure $p_t = p_2 + p_f$. Fingers arise from colinear interaction and have nothing to do with noncolinear scattering of sound by sound which does not exist.

2pPA2. Acoustic solitons in elastic media. Lev A. Ostrovsky (Univ. of Colorado, CIRES/NOAA Environ. Res. Lab., 325 Broadway R/E/ET1, Boulder, CO 80303)

Theoretical models involving acoustic solitons in elastic media with dispersion were first developed for nonlinear longitudinal waves in rods [Ostrovsky and Sutin (1975); Samsonov (1984)]. Recently, the interest to this problem has considerably grown. At the same time, experimental realizations of elastic solitons are scarce [e.g., the works by Lazaridi and Nesterenko (1985) where strong (nonacoustic) solitons were created in contacting spheres, and by Dreiden *et al.* (1989), where a localized elastic pulse was observed in a rod]. Here, possibilities of soliton existence in

different elastic media and some practically interesting values of their parameters are discussed. Three basic models are considered: (1) a thin elastic rod in which dispersion is due to the finiteness of its diameter; (2) a porous medium where dispersion is due to resonance properties of pores (similar to the known mechanism in a liquid with bubbles); and (3) a grainy medium with random packing of spherical grains in which nonlinearity is due to grain contacts and dispersion is due to finite grain sizes. It is shown that in all these cases, acoustic solitons may exist under realistic conditions. Possible realizations of such waves and their applications for medium-structure diagnostics are also evaluated.

2:00

2pPA3. Acoustic streaming at high Reynolds numbers due to focused piston radiation in real fluids. S. J. Younghouse, M. F. Hamilton, Yu. A. Il'inskii, and E. A. Zabolotskaya (Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78712-1063)

In an earlier presentation [Hamilton *et al.*, J. Acoust. Soc. Am. **97**, 3376(A) (1995)], a theoretical model for acoustic streaming at high Reynolds numbers (based on streaming velocities) in thermoviscous fluids was used for analytical and numerical predictions based on the assumption of weakly nonlinear focused Gaussian beams. Here, numerical calculations of streaming are presented for radiation from circular pistons, both focused and unfocused, and at amplitudes for which shock formation occurs. The forcing function for the streaming equations is evaluated with time-domain numerical solutions of the nonlinear parabolic wave equation. Effects of thermoviscous dissipation on the acoustic waveforms, and in particular on shock structures, are taken into account explicitly. The calculations provide a consistent description of acoustic streaming generated in real fluids by sound beams at moderate amplitudes, either with or without shocks. Because the shocks are not considered to be discontinuous features of the solutions, uniformly valid predictions of streaming are obtained in the transition region encompassing the birth and death of shocks amid the competing effects of dissipation and diffraction. The model accommodates both continuous and pulsed sources. Comparisons are made with experiments reported in the literature. [Work supported by the Office of Naval Research.]

2:15

2pPA4. Laser-generated nonlinear Rayleigh waves with shocks. A. Lomonosov, V. G. Mikhalevich (General Phys. Inst., Russian Acad. of Sci., 117942 Moscow, Russia), P. Hess (Univ. of Heidelberg, 69120 Heidelberg, Germany), E. Yu. Knight (Univ. of California, Berkeley, CA 94720-7300), M. F. Hamilton, and E. A. Zabolotskaya (Univ. of Texas, Austin, TX 78712-1063)

Intense laser radiation was used to generate a Rayleigh wave pulse of finite amplitude in fused quartz [A. Lomonosov and P. Hess, Proc. 14th ISNA, edited by R. J. Wei (Nanjing, 1996), pp. 106–111]. The laser beam was focused with a cylindrical lens to form a strip having dimensions 6 mm × 50 μ m on the solid. Measurements of the vertical (normal) component of the particle velocity at the surface were made with a second laser beam at distances 2.3 and 18.3 mm away from the irradiated strip. The horizontal (in-plane) component is related through the Hilbert transform. A Rayleigh wave pulse of duration 50 ns, velocity amplitude 30 m/s, was measured at 2.3 mm and used as the initial condition for numerical calculations based on a theoretical model for nonlinear Rayleigh waves in isotropic solids [E. A. Zabolotskaya, J. Acoust. Soc. Am. **91**, 2569(A) (1992)]. Numerical predictions for the waveform at 18.3 mm are in close quantitative agreement with the measurement. Nonlinear propagation effects produced well-defined shocks with cusped spikes in the horizontal velocity waveform, and a nearly doubling of the pulse duration due to the different propagation speeds of the head and tail shocks. [Work supported by Volkswagen, ISF, NSF, and ONR.]

2:30

2pPA5. Nonlinear interactions of acoustic, vorticity, and entropy modes, with applications to computational aero- and hydro-acoustics. Sheryl M. Grace and Allan D. Pierce (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02115)

The paper deals with fluid dynamical disturbances governed by the full nonlinear set of equations for a compressible fluid with viscosity and thermal conductivity included with arbitrary Prandtl number. The field governed by these equations is regarded as being a set of five functions of space and time, these functions being the three fluid velocity components, the pressure, and the entropy. A projection operator formalism is defined which uniquely separates the five-component field at any instant into three component fields, labeled for convenience as acoustic, vorticity, and entropy modes. The equations coupling these can be developed to any order of nonlinearity, but the superposition principle applies regardless of order. Lighthill's formulation of sound generated by flow falls easily out of the formulation without any intrinsic arbitrariness. Applications to computational aero-acoustics are discussed, with some numerical results for the case of sound generated by two adjacent line vortices.

2:45

2pPA6. The influence of aperture shape on Rayleigh conductivity for one-sided grazing flow. Sheryl M. Grace and Kelly P. Horan (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215)

It is well known that flow past wall apertures and wall cavities can generate sound. This resulting sound is most often undesirable. For underwater applications, such as flow past cavities in a ship's hull, the constancy of the phenomenon warrants passive control to reduce the acoustic radiation. In this research aperture shape is studied as a passive control mechanism. The Rayleigh conductivity for apertures of different shapes in the presence of one-sided grazing mean flow subject to an externally applied pressure perturbation is investigated numerically. Physically, the external pressure perturbation can result from an external sound source or large scale structural vibration. Because the linear stability of the shear flow over the aperture is governed by the properties of the Rayleigh conductivity, studies of the Rayleigh conductivity as a function of Strouhal number determine the frequency and flow speeds which correspond to self-sustaining oscillations of the flow and acoustic radiation. It is shown that all symmetric apertures with smooth leading edges behave similarly. For example, circular, square, triangular, and cross-shaped apertures all show similar conductivity patterns. However, results show that the conductivity can be altered by making either the aperture asymmetric or the leading edge of the aperture jagged.

3:00–3:15 Break

3:15

2pPA7. Numerical simulations of the absorption of sound by noise in one dimension. Vivek Mital, Bruce Denardo (Dept. of Phys. and Astron. and Natl. Ctr. for Phys. Acoust., Univ. of Mississippi, University, MS 38677), Hyeon Jang, and Andrés Larraza (Naval Postgrad. School, Monterey, CA 93943)

The absorption of a small-amplitude monofrequency signal by finite-amplitude shockless noise in one dimension has been predicted [O. Rudenko and A. Chirkin, Sov. Phys. JETP **40**, 945 (1975)] and experimentally measured [Larraza *et al.*, J. Acoust. Soc. Am. **100**, 3554 (1996)] to cause the signal amplitude to decrease as a Gaussian in the distance from the source. Numerical simulations based on Riemann's exact implicit solution for the unidirectional propagation of sound are presented, where the noise at the source is composed of 50 equally spaced equal-amplitude frequency components with random phases. The Gaussian attenuation is confirmed and is extended to the case where the signal is injected downstream from the noise source. When the phases of the spectral components of the noise are equal, the attenuation of the signal is dramatically reduced.

The extent to which the Gaussian attenuation is restored due to small random variations of the amplitudes, phases, and frequencies of the spectral components of the noise is investigated. Preliminary results show that this is specified by the statistical nature of the noise at the source; specifically, by the degree to which the instantaneous amplitude is normally distributed. [Work supported by ONR.]

3:30

2pPA8. The development of self focused pulses of nonlinear second sound in helium II from delayed edge waves. L. C. Krysac and J. D. Maynard (Dept. of Phys., Penn State Univ., 104 Davey Lab., University Park, PA 16802)

Second sound entropy waves in helium II are considered as a scalar temperature field imposed on an underlying counterflow velocity vector field. This unusual viewpoint offers a conceptual explanation of the form of nonplanar pulses of second sound, as well as a tool for understanding, and perhaps solving, the evolution of nonlinear pulses of second sound. A simple model of two point sources will be used to show how the large and negative finite amplitude nonlinearity near the phase transition of helium II is responsible for some of the observed features of the recently observed self-focusing of second sound.

3:45

2pPA9. Nonanalytic nonlinear oscillations: Christiaan Huygens, quadratic Schrödinger equations, and solitary waves. Bruce Denardo (Dept. of Phys. and Astron. and Natl. Ctr. for Phys. Acoust., Univ. of Mississippi, University, MS 38677)

An example of an oscillatory system with a time-reversible nonanalytic nonlinearity is shown to be a pendulum with a flexible cord sandwiched between two identical circular disks, in contrast to the analytic case of a pendulum interrupted by a single circular disk. The amplitude-dependent frequencies of both cases are perturbatively calculated, and are compared to numerical simulations. The nonanalyticity causes the unusual effect of the frequency to vary *linearly* with amplitude for small amplitudes, which has also been observed in compressional standing waves in sandstone. The general condition for a nonanalytic nonlinearity to yield this behavior is presented. The amplitude-dependent frequency for the double-interrupted pendulum allows an explanation for Huygens' observation that circular interrupters were as effective as cycloidal interrupters in achieving isochronous motion. A lattice of linearly coupled double-interrupted pendulums is described near the lower and upper cutoff modes by *quadratic* nonlinear Schrödinger (NLS) equations, in contrast to cubic NLS equations which arise for analytic pendulum lattices as well as typical acoustic and surface wave guides. Solitary breather and kink solutions to the quadratic NLS equations are presented, and the quadratic NLS breather is predicted to occur in a bar of sandstone. [Work supported by ONR.]

4:00

2pPA10. Acoustic wave propagation and reflection on nonlinear boundaries. Igor Yu. Solodov (Dept. of Phys., Moscow State Univ., Moscow, 119899 Russia)

Unlike optics reflection phenomenon in acoustics has long been considered a linear process. In the present paper a general approach to acoustic reflection problems at nonlinear solid-solid and liquid-solid interfaces will be discussed in terms of perturbation theory. Fresnel formulas for the

second-order reflection waves were derived for ideally bonded contacts of isotropic solids. Nonlinear bulk wave reflection was shown to result in efficient second harmonics generation, particularly, for overcritical angles of incidence. Unlike the linear case, nonlinear reflection generally leads to the rotation of the polarization plane for the harmonics of shear waves. Reflected harmonics are to be observed even when acoustic impedance is continuous across the boundary forbidding linear reflection to exist. Based on the nonlinear sound reflection problem, the method for analysis of boundary wave nonlinear propagation has been developed. It uses the idea of partial mode decomposition in wave reflection and takes into account self- and cross-interactions of acoustic waves reflected and transmitted at complex angles. The examples of analytical solutions for the second harmonics, as well as numerical calculations of nonlinear wave profiles, will be demonstrated for the Rayleigh, Stoneley, and Stoneley-Sholte wave propagation along nonlinear boundaries. The results obtained permit general characteristics of boundary acoustic nonlinearity that were shown to be different from those in the bulk of nonlinear medium to be defined.

4:15

2pPA11. Modulation instability of wave packets in media with resonant dispersion. Oleg M. Zozulya (N. N. Andreyev Acoust. Inst., 117036 Shvernika St. 4, Moscow, Russia)

In this work the modulation instability of quasimonochromatic wave packets propagating in conservative media with quadratic nonlinearity and linear nondamped oscillators attached to each point of a continuum is analyzed. Various density distributions of oscillators on natural frequencies are considered and carrier frequency ranges of unstable wave packet propagation are determined. It is shown that instability is closely related to the fact that the resonant dispersion relation can have many branches and thus permit phase-group synchronism between long and short waves. A wave packet propagating under the synchronism condition generates long-wave disturbances (with wavelengths near envelope width) and transmits up to one-half of its initial energy to them. The possible application of this effect for evaluation of microcavities size distribution in imperfect solids is considered.

4:30

2pPA12. Sound amplification by a supersonic jet under the action of saw-toothed sound waves of finite amplitude. Valerii G. Pimshtein (Ctr. Aerohydrodyn. Inst., Acoust. Div., Radio St., 17, 107005 Moscow, Russia)

At sound excitation of a supersonic shear layer, disturbances arise there which can radiate Mach waves at supersonic values of the convective movement velocity. In this paper, results are presented of experimental investigations of the action of the finite-amplitude saw-toothed sound waves on the mixing layer of a supersonic air jet issuing from a conical supersonic nozzle designed for $M=2.0$ at different sound incidence angles onto the jet boundary. It is shown that sound radiation by the excited jet at the external excitation frequency is associated with Mach wave radiation, its directivity is independent of the sound incident angle, and its intensity may exceed the sound intensity in the incident wave by the value up to 20 dB. The amplification value depends on the total pressure in the jet and the sound incidence angle depends on the jet boundary.

2p TUE. PM

Session 2pPP

Psychological and Physiological Acoustics: Physiological and Psychological Aspects of Amplitude Modulation Processing II; Binaural Processing and Sound Localization I

Constantine Trahiotis, Chair

University of Connecticut Health Center, Surgery and Neuroscience L4071, Farmington, Connecticut 06032

Contributed Papers

1:00

2pPP1. Detection of amplitude modulation, frequency modulation, and quasifrequency modulation by the budgerigar (*Melopsittacus undulatus*). Jian-Yu Lin and Robert J. Dooling (Dept. of Psych., Univ. of Maryland, College Park, MD 20742, dooling@bss3.umd.edu)

In budgerigars, as in humans, the detection of amplitude modulation (AM) remains relatively constant as modulation frequency increases while detection of frequency modulation (FM) improves. The point at which FM and AM are equal defines the critical modulation frequency (CMF). The CMF is approximately half the size of the critical band in humans because phase information is lost outside the critical band. At small modulation indices, the power spectrum of FM is almost identical to the power spectrum of AM with the difference being the relative phase of the components. The power spectrum of quasifrequency modulation (QFM) is exactly the same as AM even at high-modulation indices. In this experiment, two budgerigars were trained by operant conditioning to detect AM, FM, and QFM at several modulation rates at three carrier frequencies. Budgerigars show nearly identical thresholds for detecting modulation in FM and QFM tones at low-modulation rates and similar thresholds for detecting modulation in FM, AM, and QFM tones at higher modulation rates. These results argue for an insensitivity to phase differences in budgerigars when they fall outside the frequency bandwidths of the auditory system. [Work supported by NIH Grant Nos. DC-00198 and MH-00982 to RJD.]

1:15

2pPP2. Increment detection and sensitivity to amplitude modulation. M. Wojtczak and N. F. Viemeister (Dept. of Psych., Univ. of Minnesota, 75 East River Rd., Minneapolis, MN 55455)

It is clear that the ability to detect amplitude modulation (AM) is related to intensity resolution but the nature of this relationship is unclear. To address this issue, increment thresholds and modulation detection thresholds were measured under comparable conditions using the same subjects. The pedestal/carrier was a 1-kHz tone presented continuously throughout each block of trials. To ensure a substantial range of thresholds, the level of the pedestal/carrier was varied over a wide range (6–85 dB SPL). The duration of the increment and of the AM was 500 ms. Sinusoidal AM at a modulation frequency of 4 Hz was used. At this low frequency, modulation detection presumably was not affected by temporal resolution. As expected, the observed thresholds decreased with increased pedestal/carrier level. Over the range of measured modulation thresholds (about -33 to -5 dB, in $20 \log m$), the following well-describes the data: $10 \log(\Delta I/I) = 0.5(20 \log m) + 3$, where $\Delta I/I$ is the Weber fraction for increment detection and m is the modulation index at modulation threshold. The generality of this relationship and its theoretical interpretation will be discussed. [Work supported by Grant No. DC00683 from NIDCD, NIH.]

1:30

2pPP3. The effects of virtual pitch cues on the detection of coherent and incoherent frequency modulation. Robert P. Carlyon (MRC Appl. Psych. Unit, 15 Chaucer Rd., Cambridge CB2 2EF, England)

Recent experiments [S. Furukawa and B. C. J. Moore, J. Acoust. Soc. Am. **101**, 1632–1643 (1997)] indicate that listeners are better at detecting coherent than incoherent FM imposed on two widely separated frequency components. Experiment 1 repeated one of these experiments using carrier frequencies of 1100 Hz and 1925 Hz, modulated by one cycle of a 2.5-Hz sinusoid, in a noise background. Four out of eight listeners showed superior detection for coherent than for incoherent FM. Experiment 2 investigated whether this superiority was due to coherent FM producing larger changes in virtual pitch than does incoherent FM, despite the inharmonic ratio between the two components. In the signal interval of a 3I,2AFC task, the carrier frequencies of 1100 and 1925 Hz were shifted in either the same or in opposite directions, without disrupting the pattern of FM, which was either coherent, incoherent, or absent. Whatever the pattern of modulation, sensitivity was better for “same direction” than for “opposite direction” shifts, a result consistent with sensitivity to changes in virtual pitch. It is concluded that virtual phase sensitivity may account for the superior detection of coherently modulated tones, without recourse to an across-frequency mechanism specific to FM coherence.

1:45

2pPP4. Detectability of gaps in the modulation of a 4-kHz carrier. Aleksander Sek and Brian C. J. Moore (Dept. of Experimental Psychol., Univ. of Cambridge, Downing St., Cambridge CB2 3EB, England)

This experiment tested the idea that the auditory system contains a “modulation filter bank.” It was an analog in the modulation-frequency domain of an experiment in the audio-frequency domain, whose results were explained in terms of “ringing” in the auditory filter [M. J. Shailer and B. C. J. Moore, J. Acoust. Soc. Am. **81**, 1110–1117 (1987)]. Subjects had to detect a gap in sinusoidal modulation (rates of 5, 10, 20, 40, and 100 Hz, modulation index = 1) applied to a 4-kHz 70-dB sinusoidal carrier. The modulator preceding the gap ended with a positive-going zero crossing. The modulator was turned on at the end of the gap either at a positive-going zero crossing (standard phase) or at a negative-going zero crossing (reversed phase). For very small gaps [<0.1 times the period (P) of the modulator], detectability was very poor for the standard-phase condition but was good for the reversed-phase condition. For low modulation rates, and for a gap duration of $0.5P$, the opposite pattern was observed. The results are consistent with what would be expected from “ringing” in a modulation filter bank. [Work supported by The Wellcome Trust and the MRC, UK.]

2pPP5. Modulation gap detection in multicomponent stimuli. John H. Grose and Joseph W. Hall III (Univ. of North Carolina, Chapel Hill, NC 27599-7070)

The auditory system does not appear to adopt an efficient wideband temporal analysis if the temporal event to be detected traverses frequency with concomitant spectral discontinuity across the temporal event. In order to determine whether spectral continuity across the temporal event is a prerequisite for wideband temporal analysis, a novel paradigm was implemented in which the temporal event to be detected was a gap in the modulator of a sinusoidal carrier (the carrier was present throughout the modulation gap). The two markers of the modulation gap could be carried either by the same carrier frequency or by different frequencies across the modulation gap. In the latter case, the multiple carrier frequencies were continuously present whether they were carrying the modulation or not. Modulation gap detection was measured for an 8-Hz modulator carried by multicomponent stimuli whose standard duration was 625 ms. For all listeners, modulation gap detection deteriorated when the modulation gap traversed carrier frequency relative to when it was carried by a single frequency. Results will be discussed in terms of the conditions necessary for wideband temporal analysis. [Work supported by NIH NIDCD R01-DC01507.]

2:15

2pPP6. Amplitude modulation detection thresholds and associated psychometric functions.. René H. Gifford, Timothy D. Trine, and D. Wesley Grantham (Vanderbilt Univ. School of Med. and Bill Wilkerson Ctr., 1114 19th Ave. S., Nashville, TN 37212, timothy.d.trine@vanderbilt.edu)

Amplitude modulation detection thresholds for broadband stimuli were measured as a function of modulation frequency in seven listeners with normal hearing. Two adaptive procedures (1-down, 1-up and 4-down, 1-up) were used to track the 50% and the 84% correct levels in a three-interval, three-alternative, forced-choice task. Preliminary data show that the temporal modulation transfer functions for the two criteria are parallel for lower modulation frequencies but tend to converge with increasing modulation frequency. The resultant two-point psychometric functions associated with the modulation detection data demonstrate an increase in slope with increasing modulation frequency. Time constants were calculated from a least-squares fit of the equation $y = k/(f + a)$ and further illustrate the effects of the changes in psychometric function slope. These data have implications for comparisons of temporal acuity estimates across data sets.

2:30–2:45 Break

2:45

2pPP7. The Haas effect with and without binaural differences. Brad Rakerd, Joy Hsu (Dept. of Audiol. and Speech Sci., Michigan State Univ., East Lansing, MI 48824), and William M. Hartmann (Michigan State Univ., East Lansing, MI 48824)

The precedence effect for click signals operates successfully for delays in the range 0.5–5 ms. Previous work in our lab has shown that this effect behaves similarly whether sources are in the azimuthal plane or the horizontal plane. The precedence effect for running speech can operate successfully for delays as long as 50 ms—the “Haas zone.” Experiments were performed to measure two kinds of precedence threshold for running speech, echo threshold (listeners reported a barely perceptible echo), and masked threshold (listeners could barely detect that a lagging sound was present). The results were as follows: (1) Echo threshold levels decreased with increasing delay (20–80 ms) with a slope that was the same whether sources were in the azimuthal plane or the sagittal plane; (2) echo thresholds were higher in the sagittal plane than in the horizontal plane by a constant 3 dB; (3) masked thresholds were systematically 10–15 dB lower

than echo thresholds, and there was essentially no difference between azimuthal and sagittal planes. These results support the idea that precedence is functionally similar whether or not binaural differences are present. [Work supported by the NIDCD, DC00181.]

3:00

2pPP8. Measurements of precedence phenomena in binaural and monaural conditions. R. Y. Litovsky, M. L. Hawley, R. M. Dizon, and H. S. Colburn (Hear. Res. Ctr. and Dept. of Biomedical Eng., Boston Univ., Boston, MA 02215)

The precedence effect is a compelling auditory illusion which is thought to negotiate competition for perception and localization between the direct sound and its reflections. Three perceptual phenomena that have been related to the precedence effect are known to occur: (1) the two sounds are fused into one auditory event (fusion); (2) the perceived location of the fused auditory event is dominated by the leading source (localization dominance); and (3) directional information from the lagging source is inaccessible (discrimination suppression). Performance was measured on the above tasks under both binaural and monaural conditions, in both the azimuthal and median-sagittal planes, and in listeners with profound monaural deafness. Results suggest that: (1) Fusion occurs at similar delays under monaural and binaural conditions; (2) discrimination suppression is strong under binaural conditions but weak or absent under monaural conditions; (3) localization dominance is similar in the azimuthal and median-sagittal planes. It is tentatively concluded that the precedence effect may not be mediated solely by binaural circuits, but that monaural pathways may be involved as well. [Work supported by NIH Grant Nos. DC02692 and DC00100.]

3:15

2pPP9. Observer weighting of binaural information in source and echo clicks. Mark A. Stellmack (Dept. of Psych., Univ. of Minnesota, 75 East River Rd., Minneapolis, MN 55455), Raymond H. Dye, Jr., and Sandra J. Guzman (Loyola Univ., Chicago, IL 60626)

Relative weights given to interaural differences of time (IDTs) of source and echo clicks were computed for a range of echo delays (1–256 ms) for stimuli presented over headphones. On each trial, source and echo IDTs were selected randomly and independently ($\mu = 0 \mu\text{s}$, $\sigma = 100 \mu\text{s}$). Listeners were instructed to indicate the direction of the IDT, left or right, of the source or echo (in separate conditions). Weights were obtained by computing the correlation between the source or echo IDT and the listener's binary responses. In all conditions, little weight was given to the echo at short echo delays ($< 8 \text{ ms}$), but the echo was given weight equal to or greater than that of the source at intermediate echo delays (8–32 ms). For echo delays $> 32 \text{ ms}$, listeners gave greater weight to the source or echo, as appropriate. For all echo delays, when listeners were instructed to indicate the direction of the source IDT, the percent correct was lower when the echo DT varied across trials than when it was fixed at $0 \mu\text{s}$, indicating that the binaural information of the echo was not completely suppressed in any condition. [Work supported by NIH, NIDCD, and AFOSR.]

3:30

2pPP10. The influence of two-source noise on echo suppression. Megan L. Silk, Richard L. Freyman (Dept. of Commun. Disord., Univ. of Massachusetts, Amherst, MA 01003), and Daniel D. McCall (Univ. of Massachusetts, Amherst, MA 01003)

When listeners are presented with lead and lag pairs of clicks from two loudspeakers in an anechoic room, they hear a single image near the location of the lead loudspeaker, provided that the delay of the lag is only a few ms. However, Thurlow and Parks [Perceptual and Motor Skills **13**, 7–12 (1961)] reported that the addition of a continuous background of noise from a single loudspeaker causes the lag click to become audible as a separate auditory event. The current experiment investigated whether this apparent breakdown of echo suppression would also occur if the background noise was presented from two loudspeakers with one delayed relative to the other. The signal was a pair of brief white-noise bursts presented

from loudspeakers located at 45 deg right and 45 deg left in an anechoic chamber, with delays ranging from 2 to 14 ms. A continuous background white noise presented from the same two loudspeakers disrupted echo suppression only when the direction of its time delay was inconsistent with the signal delay. One interpretation is that echo suppression was broken because stimulus and noise provided conflicting information about the acoustics of the room. [Work supported by NIH DC01625.]

3:45

2pPPP11. Measurements of binaural temporal resolution using an analog of gap detection. Michael A. Akeroyd and Quentin Summerfield (MRC Inst. of Hearing Res., University Park, Nottingham NG7 2RD, UK)

One method for estimating the temporal resolution of the monaural auditory system is to measure the shortest silent gap, placed in a noise burst, that is detectable. An analog of this method was developed to estimate the temporal resolution of the binaural auditory system. The shortest burst of interaural-uncorrelated noise, placed between two bursts of interaural-correlated noise, that listeners could detect was measured. The spectrum level (40 dB SPL), bandwidth (100 Hz), and center frequency (250 to 1000 Hz) of the three bursts were the same. Its threshold duration depended upon center frequency: 6.5 ms at 250 Hz and 35 ms at 1000 Hz (silent-gap thresholds were 15 ms at 250 Hz and 13 ms at 1000 Hz). The interaural-correlation jnd was also measured: 0.03 at 250 Hz and 0.10 at 1000 Hz. The results were modeled using a Gaussian temporal window. Its duration is determined by both the threshold duration and the jnd: The values varied little with center frequency, i.e., 260 ms at 250 Hz and 230 ms at 1000 Hz. This lack of variation differs considerably from the variation of the experimental thresholds. The results are compatible with binaural temporal resolution being independent of center frequency.

4:00

2pPPP12. Binaural sluggishness in the masked discrimination of signals with different patterns of frequency change. John F. Culling (Dept. of Biomedical Eng., Boston Univ., 44 Cummington St., Boston, MA 02215)

The binaural system helps listeners to understand speech at low signal-to-noise ratios, presumably via time-varying binaural cues. However, the binaural system is often described as sluggish, because listeners cannot follow rapidly changing binaural cues. The speech in such experiments may have been articulated slowly enough for the binaural system to temporally resolve critical formant transitions. Alternatively, changing signal spectra may be analyzed by a nonsluggish aspect of the binaural system. In the former case, binaural unmasking of speech will depend upon speech rate. In a preliminary study, listeners' discrimination of changing signals in noise was measured as a function of the rate of frequency change. Masked discrimination thresholds were taken for NoSo and NoS π using a 2I-2AFC task. The signals consisted of ascending or descending three-tone glissandos (400, 500, and 625 Hz), and presented repeatedly and continuously during a 1.6-s broadband white noise burst. Rate of frequency change was controlled by altering the duration of each individual tone (32 to 100 ms). The difference between NoSo and NoS π was smaller for rapidly changing signals. Thus the binaural unmasking of speech should be dependent upon speech rate. However, the rate of articulation sufficient to abolish masking release is not known.

4:15

2pPPP13. Efficient representation of head-related transfer functions. Noel Chateau (TNO-TM, Kampweg 5, P.O. Box 23, 3769 ZG, Soesterberg, The Netherlands) and Adelbert W. Bronkhorst (TNO-TM, Kampweg 5, 3769 ZG, Soesterberg, The Netherlands)

Head-related transfer functions (HRTFs) contain information that is vital both for the analysis of acoustic cues used in sound localization and for the application of virtual-sound synthesis techniques. As the HRTFs measured for a representative set of source positions constitute a large body of data, the application of suitable data reduction techniques is of

crucial importance. An obvious method, which takes the properties of the human ear into account, is the averaging of the frequency spectrum within equidistant intervals on a logarithmic frequency scale. This method was perceptually evaluated in an experiment in which six listeners compared original (individualized) HRTFs, obtained for ten source positions, with HRTFs averaged within 1/3-, 1/6-, or 1/12-octave bands. Sound stimuli were bursts of pink noise with either a flat spectrum or with ± 2 or ± 4 dB roving within 1/3-octave bands. It appeared that differences could only be perceived for the HRTFs averaged within 1/3-octave bands. Neither roving level nor virtual source position had a significant effect on the results. It was furthermore found that data reduction of the HRTFs using principle component analysis is somewhat more efficient when HRTFs are first averaged within 1/6-octave bands.

4:30

2pPPP14. Constrained optimization of directional transfer functions using subjective preferences. Paul R. Runkle and Gregory H. Wakefield (Dept. of Elec. Eng. and Comput. Sci., Univ. of Michigan, Ann Arbor, MI 48109, paulr@eecs.umich.edu)

The synthesis of high quality virtual acoustic sources over headphones has been shown to depend on individual differences between subjects: Sources synthesized using one's own head-related transfer functions provide more accurate localization cues than sources generated with HRTFs obtained from others. A method of perceptually optimizing a generic set of HRTFs to more closely approximate those of a given subject is discussed. In this paper, HRTFs are represented as a cascade of a common and a directional transfer function [Middlebrooks and Green (1990)]. The DTFs are modeled as low-order pole-zero systems, whose parameters are varied using an adaptive iterative method [Wakefield and Runkle (1992)] that incorporates subject preferences with a search algorithm. Unlike previous attempts where no assumptions were made about the subject's criteria for judging the quality of prospective solutions, psychophysically based constraints have been imposed on the parameter space to facilitate rapid convergence. To assess the feasibility of using subjective feedback in the optimization of DTFs, subjects were asked to "move" a binaural source from an initial position to a nearby desired position. Issues of convergence and repeatability will be discussed. [Work supported by the Office of Naval Research in conjunction with the Naval Submarine Medical Research Laboratory.]

4:45

2pPPP15. A structural model of the head-related impulse response. C. Phillip Brown (Dept. of Elec. Eng., Univ. of Maryland, College Park, MD 20742) and Richard O. Duda (San Jose State Univ., San Jose, CA 95192, rod@duda.org)

A signal-processing model is proposed for the head-related impulse response (HRIR). As in earlier work by Genuit [K. Genuit, "A model for the description of outer-ear transmission characteristics," Ph.D. thesis, Rheinisch-Westfälischen Technischen Hochschule Aachen (1984)], different components of the model account for the directionally dependent filtering by the head, shoulders, and pinnae. Interaural-polar coordinates produce a natural separation of the azimuth- and elevation-dependent effects. The parameters of the head model depend only on head diameter. The parameters of the pinna model depend on pinna shape, and were derived from an image representation of the HRIR. The relative timing of the pinna "echoes" was found to be much more critical than their amplitudes or spectra. In a formal evaluation, subjects compared the apparent locations of virtual sources synthesized using the model to virtual sources synthesized using the subjects' individualized HRIRs. The effectiveness of the model was evaluated by having three subjects match the apparent location of a sound source. With a 500-ms Gaussian white noise burst heard through Etymotic Research ER-2 in-ear phones, the average error in matching elevations was less than 23 deg. [Work supported by NSF.]

2pPP16. Analysis of phantom source localization cues. Patrick M. Zurek (Res. Lab. of Electron., MIT, Rm. 36-730, 77 Massachusetts Ave., Cambridge, MA 02139 and Sensimetrics Corp., 26 Landsdowne St., Cambridge, MA 02139) and Barbara G. Shinn-Cunningham (Boston Univ., Boston, MA 02115)

A simple and widely used method for controlling the azimuthal direction of an apparent "phantom" sound source between two frontal loudspeakers is to vary the level difference between the (otherwise identical) signals delivered to the speakers. Explanations of phantom source localization [e.g., D. M. Leakey, J. Acoust. Soc. Am. **31**, 977–986 (1959); B. B. Bauer, J. Acoust. Soc. Am. **33**, 1536–1539 (1961)] have been based on the fact that, at low frequencies, the interaural phase delay in the dual-source stimulus matches that of an actual single source from some direction within the angular bounds of the loudspeakers. However, this theory fails to account for very similar localization results with stimuli that contain only high-frequency energy. It is also inconsistent with the perceptual robustness of the phantom image, which is unitary even with wideband stimuli. Leakey (1959) proposed that an interaural group delay underlies the high-frequency phantom. The present analysis of stimulus cues suggests that the interaural level difference may be an equally potent cue at high frequencies. Taken as a whole, this analysis points up the remarkable finding that this simple phantom-image technique results in the important interaural differences being well-matched to those of a single actual source across the audio spectrum. [Work supported by NIH.]

2pPP17. Auditory units enhance their response to sinusoidal AM in the presence of low-frequency modulation. Nikolay Bibikov (N. N. Andreyev Acoustics Inst., 4 Schverniki st, Moscow 117036, Russia, bvp@asu.acoins.msk.su)

Phase-locking responses to long sinusoidally amplitude-modulated best-frequency tones were recorded in midbrain auditory units of the curarized grass frog. The reproduction of the periodic AM (with modulation depths in the range 10%–30% and modulation frequencies in the 20- to 100-Hz range) was studied in the presence of low-frequency (0–50 or 0–150 Hz) noise added to the envelope waveform. The remarkable steadiness of periodicity extraction was observed in the presence of a low-frequency masker. Moreover, in some units, both the average firing rate and the phase-locking with a periodical modulation increased in the presence of noise. The last effect was especially pronounced when the modulation frequency (70–100 Hz) was located outside of the masker frequency band (0–50 Hz). The responses of torus semicircularis units to the best frequency tone amplitude modulated by a low-frequency two-tone complex were also studied. The typical observation was an increase in the synchronization index for the modulation frequency in the 20- to 40-Hz range with a low-frequency modulation (in 2- to 5-Hz range) being added to the signal. For both effects, the enhancement was even more pronounced if the Rayleigh statistic which took into account average firing rate was computed.

TUESDAY AFTERNOON, 17 JUNE 1997

ROOM O, 1:30 TO 4:45 P.M.

Session 2pSA

2p TUE. PM

Structural Acoustics and Vibration, Signal Processing in Acoustics, and NOISE-CON: Prediction for Dynamic Systems with Evolving Damage

David C. Swanson, Chair

Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

Contributed Papers

1:30

2pSA1. Aircraft damage detection from acoustic and noise impressed signals found by a cockpit voice recorder. Ronald O. Stearman (Dept. of Aerosp. Eng. and Eng. Mech., Univ. of Texas, Austin, TX 78712), Glen H. Schulze (Data Acquisition Systems, Littleton, CO 80123), and Stuart M. Rohre (Univ. of Texas, Austin, TX 78713)

Currently, research is being conducted to detect damage through structural acoustics, signal processing, and transducer designs. The present study illustrates that damage detection may be carried out with an existing system acting as a latent signal transducer. One example involved a reliability problem in a commuter liner aircraft engine mount design where undetected crack growth created a critical whirl flutter condition destroying the aircraft. This reliability problem prompted the need for an in-place damage detection system to identify critical engine mount conditions. Signal analysis of data acquired by a cockpit voice recorder prior to and during the catastrophic aircraft whirl flutter event provided insight into critical signals that indicated the failure onset. Although regularly scheduled inspections failed to detect the problem, cockpit voice recorder signals contained a dynamic signature of this damage feature throughout the duration of the tape. It is highly probable that this damage signature existed for a much longer period of time, but due to the endless loop configuration of the cockpit voice recorder the earlier data were erased. The study indicated that even in the case of an unused cockpit voice recorder track, careful signal processing can extract surprising details about detecting potential damage.

[See NOISE-CON Proceedings for full paper.]

1:45

2pSA2. Theoretical bases of the free-oscillation method for acoustical nondestructive testing. Nadejda I. Bouraou (Dept. of Orientation and Navigation, Natl. Tech. Univ. of Ukraine, 37, Peremogy pr., Kiev, 252056, Ukraine) and Leonid M. Gelman (Natl. Tech. Univ. of Ukraine, Kiev, 252056, Ukraine)

For acoustical low-frequency nondestructive testing, the free-oscillation method is used. However, in spite of wide practical application, the theoretical bases on this method are insufficiently investigated, which does not allow an analytical estimate of the basic acoustical testing data by using the free-oscillation method. For the first time, theoretical research of the free-oscillation method for NDT is given with account damping. A theoretical investigation is carried out to differentiate the spectral density data of free oscillations of materials in the presence and the absence of cracks, on the basis of which acoustical testing data are proposed. New analytical expressions between the estimates of spectral densities of acoustical signals and relative sizes of cracks, factors of damping, and the duration of nondestructive testing are received. Experimental test results of metal materials in the presence and the absence of fatigue cracks which match with analytical results are presented.

[See NOISE-CON Proceedings for full paper.]

2pSA3. Using phase space reconstruction to track parameter drift in a nonlinear system. Joseph P. Cusumano, David Chelidze, and Norm K. Hecht (Dept. of Eng. Sci. and Mech., Penn State Univ., University Park, PA 16802, chelo@psu.edu)

A model-based experimental method for tracking parameter drift in nonlinear dynamical systems is described. Local linear tracking models are constructed using data sampled over a fast time scale. These models are used to analyze data from systems with parameters which evolve over a slow time scale according to a "hidden" rate law. The method is applied to a numerical study of a nonlinear electrical circuit with a variable resistance as the drifting parameter. The mean-square tracking model prediction error is shown to follow successfully both ramped and sinusoidal parameter variations, suggesting that, at least in the cases studied, the method provides an invertible mapping between the parameter space and the observable space. Thus it should be possible to extract *rate information* about hidden drift, a requirement for true prediction. [This work is supported by the Office of Naval Research MURI on Integrated Predictive Diagnostics, Grant No. N0014-95-0461.]

2pSA4. Spectrum analysis structure of echosignals from composite pipe lines in aircraft technique. V. Genis (Biomed. Eng. and Sci. Inst., Drexel Univ., Philadelphia, PA 19104), I. L. Oboznenko (Kiev Polytech. Inst., Kiev, 252056, Ukraine), and T. Yu. Shkvarnitskaya (Kiev Avia Inst., Kiev, 252057, Ukraine)

An acoustical diagnostics model of composite shells for aircraft systems is proposed. The peculiarity of this model emerges from taking into account shell anisotropic properties. On this basis, the decision to use a method of structure breach in the pipe line is proposed. Peculiarities that are bound up with pipe-line constructions and determine natural resonances and antiresonances are noted. By excitation of surface waves it becomes possible to identify the presence of shell norms and destruction. Qualitative corroboration of sound-speed change effect in the composite shell is obtained. It allows one to prognosticate the state of pipe-line aircraft systems.

2pSA5. Crack detection in simple and complex beamlike structures. Sevag H. Arzumanyan (Dept. of Appl. Math. and Theor. Phys., Cambridge Univ., Silver St., Cambridge CB3 9EW, England) and David C. Swanson (Penn State Univ., State College, PA 16801)

An approximate technique, inspired by Skudrzyk's mean value theory, for determining the severity of single cracks in uniform beams having arbitrary intermediate and boundary conditions is presented. It is based on a weighted sum over eigenfrequency shifts of higher-order modes—modes 2 and higher—and does not require knowledge of the details of the structure. Starting from an analysis of a uniform beam with a midspan crack, a closed form expression, valid for all ranges of crack compliance, for the eigenfrequency shifts of higher-order modes is obtained. This result is used to derive an expression, consisting of a weighted sum of eigenfrequency shifts over a fixed frequency band, that is independent of structural details that are further than half a wavelength away from crack. This sum is valid for arbitrarily located cracks on uniform beams and, if evaluated over an appropriate band, is shown to be independent of boundary conditions and beam length. Finally, Skudrzyk's mean value theory is invoked to extend this approach to truly complex beam assemblies. Analytical predictions agree well with computational results. [Work supported by ONR.]

2pSA6. Prediction of crack growth in a tensioned band. J. Michael Spencer and David C. Swanson (Appl. Res. Lab., Penn State Univ., State College, PA 16801)

An experimental method is developed to predict crack growth, and ultimately failure, in a dynamically fatigued tensioned steel band. The transfer function across the double edge notched crack is measured. The band's modal response changes as the crack progresses, due to a change in the crack's equivalent compliance and a reduction in the remaining area of the material. The frequencies of the transfer function are tracked throughout the fatigue cycle, and the changes in frequencies are used in estimating crack growth and failure. A general trend shows that the resonance frequencies drop as the crack grows. When the material starts to approach failure, the frequencies stop dropping, and in some cases rise. This is an indication that failure can occur at any time. Just before failure, the modal frequencies drop rapidly, until the band breaks. Using the general trends in the change in the bar's frequency response it is possible to predict failure.

3:00–3:15 Break

2pSA7. An algorithm to count fatigue cycles in an arbitrary loading waveform. R. P. Kendig (Energy Systems Technol. Development, Westinghouse Sci. and Technol. Ctr., Pittsburgh, PA 15235, kendig.r.p.@wec.com)

The purpose of this investigation is to introduce a different algorithm for counting fatigue cycles due to an arbitrary acoustic or vibration excitation. The scope of this study is limited to a mathematical analysis with comparisons to several standard fatigue counting methods. The proposed algorithm, based on the Hilbert transform, processes a time history representing the loading condition to generate the number of cycle-counts with their corresponding amplitudes. The approach is general, is accurate within any desirable degree of accuracy, and is amenable to modern signal processing. The utility of the algorithm is demonstrated through comparisons to standard fatigue cycle-counting practices.

[See NOISE-CON Proceedings for full paper.]

2pSA8. Detection of surface defects in bearings using electromagnetic flux. Ellen Bradley and David C. Swanson (Appl. Res. Lab., Penn State Univ., State College, PA 16801)

The electromagnetic flux of rotating ball bearings was measured by a gradiometer-configured inductance coil. Ball spin, and drive and slave ring frequencies are distinguishable throughout the testing. One bearing was tested for 22 h, 45 min without reaching the failure, and the second bearing test involved restarting a bearing that had already reached the failure level but showed minimal visible damage. The first test showed increased magnetization of the bearing via an increase in flux at the spin frequency by 14 dB. The second test showed a consistent level of flux at the spin frequency compared to the final level detected in the first test. As the surface of the bearing becomes damaged from skidding, an increase in the magnetic field is hypothesized to originate from the abraded ferrous surfaces. The spinning ball allows this feature to be detected easily by a simple coil of wire.

[See NOISE-CON Proceedings for full paper.]

2pSA9. A new technique for the numerical solution of a class of integro-differential equations arising in elastodynamic crack propagation problems. Francesco Costanzo (Eng. Sci. and Mech. Dept., Penn State Univ., 227 Hammond Bldg., University Park, PA 16802, fxc8@psu.edu) and Jay R. Walton (Texas A&M Univ., College Station, TX 77843)

A strategy for the numerical solution of a type of integro-differential equation arising from the analysis of dynamic crack propagation problems is presented. Specifically, the problem of interest is that of a mode III crack dynamically propagating in a homogeneous linear elastic medium while the crack tip consists of a nonlinear rate-dependent cohesive (or failure) zone. The mode of propagation is general, that is, not restricted to steady-state or other special regimes. Furthermore, the presented solution technique is in no way restricted to mode III problems.

2pSA10. Dispersion of circumferential creeping waves around a fluid-filled cylindrical cavity in an elastic medium. Waled Hassan and Peter B. Nagy (Dept. of Aerosp. Eng. and Eng. Mech., Univ. of Cincinnati, Cincinnati, OH 45221-0070, whassan@uceng.uc.edu)

Ultrasonic inspection techniques using circumferential creeping waves can readily detect radial fatigue cracks on the fair side of thin airframe stiffener holes in empty aircraft wings. It is desired to utilize similar techniques in detecting fatigue cracks in the case of fluid-filled fuel tanks. The dispersion equation of a fluid-filled cylindrical cavity in an infinite elastic medium has been studied numerically. It was found that two distinct waves can propagate. One is the circumferential creeping wave that is attenuated by radiating energy into both the fluid and the surrounding solid. The other is the essentially nonattenuated halo wave caused by the resonances in the fluid within the cavity. The slower halo wave is the strongest of the two, but unfortunately it is not sensitive to the existence of any cracks at the surface because it propagates mostly in the fluid. It is the Rayleigh-type creeping wave that is important in the detection of surface breaking or near-surface cracks. It can be concluded that the adverse effect of fluid loading on the inspection technique is significant but not prohibitive. Additional finite-element calculations as well as experimental measurements are being conducted.

[See NOISE-CON Proceedings for full paper.]

2pSA11. Application of resonance acoustic spectroscopy to nondestructive evaluation of clad rods. F. Honarvar and A. N. Sinclair (Dept. of Mech. and Industrial Eng., Univ. of Toronto, 5 King's College Rd., Toronto, ON M5S 3G8, Canada)

The application of resonance acoustic spectroscopy (RAS) for the purpose of nondestructive evaluation (NDE) of various properties of a clad rod is investigated in this paper. RAS is the study of the resonances of an elastic object, which are excited by an incident wave, for characterization of the material properties of the elastic object. The equations for the back-scattered pressure field from an elastic/viscoelastic clad rod at both normal and oblique incident angles are derived. Numerical studies show how RAS can effectively help in monitoring the variations of the cladding thickness and cladding delamination. Experimental measurements, conducted on copper-clad aluminum rods, are found to be consistent with numerical results obtained from the mathematical model. It is demonstrated that there exists a potential for using RAS for NDE and on-line monitoring of various properties of clad rods.

2pSA12. Phase method of ultrasonic velocity measurement for the control of elastic properties of materials. Stanislav M. Mayevskyyi (Dept. of Nondestruct. Testing, Natl. Tech. Univ. of Ukraine, 37, Peremogy pr., Kiev, 252056, Ukraine)

Precision measurement of vibrations velocity propagation in solid body (components of power constructions) is discussed to control elastic properties of material and reliability forecasting. The ultrasonics velocity measurement method is based on the determination of the full phase shift of ultrasonic vibrations, coherent to base generator electrical oscillations. Ultrasonic vibrations are formed as long radio impulses, the maximal length of which is selected from the condition of reverberation influence exception on the measured phase shift result. The quantity of full phase cycles in phase shift is done by additional phase measurement on other frequencies, which differs little from the first frequency. The complete phase shift expression and time delay of ultrasonic vibrations is given. The US-velocity measurement method was used for longitudinal and transverse waves, propagated on the same trajectory. Using this, real Poisson's ratio construction material was determined. Analyses of reliable accuracy of ultrasonic velocity measurement and the results of experimental research is given.

Session 2pUW**Underwater Acoustics, Physical Acoustics, and Signal Processing in Acoustics: Time-Reversed Acoustic Propagation**

David R. Dowling, Chair

*Department Mechanical Engineering and Applied Mechanics, University of Michigan, 2033 Autolab 2121, Ann Arbor, Michigan 48109-2121****Invited Papers*****12:30****2pUW1. Time reversal in acoustics: Background and perspectives.** Darrell R. Jackson (Appl. Phys. Lab., College of Ocean and Fishery Sci., Univ. of Washington, Seattle, WA 98105)

The use of time reversal to obtain retrodirective focusing of a diverging acoustic wave field has precedents in electromagnetic and optical literature. This history will be discussed and possible acoustic applications will be illustrated by simulated examples. Phase conjugation in underwater acoustics is distinguished by the need to obtain true, rather than approximate, time reversal and by the effects of the ocean waveguide. The remarkable invariance of time-reversed focusing with respect to scattering by the propagation medium and deformation of the array will be discussed along with the sensitivity to time variation of these same factors. [Work supported by ONR.]

Distinguished Lecture**12:50****2pUW2. Time-reversed ultrasound.** Mathias Fink (Laboratoire Ondes et Acoustique ESPCI, Univ. Denis Diderot (Paris VII), 10 rue Vauquelin, 75005 Paris, France)

Time-reversal invariance is a very powerful concept in classical and quantum mechanics. However, simple experimental evidence of this concept is difficult to obtain. In the field of acoustic waves, where time reversal invariance also occurs, time-reversal experiments may be achieved simply with large ultrasonic piezoelectric transducer arrays. Time-reversal mirrors (TRMs) are made of such reversible acoustic retina, allowing an incident acoustic field to be sampled, recorded, time-reversed, and re-emitted. Time reversal of acoustic fields allows an optimal approach to focus pulsed ultrasonic waves through inhomogeneous media in the sense that it realizes the matched filter in both the spatial and temporal dimensions to the inhomogeneous propagation transfer function. In the field of acoustics, the ability to work directly in the time domain, with piezoelectric retina, allows pulse echo mode processing in an iterative mode. In multitarget media, such an iterative process converges on the most reflective target. In the case of an extended elastic target, automatic resonances of vibration modes can be achieved. TRMs are also innovative tools in the field of fundamental physics. They may be used to study inverse scattering problems, multiple scattering processes, chaotic scattering as well as scattering by turbulent media. Several experiments illustrating these areas will be presented. Applications of TRMs will also be described and include medical applications (kidney stone destruction and hyperthermia) as well as nondestructive testing of solid media. Experimental results illustrating these applications will be presented.

1:35**2pUW3. Experimental demonstration of an acoustic time-reversal mirror in the ocean.** William A. Kuperman, William S. Hodgkiss, Hee Chun Song (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0701), T. Akal, and C. Ferla (SACLANT Undersea Res. Ctr., 19138 La Spezia, Italy)

A time-reversal mirror (TRM) was implemented in an experiment conducted in the Mediterranean Sea in April 1996. The experiment utilized a vertical source-receiver array (SRA) spanning 77 m of a 125-m water column with 20 sources and receivers and a single source/receiver transponder (SRT) colocated in range with another vertical receive array (VRA) of 46 elements spanning 90 m of a 145-m water column located 6.3 km from the SRA. The TRM demonstration consisted of transmitting a 50-ms pulse from the SRT to the SRA, digitizing the received signal, and retransmitting the time-reversed signals from all the sources of the SRA. The retransmitted signal was then received at the VRA. An assortment of runs were made to examine the structure of the focal point region and the temporal stability of the process. The process was extremely robust and stable. Iterative enhancement of the focusing process was also demonstrated by a "ping-pong" exchange of a pulse between the transponder and the TRM.

2pUW4. Time-reversal signal processing: Background, theory, and application. James V. Candy (Lawrence Livermore Natl. Lab., Univ. of California, P.O. Box 808, L-437, Livermore, CA 94551)

In this paper the concept of time-reversal signal processing is briefly reviewed and its evolution is motivated as a viable technique to “focus” energy on a source through an inhomogeneous medium using an array of transducers. Discussed (briefly) are the underlying mathematics of this process and how it is a natural extension of the concept of phase conjugate mirrors, based on *monochromatic* waves (as in optics) to temporally *broadband* waves, which occur in ultrasonics or any pulse-echo based application such as ground penetrating radar. Once the basic theory and principles motivating this tantalizingly simple (algorithmically) approach have been developed some simulations to motivate the processing are developed. Shown here, this approach can be applied to “detect” anomalies, or more appropriately, scatterers in a medium and an optimization technique to sequentially remove the dominant (strongest) to detect even the weakest scatterer is discussed. This approach is similar to the CLEAN algorithm utilized in radio astronomy.

Contributed Papers

2:15

2pUW5. Iterative time reversal in the ocean. Hee Chun Song, William A. Kuperman, William S. Hodgkiss (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0701), T. Akal, C. Ferla (SACLANT Undersea Res. Ctr., 19138 La Spezia, Italy), and Darrell R. Jackson (Appl. Phys. Lab., College of Ocean and Fishery Sci., Univ. of Washington, Seattle, WA 98105)

The efficiency of an iterative time reversal process in the ocean has been demonstrated in an experiment conducted in the Mediterranean Sea during April 1996. [See Kuperman *et al.* in this session.] The iterative time reversal process consisted of initiating a 50-ms pulse with equal amplitudes on all elements from SRA to VRA, capturing the signal at 75-m depth on a single hydrophone of VRA, and retransmitting it from SRT at the same depth to SRA, reversing the received signal in time and retransmitting it from all the SRA sources back to VRA. This process was repeated many times over an hour. Although this was not the intended mode of operation of iterative phase conjugation as in the free-field case with many scatterers [Prada *et al.*, J. Acoust. Soc. Am. **90**, 1119–1129 (1991)], it shows that an iterative process can enhance the focusing of a single target in the ocean. Comparisons of data with simulations and theoretical analysis will be presented and discussed.

2:30

2pUW6. Measurements and modeling of signal spread in the Barents Sea using time-forward and time-reversed signals. Roger C. Gauss (Naval Res. Lab., Washington, DC 20375-5350) and John R. Preston (Penn State Univ., State College, PA 16804)

Low-frequency (350–450 Hz), shallow-water propagation measurements were made in several sites in the southern Barents Sea in June 1991 using combinations of broadband signals. The experiment used a two-element vertical source array to transmit the signals, and an omnidirectional source/receiver echo-repeater system to retransmit the signals back to a horizontal line array near the original source location. (All platforms moved at 3 kt.) Measures of two-way signal spreads were collected over a variety of ranges (5 to 50 km) and environmental conditions. Considerable time spreads (up to 1 s) were observed with a characteristic constancy of peak values over the length of the spreads, while analysis of PRN and CW data showed that spectral spreads were confined to very low Dopplers. Furthermore, selected sets of signals were time reversed prior to echo repeating to measure the ability of this highly-spread environment to refocus acoustic energy back to the original source location. Simulations of two-way temporal signal dispersion using the OASES and ORCA models for both time-forward and time-reversed retransmissions are presented and compared to data results for several representative cases. [Work supported by ONR and NAWC.]

2:45

2pUW7. Modeling of signal spreads in shallow water for one-way, two-way, and two-way time-reversed signals. John R. Preston (Appl. Res. Lab., Penn State Univ., State College, PA 16804) and Roger C. Gauss (Naval Res. Lab., Washington, DC 20375-5350)

Shallow-water broadband propagation modeling and time-series simulation was done for representative ocean environments. Simulations of two-way temporal signal dispersion for both time-forward and time-reversed transmissions are presented. The modeling effort used the OASES and ORCA models to simulate acoustic conditions from horizontal line arrays and omnidirectional sources in the 30 to 600-Hz frequency region. Two-way predictions are presented and compared to data results for several representative cases. Some sets of signals were time reversed and back propagated through several perturbations of selected experimental conditions to quantify signal decorrelation effects due to the environment. Additionally, the effects of water depth on pulse shape and spread are demonstrated for two hypothetical shallow-water cases. Comparisons with one-way time series measured during the Celtic Duet sea trial conducted off the Southwest Approaches in July 1992 are discussed. Results for one-way modeling in one hypothetical case for a sandy bottom show pulse spreads gradually increasing up to 2 s as depth is varied from 50 to 350 m and ranges are increased out to 18 km. [Work supported by ONR and NAWC.]

3:00–3:15 Break

3:15

2pUW8. Time-reverse acoustic array focusing in the presence of a soliton-type internal wave. Sunny R. Khosla (Dept. of Mech. Eng., Univ. of Michigan, Ann Arbor, MI 48109, indu@engin.umich.edu) and David R. Dowling (Univ. of Michigan, Ann Arbor, MI 48109)

Time-reverse or phase-conjugate arrays (TRAs or PCAs) have been shown to retrofocus sound to the location(s) of its origin even when the acoustic medium and its boundaries are unknown. This retrofocusing is degraded by temporal variations in the acoustic environment. In this study, analytical predictions are made for the focal properties of a linear vertical time-reversing array in a stably stratified acoustic medium having sharp thermocline separating regions with two different sound speeds. The medium is assumed to be quiescent except for the presence of a deterministic soliton internal wave traveling along the thermocline. The soliton wave shape and propagation speed are obtained from the K–dV equation and relevant shallow-water oceanic parameters. Soliton motion is incorporated in the necessary two-way acoustic propagation formalism with the single-scattering approximation. Depending on the acoustic frequency and source-array range, current results suggest there is ample opportunity for successful TRA deployment in the presence of soliton internal waves. [Work sponsored by the Office of Naval Research.]

2p TUE. PM

2pUW9. Implementation of a time-reversal cavity by confining a time-reversal mirror in a waveguide. Mickaël Tanter, Jean-Louis Thomas, and Mathias Fink (Lab. Ondes et Acoust., E.S.P.C.I., Univ. Paris VII, U.R.A C.N.R.S 1503, 10 rue Vauquelin 75231 Paris Cedex 05, France)

Time reversal is an original method to focus ultrasonic waves through heterogeneous lossless media. Nevertheless, to conduct a perfect time-reversal experiment, the wave field coming from an acoustic source should be recorded on a closed surface covered with transducers and surrounding the region of interest. This led to the concept of time-reversal cavity. In practice, its implementation is not easily realizable and the time-reversal operation is only performed on a restricted area named a time-reversal mirror. Some aberrating medium may diffract a part of the wavefield coming from the acoustic source outside the limited aperture of the time-reversal mirror. Hence, focusing quality may be strongly degraded. To approach the optimal focusing obtained with a time-reversal cavity, a channel bounded by totally reflecting walls was used to guide the wave field between the source and the time-reversal mirror. The previously lost information was then recovered to focus on the source location. Therefore, the combination of a time-reversal mirror with a waveguide can simulate a time-reversal cavity with only a small number of transducers. Experimental results give good agreement with a wave propagation model using a finite differences scheme. Both show that the focusing is now the same as that in a homogeneous medium.

2pUW10. One-channel time-reversal in a chaotic two-dimensional silicon cavity. Carsten Draeger and Mathias Fink (Lab. Ondes et Acoust., Univ. Paris 7, URA CNRS 1503, ESPCI, 10 rue Vauquelin, 75231 Paris Cedex 05, France)

Acoustic time-reversal experiments usually need large arrays of transducers. An element reduction normally decreases the aperture and thus degrades reversal or focusing quality. However, one can use reflections to increase artificially the transducer aperture. It is shown that in a reflecting cavity with negligible absorption, a time-reversal of a pointlike source can be performed using a single element. The wave field is measured by the transducer over a long period of time at a point inside the cavity, and then the time-reversed signal is reinjected at the same location. The wave field thus created forms wave fronts of the same shape as the initially emitted ones, but propagating in the opposite direction. They finally collapse at the location of the initial source, leading to excellent focusing which is not aperture limited. The difference between a perfect and a one-channel time-reversal can be quantified. A chaotic shape of the system guarantees a successful time-reversal; in some regular cavities, the procedure fails. One-channel time-reversal experiments have been successfully carried out using elastic waves in a silicon wafer. Finite-difference computations simulating a time-reversal of a two-dimensional scalar wave field complete this work.

2pUW11. Time reversal array focal zone predictions using simulations in the time domain. Ibrahim M. Hallaj, Steven G. Kargl (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105), and Ronald A. Roy (Boston Univ., Boston, MA 02215)

The use of finite-difference time-domain (FDTD) computer simulations for wave equation calculations has deep roots in the electromagnetic pulse propagation literature, where many of the modern techniques were developed. Results are presented from high-resolution simulations of the linear acoustic wave equation in inhomogeneous media with absorbing boundary conditions. The results address the focusing behavior of sparse ultrasonic arrays under various propagation conditions. Focal spot attributes and size are investigated in the presence of inhomogeneities, time-varying media, and multiple scatterers. Comparison with known benchmarks, such as the diffraction limit, show that useful results can be

obtained using FDTD calculations in medical applications, for example, where the frequencies and scales allow full-sized simulations in major organs. [Work supported by the Office of Naval Research.]

2pUW12. The D.O.R.T. method. Claire Prada (Lab. Ondes et Acoust., Univ. Paris 7, ESPCI, Paris 75005, France)

The D.O.R.T. method (Décomposition de l'Opérateur de Retournement Temporel) generalizes the principle of acoustic time reversal mirrors. It applies to detection and focusing with large arrays of transducers [Prada *et al.*, J. Acoust. Soc. Am. **99** (1996)]. A transducer array insonifying a scattering medium is characterized by its frequency-dependent transfer matrix $K(\omega)$. The time-reversal process is then described by time-reversal operator $K(\omega)K^*(\omega)$. The eigenvalue decomposition of this hermitian operator provides a classification and a localization of the scattering centers in the medium. The two main steps of the D.O.R.T. method are the measurement of the interelement impulse responses of each pair of elements in the array and the diagonalization of the corresponding time reversal operator at chosen frequencies. The method is efficient for the detection and selective focusing in an inhomogeneous multiple target medium. It also applies to the recognition of shells, in particular to the detection and characterization of circumferential waves propagating around the shell. The D.O.R.T. method and recent experiments of detection through channels and through inhomogeneous media will be presented.

2pUW13. Solving the inverse problem in nonlinear acoustics by backpropagation of the received signal. Claes M. Hedberg (Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78712, claes@mech.kth.se)

The incentive for this work was to determine when and how the information in a nonlinear signal disappears when propagating. This may be done by backpropagating the evolution and comparing the back-traced signal with the original one. The Burgers equation is investigated for both the nondissipative and dissipative cases. For a nondissipative evolution the wave completely coincides with the original one when shocks have not formed, while only part of it coincides after shock formation. Although the wave no longer contains the original information, except for the original signals with all derivatives continuous, parts of the wave theoretically carry all the information. These parts grow smaller and smaller, and thus a problem lies in the numerical accuracy which sometimes causes several different signals conform, and in that way the wave in practice has lost information. A semi-analytical model has been developed which automatically determines the distance to be backpropagated for modulated waves and gives one or more alternatives for the original signal. This may be used as a signal processing tool. [Work supported by TFR, the Swedish Research Council for Engineering Sciences.]

2pUW14. Three-dimensional modeling of a time reversal mirror to image objects buried in an inhomogeneous seabed. Altan Turgut (Naval Res. Lab., Acoust. Div., Washington, DC 20375)

The time reversal mirror (TRM) imaging and self-focusing concept was tested numerically by using a newly developed 3-D full-wave propagation model and a stochastic model of bottom heterogeneity. The 3-D propagation model was used to calculate pulse scattering from an attenuating inhomogeneous seabed and buried objects. Two-dimensional seabed roughness and three-dimensional sediment volume inhomogeneities were modeled by using the Von Karman type of spectra with proper statistical parameters which were obtained from stereo photography, sediment tomography, and chirp sonar measurements. A parametrical study was performed to identify proper choice of pulse shape, time gating, phase-screen backpropagation, and amplitude correction for selected distributions of multiple scatterers, burial depths, sediment layering, and statistical prop-

erties. Numerical simulations indicated that the TRM should perform well for detection of 3-D objects buried in the seabed. Finally, based on the results of the parametrical study, design criteria of a TRM field experiment are also discussed.

5:00

2pUW15. Imaging with eigenfunctions of the scattering operator.

Robert C. Waag (Dept. of Elec. Eng. and Radiol., Univ. of Rochester, Rochester, NY 14627), D.-L. (Donald) Liu (Univ. of Rochester, Rochester, NY 14627), T. Douglas Mast (Penn State Univ., University Park, PA 16802), and Adrian I. Nachman (Univ. of Rochester, Rochester, NY 14627)

An inverse scattering method that uses eigenfunctions of the scattering operator has been applied to obtain quantitative images of various scattering objects. The eigenfunctions of the scattering operator constitute a basis of an operator that is essentially equivalent to the time-reversal operator previously defined by others. The scattering objects span regions that are large compared to the wavelength of the acoustic illumination. Among the images are results for patterns of wires, homogeneous cylinders, and cylinders with wire inclusions. The eigenfunction method is compared to the method of filtered backpropagation implemented in each of two ways: numerical quadrature and fast Fourier transform. The results show the capability of the eigenfunction method to image objects with large ka . The results also show the eigenfunction method can be more efficient than the filtered backpropagation method when the scattering operator has few eigenvalues. [Work supported in part by NIH and USAMRMC.]

5:15

2pUW16. Experimental investigation of modal structure of field generated by a pulse-source in a laboratory hydroacoustic waveguide.

Sergei A. Egorychev, Vasily V. Kurin, Nikolai V. Pronchatov-Rubtsov (Univ. of Nizhny Novgorod, 23 Gagarin Ave., Nizhny Novgorod 603600, Russia), and Anatoly L. Virovlyansky (Inst. of Appl. Phys., Nizhny Novgorod 603600, Russia)

This work contains the results of an experimental investigation of the modal structure of an acoustic field in a model oceanic waveguide. Due to limited propagation range, it is impossible to resolve modes either by arrival times of modal pulses or by Doppler spectra of the received signal. The experiment was done in a laboratory two-layer waveguide where 32 normal waves had propagated at ranges up to 260 cm. As a result of intermodal dispersion, original short pulses radiated at 200 kHz were destroyed, and different parts of the received signal carried information about different modes forming the field in the waveguide. The beginning of the received pulsed signal was formed by the low-number modes, while the area near the back front of the pulse was constructed by high-number modes. Selection of modes was done using Doppler spectra of the received signal by a receiver uniformly moving along the waveguide and information about interference structure was read in the beginning, in the middle, and in the end of each received pulse. Such signal processing allowed to "impoverish" the original modal spectrum and, afterward, to analyze the modal structure of the acoustic field in the waveguide in more detail. [Work supported by RFBR Grant No. 95-02-04565.]

2p TUE. PM

Equipment Exhibition**See page 3030 for list of exhibitors**

WEDNESDAY MORNING, 18 JUNE 1997

ROOM H (POSTERS); ROOMS G AND I (DEMONSTRATIONS), 8:30 TO 11:30 A.M.

Session 3aAA**Architectural Acoustics: Acoustics and Sound Reinforcement for Large Audience Spaces
Posters and Demonstrations**

Robert C. Coffeen, Cochair
9214 Manor Road, Leawood, Kansas 66206

Jason T. Weissenburger, Cochair
Engineering Dynamics International, 8420 Delmar Boulevard, St. Louis, Missouri 63124

Edward Logsdon, Cochair
David L. Adams Associates, Inc., Denver, Colorado 80211

Contributed Poster Papers

All posters will be on display from 8:30 a.m. to 11:30 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:30 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 11:30 a.m.

This session will relate to room acoustics, noise control, and sound reinforcement considerations and solutions for large audience spaces such as arenas, stadiums, exhibit halls, large outdoor or partially enclosed music performance spaces, gymnasiums, field houses, large auditoriums, and other venues that seat in excess of about 4000 persons. Facilities to be described and subjects to be discussed in poster format will include papers 3aAA1 to 3aAA18 plus other venues and subjects described by contributed posters received prior to the submission deadline of 10 June 1997. Two demonstration papers, 3aAA19 and 3aAA20, will also be presented.

- 3aAA1. Great Woods Center for the Performing Arts, Mansfield, Massachusetts
Cavanaugh Tocci Associates, Inc., Consultants
- 3aAA2. Marine Midland Arena, Buffalo, New York
Acoustic Dimensions, Consultants
- 3aAA3. Utah Valley State College Event Center, Orem, Utah
David L. Adams, Consultants
- 3aAA4. Tokyo International Forum, Tokyo, Japan
Jaffe Holden Scarbrough, Consultants
- 3aAA5. Concord Pavilion, Concord, California
Jaffe Holden Scarbrough, Consultants
- 3aAA6. Marine Messe, Fukuoka City, Fukuoka Prefecture, Japan
Nagata Acoustics, Inc., Consultants
- 3aAA7. Oita Convention Center, Beppu City, Oita Prefecture, Japan
Nagata Acoustics, Inc., Consultants
- 3aAA8. Nashville Arena, Nashville, Tennessee
Wrightson, Johnson, Haddon & Williams, Inc., Consultants

- 3aAA9. The Rose Garden Arena, Portland, Oregon
Wrightson, Johnson, Haddon & Williams, Inc., Consultants
- 3aAA10. Speed Skating Arena for 1998 Winter Olympic Games, Nagano, Japan
Coffeen Fricke & Associates and Kajima Construction Technology Institute, Consultants
- 3aAA11. Kiel Center Arena, St. Louis, Missouri
Coffeen Fricke & Associates, Consultants
- 3aAA12. The Astrodome, Houston, Texas
Randorff & Associates, Consultants
- 3aAA13. Miller Outdoor Theatre, Houston, Texas
Randorff & Associates, Consultants
- 3aAA14. Coors Field, Denver, Colorado
Acentech, Consultants
- 3aAA15. The Pyramid, Memphis, Tennessee
Acentech, Consultants
- 3aAA16. B. C. Place Stadium, Vancouver, British Columbia
Acentech, Consultants
- 3aAA17. Contrasts between acoustics of arenas and closed stadiums
Pelton Marsh Kinsella, Consultants
- 3aAA18. Using multiple driver arrays for directivity control in arenas and stadiums
Pelton Marsh Kinsella, Consultants

Contributed Demonstration Papers

- 3aAA19. Speech intelligibility simulator using real time convolution to demonstrate the effects of various parameters on speech intelligibility
John van der Werff (Peutz & Associes B.V., Postbus 66, 6585ZH Mook, The Netherlands)
- 3aAA20. Improved speech intelligibility in large reverberant environments using DSP-controlled "steerable" loudspeaker line-array system
Gerrit Duran and Gert de Vries (Duran Audio B.V., P.O. Box 2050, 5300CB Zaltbommel, The Netherlands)

WEDNESDAY MORNING, 18 JUNE 1997

ROOM O, 8:15 TO 11:45 A.M.

Session 3aEA

Engineering Acoustics: New Transducer Materials

Thomas R. Howarth, Chair
Naval Research Laboratory, Code 7135, 4555 Overlook Avenue, S.W., Washington, DC 20375-5350

Chair's Introduction—8:15

Invited Papers

8:20

3aEA1. Piezogran™: High sensitivity/low-cost composites for wide aperture arrays. Todd L. Jessen (Mech. of Mater., US Naval Res. Lab., Washington, DC 20375-5343, jessen@anvil.nrl.navy.mil), Manfred Kahn, and Mark T. Chase (Potomac Res. International, Fairfax, VA)

A shift in US Navy emphasis from blue water to littoral regions has resulted in a need for improved acoustic devices. Wide aperture arrays will be hull-mounted, thereby making device weight critical. The Naval Research Laboratory has developed a unique, low-weight composite, transducer design, Piezogran™, to address this need. Piezograns consist of discrete, polycrystalline piezoelectric-ceramic elements dispersed in an epoxy matrix between two pressure plates. These plates make a direct mechanical as well as electrical contact to the piezoelectric elements. This design approach uses a 1-3 composite structure for high-gain operation with the flexibility of variable element spacing and ceramic volume fraction while achieving both low tooling and fabrication costs. Independent tests of 3.8-cm diameter Piezogran™ prototypes have yielded high hydrostatic response ($d_h > 150$ pC/N) and capacitance of 1.18 nF. Ten-centimeter-square array panels have also been assembled and are currently undergoing testing. This presentation will focus on fabrication and assembly of the Piezogran™ transducers as well as material properties.

3aEA2. Application of piezoelectric composites and net-shape PZT ceramics as acoustic sensors and actuators. C. Near, R. Gentilman, B. Pazol, D. Fiore, W. Serwatka, H. Pham-Nguyen, K. Markowski, P. McGuire, and L. Bowen (Mater. Systems, Inc., 521 Great Rd., Littleton, MA 01460)

A ceramic injection molding process is used to custom manufacture piezoelectric composites for both sensors and actuators with operational frequencies from 100 Hz and 10 MHz. These piezoelectric composites can be specifically tailored to couple acoustic energy into and from a wide variety of mediums including air, water, biological tissue, polymers, composites, and metals. Both 1-3 and 2-2 piezoelectric composites have been made with different lead zirconate-titanate (PZT) formulations, element dimensions, PZT volume fractions, and polymeric matrices. Piezoelectric composite transducers are easily conformed to various substructures, shaded to adjust beam patterns, and sectioned into arrays for imaging. Examples include face-plated SonoPanels (and stiff matrix 1-3 piezoelectric composites for underwater actuation, sensing, and imaging. SmartPanels, capable of both sensing and actuating, have been recently introduced for active surface control. These devices integrate 1-3 piezoelectric composite actuators and pressure sensors and net-shape molded PZT velocity sensors in a large-area, low-profile panel. Multiple net-shape molded accelerometer elements have been designed, fabricated, and tested. Both single and double layer 100×100-mm SmartPanels have been characterized for sensor sensitivities, actuator authority, surface displacement uniformity, and sensor-actuator coupling. Also, recent developments of injection molded PZT bender arrays for low-frequency air acoustic transducers will be presented. [Work supported by ONR and DARPA.]

3aEA3. Capped ceramic underwater sound projector. James F. Tressler, Wenwu Cao, Kenji Uchino, Robert E. Newnham (Mater. Res. Lab., Penn State Univ., University Park, PA 16802, jft104@psu.edu), and W. Jack Hughes (Penn State Univ., University Park, PA 16802)

The cymbal-type transducer developed at the Materials Research Laboratory at Penn State Univ., which is similar to the ring/shell flexensional transducer, is currently being investigated as a shallow-water (<250 m) sound projector for the 10–75 kHz frequency range. The standard size cymbal transducer is 12.7 mm in diameter and approximately 2 mm in total thickness. It consists of a poled lead zirconate-titanate (PZT) disk capped symmetrically on each face by identical thin metal caps. Each cap is shaped with a cavity on its inner surface so that an air space exists between the PZT disk and the top of the cap. The size of the air space, the thickness of the cap, and the type of cap material all influence the operating performance and allowable operating depth of the device. Parameters necessary to characterize an underwater projector such as resonance frequency, mechanical Q , electroacoustic efficiency, transmit voltage response (TVR), and source level will be presented. Experimental data will be compared with values calculated from the Ansys® finite-element program. Results from single elements as well as cymbal arrays (both potted and unpotted) will be shown. [This work was funded by the Office of Naval Research Contract No. N00014-96-1-1173.]

3aEA4. Relaxor-based single-crystal materials for ultrasonic transducer applications. Thomas R. Shrout, Seung-Eek Park (Mater. Res. Lab., Penn State Univ., University Park, PA 16802), Patrick D. Lopath, and Kirk K. Shung (Penn State Univ., University Park, PA 16802)

Relaxor ferroelectric single crystals of $\text{Pb}(\text{Zn}_{1/3}\text{Nb}_{2/3})\text{O}_3$ (PZN), $\text{Pb}(\text{Mg}_{1/3}\text{Nb}_{2/3})\text{O}_3$ (PMN) and their solid solutions with normal ferroelectric PbTiO_3 (PT) were investigated for ultrasonic transducer applications. Longitudinal coupling coefficients (k_{33}) as high as 94% and dielectric constants (K_3^T) in the range of 3500–6000 were achieved with low dielectric loss (<1%) using (001) oriented rhombohedral crystals of $(1-x)\text{PZN}-x\text{PT}$ and $(1-y)\text{PMN}-y\text{PT}$, where $x < 0.09$ and $y < 0.35$. Dicing direction as well as poling direction was critical for high coupling under laterally clamped conditions. Dicing parallel to the (001) yields 90% of laterally clamped coupling (k_{bar}) out of 94% longitudinal coupling (k_{33}) for PZN-8%PT. Thickness coupling (k_T) as high as 64% and low dielectric constant (K_3^T) < 600 with low loss (<1%) could be achieved using tetragonal crystals of $(1-x)\text{PZN}-x\text{PT}$ and $(1-y)\text{PMN}-y\text{PT}$, where $x > 0.1$ and $y > 0.4$. The performance gains associated with these ultra-high coupling coefficients and range of dielectric constants are evident in relations to broader bandwidth and electrical impedance matching. Using the KLM model, a pulse/echo response simulated a bandwidth of ~124% for a subdiced array element with a center frequency of 10 MHz. An optimized array design of the same geometry constructed of PZT 5H displayed 87% bandwidth. [Work supported by ONR, DARPA.]

3aEA5. Improved giant magnetostrictive materials: New and improved Terfenol-D. Jonathan D. Snodgrass, Frederick T. Calkins (ETREMA Products, Inc., 2500 N. Loop Dr., Ames, IA 50010), Marcelo J. Dapino, and Alison B. Flatau (Iowa State Univ., Ames, IA 50011)

The giant magnetostrictive terbium-iron alloy material, Terfenol-D, offers the highest strain of any magnetostrictive material at room temperature using practical drive fields. The increasing use of Terfenol-D in a wide variety of production devices and systems has driven the development of new, advanced manufacturing techniques in material production to provide greater strain performance, uniformity, rod size, and production capacity. As more applications for this smart material are developed, the demand for reliable design data to enable applications engineers to take advantage of Terfenol-D's superior properties has also grown. Recent developments in production techniques have made strides to fulfill needs in both areas. The ETREMA crystal growth (ECG) system, based on the modified Bridgman process, is capable of producing larger rod sizes (65-mm diameter and 175-mm length) with greater quality control, performance, material homogeneity, and uniformity. These improvements are quantified with material property data. Terfenol-D samples are characterized in a transducer under a wide variety of real-world operating conditions. Results show greatly improved consistency of material performance, in terms of material properties, from test to test.

Contributed Papers

10:15

3aEA6. A multilayer 1-3 piezoelectric composite underwater projector. Thomas Howarth (Naval Res. Lab., Code 7130, Washington, DC 20375-5350, howarth@nrl.navy.mil) and Robert Ting (Naval Undersea Warfare Ctr., Orlando, FL 32856-8337)

An underwater acoustic projector for synthetic aperture SONAR (SAS) application on an AUV has recently been developed. The transducer features a two layered 1-3 piezoelectric composite material stacked in mechanical series and electrically wired in parallel. This arrangement was selected in order to maximize the source level output. The transducer is resonant at 100 kHz but has been designed to deliver high sound-pressure levels for operation over the frequency range of 10 to 100 kHz. The center electrode of the monolithic 1-3 piezoelectric composite layers has been segmented to offer four individual elements such that combinations of the sectors offer the ability to access nine different apertures. The basic design of the piezoelectric composite material will be shown along with the methods used to accomplish the successful multilayering effect. The fabrication of the prototype array and the measured transmitting voltage responses (TVRs), source levels (SL), and beam patterns will also be presented. [Work supported by the NSWC/Coastal Systems Station.]

10:30

3aEA7. An affordable, acoustically transparent, two-dimensional, high-resolution, acoustic imaging array. Kim Benjamin, James Powers, Fred Nussbaum, and Gerald Stevens (Naval Undersea Warfare Ctr., New London Detachment, Newport, RI 02481)

The design, fabrication, and measured results for a prototype, two-dimensional, ultrasonic imaging array will be discussed. The active sensor consists of virtually three layers: (1) a piezoelectric 1-3 composite layer; (2) a flex-circuit component; and (3) a special z -axis conductive adhesive which joins them. The sensor construction utilizes electroplated, injection-molded, 1-3 piezoelectric composite as the active layer. Four acoustically thin, flex circuits provide the electrical connections required for addressing the 468 individual array elements. Of key importance to the fabrication process is the use of a composite, b -stage, adhesive film layer, which combines both conductive and nonconductive regions in a pattern-specific orientation within the plane comprising the bond line interface. The conductive regions of the adhesive film are registered with respect to the electroplated sections on both the flex circuit and the 1-3 piezoelectric composite substrate. The transducer array is a reciprocal device operable in both transmit and receive modes. Measured results will include individual element and array calibrations (i.e., transmit and receive responses, beam patterns, and relative phase). [Work supported by ONR 333, Kam Ng.]

10:45

3aEA8. 1-3 piezoelectric composite materials for ultrasonic transducer applications—A comprehensive theoretical and experimental study. Xuecang Geng and Q. M. Zhang (IMRL, Penn State Univ., University Park, PA 16802, zxg1@psuvm.psu.edu)

1-3 piezoelectric composites are widely used as the transduction materials for underwater and ultrasonic transducers. In order to guide the design of the composite material and the devices based on it, a dynamic model is developed to precisely predict the dependence of the thickness drive on various parameters. For example, the velocity of the longitudinal wave, the dispersion curves of the wave propagation in 1-3 piezoelectric composite, and the antiresonance frequency, as well as the electromechanical coupling coefficient of the fundamental thickness mode, can be exactly predicted. The effect of the loss of the polymer phase on the properties of the 1-3 piezoelectric composite is also included in this theory; as a consequence, the quality factor of the 1-3 piezoelectric composite and its dispersion can be described. The origin of the lateral modes in the 1-3

composite plate can also be elucidated from this model, which shows quite different behavior from the early models such as the stop-band edge resonance model. Parallel to the analytic model, a systematic experimental study was also conducted to examine the mode coupling and the mode evolution with the thickness of a 1-3 composite plate.

11:00

3aEA9. High power piezoelectric characteristics and loss mechanisms of PMN-PT based ceramics. YunHan Chen (IMRL, Penn State Univ., University Park, PA 16802, yxc139@psu.edu) and Jin-Soo Kim (Penn State Univ., University Park, PA 16802)

Application of ultrasonic transducers for high-power piezoelectric devices such as motors and actuators has been vigorously studied in recent years. Those devices are usually driven at a rather high vibration level. However, it was shown that a drastic change in the electromechanical properties, as well as a great amount of heat generation, occurred above certain vibration velocity due to the large electrical and mechanical loss. In this research, lead magnesium niobium oxide-lead titanium (PMN-PT) based material were prepared and studied. The PMN-based ceramics is intriguing due to their extremely high electro-mechanical coupling factors. Combined with lead titanium and modified by some dopants, the mechanical quality can be much improved. The properties under high level driving and loss mechanisms of PMN-PT with different dopants will be reported in this paper. [Work supported by ONR.]

11:15

3aEA10. Space-charge-enhanced, electric-field-induced strain responses in electrostrictive polyurethane elastomers. J. Su and Q. M. Zhang (Intercollege Mater. Res. Lab., Penn State Univ., University Park, PA 16802, jxs65@psuvm.psu.edu)

Polymers have become more and more attractive to acoustical science and technology. Recently, large electric-field-induced strain responses in segmented polyurethane elastomers have drawn the wide attention of many researchers due to their promising applications in sensor, actuator, and microelectromechanical system (MEMS) devices. Mechanisms of the large strain responses were investigated by studying elastic, dielectric, and electromechanical properties as functions of temperature, frequency, and sample thickness. The effects of the injected charges on the large electric-field-induced strain coefficient in electrostrictive polyurethane thin films were also studied using the thermally stimulated discharging current (TSDC) technique. It was found that the contributions to the field-induced strain responses are associated with intrinsic electrostriction and the Maxwell stress effect. These effects can be enhanced by distribution of the injected charges, which results in a modified nonuniform internal electric field profile in the region near the polymer-electrode interface. In addition, transitional phenomena related to the molecular motions of soft segments and hard segments of the segmented polyurethane play an important role. The experimental results will be presented and discussed. The achievements will supply a guideline to develop new and better polymer materials for electromechanical and acoustical applications.

11:30

3aEA11. Dynamic performance of 1-3 type piezoelectric ceramic composites. Yushieh Ma (Inst. of Appl. Geophys., Natl. Chung Cheng Univ., Ming-Hsiung, Chia-Yi, Taiwan, ROC)

Expressions of the frequency-dependent displacement for the 1-3 type piezoelectric ceramic composite unit cell are derived using Navier's equation in the cylindrical coordinates. It does not require an assumed form for the displacement field before the derivation. The formalism also covers that of the static cases. The thickness-drive vibration problem is analyzed under a uniaxial as well as a hydrostatic load. The frequency-dependent

displacement, which is along the axis of the unit cell, depends on the aspect ratio, the relative content of both the discrete (ceramics) and continuous (matrix) phases, and also on the boundary conditions of both ends of the unit cell. In addition, Poisson's ratio plays an important role on the dynamic response of the displacement. Numerical results are presented as normalized stress amplification parameters as a function of frequencies.

The frequency-dependent hydrostatic piezoelectric constant and figure of merit are also computed, and they are depicted in figures under various configurations of the composites. From the computed resonance frequencies of the unit cell and the information provided, the 1-3 piezoelectric ceramic composites can be tailored to best suit for their medical and underwater applications over a wide frequency range.

WEDNESDAY MORNING, 18 JUNE 1997

FOYER OUTSIDE ROOMS O AND N, 10:45 A.M. TO 12:15 P.M.

Session 3aED

Education in Acoustics: Undergraduate Research Poster Session

Victor W. Sparrow, Chair

Graduate Program in Acoustics, Pennsylvania State University, 157 Hammond Building, University Park, Pennsylvania 16802

Contributed Papers

All authors will be at their posters from 10:45 a.m. to 12:15 p.m.

Posters will be on display from 10:45 a.m. to 3:30 p.m.

3aED1. On-line contactless measurement of the elastic modulus of a moving sheet by ultrasonic method. Mazen Khoury (E.F.P.G, Lab. de Genie des Procédés Papetiers, UMR 5518 CNRS, Domaine Univ. 461, rue de la Papeterie, 38402 Saint Martin d'Heres, France), Andre Schroder (Elfatochem-service UMR, 92303 Levallois Perret, France), and Guy E. Tourtollot (Domaine Univ., 38044 Grenoble, France)

There is still a strong interest in the paper industry for a contactless method that would allow the measurement of the elastic modulus of paper and of the anisotropy of this modulus in the sheet plane. A contactless method is presented that enables the measurement of the transmission coefficient of sound through a paper sheet, in the frequency range 200–800 kHz, for incidence angles between -66° and $+66^\circ$. These measurements can be performed on a moving sheet at velocities in the same range as those observed on a paper machine. From the state equation of plate waves, the possibility to calculate the elastic modulus of paper was demonstrated by using only the equation of the first antisymmetric mode, without neglecting any term of it, contrary to what is commonly used in the literature. This model fits very well with measurements made and leads to consideration of the development of an on-line version for use on industrial paper machines.

3aED2. Effect of acoustic cavitation on genetically engineered bacteria. Sylvia D. S. Kwakye and E. Carr Everbach (Dept. of Eng., Swarthmore College, Swarthmore, PA 19081)

The interaction of a complex acoustic cavitation field and suspended cells *in vitro* has been explored by many previous investigators. Of greatest interest is the effect of collapsing cavitation bubbles on the membrane structure and DNA of cells. By employing *E. coli* bacteria that have been genetically engineered to emit visible light when they are stressed in prescribed ways, the factors influencing bubble-cell interactions can be better explored. These factors include insonation frequency (1–5 MHz), duty cycle, PRF, initial bubble size distribution, and fluid viscosity. Results from this ongoing study will be presented. [Work supported by NSF.]

3aED3. Vowel production by a child with developmental apraxia of speech. Christina F. Famoso (Dept. of Speech, Commun. Sci., & Theatre, St. John's Univ., 8000 Utopia Pkwy., Jamaica, NY 11439) and Fredericka Bell-Berti (St. John's Univ., Jamaica, NY 11439 and Haskins Labs., New Haven, CT 06511)

Vowel production is only rarely considered in discussions of articulation disorders, and the most widely used clinical tests of articulation provide only minimal evaluation of vowel articulation. Indeed, even the literature devoted to the mastery of vowel production by children without articulation disorder is limited, although several recent studies have demonstrated that vowel acquisition does not occur as easily or as early as had been suggested. The subject for the study reported here is an 8-year-old boy diagnosed with developmental apraxia of speech who has been treated for his speech difficulties since the age of 2 years 10 months. At this time his speech is intelligible, although still distorted. This study will evaluate his vowel production in two ways. First, listeners will identify words from minimally contrastive word sets; this to allow us to examine his ability to produce distinctively different and identifiable English vowels. Second, measurements will be made of vowel duration for English long-short vowel pairs and formant frequency for the ten English simple vowels. The latter will be compared to the children's formant frequency data reported by Peterson and Barney [G. E. Peterson and H. L. Barney, *J. Acoust. Soc. Am.* **24**, 175–184 (1952)].

3aED4. Individual differences of transient evoked otoacoustic emissions: Effects of gender and ear canal volume. Lisa Jayne Romano and Tina Jupiter (Dept. of Speech, Commun. Sci., & Theatre, St. John's Univ., 8000 Utopia Pkwy., Jamaica, NY 11439)

Transient evoked otoacoustic emissions (TEOAE) are an objective means for measuring cochlear hearing loss in newborn infants. TEOAEs are low-intensity sounds produced by the normal cochlea that are elicited by acoustic stimuli and then recorded in the outer ear canal. Protocols are being developed for clinical use, and pass/fail criteria are usually based on the size of the emission. The common practice is to wait until an emission

of 6 dB is achieved before recording a positive result. Within subject, stimulus-related characteristics may influence the results. Quantifying the sources of variability can effect the analysis of the results. This study will investigate the effects of gender and ear canal volume on TEOAE size. Five male and five female young adult volunteers will serve as subjects. None of the volunteers will have a history of otologic disease or use of medication. All subjects will have normal hearing and normal middle ear function. Pure-tone audiometry and single frequency immittance testing (226 Hz probe tone) will be made. TEOAE will be measured using ILO88 hardware controlled by an IBM PC.

3aED5. Temporal measures of tongue twisters produced by children. Patricia N. Schwartz (Dept. of Speech, Commun. Sci., & Theatre, St. John's Univ., 8000 Utopia Pkwy., Jamaica, NY 11439) and Nancy S. McGarr (St. John's Univ., Jamaica, NY 11439)

The production of tongue twisters has been used to study the planning and organization of speech production in both adults and children. In general, temporal measures have differed as a function of age with older children (10+ yr) showing more adult-like productions than much younger children. In this study, five children, 8 years of age, were audio-tape recorded producing tongue twisters and control stimuli following Smith [B. Smith, *Appl. Psycholinguist.* **11**, 131–144 (1990)]. Phonetic transcriptions of the children's productions will be presented. Measures of VOT, stop consonant closure, and vowel duration will also be presented. Preliminary analysis of VOT, stop consonant closure, and vowel duration will also be presented. Preliminary analysis of VOT measures show that the twister VOT means were longer than the corresponding control VOT means, as have been reported for adults. These results will be compared to the data reported by Smith.

3aED6. Design and development of PC-IMAT: Teaching strategies for acoustical oceanography. Jacob A. Foret, Murray S. Korman (Dept. of Phys., U.S. Naval Acad., Annapolis, MD 21402), John W. Schuler (Armstrong Lab., Brooks AFB, TX 78235), and Eleanor Holmes (Integrated Performance Decisions, Inc., Arlington, VA 22202)

PC-IMAT (Interactive Multisensor Analysis Training) was developed in 1994 by NPRDC as a set of software tools to satisfy the initial need — to enhance the training of aviation ASW operators. While Navy training opportunities have decreased (reductions in deployments, encounters, and shore-based training facilities), training requirements have increased. The PC-IMAT 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. Students 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). Recent research on “scientific visualization” (to enhance comprehension and retention) and modeling from student feedback will be used to help develop and evaluate existing training materials. Questions arise naturally. What are the issues of the educational platforms to successful learning? What is the best computer utilization so that there is efficient evolution in the distributed training, whose contents have been reviewed, overhauled, and accepted by experts in the field? Finally, shipboard training is limited by the senior person's knowledge.

3aED7. A large excursion planar magnetic loudspeaker. Kelton D. Mitchell (Appl. Res. Lab., Penn State Univ., State College, PA 16803)

Planar magnetic loudspeakers with wide diaphragms, such as woofers and midranges, are usually constructed using vertical rows of magnets with each row alternating between north and south poles. Narrow planar magnetic loudspeakers, usually called ribbons, are constructed with the diaphragm directly between the north and south poles of the magnet so it is totally immersed in the magnetic field. In this project, a loudspeaker with a $3\frac{1}{2}$ -in.-wide diaphragm was designed and constructed with the ribbon configuration since the magnetic field is more linear with diaphragm ex-

cursion and is distributed more evenly over the surface of the diaphragm. A smaller emphasis was placed on the efficiency so that the speaker would have a diaphragm surface area and excursion that is comparable with cone-type speakers.

3aED8. Automated reverberation time experiments employing artificially generated noise. Timothy W. Lancey and Jon Marine (California State Univ., Fullerton, CA 92634)

Using an IBM compatible PC, an inexpensive and simple method was developed to conduct acoustic absorption measurements in a reverberant chamber. The technique provided random noise distributed around 1/3 octave band center frequencies from 200 to 1000 Hz, through a D/A converter, to speakers within the chamber. Turbo-Pascal software was written to create a random noise file for each of the 1/3 octave bands. As a consequence, while a random noise signal was sent by the computer, that signal was identical for each experiment, for a specific 1/3 octave band. The noise signal, of 3-s duration, was programmed to increase to maximum intensity, and decrease to zero over a period of 25 μ s, eliminating a steplike transient to the dynamic system at termination. The decaying noise was sampled 1000 times per second using an A/D converter, and the decay rate calculated. This technique was initiated with a keystroke and produced a sound-pressure level versus time plot file, allowing determination of the reverberation time and absorption coefficient of the chamber.

3aED9. Abstract withdrawn.

3aED10. Fatigue monitoring of a thin laminated composite using nonlinearity parameter and ultrasonic C-scan. Corinne M. Darvennes, Bryant K. Johnson, and Michael J. Woodward (Dept. of Mech. Eng., Tennessee Tech. Univ., Cookeville, TN 38505, cmd2221@mttech.edu)

Previous research [X. Hou and C. Darvennes, *J. Acoust. Soc. Am.* **100**, 2568(A) (1996)] has shown that fatigue cycles can affect the nonlinearity parameter of laminated composites. In this research a thin laminated composite plate is put under fatigue until it fails. The fatigue test is stopped at regular intervals, when a C-scan picture is taken and the material nonlinearity is measured. To determine the nonlinearity parameter, the growth of a second harmonic is measured in the material. A monochromatic ultrasonic signal is sent into the sample via a contact transducer placed on the top surface. The growth of the second harmonic is recorded, with a second transducer placed on the bottom face, as the amplitude of the input signal is gradually increased. The nonlinearity parameter is then plotted as a function of the number of fatigue cycles. The C-scan images provide a visual clue about changes in the material structure that can be correlated to changes in nonlinearity. Hopefully this method will provide a means of predicting failure in advance. [Work supported by TTU Research Planning Grant program.]

3aED11. Underwater sound radiation from a spherical bubble cloud encapsulated by a thin spherical shell. Sandra Lee Koslowski and Murray S. Korman (Dept. of Phys., U.S. Naval Acad., Annapolis, MD 21402)

Measurements of sound transmission versus frequency are performed (in the Rickover Hall Hydrodynamics Tow Tank Facility) for a submerged spherical bubble cloud that is encapsulated by a thin spherical shell. The bubble cloud is driven by a miniature spherical PZT transducer unit that is suspended in the center. A theoretical model has been developed which predicts the resonant frequencies for the case where the shell thickness has

negligible effect. Therefore, the resonances will depend on the radius of the cloud and the sound speed in the bubbly media (which is a function of the air volume void fraction). A shell of urethane elastomer is molded as a single piece by using the "lost wax technique." Hemispherical molds are used. The first set is for making the "inner" spherical wax ball. The second set forms the "outer" boundary of the casting for the spherical shell with filling ports on the north and south poles. The shell's inner wax material is then melted out to complete the process. The shell is nearly acoustically transparent since $\rho = 1.03 \text{ g/cm}^3$, $c_{\text{long}} \sim 1450 \text{ m/s}$ and $c_{\text{tran}} \sim 70 \text{ m/s}$. The bubbly fluid consists of castor oil that has been whipped in a microblender.

WEDNESDAY MORNING, 18 JUNE 1997

ROOM E, 8:30 TO 11:05 A.M.

Session 3aMU

Musical Acoustics and Education in Acoustics: Role of Musical Acoustics in Teaching Acoustics, Mathematics and Engineering I

Robert D. Collier, Cochair
Creare, Inc., Etna Road, Hanover, New Hampshire 03755

Ronald A. Roy, Cochair
Department of Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Chair's Introduction—8:30

Invited Papers

8:35

3aMU1. Musical acoustics: A bridge between physics and music. Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115)

For more than 30 years, I have taught elementary courses in acoustics to students in the liberal arts. Although these courses were designed especially for music students, the subject of musical acoustics has broad appeal to students with a wide variety of backgrounds and with a broad spectrum of professional and vocational plans. One can easily argue that musical acoustics is "good physics," in that it applies the principles of mechanics, wave motion, thermodynamics, optics, and electricity and magnetism to understanding musical instruments and the way in which musical sound is recorded and reproduced. I have written two textbooks, two reprint books, a book of acoustics laboratory experiments, a book of demonstration experiments, and other suggestions for teaching the science of sound. Some of my experiences and ideas for teaching physics through musical acoustics will be shared.

9:05

3aMU2. Music: A vehicle for science—Workshop for teachers. Uwe J. Hansen (Dept. of Phys., Indiana State Univ., Terre Haute, IN 47809) and Thomas D. Rossing (Northern Illinois Univ., DeKalb, IL 60115)

With ASA technical initiative support, approximately ten acoustics workshops for teachers have been conducted by members of the technical committee on musical acoustics in the past six years. Workshop content emphasizes laboratory and demonstration skills for high school and college level teachers. The workshops for elementary teachers have been more concerned with generating excitement for science. Music is an ideal vehicle to accomplish that. Elementary teachers in the workshops participate in monochord construction and receive training in the use of a number of pieces of equipment which they take back to the classroom. Basic wave properties are illustrated and related to string harmonics and musical intervals. Workshop content and training approaches will be discussed.

9:35

3aMU3. Mathemusical connections: The role of musical acoustics in teaching mathematics. Stephen G. Hinthorne (Dept. of Math., Principia College, Elmhurst, IL 60120)

The purpose of this established liberal arts course, *A survey of mathematics*, is to change student attitudes towards mathematics through an innovative introduction to musical acoustics. This approach is particularly effective for those students who may not like mathematics or may have some phobia towards it. The topics and structure of the course are built around the quadrivium of mathematics attributed to Pythagoras. This includes four disciplines: arithmetic (today's number theory); music; geometry; and astronomy. Added to these topics are: numeration systems; algebra; finance; and probability and statistics. Sequences of musical scale ratios associated with lengths of vibrating strings and the frequencies of related sounds in the "major" scale were invented by Pythagoras, Claudius Ptolemy, and musical theorists of the sixteenth-century Renaissance. The study of the development of these three

scales and their comparison, together with Mersenne's law of vibrating strings and some work with graphing sine functions, provide students with their first look at the mathematics of music. Music and acoustics provide the vehicle for students to gain a better understanding of the often troublesome areas of ratios, proportions, formulas, and graphs. This paper includes the course outline and structure, examples of topics and assignments, and student response.

10:05

3aMU4. Musical acoustics demonstrations play a role in teaching underwater acoustics and sonar. Murray S. Korman and Samuel A. Elder (Dept. of Phys., U. S. Naval Acad., Annapolis, MD 21402)

Live demonstrations with apparatus will be used to present a variety of effects in physics that involve musical acoustics. Musical acoustics has been taught as part of the physics majors course, called Acoustics, for over 30 years now. (Seniors take this four-credit elective in the fall semester, which includes a laboratory.) Also, many classroom demonstrations involving musical acoustics have been done in the Underwater Acoustics and Sonar course. (This three-credit course is taken by Oceanography and General Science Majors, in separate semesters, in their senior year.) These demos help motivate certain topics and help keep the student interested. They are fun and contribute to a lively class participation. It is difficult for the student to forget about the detection process when the signals and background noise become impromptu music that they have participated in generating. Fourier analysis of the voice or musical instruments suggest a structure that can help motivate what is meant by a sonar ping or the "orchestration" of background noise. Student involvement in the modal analysis of a guitar surface is certainly more effective and intriguing than a study of a submarine hull. A short video of some of the projects will supplement this presentation.

Contributed Paper

10:35

3aMU5. A course in acoustics of musical instruments. Courtney B. Burroughs (Graduate Program in Acoust., Penn State Univ., State College, PA 16802)

Graduate studies and a career in research and/or teaching in acoustics are often the result of an interest in acoustics started with the playing of a musical instrument. In teaching and research, there is a need for the application of acoustic theory to real sources, which often cannot be treated effectively with academic idealizations. Taking advantage of the interest in

and familiarity with musical instruments possessed by most students in acoustics, and the need for application of acoustic theory to real sources, a course in the acoustics of musical instruments has been developed for graduate students in acoustics. This source is not a course in musical acoustics designed for musicians. Acoustic theory and mathematical models are applied to the analyses of the mechanisms of sound generation and radiation by musical instruments. A description of and philosophy behind this course are presented. The instruments treated are discussed. The principles revealed by each instrument are outlined.

10:50–11:05

Discussion

3a WED. AM

WEDNESDAY MORNING, 18 JUNE 1997

ROOM P, 8:00 TO 10:15 A.M.

Session 3aNSa

Noise: Combined Exposure to Noise and Other Hazards

William J. Murphy, Chair

Bioacoustics and Occupational Vibration Section, National Institute for Occupational Safety and Health, MS C-27, 4676 Columbia Parkway, Cincinnati, Ohio 45226-1998

Chair's Introduction—8:00

Invited Papers

8:10

3aNSa1. Effect of noise and solvents on hearing: Findings from three field studies. Thais C. Morata and John R. Franks (Bioacoust. and Occup. Vib. Sect., Natl. Inst. for Occup. Safety and Health, MS C-27, 4676 Columbia Pkwy., Cincinnati, OH 45226-1998)

A holistic approach to studying the workplace as a combination of physical, chemical, biological, and organizational factors is receiving growing attention. This approach includes initiatives to investigate the potential interaction between noise and chemicals. NIOSH has conducted three cross-sectional studies on the effects of solvents on hearing, alone or in combination with noise. Audiometric thresholds were compared with ISO 1999 estimates, and the prevalence, risk ratios, and predicted probability of developing a hearing loss were calculated. In all of the investigations, solvent mixtures which had toluene as a main component were found

to affect the hearing of workers exposed to low noise levels. Considering the multiplicity of chemicals that are used occupationally and evidence that some are ototoxic, it is conceivable that numerous populations are being underserved regarding hearing loss prevention since the permissible exposure levels for chemicals do not take into account their potentiation of hearing loss. Workers who are exposed to A-weighted noise levels below 85 dB time-weighted average are not required to be included in hearing conservation programs. Furthermore, methods currently used in hearing conservation (e.g., hearing protectors and noise control) may be insufficient, or even inappropriate, for workers exposed to both chemicals and noise.

8:35

3aNSa2. A review of the effects of exposure to noise in combination with organic solvents and metals. William J. Murphy, John R. Franks, and Thais C. Morata (Bioacoust. and Occup. Vib. Sect., Natl. Inst. for Occup. Safety and Health, MS C-27, 4676 Columbia Pkwy., Cincinnati, OH 45226-1998)

The effect of noise exposure in combination with other ototraumatic agents (e.g., solvents, metals, and pharmacologic agents) is an emerging area of concern for the development of comprehensive standards for hearing loss prevention. Noise exposure primarily damages the organ of Corti, while solvent exposures exhibit both central and peripheral effects in rats. Organic solvent-induced hearing loss presents impairment which is similar to noise-induced hearing loss in humans. Exposures to metals tend to affect the central nervous system. The effects of lead exposure on auditory brainstem response exhibit an increased threshold across all frequencies and increased latencies in humans. However, other metals, such as zinc and copper, do not produce similar effects. Combined exposures of noise with metals or solvents can produce nonadditive interactions both in humans and animals. This paper reviews the effects of exposure to commonly used organic solvents and metals found in occupational settings. The interactions of several solvents and metals will be classified according to their effects on the peripheral or central auditory system both in laboratory animals and humans.

9:00

3aNSa3. The combined effects of ototoxic drugs and noise. Leonard P. Rybak (Dept. of Surgery, SIU School of Medicine, 800 N. Rutledge, P.O. Box 19230, Springfield, IL 62794-1312)

Cisplatin is a widely used chemotherapeutic drug which causes ototoxicity. Previous studies from this laboratory using the rat have shown that cisplatin ototoxicity may be mediated by the production of free oxygen radicals. The ototoxic effects of cisplatin can be reduced by treatment with reducing agents such as diethyldithiocarbamate, methylthiobenzoic acid, ebselen, D-methionine, or sodium thiosulfate. The injury to the inner ear caused by cisplatin is potentiated by noise exposure. The greatest degree of cochlear damage in the guinea pig was observed when noise exposure precedes the administration of cisplatin [G. Laurell, Ann. Otol. 101, 969-976 (1992)]. These findings could be explained by an increase in oxidative stress produced by noise exposure, which could predispose to further injury by subsequent cisplatin administration. [Work supported by NIH.]

9:25

3aNSa4. Pharmacological intervention with noise-induced hearing loss. Donald Henderson and Bo-Hua Hu (Ctr. for Hearing and Deafness, SUNY, 215 Parker Hall, Buffalo, NY 14214, donaldhe@acsu.buffalo.edu)

High-level noise has the capability of damaging the cochlea via several mechanisms, including the generation of reactive oxygen intermediaries with their capacity of cytotoxicity. Several laboratories have reported that prior exposure to conditioning noise (moderate levels of noise for several days) can render the cochlea more resistant to future traumatic exposures. This lab in collaboration with the Albert Einstein College of Medicine, reported that antioxidant enzymes: catalase, glutathione reductase, and γ -glutamyl cysteine synthase, are upregulated with noise exposure—the greatest increase being found in animals exposed to a traumatic noise with previous conditioning noise exposure. This increase, with conditioning noise, suggests the antioxidant system may be involved in protection from noise exposures. To better understand the role of antioxidants, a second set of experiments was conducted that manipulated their level. The applications of R-N6-phenylisopropyladenosine (R-PIA) (a drug that upregulates the antioxidant system) decreases the effects of noise; and conversely, applications of buthione-sulfoximine (a drug that suppresses the antioxidant system) increases the effects of noise. These results will be discussed in terms of the balance of free radicals/antioxidants as a common factor in noise and other ototoxic reactions. Also, the R-PIA studies will be discussed in terms of potential therapeutic value.

9:50

3aNSa5. Influences of aging and calcium-channel blockers on noise-induced hearing loss. Flint A. Boettcher (Dept. of Otolaryngol., Medical Univ. of South Carolina, 171 Ashley Ave., Charleston, SC 29425-2242)

Two issues will be reviewed: (1) combined effects of noise and aging on hearing and (2) protective effects of calcium-channel blockers (CCBs). There are no empirical data on the auditory effects of noise exposure and the aging process in humans. The only sources of information on combined effects of noise and age have been epidemiological studies. In recent experiments, gerbils were exposed monaurally at middle age to an intense tone which caused approximately 20 dB of noise-induced permanent threshold shift (NIPTS). At 36 months (lifespan of the gerbil), the threshold in the unexposed ear was considered the age-related loss and that in the exposed ear was considered the loss caused by aging plus noise. Data were examined in terms of published models which attempt to discriminate between NIPTS and presbycusis in older individuals. Most models significantly overestimated the noise-induced component of the hearing loss in the aged subjects. In the second series of experiments, gerbils were exposed to an intense noise while concurrently receiving CCBs or placebos. Nimodipine (a potent CCB) reduced the NIPTS. Furthermore, epidemiological data suggest that aged women taking CCBs have hearing levels 10–12 dB better than average. [Work supported by NIH-NIDCD and NOHR.]

Session 3aNSb**Noise: Product Noise Labeling Issues**

Julia D. Royster, Chair

*Environmental Noise Consultants, Inc., P.O. Box 30698, Raleigh, North Carolina 27622-0698****Invited Papers*****10:20****3aNSb1. On the need for noise labeling standards.** Paul D. Schomer (U.S. Army Construction Eng. Res. Labs., P.O. Box 9005, Champaign, IL 61826-9005)

Currently, "quiet" is being advertised as an attribute of certain products. For example, one can purchase "quiet" dishwashers. Other products have had their own labels for years. For example, some home bathroom fans have carried a rating in sones. But no standard, uniform method exists to label such products. This paper discusses the need for such standards and some of the attributes such standards might possess.

10:35**3aNSb2. Product noise labeling and public information: Recommendations from the 1991 National Noise Strategy Conference.** Elliott H. Berger (E-A-R/Aearo Co., 7911 Zionsville Rd., Indianapolis, IN 46268-1657)

In 1990, a symposium "Combating Noise in the '90s: A National Strategy for the United States," was initiated by Congressman Richard J. Durbin of Illinois and funded by the American Speech-Language-Hearing Association in cooperation with co-sponsoring organizations, the Acoustical Society of America and the American Academy of Otolaryngology—Head and Neck Surgery. The key participants were divided into nine working groups covering topics dealing with the effects of noise, the control of noise, and education and regulation. Two working groups, VI *Consumer Noise Sources and Hearing Protection*, and VII *Public Information and Education*, examined issues relating to the labeling of noisy products and also noise-reducing products, how to inform the public about the need for such ratings, and the meaning and application of such numbers. The products requiring rating and/or classification that were examined were categorized as guns, music reproduction equipment, equipment and appliances, toys, hearing protection, and building spaces. The author, who was the chair of Working Group VI, will review the recommendations of both working groups VI and VII in the context of current events.

10:50**3aNSb3. On-product noise labeling: Design and efficacy.** Harry L. Snyder (Snyder and Assoc., Inc., P.O. Box 7028, Sunset Beach, NC 28470)

Labeling issues pertinent to both nonhazardous and hazardous noise levels are considered and discussed. Among these are pertinent ANSI and engineering standards for warning label design; recommended label content and design criteria; format and presentation variables that affect the likelihood of noticing, reading, and complying with the labels; the effect of user cost of compliance (e.g., comfort) on the efficacy of labels; and conditions that indicate high versus low likelihood of label effectiveness. Particular emphasis will be placed on the experimental literature that has addressed warning label efficacy.

11:15**3aNSb4. HPD labeling issues.** John G. Casali and Gary S. Robinson (Dept. of Industrial and Systems Eng., Virginia Tech., Blacksburg, VA 24061-0118)

Issues related to label content are analyzed from a human factors perspective. Current labeling requirements are critiqued, and the need for additional information or the presentation of existing information in a different form is examined. Some of the specific questions addressed include the following. Who are the users and how do the information requirements of the individual selecting the hearing protection devices (HPDs) for the employer differ from those of the end user/wearer? What are the responsibilities of the government regulatory agencies (i.e., EPA, OSHA), the HPD manufacturer, and the employer, relating to the accuracy of the reported attenuation performance of the HPD? Should comfort data be obtained and reported? How should the fitting instructions be presented? What special workplace conditions (signal detection, speech intelligibility, dust, dirt, temperature, etc.) should be considered in the HPD selection process, and should these factors be addressed in the HPD label?

3aNSb5. Important labeling considerations in production environments. Larry H. Royster (Dept. of Mech. Eng., North Carolina State Univ., Raleigh, NC 27695-7910, royster@eos.ncsu.edu)

The use of labels to warn workers of potential hazards (noise and vibration) in production environments has been of interest to many professionals for several years. At the present time, at least in the USA, the use (or nonuse) of labels has been strongly influenced by the legal process. This presentation will focus mainly on the use of, and potential effectiveness of, labels as a warning tool in real-world production environments. In addition, a limited discussion of warning label activities in other countries will be presented.

Contributed Paper

12:05

3aNSb6. How should we measure hearing protector attenuation? John P. Barry (U.S. Dept. of Labor—OSHA, 3535 Market St., Ste. 2100, Philadelphia, PA 19104)

There has been an ongoing controversy relative to hearing protector effectiveness. The American National Standards Institute (ANSI) sets consensus guidelines to measure hearing protector attenuation (S3.19-1974 through S12.6-1990). Beginning with Padilla in 1976, field studies have consistently shown that laboratory data greatly overestimate the attenuation typically received in the workplace. The ubiquitous NRR (noise re-

duction rating) also overpredicts actual workplace attenuation because it is derived from laboratory data. Occupational fit is usually poorer than laboratory fit. Reusable hearing protectors deteriorate due to repetitive use and aging of materials. Interaction with other protective equipment can reduce hearing protector attenuation. Both the Occupational Safety and Health Administration (OSHA) and the National Institute for Occupational Safety and Health (NIOSH) have proposed workplace attenuation derating factors for the NRR. However, the development of employer hearing protector fit testing procedures analogous to respirator fit testing in concert with a fully effective hearing conservation program will be the only effective answer to this question.

12:20–12:40

Panel Discussion

WEDNESDAY MORNING, 18 JUNE 1997

ROOM N, 9:00 TO 11:15 A.M.

Session 3aPA

Physical Acoustics: Atmospheric Acoustics

David L. Gardner, Chair

Los Alamos National Laboratory, Material Division, MS/K764, Los Alamos, New Mexico 87545

Contributed Papers

9:00

3aPA1. Analysis of the complex acoustic signal for use in identifying outdoor propagation modes and effects of turbulence. David E. Norris (Graduate Program in Acoust., Penn State Univ., University Park, PA 16802) and Dennis W. Thomson (Penn State Univ., University Park, PA 16802)

In outdoor sound propagation; refraction, diffraction, and scattering may all affect the properties of the propagated signal. The wind and temperature profiles characterize the refractive effects, the boundary conditions determine the effects of diffraction, and the presence of atmospheric turbulence drives the scattering. Signal measurements recently conducted at Penn State Univ.'s Rock Springs micrometeorological field site were focused on measuring the properties of the received complex acoustic signal. A source was configured to radiate continuous tones at five frequencies ranging from 90 to 660 Hz. Ten- to fifty-min recordings were made at distances ranging from 46 to 460 m. Both the magnitude and phase of the received signal were evaluated. The results suggest that use of the complex signal, rather than considering only the sound pressure levels, is valuable in characterizing the propagation. Recorded phase differences exhibit strong coherence between the source frequencies, suggesting refraction dominated propagation. Recorded fluctuations in magnitude exhibit much weaker coherence and appear to be independent of source

frequency. Observed travel time fluctuations were on the order of 0.2%. A model is presented that explains these fluctuations by considering the influence of atmospheric turbulence over a fixed propagation path.

9:15

3aPA2. Atmospheric multipath resolution using spread spectrum acoustic signals. David C. Swanson, T. Douglas Mast, Mark P. Mahon, and David Norris (Appl. Res. Lab. and Dept. of Meteorol., Penn State Univ., P.O. Box 30, State College, PA 16804-0030)

An investigation into multipath resolution for outdoor sound propagation is done using spread spectrum acoustic signals. A low-frequency sinusoid is bi-phase modulated using a pseudorandom sequence with period $2^M - 1$. The "chip rate" fc in Hz is the number of bits per second used in the bi-phase modulation. The shape of the spread spectrum sinusoid is in the form of $\sin(x)/x$, where the peak is at the sinusoid frequency and the adjacent nulls are at $\pm fc$ in Hz. Faster chip rates result in broader spectra. Cross correlating the received acoustic signals with the transmitted signals provides a time delay and amplitude estimate for each path in the propagation channel. A field experiment in a strong 15-kn wind revealed a clear direct path and ground reflection appropriate for the profile conditions and confirmed by ray tracing. [Work supported by DARPA.]

3aPA3. Atmospheric boundary layer sound velocity profile model for wind and temperature. David C. Swanson, Karl M. Reichard, Kathy T. McClintic, and David Norris (Appl. Res. Lab. and Dept. of Meteorol., Penn State Univ., P.O. Box 30, State College, PA 16804-0030)

Effective outdoor sound propagation modeling requires an accurate knowledge of the wind and temperature profiles to elevations at least on the order of 1/10th the propagation distance of interest. Using similarity theory and ground-based wind and temperature sensors, the profile near the ground is modeled. However, extending a similarity-based profile above the surface layer is not physical and leads to large errors. This work presents a new profile model which fits very closely to similarity theory within the surface layer, and smoothly transitions to fit upper elevation atmospheric data available worldwide through aviation weather resources. Field experiments confirm the new profile's validity in stable atmospheres, both from SODAR and RASS measurements, as well as very long-range acoustic detection and parabolic equation modeling for low-frequency sinusoids. [Work supported by DARPA and U.S. Army CE-COM.]

3aPA4. Application of a turbulence generation model for atmospheric propagation calculations. Michael R. Stinson and Gilles A. Daigle (Inst. for Microstructural Sci., Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

Realizations of atmospheric turbulence, used as input for parabolic equation (PE) calculations, are commonly generated using a two-dimensional Fourier representation of the spatial fluctuations in refractive index, i.e., the turbulence field. A number of factors affect the nature of the generated turbulence and determine the character of the acoustic propagation. These factors include the turbulence spectrum (e.g., Gaussian, Kolmogorov, or von Karman spectra) and the number of Fourier modes included in the representation. The Green's function PE is used to calculate sound field characteristics in the case of upward refraction above an impedance plane using different turbulence fields.

10:00–10:15 Break

3aPA5. Incorporation of source motion into the fast-field program. D. Keith Wilson (U. S. Army Res. Lab., Battlefield Env. Div., 2800 Powder Mill Rd., Adelphi, MD 20783, dkwilson@arl.mil)

The fast-field program (FFP) has proven to be a useful tool for determining low-frequency sound fields generated by stationary sources in horizontally stratified atmospheres and oceans. Extension of the FFP to the case of moving sources is the topic of this paper. It is shown that the exact moving source solution involves a three-dimensional inverse Fourier transform (with respect to frequency and two horizontal wave numbers), that unfortunately is computationally infeasible. By shifting to source-following, Lorentz coordinates, however, the solution can be reduced to an approximate, one-dimensional, inverse transform involving only the horizontal, radial wave number. The result is an efficient computational method that works well at any subsonic Mach number, and includes accurate calculation of frequency Doppler effects in the presence of refraction. An example simulation of the noise field generated by a landing aircraft will be provided.

3aPA6. Sonic boom scattering from anisotropic turbulences. Mark Kelly, Richard Raspet, and Henry E. Bass (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677)

The single-scatter first Born approximation has previously been used to describe sonic boom scattering from spherically shaped turbulences. Boulanger *et al.* [J. Acoust. Soc. Am. **98**, 3412–3417 (1995)] used it to model the turbulent atmosphere (assuming isotropic turbulence) with reasonable success. An adaptation of this method to calculate the scattered wave from

an ellipsoidally shaped turbule, using a Cartesian coordinate system centered at the turbule, gives more realistic results than those obtained from a spherical scatterer. The ellipsoidal method produces directionally dependent rise times, allowing better description of turbulence with mean wind flow, or anisotropic turbulence. This new approximation was applied to simulated booms from a T-38 flying at an altitude of 10 000 m, with a collection of turbulences located in the planetary boundary layer (PBL). The size distribution and number of turbulences used to simulate the atmosphere in the PBL were calculated using recorded meteorological data and Monte Carlo routines. This method of simulating the turbulent atmosphere has produced rise time distributions in good agreement with data empirically obtained from the Joint Acoustic Propagation Experiments performed at White Sands in 1991.

3aPA7. Effect of aircraft maneuvers on sonic boom penetration into a flat ocean. Tracie J. Ferguson and Victor W. Sparrow (Graduate Program in Acoust., Penn State Univ., 157 Hammond Bldg., University Park, PA 16802)

To fulfill its ongoing mission of evaluating the environmental noise effects of training operations, the U. S. Air Force Materiel Command has initiated research studies on the penetration of sonic boom noise into seas, bays, and lakes. All of the previous research on sonic boom noise penetration into water from supersonic aircraft have assumed steady flight. In realistic training, however, aircraft accelerate and decelerate and are involved in maneuvers such as climbing, diving, and turning. The present research will report the results of initial studies predicting the penetrating sonic boom noise created by maneuvering supersonic aircraft. The commercial package PCBOOM3 [Wyle Laboratories, Arlington, VA] is used to calculate the ray paths taken from the supersonic aircraft to the water surface. The sonic boom waveforms and respective incident angles on the surface are obtained from PCBOOM3 and are input to an algorithm for calculating the underwater sound field recently presented by Sparrow and Ferguson [AIAA Paper 97-0486 (1997)]. Given an arbitrary shaped incident waveform, the algorithm provides the resulting underwater sound field assuming the water surface is flat. [Work supported by Armstrong Laboratory, Air Force Materiel Command, USAF, under Grant No. F41624-96-1-0003.]

3aPA8. Effects of wind-wave generated bubbles on sonic boom noise penetration into the ocean. Judith L. Rochat and Victor W. Sparrow (Graduate Program in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804, rochat@sabine.acs.psu.edu)

Supersonic aircraft, particularly the projected high speed civil transport (HSCT), fly over the ocean to minimize the sonic boom noise impact on land. This noise interacts with the ocean surface, causing an underwater pressure disturbance. In steady, level supersonic flight at less than Mach 4.4 over a homogeneous ocean, the underwater disturbance is a decaying evanescent wave persisting in depth to 100 m or more. Realistically, however, the ocean is not homogeneous and contains wind-wave generated bubble clouds near the surface. In the current research, a two-dimensional finite difference simulation is used to predict underwater sound levels due to a sonic boom impinging on an ocean with bubble layers near the surface. Each of the bubble layers has an approximate sound speed appropriate for its corresponding depth; bubbles in the ocean can cause the sound speed to drop as low as the speed of sound in air. In this presentation the effects of bubble layers on the sonic boom noise penetration will be described. Preliminary results show that one can get propagating waves in the upper bubble layers but only evanescent waves below the layers. [Work supported by NASA Langley Research Center, under Grant NAG 1-1638.]

Session 3aPP

Psychological and Physiological Acoustics: Binaural Processing and Sound Localization II; Processing of Complex Sound Streams (Poster Session)

Richard M. Stern, Chair

*Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213***Contributed Papers**

All posters will be on display from 8:30 a.m. to 11:30 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:30 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 11:30 a.m.

3aPP1. A comparison of spectral correlation and local feature-matching models of pinna cue processing. Ewan A. Macpherson (Dept. of Psych. and Waisman Ctr., Univ. of Wisconsin, Madison, WI 53705-2280, macpherson@waisman.wisc.edu)

Two models of spectral cue processing were compared to determine how the auditory system might extract directional information from pinna filtering. Previous experiments [E. Macpherson, J. Acoust. Soc. Am. **99**, 2515 (1996)] suggest that the processing is robust to irregularity in source spectra, making effective use of pinna cues without utilizing prior knowledge of the source. The two candidate models attempt to achieve this by different means. The first uses correlation to find the best match between eardrum spectra (or their derivatives) and a set of stored templates associated with particular directions. The second makes the match based on prominent local spectral features such as peaks and notches. The models' predictions were compared to listeners' responses in localization experiments involving free-field presentation of stimuli with a variety of nonflat spectra. Correlational matching of the spectral gradient gave the best agreement overall, while feature-matching proved unrealistically sensitive to spectral irregularities. Neither model consistently predicted specific mislocalizations or stereotypical front-back confusions. Despite serious shortcomings, the correlational model demonstrates that a simple, physiologically plausible process can exhibit the required general robustness. The results do not support the hypothesis that local features are the spectral cues for sound direction. [Work supported by NIDCD.]

3aPP2. Effects of masker envelope distribution on binaural masking at high frequencies: Evidence for the influence of peripheral compression? Steven van de Par and Armin Kohlrausch (Inst. for Perception Res. (IPO), P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands)

NoS π thresholds were measured using a 3IFC paradigm for masker levels between 10 and 80 dB SPL. Three types of narrow-band maskers were used to investigate the influence of the envelope probability distribution on signal detectability. The three types were multiplied noise (MN), Gaussian noise (GN), and low-noise noise (LNN) each with a bandwidth of 25 Hz. Due to the rather low NoSo thresholds for LNN, both monaural and binaural detection cues might be used by subjects in the NoS π condition. In order to avoid this, for LNN maskers, an NoSo stimulus was presented in the reference intervals at the same signal-to-noise ratio as in the NoS π test interval. The results show that at medium levels (40–70 dB), thresholds in MN are about 10 dB lower than for the other maskers. This difference decreases or even disappears at low and high levels. It is proposed that these effects are related to the compressive transformations

which occur in the cochlea. These transformations decrease the size of the interaural intensity differences present in the envelopes of the maskers, and can be shown to do so more effectively for GN and LNN maskers than for MN maskers.

3aPP3. Sound localization in noise in normal-hearing listeners. Christian Lorenzi, Stuart Gatehouse, and Catherine Lever (Inst. of Hearing Res., MRC, Glasgow Royal Infirmary, 16 Alexandra Parade, Glasgow G31 2ER, UK)

The ability to localize a click train (100-Hz click rate, 300-ms duration, 70 dB SPL) was measured in the presence of a white noise masker, using an array of 11 speakers located in the horizontal plane at azimuth angles ranging from -90 to 90 deg. All clicks were low-pass filtered at 11 kHz. Clicks were low-pass filtered at 1.6 kHz in the low-pass condition and high-pass filtered at 1.6 kHz in the high-pass condition. Six signal-to-noise ratios (S/N) ranging from -9 to 18 dB (plus a silent condition) and three masker locations (-90 -, 0 -, and 90 -deg azimuth) were tested. Results obtained with four normal-hearing listeners show that, in the broadband and high-pass conditions, localization performance is not affected by noise when S/N is greater than 0 dB. Below this value, performance degrades slowly when S/N decreases and when noise is presented straight ahead; performance degrades rapidly when S/N decreases and when noise is presented at the sides. The same pattern of results is observed in the low-pass condition when noise is presented straight ahead. When noise is presented at the side, performance measured with low-pass filtered clicks starts to worsen at higher S/N's and reaches a lower level.

3aPP4. Sound localization in noise in hearing-impaired listeners. Christian Lorenzi, Stuart Gatehouse, and Catherine Lever (Inst. of Hearing Res., MRC, Glasgow Royal Infirmary, 16 Alexandra Parade, Glasgow G31 2ER, UK)

The present study assesses the ability of four listeners with high-frequency, bilateral symmetrical sensorineural hearing loss to perform horizontal localization of a click train in the presence of a white noise masker. The speaker array and stimuli are identical to those described in the preceding abstract. The localization performance of hearing-impaired listeners is inferior to that of normal-hearing listeners in all conditions. Both groups of listeners show the same pattern of performance when noise is presented straight ahead: Performance is not affected by noise when the signal-to-noise ratio (S/N) is greater than 0 dB; below this value, performance degrades when S/N decreases. When noise is presented at the side, the performance of impaired listeners starts to worsen at higher S/N's (6 to 12-dB S/N) than normal-hearing listeners and reaches a lower level. The

normal-hearing listeners in the low-pass conditions reproduce the general form of the effects of noise on hearing-impaired listeners, but not the absolute level of performance. Parallel measures of clicks detectability are available to assess the extent to which audibility is responsible for the lower absolute level of performance observed in impaired listeners.

3aPP5. Sound localization in the presence of multiple distracters.

Frederic L. Wightman (Waisman Ctr. and Dept. of Psych., Univ. of Wisconsin, Madison, WI 53705, wightman@waisman.wisc.edu) and Doris J. Kistler (Univ. of Wisconsin, Madison, WI 53705)

Listeners gave verbal apparent position estimates in response to a frozen 250 ms wideband noise burst presented from one of 36 locations in virtual free-field. A click train and a modulated harmonic complex, both presented simultaneously with the noise target, served as distracters. On each trial, the distracters were presented in one of six randomly chosen configurations in virtual space. All virtual sources were synthesized from the listeners' own HRTFs. The impact of the distracters on the apparent position of the target varied considerably across listeners. For all listeners, when the distracters were presented on the horizontal plane at +90 degrees and -90 degrees (directly opposite the two ears), the apparent elevation and front-back position of the target were disrupted. For some listeners, this was the only effect of the distracters. In a separate condition, the stimuli were presented ten times in succession and listeners were encouraged to move their heads during stimulus presentation. Synthesis of the virtual sources was coupled to head position in order to simulate stationary virtual sources. In this condition the distracters had no effect on the apparent position of the target. [Work supported by ONR.]

3aPP6. Auditory localization acuity in free field and simulated environments: Reaction time measures.

Thomas N. Buell and Thomas E. Hanna (Naval Submarine Medical Res. Lab, Groton, CT 06349-5900)

A reaction time measure was used to compare localization acuity in a free field and a simulated environment (where headphone-presented stimuli are processed so as to approximate the waveforms that would occur at each ear canal for a free-field source). Four listeners discriminated the relative direction of a broadband noise (whether it was to the left or right of midline or whether it was in front or back of the interaural axis). The results across four listeners demonstrated that psychometric functions could be plotted for both reaction time as well as percent correct versus angle. Localization was better in the free-field conditions (higher percent correct scores as well as faster reaction times) for front-back discrimination; however, the free-field advantage was negligible for left-right discrimination. Under headphone presentation, there was large inter-subject variability in reaction times for front-back discrimination, but not for left-right discrimination. Although the above findings were not especially surprising, the two metrics did mirror one another in a predictable fashion. The strengths of each procedure will be discussed. [Work supported by NMRDC and ONR, 61153N.MR04101.OOH-5601.]

3aPP7. Minimum audible angle and minimum audible movement angle in the median sagittal plane as a function of stimulus spectrum and reference elevation.

Jamie M. Figueiredo and D. Wesley Grantham (Div. of Hearing and Speech Sci., Vanderbilt Univ. School of Medicine and the Bill Wilkerson Ctr., 1114 19th Ave. South, Nashville, TN 37212)

Minimum audible angle (MAA) and minimum audible movement angle (MAMA) were measured employing a pseudo-virtual method of stimulus presentation. All stimuli were digitally recorded through KE-MAR's two ears while he was positioned in an anechoic chamber with fixed and moving loudspeakers presenting sounds in his median sagittal plane; these recorded binaural stimuli were subsequently played through insert earphones to subjects seated in a darkened sound-insulated room.

For the measurement of MAA, subjects had to say which of two stationary 300-ms sounds was higher in elevation; for the measurement of MAMA, subjects had to say whether a single moving sound moved up or down. Velocities in the MAMA experiment ranged from 20°/s to 150°/s. Stimuli were Gaussian noise bursts that were either wide band, 4-kHz low passed, or 4-kHz high passed. Reference elevation was either 0° or 45° elevation. Results will be discussed in terms of static and dynamic pinna cues that underlie elevation discrimination for high-frequency stimuli and the possible shoulder-bounce cue that may enable elevation discrimination for low-frequency stimuli. [Work supported by NIDCD 00185.]

3aPP8. Auditory motion aftereffects with varying interaural time differences.

Hisashi Uematsu, Makio Kashino, and Tatsuya Hirahara (NTT Basic Res. Labs., 3-1, Morinosato Wakamiya, Atsugi, Kanagawa, 243-01 Japan)

After prolonged listening to a sound moving across the horizontal plane (adapter), a stationary sound (test) can be perceived as moving in the opposite direction. In experiment 1, the magnitude of the auditory motion aftereffect for 500-Hz tones was measured as a function of adapter velocity using the method of constant stimuli. Apparent sound movement was produced by varying only interaural time differences (ITDs). In each trial, an adapter whose ITD varied in the range of $\pm 600 \mu\text{s}$ with a given rate (velocity) was presented for 10 s, followed by a 1-s test whose velocity was chosen randomly from seven values surrounding zero. Subjects judged whether the test moved to the left or right, and subjective stationary points were obtained. Negative aftereffects were observed when the adapter velocity ranged between 300 and 600 μs per second. In experiment 2, the magnitude of the aftereffect was measured as a function of the frequency difference (Δf) between pure-tone adapter and test, using the same procedure as experiment 1. The aftereffect was reduced with increasing Δf , and disappeared when Δf was 1/2 oct. The results will be discussed in terms of ITD-change detectors each tuned to a specific range of ITD change and frequency.

3aPP9. Head movements in localization.

William A. Yost and Dan Mapes-Riordan (Parmly Hearing Inst., Loyola Univ., Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, wyost@luc.edu)

Head movements in all three planes were measured while listeners localized a pulse of transient sounds (train of 100 μs broadband impulses) located in the frontal azimuth plane. The eight loudspeakers were at the same height as the head of the seated listener. They were located every 25.7° from +90° to -90°, where 0° is straight ahead. Some listeners were able to see the eight sound sources, while for other listeners the sound sources were hidden from view. In addition, listeners were instructed differently in terms of the necessity for high accuracy or fast reaction. In all cases, the primary means of the listener locating the sound source was head movements. Listeners were instructed to face the source of the sound so that their nose was facing directly at the perceived sound source. The results will be discussed in terms of processing sounds in dynamic listening situations. [Research supported by the Air Force Office of Scientific Research and a Program Project Grant from the NIDCD, DC00293.]

3aPP10. Scaling perceived distance of virtual sound sources.

Pavel Zahorik (Dept. of Psych. and Waisman Ctr., Univ. of Wisconsin, Madison, WI 53705, zahorik@waisman.wisc.edu)

Magnitude estimation and paired-comparison scaling methods were applied to the problem of measuring perceived auditory distance. Stimuli consisted of binaural recordings made in a reverberant hallway at distances of 1, 2, 4, 8, 16, 32, and 64 ft. A loudspeaker produced broadband noise burst, 100 ms in duration, served as the distal stimulus. Insert-style microphones were placed in the ears of a single individual, with measurements made from two directions, extending in distance either from the individual's front or left side. Construction of virtual sound source stimuli in this

fashion insured that the stimulation reaching the listener was naturalistic and therefore possessed the acoustic cues thought to be important in the real-world perception of auditory distance. Five listeners were presented these stimuli in several variants of both the magnitude estimation and paired-comparison paradigms. Though relatively large individual variability exists, the resulting scales are all compressed when compared to the corresponding scale of physical distance. The amount of compression is shown to be approximately constant across scaling method, thereby providing validation for each method's measurement of a similar underlying psychological process. [Work supported by NASA.]

3aPP11. Three-dimensional sound: Distance identification of sounds synthesized using an atmospheric propagation model including turbulence. G. L. Gibian, K. M. Pepe, W. S. Koroljow (Planning Systems, Inc., 7923 Jones Branch Dr., McLean, VA 22102, ggibian@plansys.com), and W. E. McBride (Planning Systems, Inc., Stennis Space Center, MS 39529)

The perception that an electronically generated sound source originates from outside the head at a particular direction can be produced by filtering the sound with head-related transfer functions (HRTFs). The perception that the source is at a particular distance is more difficult to produce, particularly the first time an unfamiliar sound is heard. In this study, listeners tried to identify source distances when a 300-ms white noise burst was filtered with transfer functions from a physically based atmospheric model including scattering by turbulent eddies in the atmosphere ("turbules"). Distances of 200, 400, 800, 1600, and 3200 m were simulated using a fixed distribution of turbules. Listeners' performance with and without inclusion of scattering by turbules was compared. Results indicate that: (1) when the subjects were given known sounds for reference or had become familiar with the test sounds, they correctly identified the distances about 80% of the time, despite changes in playback gain or reduced dynamic range, and (2) distance identification improved by about 10% when the effects of scattering were included. The improvement in distance identification was produced without the need for increased real-time computation over simple atmospheric absorption filtering. Larger improvements are expected with time-varying and spatial filters.

3aPP12. Positional dependence on localization dominance in the median-sagittal plane. Robert M. Dizon, Ruth Y. Litovsky, and H. Steven Colburn (Hearing Res. Ctr., Boston Univ., Boston, MA 02215)

Previous experiments by the authors investigated aspects of the precedence effect in both the azimuthal and median-sagittal planes in a virtual acoustic environment. Among the aspects studied was localization dominance; where the location of a lead-lag combination is dominated by the leading source. It was apparent that the virtual presentation was not providing an adequate elevation percept in the median plane, possibly due to the use of nonindividualized head-related transfer functions. Nonetheless, subjects, who were given feedback, were able to use timbral differences in the fused lead-lag pairs to perform the task. The results of similar median plane studies in a free-field anechoic condition are reported on, where location percepts are more robust. Preliminary results show position-dependent biases, where some lead positions (front) show greater amounts of localization dominance over certain lag positions (overhead). Frontal dominance has been shown in previous work in the median plane (Rakerd and Hartmann, personal communication). It is postulated that some positions in the median plane are dominant, and that in a competition paradigm such as localization dominance, the amount of dominance given to the lead would be dependent on the locations of the leading and lagging stimuli. [Work supported by NIH Grant Nos. DC02692 and DC00100.]

3aPP13. Collecting localization responses with a virtual acoustic pointer. Erno H. Langendijk and Adelbert W. Bronkhorst (TNO Human Factors Res. Inst., Soesterberg, The Netherlands, langendijk@tm.tno.nl)

A virtual acoustic pointer may provide a natural way of indicating the apparent position of virtual auditory stimuli. In a localization experiment, listeners were instructed to place an acoustic pointer at the same apparent position as a virtual target. The pointer was generated by filtering a click train (100 ms) with equalized minimum-phase head-related transfer functions (HRTFs). Listeners controlled the position of the pointer with a hand-held stick that rotated about a fixed point. In the baseline condition the target, a 100 ms noise burst, was filtered with the same HRTFs as the pointer. In other conditions, the HRTFs of the target were modeled with three, five, or seven principal components. Target and pointer were never presented simultaneously. Listener's judgments with the pointer were more accurate than those reported in previous studies requiring verbal responses. The acoustic pointer also seemed to help listeners to resolve front-back confusions. With fewer than seven principal components used to synthesize stimuli, subjects made substantially more front-back confusions and many more errors in elevation, in agreement with previous studies. The results suggest that an acoustic pointer allows more subtle cues for sound localization to be studied. [Work supported by TNO.]

3aPP14. Auditory scene analysis-constrained array processing for sound source separation. Laura Drake and Janet Rutledge (Dept. of Elec. and Comp. Eng., Northwestern Univ., Evanston, IL 60208-3118)

In this work, techniques are developed and studied for the extraction of single-source acoustic signals out of multi-source cacaphony. Such extracted signals can be used in a variety of applications including: automatic speech recognition, teleconferencing, and robot auditory systems. Typical approaches fall into two categories: auditory modeling and array signal processing. The approach taken here is to combine these complementary techniques into an integrated one: auditory scene analysis-constrained array processing. Theoretically, this integrated approach should provide a performance gain since the information used by array processing, i.e., direction of propagation through a wave field, is independent of the signal structure information used by auditory scene analysis. One difficulty encountered by auditory scene analysis that can be overcome by array processing is the sequential grouping of spectrally dissimilar phonemes in a speech signal, such as a fricative followed by a vowel, or nasal.

3aPP15. Magnitude estimation of perceptual fusion of multiple sound sources. Sandra J. Guzman and William A. Yost (Parnly Hearing Inst., Loyola Univ., Chicago, IL 60626, sguzman@luc.edu)

The precedence effect states that the first sound wave arriving at a listener has more influence on the perception of location than later arriving echoes. When a source plus echoes is perceived approximately at the location of the source, the image is said to be "fused." The following set of experiments introduces a new procedure to estimate fusion of clicks. Listeners are presented with a source click plus a single echo click at a variety of delays. Stimuli are rated between a value of "0" and "100," with "0" indicating a single source and "100" indicating more than one source. These ratings are used to determine delays which are perceived as a single source (receive a rating of "0") and delays which are perceived as multiple sources (receive a rating of "100"). These anchor stimuli are then presented on each trial. The stimulus which received a rating of "0" is presented first, followed by the sound to be rated. The stimulus which received a rating of "100" is presented last. This procedure was used to determine fusion as a function of location and number of echoes presented. [Research supported by a program project grant from NIDCD.]

3aPP16. Detection of simultaneous streams of running speech. Jennifer Tufts and Tom Frank (Dept. of Commun. Disord., Penn State Univ., University Park, PA 16802)

The purpose of this study is to determine how many simultaneous streams of running speech a listener can detect and correctly enumerate. Thirty normal-hearing subjects will participate. Stimuli will consist of three sets of digitally mixed recordings of female talkers where the number of talkers in each set ranges from two to ten. Subjects will be asked to identify how many speakers they hear and to rate the certainty with which they make their judgments. Each subject will be tested in the left ear, right ear, and binaurally. Results are expected to demonstrate the limits of the auditory system in resolving individual speakers or streams of speech without localization cues.

3aPP17. Effects of contralateral presentation and interaural intensity differences on the detection of mistuning. Hedwig Gockel and Robert P. Carlyon (MRC Appl. Psych. Unit, 15 Chaucer Rd., Cambridge CB2 2EF, UK, hedwig.gockel@mrc-apu.cam.ac.uk)

Listeners were required to detect a mistuning imposed on the center ("target") component of a harmonic complex consisting of the first seven harmonics of a 500-Hz fundamental. In the standard interval all components were frequency modulated in phase by a 5-Hz sinusoid. In the signal interval the frequency modulation of the target component was inverted in-phase, thereby introducing a mistuning proportional to the depth of FM. For monaural presentation, Carlyon [J. Acoust. Soc. Am. **95**, 2622–2630 (1994)] reported a substantial elevation of thresholds in the presence of an unmodulated asynchronous interferer whose frequency was identical to that of the target component. He explained this by assuming that the interferer caused the target component to be perceptually segregated from the remainder of the complex, thereby impairing across-frequency com-

parisons. The present study investigated the effect of lateralization on the detection of mistuning. Preliminary results of experiment 1 suggest that in the absence of an interferer, presenting the target component contralaterally to the remainder of the complex raises thresholds, but that the task is still possible. Results from experiment 2 indicate that a contralaterally presented interferer impairs performance, a result which is interpreted in terms of auditory streaming.

3aPP18. Auditory guided visual search with visual distracters. William R. D'Angelo (AL/CFBA, Bldg. 441, 2610 Seventh St., Wright-Patterson AFB, OH 45433-7901, wdangelo@biocom.al.wpafb.af.mil), Robert S. Bolia (Systems Res. Labs., Dayton, OH 45433), Richard L. McKinley (Wright-Patterson AFB, OH 45433), and David R. Perrott (California State Univ., Los Angeles, CA 90032)

In order to study the interaction of the auditory and visual systems, a series of experiments has been designed. The first experiment was completed in 1994 and recently published [Perrott *et al.*, Human Factors **38**(4), 702–715 (1996)]. This initial experiment explored auditory-aided visual search with a single high-contrast target and a dark background. The results demonstrated a distinct advantage in visual target acquisition with the addition of a correlated directional audio cue. The objective of this experiment was to develop a function describing the effects of directional auditory cueing on search times in visual fields of varying complexity. Visual conditions included 1, 5, 10, 25, or 50 distracters randomly distributed. The locations of the targets and distracters ranged between $\pm 180^\circ$ in azimuth and between -75° and $+90^\circ$ in elevation. The three auditory conditions included no auditory cue, a free-field cue, and a virtual cue. In each case, the stimulus was pink noise (20–20 000 Hz). Results will be presented which show the effect of the number of distracters on search times with and without auditory cues. [Work supported by AFOSR.]

WEDNESDAY MORNING, 18 JUNE 1997

ROOM F, 8:30 A.M. TO 12:00 NOON

Session 3aSA

Structural Acoustics and Vibration: Active Control of Structures and Radiation

Scott D. Sommerfeldt, Chair

Department of Physics, Brigham Young University, 277 Fletcher Building, Provo, Utah 84602

Contributed Papers

8:30

3aSA1. Active control of noise transmission through aeroelastic plates. Kenneth D. Frampton and Robert L. Clark (Dept. of Mech. Eng., Duke Univ., Durham, NC 27708-0302, kdf1@acpub.duke.edu)

This presentation investigates the application of feedback control toward the problem of decreasing noise transmission through aeroelastic plates. The modeling method is discussed for a plate with pinned boundary conditions, coupled to convected fluid loading on one side, and backed by a nonconvecting acoustic environment on the opposite. A static output feedback compensator is designed to minimize the sound transmitted to the nonconvecting medium. Results from the application of this control system are compared with levels of attenuation predicted by state feedback control. Results indicate that significant reductions in transmitted noise can be achieved; however, variations in the fluid convection velocity can alter the level of performance.

8:45

3aSA2. Design of adaptive panels with high transmission loss characteristics. Richard L. St. Pierre, Jr., Gary H. Koopmann, and Weicheng Chen (157 Hammond Bldf., Penn State Univ., University Park, PA 16802)

This paper presents a design method for minimizing the sound transmitted through panels. This unique method is based on minimizing the volume velocity of individual segments of a vibrating panel using control loudspeakers. In this manner, each of the control segments of the panel can be controlled separately. The basic acoustic equations governing volume velocity control are presented, and it is shown that, at low frequencies, this method will achieve global sound power reductions in the far field. Further analysis shows that this method will achieve significant sound power reductions (10 dB or more) for values of kL of 3.0 or less, where L is the characteristic dimension of each control segment. Tests conducted in a transmission loss chamber verify the efficacy of volume velocity control using a controller based on the filtered- x control algorithm. This controller

allows for control over a frequency band, and results show sound power reductions of 9 dB over a 60-Hz bandwidth. These tests are among the first attempts to use the volume velocity control method over a band of frequencies. The experimental results are verified with numerical simulations. The results show promise that this methodology is an effective way to control broadband, low-frequency sound.

9:00

3aSA3. The use of a piezoelectric double amplifier active-skin in the control of panel radiation. Brody D. Johnson, Chris R. Fuller, and Ricardo A. Burdisso (Vib. and Acoust. Labs., Dept. of Mech. Eng., Virginia Tech., Blacksburg, VA 24060)

The potential of a piezoelectric active-skin is analytically demonstrated for the reduction of broadband acoustic radiation from a vibrating panel in the frequency range from 0 to 750 Hz. The active-skin is comprised of a number of independently controllable piezoelectric double amplifier elements arranged in a contingent surface over the vibrating panel. Both the finite-element method (FEM) and the boundary element method (BEM) are employed in the development of numerical models for relation of the piezoelectric voltage excitation to vibrational and acoustic responses of these active-skin elements. The adaptive feedforward filtered- x least-mean-squares (LMS) algorithm is employed in the time domain simulation of control with the active-skin. Both FIR and IIR compensator types are investigated. ANC and ASAC perspectives are examined for the realization of the active-skin. Control performance is quantified using both sound power level (L_w) and sound pressure level (SPL) data. Simulation shows that an attenuation of 10 dB in the radiated power is possible over the investigated frequency range, which encompasses six modes of panel vibration. Conclusions regarding the potential for the active-skin for the broadband attenuation of structurally radiated sound are then made from the analytical results. [Work supported by ONR.]

9:15

3aSA4. Finite-element analysis of a laminated plate including elastic, viscoelastic, and piezoelectric layers. Louis Guertin, Francois Charron, and Alain Berry (GAUS, Mech. Eng. Dept., Sherbrooke Univ., PQ J1K 2R1, Canada)

A finite element formulation for the vibration of a laminated plate including elastic, viscoelastic, and piezoelectric layers is presented. The model takes into account the mass, stiffness, and damping of each layer. A simplified discrete-layer plate theory and Hamilton's principle are used to formulate the equation of motion of the system. The continuity of the transverse stresses and the displacements are imposed at interlaminar layer interfaces. Only five degrees of freedom per node are necessary to describe the deformation of the plate. The discrete element is a quadrilateral shape with eight nodes and a total of 40 degrees of freedom. Extensional, bending, and transverse shear deformation are included in the formulation. One additional voltage degree of freedom per element is necessary for each piezoelectric layer. This finite-element formulation can be used to modelize the dynamic as well as the static response of a multilayered isotropic plate subjected to both mechanical and electrical excitations. The multilayer elastic-viscoelastic-piezoelectric element formulation will ultimately be used to study smart structures integrating piezoelectric layers acting as distributed actuators and sensors. Comparison with published experimental, analytical, and numerical results will be presented to validate the model.

9:30

3aSA5. Smart foam for active/passive plate radiation control. C. A. Gentry, C. Guigou, and C. R. Fuller (Dept. of Mech. Eng., Virginia Polytechnic Inst. & State Univ., Blacksburg, VA 24060-0238, cgentry@vt.edu)

A new type of noise suppression device known as "smart foam" that integrates a distributed piezoelectric polymer actuator (the active component) between individual layers of sound-absorbing foam (the passive component) is discussed. Smart foam acts as a hybrid active/passive sur-

face coating as it adaptively modifies the acoustic impedance of a vibrating surface. The potential of such a device to reduce low- and high-frequency sound radiation from a simple acoustic source (piston) has been demonstrated. In this experimental study, an array of smart foam modules are implemented to yield a net decrease in the sound power radiated by a vibrating, baffled plate in an anechoic environment. A MIMO feedforward filtered- x LMS algorithm is used to minimize the error signals provided by microphones located in close proximity of the smart foam array for harmonic and broadband plate excitations. Two different reference signals, i.e., the voltage sent to the plate disturbance actuator and the signal from an accelerometer directly mounted on the plate, are investigated and compared in terms of the sound attenuation achieved. The capability of smart foam for reducing plate sound radiation is demonstrated. Smart foam is relatively lightweight, which makes it a good candidate for interior aircraft noise control applications. [Work supported by NASA LaRC.]

9:45

3aSA6. Active cabin noise control in the time domain. Xiao-qi Bao, Vasundara Varadan, and Vijay Varadan (Penn State Univ., 149 Hammond Bldg., University Park, PA 16802)

Active control of noise inside an enclosure from exterior noise sources is accomplished by the use of piezoelectric actuators and either piezoelectric or MEMS sensors placed on the enclosure wall. It is generally known that this approach provides very good results at the resonant frequencies of the cabin walls and if the excitation noise is at a single frequency. In the present study, broadband excitation is looked at from the exterior, in the time domain. A sensor mounted on the cabin wall provides the reference signal for the excitation; the error sensor can either be the piezoelectric sensor on the wall or a sound-level meter placed at an appropriate position inside the cabin. Different types of closed loop control are attempted using both analog and digital strategies. The study examines the number and placement of actuators on the wall relative to bandwidth of the external source, as well as control of wall vibration as a means of controlling the radiated noise.

10:00

3aSA7. Active control of sound radiation from cylinders with piezoceramic actuators and structural acoustic sensing. Julien P. Maillard and Chris R. Fuller (Vib. and Acoust. Labs., Virginia Tech., Blacksburg, VA 24061)

In this paper, analytical and experimental results of an investigation of active control of sound radiated from cylinders are presented. The aluminum cylinder is 1 m in length, 25 cm in diameter, and 2.4 mm in thickness with two rigid endcaps at both ends. The excitation is a bandlimited random noise encompassing the first five modes of the cylinder and the control actuators are surface mounted piezoceramic transducers. Since it is desired to integrate the error sensors into the structure, the recently developed structural acoustic sensing (SAS) approach is extended to cylindrical coordinates and implemented using 12 accelerometers mounted on the cylinder. The SAS approach provides time-domain estimates of far-field radiated sound at predetermined radiation angles. The controller is a 3×3 filtered- x LMS paradigm implemented on a TMS320C30 DSP. The results show excellent global control of the radiated sound over the frequency bandwidth of excitation. The SAS approach is shown to yield similar performances as error microphones located in the far field. [Work supported by the Office of Naval Research.]

10:15–10:30 Break

10:30

3aSA8. Test of a piezoceramic actuator for the active control of power transformer vibrations. Philippe Micheau, Alain Berry, Patrice Masson, and Francois Charrette (Dept. of Mech. Eng., Univ. of Sherbrooke, Sherbrooke, PQ J1K 2R1, Canada, philippe.micheau@gme.usherb.ca)

The sound radiated by large power transformers found in electric utility substations is an annoyance to nearby residents. Noise barriers are not effective at 120 Hz, the most prominent harmonic radiated by the transformers, and active control with loudspeakers and microphones is very

sensitive to environmental conditions. Thus the active structural acoustic control is an attractive alternative because this strategy controls the sound at the source. However, efficient actuators are needed to generate anti-vibrations. This paper describes an experimental investigation conducted in order to select and test a piezoceramic actuator for the active control of large power transformer vibrations. It presents experimental measurements of transverse displacements on the various sides of a transformer, together with simulation and experimentation of a selected piezoceramic actuator bonded to a test structure and to the transformer. Vibration levels obtained with a $(0.0075 \times 2.5 \times 1.5 \text{ in.})$ PZT piezoceramic actuator bonded to the transformer are generally larger than the primary vibration field at 240 Hz and at higher harmonics, but smaller at 120 Hz. Guidelines for the design of appropriate piezoceramic actuators to control power transformer vibrations are given. [Work supported by Hydro-Québec.]

10:45

3aSA9. Control of structural sound radiation using multiple detuned vibration absorbers. R. I. Wright, F. Charette, and C. R. Fuller (Vib. and Acoust. Labs., Dept. of Mech. Eng., Virginia Polytech. Inst. and State Univ., Blacksburg, VA 24061)

The results of an experimental investigation into using multiple tunable vibration absorbers to minimize sound radiation from vibrating plates will be presented. The vibration absorbers consist of lumped mass-spring systems whose characteristics (i.e., resonance frequency) can be electronically varied. A multiple channel, gradient search algorithm, implemented on a DSP, is used to adapt the characteristics of each of the absorbers (i.e., detune them) so as to minimize the radiated sound at selected multiple positions in the farfield. Experimental results for the detuned case are presented and compared to those obtained by tuning the resonance frequency of the vibration absorbers to the disturbance frequency. Preliminary results indicate that attenuation of sound radiated from the plate can be significantly improved by detuning the vibration absorbers away from the disturbance frequency via a gradient search. Thus, an adaptive-passive approach to minimizing sound radiation has been demonstrated to obtain good attenuation with very little electrical power requirement while adapting. This result is in contrast to active structural acoustic control which typically requires a significant control power. [Work supported by NASA LaRC and ONR.]

11:00

3aSA10. A new hybrid model for electrorheological fluid dampers and application to dynamic vibration absorption. Andrew N. Vavreck (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804, avx2@psu.edu) and Kon-Well Wang (Penn State Univ., University Park, PA 16802)

Dampers incorporating electrorheological or magnetorheological fluid (whose damping parameters can be adjusted by altering an imposed electric or magnetic field, respectively) as part of their structure hold great promise as adaptive-passive vibration reduction devices. Extant models of electrorheological dampers are insufficient, because of excessive complexity or the lack of broad experimental validation, to allow the prediction of the response of the damper over a wide range of operating conditions. The subject work proposes a new, simple lumped parameter model for electrorheological dampers, validates the new model against empirical data, and explores the application of an electrorheological damper in a tuned dynamic vibration absorber. The vibration absorption capacity of an electro- or magnetorheological damper in a dynamic vibration absorber is found to depend on the displacement amplitude and frequency of the vibration, but a substantial reduction in vibration is achievable with only adaptive-passive control. [Work supported by U.S. Army Research Office.]

11:15

3aSA11. Experimental results on active structural acoustic control using fiber-optic strain sensors. Patrice Masson and Alain Berry (G.A.U.S., Dept. of Mech. Eng., Univ. of Sherbrooke, Sherbrooke, PQ J1K 2R1, Canada)

A new strategy for the active control of the sound power radiated from a flexural vibrating beam is experimentally investigated. This strategy involves sensing the discrete structural strain using white-light interferometric multimode fiber-optic strain sensors. Two approaches are used to relate the radiated sound power to the structural strain. The first approach is based on the reconstruction of the transverse structural displacement field from the structural strain field using a finite-difference scheme, and a wave-number transformation of the structural displacement to obtain the radiated sound field. The second approach is based on an exact expression directly relating the radiated sound field to the wave-number transform of the structural strain. In both approaches, the wave-number transform is performed over the supersonic (radiating) components of the structural information. The dynamic strain measurements are first compared with polyvinylidene film strain sensors. As a second step, a straightforward control scheme is implemented into a multichannel filtered-x LMS algorithm to control the strain amplitudes on the beam actively. Finally, both strategies for using the strain information are implemented to perform the active control of the sound power radiated from the beam. Preliminary performance results of both strategies using four strain sensors are then presented.

11:30

3aSA12. Frequency dependence of decoupling filters for multichannel feedforward active noise control. Randolph H. Cabell (Virginia Consortium of Eng. and Sci. Univ., VPI & SU, 303 Butler Farm Rd., Ste. 101, Hampton, VA 23661, r.h.cabell@larc.nasa.gov) and Chris Fuller (VPI & SU, Blacksburg, VA 24061)

The performance of multichannel feedforward LMS-type control algorithms is improved if control is implemented in noninteracting coordinates. A decoupling matrix can be applied to sensor inputs and actuator outputs to transform the harmonic feedforward control problem into an orthogonal (i.e., principal component) coordinate system. Convergence of the adaptive algorithm is faster in the principal coordinate system, and sensitivity to sensor noise can be reduced by eliminating components associated with ill-conditioning. One question that arises is how quickly the decoupling matrix changes with frequency. This is important for control systems where the excitation frequency might wander about a nominal value. Results of a numerical study of this frequency dependence show that the decoupling matrix is stationary around a resonance, and has the highest rate of change between resonances. However, in some cases the matrix remains stationary between resonances. The frequency dependence of the decoupling matrix will be discussed for three control systems: a simply supported plate, a plate radiating into a half-space, and an enclosed cavity. [Work supported by NASA Langley Research Center.]

11:45

3aSA13. Comparison of controller power requirements for output feedback control of broadband structural acoustic radiation. G. Clark Smith, Jr. and Robert L. Clark, Jr. (Dept. of Mech. Eng. and Mater. Sci., Duke Univ., Box 90302, Durham, NC 27708-0302, gcs@acpub.duke.edu)

Performance comparisons of active structural-acoustic control system designs with respect to controller power requirements are important for practical implementation. In this work, a broadband measure of controller energy usage per unit of input disturbance energy is developed. This metric is the H_2 norm of the system transfer matrix from the disturbance signals to the control actuator signals. Comparisons of acoustic performance are made for three output feedback controllers, each designed to reduce one of the following variables: (1) sound power radiation; (2) structural modal velocity; and (3) structural velocity at twelve discrete points. All three controllers were designed such that their controller power requirements were equivalent. An array of five, exogenous, point-force disturbances

were used to drive a plate model, and each control system utilized the same three colocated point-velocity sensor and point-force actuator pairs. Analytical results demonstrate that the cost functional based upon sound power places a greater emphasis in controlling the low-frequency bandwidth.

This controller provides 2.5-dB greater sound reduction than design #2 in the third octave band including the first radiation mode and 2.7-dB greater reduction than #3 in the same band. [Work supported by NSF Career Program CMS-9501470.]

WEDNESDAY MORNING, 18 JUNE 1997

ROOM Q, 8:30 TO 11:30 A.M.

Session 3aSC

Speech Communication: Perception of Vowels and Consonants (Poster Session)

Terrance M. Nearey, Chair

Department of Linguistics, University of Alberta, 4-32 Assiniboia Hall, Edmonton, Alberta T6G 2E7, Canada

Contributed Papers

All posters will be on display from 8:30 a.m. to 11:30 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:30 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 11:30 a.m.

3aSC1. Comparing listener's perception of place coarticulation in voiceless stops with discriminant analysis. Shu-hui Peng and Terrance Nearey (Dept. of Linguist., Univ. of Alberta, Edmonton, AB T6G 2E7, Canada)

Previous EPG study of place coarticulation in Taiwanese stop consonants suggested that place gestures of two adjacent stops separated by a word boundary (CVC1#C2V(C)) overlapped each other to different degrees depending on speech rate. The place of articulation of C1 and C2 was either dental or velar. As speech rate increased, the perceptual categorization of C1 was affected. In faster speech rate, when the first word was heard in isolation, C1 tended to be perceived with the characteristics of C2 place gesture. However, there was no simple or consistent correlation between the patterns of gestural overlapping and the perception of C1. This lack of transparent relationship between articulatory gestures and perception motivates the current investigation on the acoustic dimension of place coarticulation. The purpose of the current study is to investigate the effect of place coarticulation on the spectral properties of CVC1. Discriminant analyses based on acoustic measures will be used to discriminate the dental stop from the velar stop. Estimated (posterior) probability of class membership of the stop consonants will be compared with the previous perception study to reveal how accurately the predicted probability of responses generated from discriminant analyses reflects the perceptual performance of listeners.

3aSC2. Matching frequency glides with two steady tones. M. E. H. Schouten and W. J. M. Peeters (Utrecht Inst. of Linguist. OTS, Utrecht Univ., Trans 10, 3512 JK Utrecht, The Netherlands, bert.schouten@let.ruu.nl)

Subjects were asked to match sinusoidal tone glides, of various durations (30–210 ms) and sweep widths (50–600 Hz) around a center frequency of 1000 Hz, with two connected steady-state tones whose frequencies could be controlled independently. The expectation was that short glides would give rise to single perceptual events without any movement in a particular direction, so that the two matching steady-state tones would not show any frequency direction either; long glides, on the other hand, were expected to be perceived as rising or falling tones and matched accordingly. It was hoped that this would provide a basis for an explanation of diphthongs as unitary percepts. However, no effect of duration was

found at all; on average, glides were matched with tones whose frequencies corresponded to those reached by the glides after one-third and two-thirds of their course had been completed, regardless of duration.

3aSC3. Comparison of native English speakers' perception and production of the English vowel /i/. Elaina M. Frieda (Dept. of Psych., Univ. of Alabama, Birmingham, AL 35294)

This study examined the relationship between speakers' production and perception of the English vowel /i/. Participants were 24 male native English speakers representing diverse dialects. First, subjects produced /i/ in two speaking conditions: In condition one, subjects were instructed to speak normally; in condition 2, subjects produced exaggerated/hyperarticulated speech. Second, subjects completed 30 trials using a method of adjustment procedure where they selected their perceptually "preferred or prototype /i/" by sampling from a grid of 330 synthetic vowel stimuli that differed in *F1* and *F2*. Speakers differed substantially in their productions, as well as their prototype selections. Preliminary results demonstrate that subjects' prototypes were generally higher and fronter in the vowel space (i.e., had lower *F1* and higher *F2* values) than the vowels they actually produced in the normal condition. Moreover, for individual subjects, there was a negative correlation in *F2* values for these two conditions ($r = -0.40, p < 0.05$). However, the *F2* of the prototype and the hyperarticulated speech did not differ. In general, subjects both perceptually prefer and produce hyperarticulated vowels with extreme *F2* values. [Work supported by NIH.]

3aSC4. A model of silence duration adaptation in the perception of consonant cluster pairs. Michael A. Cohen (Cognit. and Neural Systems Dept., Boston Univ., Boston, MA 02215) and Krishna K. Govindarajan (MIT, Cambridge, MA 02139)

Listeners are able to adapt to the silence duration distribution between two different stop consonants $VC_1 - C_2V$ [K. K. Govindarajan and M. A. Cohen, J. Acoust. Soc. Am. **95**, 2978(A) (1994)]. This adaptation affects the likelihood of hearing just one stop consonant or the consonant cluster. An adaptation model is presented that emulates most of the individual subject data reported earlier; and based on Chi square goodness-of-fit tests, is capable of accounting for 98.5% of the variance of the psychometric functions. The model consists of a one-stage leaky integrator that averages

the silence duration of prior tokens. This average is subtracted from the current silence duration which is then corrupted by Gaussian noise of constant mean and variance. If the result is positive, the predicted response is hearing two stops, otherwise one stop is heard. The decay rate, the initial average, and the mean and variance of the noise were fit to the six subjects in a variety of experiments and conditions using maximum likelihood estimation. Reasons for the reliability of the constructed fit, and variability of parameters across subjects and conditions will be discussed. [Work supported by AFOSR, NIH.]

3aSC5. Perception of features in the identification of English consonants. Tobey L. Doeleman (Dept. of Linguist., Morrill Hall, Cornell Univ., Ithaca, NY 14853)

The perception of voicing, manner, and place of articulation was investigated in a series of gating experiments in which subjects identified English consonants in six different duration conditions. The stimuli were created from natural speech consonant tokens excised from C[a] syllables. The consonants were then gated to include 10% to 50% of the total consonant duration from the onset, to which vowel babble was appended in order to make listeners less sensitive to the absolute duration of the stimuli. In experiment 1, subjects were presented with eight consonants [p, t, b, d, f, s, v, z] and results were analyzed in terms of correct identification for the following features: voiced or voiceless, stop or fricative, and labial or alveolar. In experiment 2, subjects were presented with two mutually exclusive sets of stimuli (subsets of the stimuli from experiment 1) such that the information for a particular feature was given (e.g., voicing given in the sets [p, t, f, s] and [b, d, v, z]). A comparison of the results of the two experiments will show whether certain feature information facilitates the perception of any other feature. Results will be discussed in relation to a psychological model of speech perception.

3aSC6. An acoustic and perceptual study of Swiss German stops. William H. Ham (Cornell Univ., Ithaca, NY)

Oral stops in Swiss German show a three-way contrast, traditionally characterized as lenis /b, d, g/, fortis /p, t, k/ and long fortis /p:, t:, k:/. Fortis occur after long vowels, long fortis after short vowels, and lenis after both. Spectrographic analysis was performed on all stops in medial and final position for three Bernese speakers. Absence of voicing and lack of significant differences in voice onset times were noted, suggesting closure duration alone accounts for the three-way contrast: lenis are on average 57% shorter than fortis, and fortis 70% shorter than long fortis. Using the minimal set [ra:de] ('spin'), [ra:te] ('guess') and [ra:e] ('rat'), the salience of preceding vowel length relative to closure duration was examined in two perceptual experiments: The lenis stop was lengthened in 20-ms increments to long fortis length, and the long fortis stop was shortened to lenis length. The lengthened continuum shows a clear shift from lenis to fortis percept, but no shift to fortis long. The shortened continuum shows chance performance. Robust crossover effects in the lengthened continuum suggest vowel and consonant length are relevant to stop identification, while the shortened continuum suggests vowel length is primary.

3aSC7. Effects of preceding liquids and fricatives of different places of articulation on perception of stop consonants. Lawrence Brancazio (Haskins Labs., 270 Crown St., New Haven, CT 06511 and Univ. of Connecticut, lab93006@Uconnvm.uconn.edu)

In listener classifications of tokens along a /da/-/ga/ continuum, more /g/ judgments are given if the syllables are preceded by /al/ or /as/ than /ar/ or /a/ [Mann, Percept. Psychophys. 28, 407-412 (1980); Mann and Repp, J. Acoust. Soc. Am. 69, 548-558 (1981)]. These effects may reflect perceptual compensation for coarticulatory effects of consonants on a following stop: A front constriction causes fronting in the stop. Thus fronted preceding consonants induce more /g/ responses. These effects were investigated by contrasting /r/ with a "front" /l/ with an exaggerated tongue-

tip gesture and a "back" /l/ with no tongue-tip gesture. /s/ and /ʃ/ were contrasted with the velar /x/. These consonants were produced in /aC/ or /eC/ syllables, and were followed by synthetic /da/'s and /ga/'s. Both /al/ variants resulted in more /g/ judgments than did /r/. However, when the vowel was /e/, only the back /l/ produced this effect, contradicting the compensation explanation. The effect of /x/ relative to /s/ and /ʃ/ depended on the syllables' original context (preceding /da/ or /ga/). Finally, the syllables were tested with F2 /ba/-/da/ and /ba/-/ga/ continua. Results will be discussed with regard to gestural and acoustic explanations.

3aSC8. The effect of masking and filtering of envelope modulations on the recognition of consonants. René van der Horst (Dept. of Clinical and Experimental Audiol., Univ. of Amsterdam, Meibergdreef 9, Amsterdam, 1105 AZ, The Netherlands, r.vanderHorst@amc.uva.nl)

Consonant identification scores were obtained for 18 Dutch consonant phonemes in VCV context, in two experiments in which the narrow-band envelopes of the speech stimuli were manipulated by masking or filtering. For this purpose, the speech stimuli were split up into 26 frequency bands, each one critical band wide. The amplitude envelope of each band, as obtained with a Hilbert transform, was disturbed in two ways: (1) by a narrow-band envelope masker around 4, 8, and 12 Hz, and (2) by notch-band filtering around the same frequencies. Next the speech stimuli were synthesized by using narrow-band noise carriers, each a critical band wide, which were modulated by the aforementioned envelopes. Ten normal-hearing subjects participated in the experiments. The confusion matrices were analyzed by means of multidimensional scaling techniques and compared between the two manipulation methods. The results of the filtering experiment show that consonant recognition scores behave like a low-pass function of notch-width when expressed in octaves. The differences with the results of modulation masking will be discussed, especially with respect to modulation transfer function and consonant recognition cues.

3aSC9. Similarity scaling for consonants and consonant clusters in initial position. Rochelle S. Newman, James R. Sawusch, and Paul A. Luce (Psych. Dept., Park Hall, SUNY, Buffalo, NY 14260)

Knowing the perceived similarity between different consonants and consonant clusters in different syllable positions is important for models of word recognition that involve activation of neighborhoods of similar-sounding words. Similarity data may also help in understanding whether clusters are single entities, or are decompositional. Furthermore, the extent to which similarity is dependent on the particular talker is relevant for issues of talker normalization. In this study, participants rated the similarity of pairs of consonant-vowel and consonant cluster-vowel syllables in two different voices (one male, one female) over the course of five sessions. The similarity spaces will be discussed in relation to the issues described above. In general, there was no evidence for the decompositionality of consonant clusters. There was a high correlation ($r=0.81$) in the similarity spaces between the two voices. However, the results from a subset of the items presented in a third voice showed much lower correlations than did the results from the original two voices. This highlights the need to control factors such as the corpus of items, voices, and context when examining similarity scaling data. [Work supported by NIDCD Grant Nos. R01 DC00219 and R01 DC0265801-A1 to SUNY at Buffalo.]

3aSC10. The role of vowel duration in the perception of /E/ and /æ/. James R. Sawusch and Nancy J. Palmer (Dept. of Psych., SUNY at Buffalo, Amherst, NY 14260, jsawusch@acsu.buffalo.edu)

Perceptual studies have consistently shown that the dominant acoustic correlates of vowel perception are the frequencies of the first three formants. The present studies used natural syllables to explore the role of vocalic duration in vowel perception. The first study examined the vocalic duration and frequencies of the first three formants of American-English vowels. Out of 48 talkers, 15 were identified for whom the formant fre-

quencies were virtually identical in the CVC words “had” and “head” at the vowel mid-point. These tokens differed in duration and were highly intelligible. The second study used edited natural tokens from 4 of the 15 talkers to investigate the influence of vowel duration in perception. Pitch pulses from “had” were removed to create tokens varying in vowel duration from the natural “had” to a shorter duration similar to “head.” Pitch pulses from “head” were reduplicated to create a second series for each talker. Duration was used by some listeners for some talkers, but was not consistently used to differentiate these vowels in perception. This implicates other acoustic qualities, such as formant movements, as determinants of perception. [Work supported by NIDCD Grant No. R01 DC00219 to SUNY at Buffalo.]

3aSC11. Vowel stability may constrain selective adaptation. Jeffry A. Coady and Richard N. Aslin (Dept. of Brain and Cognit. Sci., Univ. of Rochester, Rochester, NY 14627, coady@bcs.rochester.edu)

Current models of selective adaptation for vowels implicate acoustic or auditory mechanisms rather than phonetic mechanisms. However, inconsistencies have been reported for /i/, /u/, and /ε/ [Morse *et al.* (1976)] in that the middle vowel /u/ did not elicit significant adaptation effects. These inconsistencies may reflect category goodness effects [Miller *et al.* (1983); Samuel (1982)] in that adaptation is more robust for tokens further from the category boundary. In the present study, category goodness was determined precisely for /i/, /u/, and /ε/ by obtaining within-subject rating data. Two continua were then constructed for each subject with these three vowels as the endpoints. Vowel adaptation was then assessed for the two continua using each subject’s three best exemplars as the adapting stimuli. Despite the fact that each subject’s adapting stimulus was, by definition, the “best” exemplar of that category, only /u/ and /ε/ showed significant adaptation. The absence of adaptation for /i/ may reflect its status as a point vowel, suggesting that extreme vocal tract positions, and their acoustic correlates, are remarkably stable. Research in progress on the point vowel /a/ will reveal whether this hypothesis is correct. [Work supported by NSF SBR94-21064.]

3aSC12. The spectral difference of different vowels: Toward a new acoustical concept. Dieter Maurer (Neurological Clinic, Univ. Hospital of Zurich, Haldenbach D1, CH-8091 Zurich, Switzerland, dmaurer@npsy.unizh.ch) and Andreas Klinkert (Univ. of Fribourg, CH-1700 Fribourg, Switzerland)

When isolated vowels were produced beyond F_0 of speech ($F_0 > 150$ Hz for men, $F_0 > 300$ Hz for women and children), the related formant patterns were found to deviate substantially from the values of formant statistics, and the formant frequency variations proved to be non-systematic (for details, see poster Klinkert and Maurer). Moreover, formant frequency estimation for vocalizations with $F_0 > 400$ Hz is highly problematical, although many vocalizations remain unambiguous in the vowel identity up to $F_0 = 700$ Hz and even above. The nonsystematic formant variations and the methodological problem of formant frequency estimation hardly allow for a normalization of formant patterns. Dynamic spectral properties do often not appear in isolated vowels. Thus, for the acoustic theory of vowel sounds, the question arises as to whether there is a concept to determine the physical correlates of the sounds apart from formants. A first approach toward such a concept is described for voiced vowels: A hypothesis is presented which predicts characteristic differences in the harmonic spectra of different vowels. Implications for vowel perception are discussed.

3aSC13. Fourier spectra and formant patterns of German vowels produced at F_0 of 70–850 Hz. Andreas Klinkert (Inst. of Informatics, Univ. of Fribourg, Rue Faucigny 2, CH-1700 Fribourg, Switzerland, andreas.klinkert@unifr.ch) and Dieter Maurer (Univ. Hospital of Zurich, CH-8091 Zurich, Switzerland)

Previous studies of vowel synthesis and of natural vowels have indicated: (1) a correlation between the lower formants and F_0 for $F_0 > 175$ Hz; (2) intelligibility and spectral differences of vowels with high $F_0 > 500$ Hz; (3) severe methodological problems of formant frequency estimation for $F_0 > 300$ Hz; (4) formant number alterations (appearance of different numbers of relevant formants for one vowel); (5) formant pattern ambiguity (appearance of the same formant pattern for different vowels). The purpose of this study was to replicate these findings. A large sample of the German vowels /i, y, e, ø, ε, a, o, u/ produced in isolation as well as in CVC context was investigated. There were 18 674 recordings made of 35 men, 44 women, and 20 children. Isolated vowels were produced at different levels of F_0 , and with different intensities. A listening test was performed. Fourier spectra, LPC spectra and spectral constancy were analyzed. The results strongly support the five indications mentioned. In particular, for isolated vowel sounds the correlation between the formants and F_0 , the appearance of “one-formant back vowels” (only one apparent formant < 2 kHz), and the formant pattern ambiguity for nonadjacent vowels proved to be very pronounced. Consequences for normalization approaches are discussed.

3aSC14. A phonation-type effect in vowel perception. Keith Johnson and Maria Paola D’Imperio (Dept. of Linguist., Ohio State Univ., 1712 Neil Ave., Columbus, OH 43210-1298)

Vowel perception is influenced by the fundamental frequency of voicing (F_0 normalization). This paper describes a similar effect resulting from a manipulation of phonation type. Acoustic characteristics of two phonation types were produced in formant-synthesized and LPC resynthesized stimuli. In the formant-synthesized stimuli, spectral tilt and open quotient were used to produce a flat-spectrum continuum (low-amplitude H_1 and high-amplitude spectrum above 1000 Hz) and a tilted-spectrum continuum (high-amplitude H_1 and low-amplitude higher frequencies). The flat-spectrum tokens sound like a pressed or creaky voice while the tilted-spectrum tokens sound breathy. In the LPC resynthesized stimuli a similar effect was achieved by replacing the LPC residual with voices synthesized using the LF model of voicing. Identification responses to both sets of stimuli showed reliable boundary shifts on an F_1 “hood”–“hud” continuum. There were more “hood” responses to the flat-spectrum continua than to the tilted-spectrum continua (higher F_1 boundaries). This voice quality effect is reminiscent of F_0 speaker normalization, but the creaky voice (which sounded more like a male speaker) was associated with a higher F_1 boundary. Data on perceived speaker characteristics of the stimuli will be reported. [Work supported by NIDCD R29-DC01645.]

3aSC15. Perceptual evaluation of alternate formant normalization models. Terrance M. Nearey and Debora A. Galama (Dept. of Linguist., Univ. of Alberta, Edmonton, AB T6G 2E7, Canada, nearey@nova.ling.ualberta.ca)

Two current theories of vowel normalization imply that vowel formant frequencies of one speaker can be related to those of another by a simple translation along a tonotopic frequency axis [T. M. Nearey, J. Acoust. Soc. Am. **85**, 2088–2113 (1989)]. Theory A uses a Bark scale, while theory B uses a log-frequency (musical) scale. To compare these theories, a baseline continuum is first constructed that spans the range appropriate for the front vowels of an adult male speaker. In test series A, the formant frequencies are translated upward by 1.2 Bark, corresponding to the difference between average adult male and average female and child values from a vowel data base. In test series B, an upward translation of 0.2 natural log units is made, based on the corresponding average values in a log space. (Fundamental frequencies are also changed by the appropriate average value in the appropriate scale.) Each theory predicts its associated continuum will

be identical to the baseline set, while that of the rival theory will show systematic distortion. Results of perceptual experiments will be presented that evaluate the phonetic distortion of the two test continua empirically.

3aSC16. Binaural enhancement of simultaneous vowels. D. Dwayne Paschall (Dept. of Commun. Disord., Texas Tech Univ. Health Sci. Ctr., Lubbock, TX 79409-2073, pdddp@ttacs.ttu.edu)

Previous investigations have shown that for monaural presentations, identification accuracy of simultaneous vowels improves when a precursor with the same F_0 and spectrum envelope is present which draws attention to a particular voice. In normal listening situations, however, one likely makes use of both ears rather than one ear only. The present investigation examined the ability of a dichotic precursor to enhance simultaneous vowel sounds presented to the left ear. A precursor was presented to the listeners whose intensity level in the left ear was the same as that used in previous investigations. However, the intensity of the precursor in the right ear was varied to produce different perceptual locations of the precursor: left ear only, 45° left of midline, midline, 45° right of midline, and right ear only. Thus the same monaural stimulus that has been shown to produce enhancement in previous investigations was present in each precursor location. However, the perceptual location of the precursor was different in every condition. Results from this investigation suggest that the perceived location of the precursor can have a significant effect on whether the precursor does or does not produce enhancement. If the perceived location

of the precursor occurs in a different place from the simultaneous vowel stimulus, the precursor is not effective in enhancing the identification performance of the added vowel.

3aSC17. Visual and auditory factors in vowel perception. Elizabeth A. Strand and Keith Johnson (Dept. of Linguist., Ohio State Univ., 1712 Neil Ave., Columbus, OH 43210-1298)

This study investigated four factors in vowel perception. Two of these factors are linguistic in nature: the $F1$ value and the identity of the visual word. The other two factors are speaker-related: the $F0$ value and the visual gender of the speaker. The auditory stimuli used in this experiment were tokens in continua ranging from "hood" to "hud" for both a male and a female voice which were produced using LPC resynthesis, maintaining the original male or female voice source and altering the formant values. These continua were then matched with movies of male and female speakers saying either "hood" or "hud," to create four series of stimuli. From identification judgments made by 20 listeners, we calculated 50 crossover points on the $F1$ "hood"–"hud" continuum. The boundaries were analyzed with a three-factor ANOVA (visual gender, visual word, and original voice). There were three main effects and no interactions. The visual word effect shows that visual and auditory cues are integrated in determining vowel quality. The original voice effect is essentially an $F0$ normalization effect. The visual face effect shows auditory-visual integration in the perception of speaker identity which, we hypothesize, has an impact on perceived vowel quality.

WEDNESDAY MORNING, 18 JUNE 1997

ROOM D, 8:30 TO 11:00 A.M.

Session 3aSP

Signal Processing in Acoustics: Applications of Spatio-Temporal Signal Processing

Roger Dwyer, Chair

Automated Systems Research, 43 South Cobblers Court, Niantic, Connecticut 06357

Contributed Papers

8:30

3aSP1. Optimal use of spatial bandwidth in underwater acoustic communications to enhance bandwidth efficiency. Daniel Kilfoyle, James Priesig, and Josko Catipovic (Woods Hole Oceanograph. Inst., Woods Hole, MA 02543)

The operation of underwater acoustic communication systems is often directly or indirectly limited by the available temporal bandwidth with performance measures, such as range and information rate, closely tied to the system bandwidth efficiency. Spatial diversity has typically been viewed as a secondary channel feature, either introducing undesirable intersymbol interference or, more favorably, offering a mechanism to offset fading. It is proposed to treat spatial diversity as spatial bandwidth and demonstrate higher bandwidth efficiency with existing power and temporal bandwidth constraints. Several spatial modulation algorithms are explored, including time variant transmitter beamforming where both waveform and choice of projector weights denote information, a spatial analog to multi-carrier modulation where each propagation path corresponds to a carrier, and a coordinated multi-user scenario where each projector emits independent information. Performance of each algorithm is shown to be superior to conventional phase-coherent modulation techniques that make no explicit use of spatial bandwidth. Results of a small-scale field experiment will also be given.

8:45

3aSP2. Array processing techniques for broadband mode estimation. Kathleen E. Wage, Arthur B. Baggeroer (Res. Lab. of Electron., MIT, Cambridge, MA 02139 and Appl. Ocean Phys. and Eng. Dept., Woods Hole Oceanograph. Inst.), and James C. Preisig (Woods Hole Oceanograph. Inst. and Northeastern Univ.)

Normal mode representations are useful in many aspects of underwater acoustics, e.g., modal acoustic tomography uses precise measurements of the travel time of broadband pulses to invert for environmental parameters. A thorough understanding of the resolution of mode estimators is crucial for tomographic applications. Typically, acoustic modes are used for describing and analyzing the temporal and spatial structure of narrow-band signals. As source bandwidth increases, however, mode estimation algorithms can no longer ignore frequency variations of the modal wave numbers, which result in temporal dispersion and frequency-dependent mode shapes. To address the issue, this study establishes an array processing framework for estimating the modal composition of broadband receptions. Within this framework, the time/frequency resolution tradeoffs inherent in the processing of modal pulse arrivals are explored. Specifically, the effects of mode shape frequency variations on standard modal beamformers are investigated. In addition to characterizing the performance of spatial filters for mode separation, this study considers the problem of designing receivers to compensate for the dispersion (pulse-spreading) introduced by the waveguide. Techniques developed through theoretical work have been

applied to data from the Acoustic Thermometry of Ocean Climate (ATOC) project. Some results of modal time series analysis for 1996 ATOC receptions are presented.

9:00

3aSP3. A spatio-temporal approach to acoustical imaging of laser-generated ultrasound. James V. Candy, R. Huber, Diane Chinn, and Graham H. Thomas (Lawrence Livermore Natl. Lab., Univ. of California, P.O. Box 808, L-437, Livermore, CA 94551)

In this paper an application of spatio-temporal array signal-processing techniques applied to broadband ultrasonic data gathered from a pulsed laser system is discussed. Using a laser source to heat a material specimen under test for flaws, a spatio-temporal processor capable of estimating the displacement field of the specimen is applied. The peak surface displacement is displayed as an image showing the initial source (displacement field) propagating throughout the material as well as any flaws (scatterers) that may be present within the specimen. Clearly, this method of imaging enables a unique methodology for nondestructive evaluation (NDE). Here, a pulsed laser generates an acoustic (ultrasonic) wave by heating the material and causing thermoelastic expansion. The resulting ultrasonic wave propagates throughout the material and is received by an array of interferometers created synthetically. Assuming a spherically propagating wave field, the processor creates an image of the field by estimating the peak surface displacement at a given location. The resulting image displays valuable information about the material and its inherent flaws, providing an effective method of nondestructive evaluation. In this work, the results are shown of applying this method to data synthesized from a sophisticated thermoelastic simulation and specimen available in the laboratory are tested.

9:15

3aSP4. Pulsed beam pre- and post-processing for localized inverse scattering. Timor Melamed (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215), Ehud Heyman (Tel-Aviv Univ., Tel-Aviv, Israel), and Leopold B. Felsen (Boston Univ., Boston, MA 02215)

Well-collimated short pulse wave packets, also termed pulsed beams (PB), provide new options for local forward and inverse probing of targets or of a propagation environment. The spatial-temporal resolution achieved under the PB excitation conditions furnishes an unambiguous measure of where the "physical" signal resides, in contrast to frequency-domain procedures that must rely on more intricate phase discrimination. In this paper, PB forward and inverse modeling is applied to scattering by three-dimensional weak (Born-type) inhomogeneities with finite support in an otherwise homogeneous background fluid. By performing space-time and wave number-frequency analysis and synthesis in a phase space setting, conventional slant-stack tomography is pushed to its ultimate localization by PB pre- and post-processing. The implications of this scenario with respect to resolution and related issues are discussed. It is shown that this strategy reduces the inversion of the illuminated space-time scattering cell to a *pseudo-one-dimensional problem* determined by the orientation of the incident and scattered beam axes. Illustrative examples are included.

9:30-9:45 Break

9:45

3aSP5. Optimum microphone placement for array sound capture. Daniel V. Rabinkin, Richard J. Renomeron, Joseph C. French, and James L. Flanagan (CAIP Ctr., Rutgers Univ., P.O. Box 1390, Piscataway, NJ 08855)

Microphone arrays can be used for high-quality sound pick up in reverberant and noisy environments. The beamforming capabilities of microphone array systems allow highly directional sound capture, providing superior signal-to-noise ratio (SNR) when compared to single microphone performance. There are two aspects in microphone array system performance: the ability of the system to locate and track sound sources, and its

ability to selectively capture sound from those sources. Both aspects of system performance are strongly affected by the spatial placement of microphone sensors. A method is needed to optimize sensor placement based on geometry of the environment and assumed sound source behavior. The objective of the optimization is to obtain the greatest average system SNR using a specified number of sensors. A method is derived to evaluate array performance for a given array configuration defined by the abovementioned metrics. An overall performance function is described based on these metrics. A framework for optimum placement of sensors under the practical considerations of possible sensor placement and potential location of sound sources is also characterized.

10:00

3aSP6. Applications of binary noiselike sequences to the simultaneous measurement of multichannel system functions. Douglas Rollow and David C. Swanson (Graduate Program in Acoust., Penn State Univ., P.O. Box 30, University Park, PA 16802)

A variety of binary sequences exhibit known correlation properties which can be exploited to simultaneously measure system functions in a multiple-input, multiple-output system. This work considers the practical application of codes such as linear maximal sequences and Gold codes for these measurements, and experimental data are presented for a fluid-structure interaction problem. Applications in multichannel active noise control systems and to the measurement of mutual coupling factors in transducer arrays are discussed. [Work supported by the Exploratory and Foundational fund, Applied Research Laboratory.]

10:15

3aSP7. Comparison of acoustic reflectance measurement techniques. Douglas H. Keefe (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68124, keefe@boystown.org)

Acoustic reflectance measurement techniques have been investigated that are calibrated using pressure responses measured in one or more cylindrical tubes and viscothermal models of wall loss. Of clinical interest as a test of hearing is a single-microphone technique to measure the inherently nonlinear ear-canal reflectance at various stimulus levels. Calibration of the probe containing the sound source/receiver consists of measuring the source reflectance of the probe and the incident pressure response at each stimulus level. A single-tube technique relies on separability of the tube response into incident and reflected signals, while multi-tube techniques rely on solving an overdetermined matrix equation of responses. The matrix includes weighting factors that differ across tubes by reflectance or SPL to control for inadequate signal-to-noise ratios at isolated frequencies, and a regularized matrix inverse solution is compared to the pseudoinverse solution. The single-tube technique functions as well as the best performing multi-tube technique, and requires fewer calibration steps. However, there is an ambiguity between the tube length chosen in the tube response model and the measured phase of the source reflectance. This ambiguity is addressed by *a priori* knowledge of the phase response of the probe averaged across frequency.

10:30

3aSP8. Identifying cross-correlation peaks with autocorrelation functions. John L. Spiesberger (Dept. of Meteorol. and the Appl. Res. Lab., 512 Walker Bldg., Penn State Univ., University Park, PA 16802, jspies@gamot.psu.edu)

Crosscorrelating the output of two receivers may fail to provide a means to estimate the arrival time difference of the first arriving signals when the crosscorrelation of accompanying multipaths have amplitudes that may be confused with that between the first arrivals. Based on information in two autocorrelation functions, a new method is described for identifying which peak in the cross-correlation function corresponds to the arrival time difference between the first arrivals [J. L. Spiesberger, J. Acoust. Soc. Am. **100**, 910-917 (1996)]. The technique relies on (1) a high signal-to-noise ratio following auto and crosscorrelation and (2) time series which have been preprocessed to suppress all signals except those ema-

nating from a single source. The technique works even when (1) the number of multipaths are unknown, (2) the received signal is not visible above the noise, (3) the transmission time is unknown, (4) the transmitted waveform is unknown, and (5) the spatial coordinates of the multipaths are unknown or are impractical to estimate. These techniques may be useful for localizing transient signals.

10:45

3aSP9. Compensation for array nonidealities by electronics. D.-L. (Donald) Liu, Robert C. Waag (Dept. of Elec. Eng., Univ. of Rochester, Rochester, NY 14627), and Clyde Oakley (Tetrad Corp., Englewood, CO 80111)

Array elements usually have nonuniform characteristics in terms of element position, sensitivity, and waveform shape. Crosstalk caused by acoustic and electrical coupling limits the directivity of individual elements. These array nonidealities can be compensated by using pulser elec-

tronics with programmable transmit waveforms. Experiments were performed with a 2048-element ring transducer and a 3×3 test block for a 64×64 -element 2-D array. With the ring transducer, a thin wire was suspended at the center of the ring, and single-element, pulse-echo data were collected for each element. To derive the one-way inverse response for each element, the two-way response was deconvolved using FFT, with special attention to phase unwrapping. Implementation of the inverse filtering reduced element arrival time fluctuation from 47 to 0.9 ns, energy level fluctuation from 1.2 to 0.4 dB, and increased the waveform similarity from 0.94 to 0.99. With the 3×3 test block, cancellation of crosstalk was performed by exciting neighboring elements with an opposing signal. The opposing signal was obtained by multiplying the original signal with a decoupling coefficient. Experimental results show that, with the particular test block, a decoupling coefficient of -0.08 extended the -6 dB half-angle from 10.8 to 14.8 deg, with a corresponding decrease of 4.5 dB in sensitivity.

WEDNESDAY MORNING, 18 JUNE 1997

ROOM R, 8:00 TO 11:35 A.M.

Session 3aUW

Underwater Acoustics: Chaos and Predictability in Long-Range Sound Propagation I

Frederick D. Tappert, Cochair

Applied Marine Physics, University of Miami, RSMAS, 4600 Rickenbacker Causeway, Miami, Florida 33149-1098

Michael G. Brown, Cochair

Applied Marine Physics, University of Miami, RSMAS, 4600 Rickenbacker Causeway, Miami, Florida 33149-1098

Invited Papers

8:00

3aUW1. An overview of ray chaos. Michael G. Brown and Frederick D. Tappert (RSMAS-AMP, Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149)

Important aspects of ray chaos will be reviewed, emphasizing topics which relate to the Hamiltonian structure of the ray equations. Topics to be discussed include integrable and nonintegrable ray systems, action-angle variables, the breakdown of canonical perturbation theory, the KAM theorem, area-preserving mappings, power spectra, Lyapunov exponents, and the notion of a "predictability horizon." Connections between ray chaos and chaos in the classical limit of quantum mechanics will be discussed. The notion of "wave chaos" will be introduced. The important, and as yet unresolved, question of whether ray-based wavefield representations break down under chaotic conditions at a range which is proportional to the logarithm of the acoustic frequency will be discussed. [Work supported by ONR.]

8:20

3aUW2. Chaotic light: A theory of asymmetric resonant cavities. A. Douglas Stone and Jens U. Noeckel (Yale Univ., Dept. of Appl. Phys., New Haven, CT 06520-8284)

Spherical and cylindrical dielectric cavities support high- Q modes due to total internal reflection of the trapped light. These modes are analogous to the well-known whispering gallery modes of acoustics. When such cavities undergo smooth convex deformations, the ray dynamics corresponding to these modes becomes chaotic in a universal manner determined by the Kolmogorov-Arnold-Moser theory of classical Hamiltonian mechanics. Using concepts from the theory of wave/quantum chaos, it is argued that the corresponding resonances broaden and become highly anisotropic in a universal manner, independent of the wavelength. Comparison with exact numerical results [J. U. Noeckel and A. Douglas Stone, *Nature* **385**, 45 (1997)] shows that the predicted behavior describes the emission anisotropy extremely well, but the resonance lifetime has important corrections which are conjectured to arise due to the dynamical localization of photons and to chaos-assisted tunneling. The theory is able to explain the lasing intensity profile observed in microdroplets and related experiments, and may eventually have applications to microlasers and in fiber-optic communications networks.

8:40

3aUW3. Long-time semiclassical dynamics of chaos. Steven Tomsovic (Dept. of Phys., Washington State Univ., Pullman, WA 99164)

Semiclassical methods, long used in most realms of wave physics, have recently experienced a surge of interest in quantum mechanics, especially in the context of highly chaotic systems. In this talk, discussion will center on a demonstration that accurate approximations of long-time quantum wave propagation can be constructed using the appropriate chaotic trajectories of the classical analog. It shows that pessimistic assessments of a breakdown time of $\tau = \ln(1/\hbar)/\lambda$ in the approximation, where λ measures the rate of exponential separation of trajectories, are incorrect. Recent advances will be briefly summarized. [Support from the National Science Foundation under Grant PHY-9421153 is gratefully acknowledged.]

9:10

3aUW4. Quantum chaos in a stochastic model. F. D. Tappert (Appl. Marine Phys., Univ. of Miami, RSMAS, 4600 Rickenbacker Cswy., Miami, FL 33149) and M. A. Wolfson (Penn State Univ., University Park, PA 16802)

An idealized problem in the theory of long-range ocean acoustic propagation through randomly fluctuating mesoscale structure is reformulated as a time-dependent quantum mechanics problem with a time-varying Markovian potential function. The effective Planck constant is $\hbar = \delta \varepsilon^{-2/3}$, where $\delta = (k_0 l)^{-1} \ll 1$ measures diffraction, and $\varepsilon \ll 1$ is the strength of the mesoscale structure. Also k_0 is the acoustic wave number and l is the scale length of the mesoscale structure. In the formal classical limit ($\hbar \rightarrow 0$), the parabolic ray equations are nonintegrable and the relative motion of two particles for small initial separations exhibits chaotic behavior characterized by a positive Lyapunov exponent, $\nu \sim \varepsilon^{2/3}/l$. It is shown that this physically relates to the exponential proliferation of caustics. A generalized wave kinetic equation (GWKE) is derived for the evolution in phase space of the two-particle Wigner function. The GWKE is analytically examined for $\hbar \ll 1$ by a novel boundary layer method, called the "extended quantum notch method," which yields several important results: First, the "log breakdown time (range)" $t_b \sim \nu^{-1} \log(\hbar^{-1})$ is obtained where semi-classical theory breaks down due to the saturation of caustics, and it is shown that this is the range where the scintillation index saturates to unity; and, second, a wave (quantum) manifestation of classical chaos is found to be the exponential decay of the scintillation index beyond its peak value. [Work supported by ONR.]

9:30

3aUW5. A review of data obtained from long-range acoustic propagation experiments in the ocean. Brian D. Dushaw (A.P.L.—Univ. of Washington, 1013 N.E. 40th St., Seattle, WA 98105-6698, dushaw@apl.washington.edu)

Broadband acoustic propagation data have been obtained at ranges up to 5 Mm and for frequencies as low as 28 Hz. The arrival patterns measured on the sound channel axis generally have well resolved steep ray arrivals, unstable or chaotic shallow angle arrivals, and a sharp SOFAR cutoff. The wander and spread of the steep ray arrivals increase with range roughly as predicted. The effect of internal waves on the propagation are evident: In most cases the shallow-angle acoustical energy has been severely scattered by internal waves, rendering the use of these data problematical. In addition, the internal waves cause a diffusion of energy away from the sound channel axis. At 28 Hz, more stability in the arrival patterns is evident, including the shallow-angle portion. The internal waves scatter the energy associated with the earlier, resolvable ray vertically and temporally, but with less effect than on the latest acoustic arrivals. Acoustic propagation at lower frequency apparently reduces the detrimental effect of internal waves on the coherence of the signals. Acoustical receptions at very deep receivers have "ray-like" arrivals in the shadow zone of the predicted time front. These "ray-like" arrivals are associated with the cusps or caustics of the time front, but they cannot be explained by the full wave equation or by internal wave scattering.

9:50

3aUW6. Predictability of acoustic multipaths over basin scales. John L. Spiesberger (Dept. of Meteorol. and the Appl. Res. Lab., 512 Walker Bldg., Penn State Univ., University Park, PA 16802, jspies@gamot.psu.edu), Frederick D. Tappert (Univ. of Miami, Miami, FL 33149), and Michael A. Wolfson (Penn State Univ., University Park, PA 16802)

The stabilities and instabilities of acoustic multipaths have been studied over basin scales as part of the development of studying climatic temperature changes in the ocean. In 1983, such transmissions began from the Kaneohe source where the propagation was complicated by the interaction of sound with the bottom near Oahu, Hawaii. For years, these transmissions exhibited stable arrivals only for those multipaths that were quite steep at the axis of the sound channel, for example, about 15 deg. Models indicate that the steeper multipaths may form "bounded-chaotic seas" that are useful for thermometric studies because their upper turning depths are insensitive to fluctuations. The sensitive nature of these paths is a consequence of the interactions with the bottom and the long distance of propagation, i.e., 3709 km. The flatter multipaths could not be used for inferring temperature variations because they were unstable. It is now understood that these paths are sensitive to fine and meso-scales which vertically scatter the energy over a kilometer, the same scale as the waveguide. [Work supported by ONR.]

10:10–10:20 Break

10:20

3aUW7. Secular behavior of chaotic ray solutions. Michael D. Collins and Joseph F. Lingeitch (Naval Res. Lab., Washington, DC 20375, collins@ram.nrl.navy.mil)

Does the topic of this session make sense? The existence of a connection between chaos and predictability in long-range sound propagation is investigated by asymptotically analyzing the ray solution. Along a regular ray, the Laplacian of the phase oscillates about zero so that the solution of the first transport equation [Jensen *et al.*, *Computational Ocean Acoustics* (American Institute of Physics, New York, 1994), p. 150] does not grow or decay exponentially. Along a chaotic ray, the Laplacian of the phase tends to be positive so that amplitude decays exponentially. The second transport equation (a correction that is not usually implemented) is identical to the first transport equation but with a forcing function that decays exponentially along a chaotic ray. This problem is closely related to the classical secular problem $y' + y = \exp(-x)$. Numerical techniques are applied to solve the transport equation and illustrate secularity, which implies that the ray solution is nonuniform and breaks down at long range when chaos exists. [Work supported by ONR.]

10:35

3aUW8. Exact rays for problems involving ray chaos. Joseph F. Lingeitch and Michael D. Collins (Naval Res. Lab., Washington, DC 20375, jfl@aslan.nrl.navy.mil)

Classical rays are useful for both solving problems and visualizing solutions. Exact rays [T. L. Foreman, *J. Acoust. Soc. Am.* **86**, 234–246 (1989)] are useful for visualizing solutions that have been obtained with other techniques. When chaos arises [Smith *et al.*, *J. Acoust. Soc. Am.* **92**, 1939–1949 (1992)], classical rays are of limited use. Exact rays should be useful even when classical rays are chaotic. This topic is investigated using parabolic equation techniques to generate solutions to problems that exhibit ray chaos. [Work supported by ONR.]

10:50

3aUW9. Application of phase screens to ray chaos. Martin A. Mazur and Kenneth E. Gilbert (Appl. Res. Lab. and Graduate Program in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804, mxm14@psu.edu)

In wave propagation in complex media, phase changes along the propagation path are often approximated by a series of abrupt phase changes at discrete planes called “phase screens.” It is shown here how phase screens can be used in an analogous way in ray-tracing problems.

The total travel time of each ray is formulated as a sum over phase screen and free space contributions. Fermat’s principle is then applied to the travel time, yielding a discrete mapping. The mapping connects the ray position at one phase screen to that at each succeeding screen. Examples of the method are given for both nonchaotic and chaotic ray tracing problems. The ray-tracing solution is compared to the wave solution calculated by the parabolic equation. By letting the separation between phase screens go to zero, the connection between continuous propagation and propagation with discrete transitions at phase screens is shown. [Work supported by the Penn State Univ. Applied Res. Lab.]

11:05

3aUW10. A numerical investigation of semiclassical predictability in the presence of chaotic ray motion. Michael G. Brown (RSMAS-AMP, Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149), Michael A. Wolfson (Penn State Univ., University Park, PA 16802), and Frederick D. Tappert (Univ. of Miami, Miami, FL 33149)

The question of whether ray-based acoustic wave-field expansions break down under chaotic conditions at a range proportional to the logarithm of the acoustic frequency is investigated numerically. In the assumed idealized environment, the ray equations define an Anosov system; all ray trajectories are chaotic. To address the log breakdown question, full wave simulations are compared to simulations based on the semiclassical Maslov integral representation of the wave-field. Because the Maslov integral remains valid at caustics, one might expect that it remains valid at ranges beyond those at which uncorrected ray theory starts to fail. [Work supported by ONR.]

11:20

3aUW11. Stochastic ray theory, chaos, and predictability. Michael G. Brown (RSMAS-AMP, Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149)

The motion of sound ray trajectories in deep ocean environments is considered, assuming that the internal wave-induced sound speed fluctuations can be treated stochastically. The ray equations then constitute a system of stochastic differential (Langevin) equations. Equivalently, the expected ray density—an approximation to the average acoustic intensity—can be shown to satisfy a Fokker–Planck equation. The Fokker–Planck equation can be interpreted in fluid mechanical terms as an advection-diffusion equation. This observation leads to the notion of a diffusion-limited (as opposed to diffraction-limited) estimate of the chaos-induced “predictability horizon.” [Work supported by ONR.]

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

*National Institute of Standards and Technology, Acoustics, Mass and Vibrations Group, Building 233, Room A147,
Gaithersburg, Maryland 20899*

D. F. Muster, Chair, U. S. Technical Advisory Group (TAG) for ISO/TC 108, Mechanical Vibration and Shock
4615 O'Meara Drive, Houston, Texas 77035

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 planning for the meeting of ISO/TC 108 with Subcommittees SC1 and SC5 (11–19 September 1997 in Berlin, Germany).

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.

WEDNESDAY AFTERNOON, 18 JUNE 1997

ROOM H, 1:30 TO 3:10 P.M.

Session 3pEA**Engineering Acoustics, Speech Communication and Signal Processing in Acoustics: Directional Microphones and Other Approaches for Improving Speech Intelligibility**

James E. West, Chair

Bell Labs., Lucent Technologies, 600 Mountain Avenue, Room 2D-338, Murray Hill, New Jersey 07974

Chair's Introduction—1:30

Invited Papers

1:35

3pEA1. Directional microphones for teleconferencing. Gary W. Elko (Acoust. Res. Dept., Bell Labs, Lucent Technologies, 600 Mountain Ave., Murray Hill, NJ 07974)

Directional microphones are spatial filters that can be designed to improve the sound quality of speech in audio telecommunication. At present, the most common type of directional microphone utilized in audio telecommunication is a simple first-order differential microphone having a hypercardioid directivity pattern. The directivity index (microphone gain in isotropic noise) for the hypercardioid microphone is 6 dB. For very noisy and reverberant environments more directional gain will be necessary. This talk will describe and discuss some of the microphone array systems that have been proposed in the last decade that address this problem.

3pEA2. The Huge Microphone Array (HMA). Harvey F. Silverman, William R. Patterson III (LEMS, Division of Eng., Brown Univ., Providence, RI 02912, hfs@lems.brown.edu), and James L. Flanagan (CAIP, Rutgers Univ., CoRE Bldg., Piscataway, NJ 08855-1390)

The Huge Microphone Array (HMA) is a collaborative effort between Brown University and Rutgers University that started in February 1994 to design, construct, debug, and test a real-time 512-microphone array system and to develop algorithms for use on it. Analysis of known algorithms made it clear that signal-processing performance of over 6 Gigafllops would be required; at the same time, there was a need for "portability," i.e., fitting into a small van, that also set an upper limit to the power required. It was essential that the array be able to be used in both large and small acoustic environments. These tradeoffs and many others have led to a unique design in both hardware and software. The hardware uses 128 fast, floating-point DSP microprocessors and is designed so that data flow is independent of data processing. This leads to an unusually simple software environment. This paper presents the full design and its justifications. Performance for a few important algorithms relative to usage of processing-capability, response latency, and difficulty of programming is discussed. [Work supported by NSF Grant MIP-9314625.]

Contributed Papers

2:25

3pEA3. A broadband microphone array technique for correlated interference rejection. Darren B. Ward (Acoust. Res. Dept., Bell Labs, Lucent Technologies, Murray Hill, NJ 07974)

A new method of broadband minimum variance beamforming is presented for use with microphone arrays in a reverberant environment. Conventional broadband algorithms (such as the well-known Frost algorithm) suffer from cancellation of the desired look-direction signal when employed in environments containing interference which is correlated with the desired signal. In this paper, a frequency invariant beamforming structure is used to perform spectral averaging. This reduces the correlation between the desired signal and the interference, thereby eliminating the signal cancellation problem. Such a technique is valuable for speech acquisition using a microphone array, since significant correlated interference will be present due to room reverberation.

2:40

3pEA4. Speech intelligibility in Grady Gammage Memorial Auditorium. Srilatha K. Kannan (Arizona State Univ., 1500 Bear Creek Pkwy., Apt. 1111, Euless, TX 76039, kkkannan@aol.com)

"The most important function of sound in theater is to ensure audience comfort in perceiving audio component of a performance" [Bracewell, *Sound Design of Theater* (1993)]. Background research has unveiled poor speech intelligibility at the Gammage Auditorium. This study aims to map speech intelligibility in different seating areas and measure the factors causing poor speech intelligibility. The instrument used to collect data will be the Techron TEF System 12 analyzer. This instrument uses modulation transfer function (MTF) and speech transmission index (STI) techniques. This is based on RASTI techniques by measuring the reduction of signal modulation in an acoustic transmission path [Keele, Jr. (1988)]. MTF is a

measure of how well the amplitude modulation of a signal is preserved when a signal is sent from one point to another. The measurement modulations at different frequencies are converted to a single value that indicates the speech intelligibility in the auditorium. This single value is STI. Measurements are also provided on data for signal-to-noise ratio and RT_{60} . A separate measurement of background noise level would indicate poor speech intelligibility due to excessive reverberation or background noise.

2:55

3pEA5. Prediction of speech recognition by Fletcher's articulation theory for normal-hearing and hearing-impaired listeners in noise-masked conditions. Christine M. Rankovic (Dept. of Speech-Lang. Pathol. and Audiol., Northeastern Univ., 133 Forsyth Hall, Boston, MA 02115)

Articulation theory is a computational model designed to predict speech recognition scores for nonideal listening channels. Fletcher's original formulation [H. Fletcher and R. H. Galt, *J. Acoust. Soc. Am.* **22**, 89-151 (1950)] has been overlooked, presumably due to its complexity and the availability of simplified versions such as ANSI S3.5-1969. The Fletcher and ANSI versions were compared on their ability to predict articulation scores using data collected by Rankovic *et al.* [*J. Acoust. Soc. Am.* **90**, 2319(A) (1991); **91**, 354-362 (1992)]. In those studies, consonant-correct scores were obtained from five normal-hearing and seven hearing-impaired subjects in the presence of octave-band noises centered on 0.5, 1, or 2 kHz and noise-masking patterns were obtained from all subjects in all conditions. The speech was amplified for the hearing-impaired subjects. For normal-hearing subjects, both models made good predictions of consonant-correct scores. For hearing-impaired subjects, the ANSI version overestimated scores for conditions with articulation indices (AIs) greater than about 0.5, whereas Fletcher AIs were smaller and predictions more accurate. Differences in the AI calculations that account for these results will be discussed. [Work supported by NIH.]

Session 3pID**Interdisciplinary: Hot Topics in Acoustics**

Gary W. Siebein, Chair

*Department of Architecture, University of Florida, 231 ARCH, P.O. Box 115702, Gainesville, Florida 32611-5702***Chair's Introduction—2:00*****Invited Papers*****2:05****3pID1. Hot topic in architectural acoustics: Auralization.** Richard H. Campbell (ECE Dept., Worcester Polytechnic Inst., 100 Institute Rd., Worcester, MA 01609)

High-speed computers and digital signal processing have brought us to the point where it is possible to hear a convincing simulation of the acoustic environment of an auditorium. A number of researchers around the world have contributed for years to this field and they have now assembled the major pieces of the puzzle to make this powerful tool, called "auralization," work well enough to judge acoustic quality. One needs a computer model of the acoustic space including placement of sources and receivers, an appropriate source signal, a binaural spatial transfer function from free-field pressure to eardrum pressure, and a headset which has a known equalization function from the electrical signal to the wearer's eardrum. A typical auralization session first runs the room computer model gathering sound ray arrival data at one receiver from a source. This receiver signal is then convolved with the binaural HRTF to produce a binaural electrical signal for the headset, which is equalized to achieve flat subjective response for a particular headset. The results are astonishing but positional accuracy in 3-D space is yet to be completely evaluated. Auralization provides an extra and independent layer of design security. Importantly, acoustic disasters can be avoided.

2:25**3pID2. Hot topics in acoustical oceanography.** Michael J. Buckingham (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0213)

Acoustical oceanography involves the use of inverse acoustic techniques to investigate physical and biological phenomena in the sea. Processes at the sea surface, including bubble formation from wave breaking, precipitation rates, and gas transfer across the air-sea interface, are all being interrogated using active and passive acoustic techniques. Novel methods for monitoring fish stocks, based on the acoustic properties of the swim bladder, are also receiving attention. The acoustic calls of marine mammals are being investigated as natural sound sources to characterize the ocean environment. New geo-acoustic models to marine sediments have been developed and incorporated into ambient noise inversion techniques for determining seabed properties. Time-reversal experiments have been conducted over ranges of several kilometers, successfully stabilizing acoustic arrivals and allowing fluctuations in the ocean to be characterized. Brief accounts of these topics will be presented.

2:45**3pID3. Hot topics in medical acoustics.** J. Brian Fowlkes (Dept. of Radiol., Univ. of Michigan Med. Ctr., Ann Arbor, MI 48109-0553)

Medical acoustics covers a wide variety of applications that are both diagnostic and therapeutic. One of the interesting aspects of acoustics is its capability to be apparently noninvasive (imaging of the fetus with ultrasound) and yet powerful enough to generate biological effects (destruction of kidney stones by shock waves). One new area which seems to be further expanding this capacity is ultrasound contrast agents. Most of the agents under development are very small, stable gas bubbles which can be injected into the bloodstream and still pass through the capillary beds. Some agents provide additional signals in medical imaging for several minutes, and some have the potential for therapeutic applications as well. Contrast agents are one of the fastest growing areas in medical acoustics, as evidenced by the increased research in this area and the number of major pharmaceutical manufacturers becoming involved. Research and development expenditures are expected to increase dramatically as a result and the potential impact on medical acoustics could be significant. Topics for discussion include detection of nonlinear signals, new imaging modes in medical ultrasonics, the "disappearance" or "reappearance" of agents when exposed to an acoustic field, potential bioeffects of agents, and their therapeutic applications such as use in drug delivery.

Session 3pMU**Musical Acoustics and Education in Acoustics: Role of Musical Acoustics in Teaching Acoustics, Mathematics and Engineering II**

Robert D. Collier, Cochair
Creare, Inc., Etna Road, Hanover, New Hampshire 03755

Ronald A. Roy, Cochair
Department of Mechanical Engineering, Boston University, 110 Cummings Street, Boston, Massachusetts 02215

Chair's Introduction—1:20***Invited Papers*****1:25**

3pMU1. Teaching musical acoustics as an introduction to engineering technology. Robert D. Collier (49 Eagle Ridge Dr., Lebanon, NH 03766)

The design and performance of musical instruments is a course which is designed to: (a) motivate first-year students to select engineering through hands-on learning, and (b) provide students in music and the liberal arts with an appreciation for its role in the enhancement of both music making and music listening. This course may be viewed as an alternative to traditional science courses. The mechanisms of sound generation illustrate the interdisciplinary factors involved in instrument design: materials and structures, aero and structural dynamics, vibration, mechanical systems and controls, and signal analysis. Most importantly, musical acoustics serves to demonstrate the significance of human factors in engineering design and performance. The core of the course is in the students' laboratory experience in which they investigate instruments of their own choosing and which is augmented by field trips and lecture demonstrations by noted musicians. This paper includes the course outline and structure, examples of student projects, and their response.

1:55

3pMU2. Musical acoustics in teaching engineering. E. Carr Everbach (Dept. of Eng., Swarthmore College, Swarthmore, PA 19081, ceverba1@swarthmore.edu)

Musical acoustics provides attractive examples that can be used to introduce engineering majors and nonmajors alike to critical concepts in engineering practice. Foremost among these is the concept of the eigenmodes and accompanying eigenfrequencies of a linear system. These can be wonderfully demonstrated via analysis of the overtone series of brass instruments (a one-dimensional system) and circular drumhead vibrations (a two-dimensional system). As an added bonus, important practical engineering concepts can be explored with these systems, such as the effect of nonlinearities and of nonideal properties on system behavior. Other natural topics include impedance, regenerative feedback, modal coupling, and frequency-dependent boundary conditions. [Work supported by NSF.]

2:25

3pMU3. Using musical acoustics to teach digital signal processing, scientific computing, and human-computer interface technology. Perry Cook (Dept. of Comput. Sci., Princeton Univ., Princeton, NJ 08540)

This talk discusses the role of musical acoustics in teaching engineering courses offered at the Stanford Center for Computer Research in Music and Acoustics, and in the Princeton University Department of Computer Science. For digital signal processing, musical acoustics is used to motivate and demonstrate mathematical concepts such as zeroes (multi-path cancellations in rooms), poles (resonant modes in a guitar), nonlinearity, and instability (what happens when one turns up the guitar amp). Simultaneous real-time visual and auditory displays build intuition and interest. For scientific computing, the pendulum provides a simple but meaningful first experience in numerical methods. By setting up differential equations from a physical description, noting the possibility of linearizing assumptions, and solving the system discretely on a computer, students discover how parameter values affect accuracy and stability. By solving both with and without the linearizing assumption, then placing a soundfile header on the output data, differences in frequency and waveform can be easily detected. Techniques are extended in plucked string and vibrating membrane simulations. For human-computer interface technology, musical instruments are perfect examples of effective human-artifact interfaces. Labs using computer sound synthesis clarify basic interaction concepts such as latency and cross-modal perception.

2:55–3:10**Discussion****3p WED. PM**

Session 3pNS

Noise: Active Noise Control

Karl M. Reichard, Chair

Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

Contributed Papers

1:00

3pNS1. An intelligent ventilation system to attenuate the intermittent transportation noise. Ramin Mohajeri, John Atkinson (Dept. of Mech. and Mechatronic Eng., Univ. of Sydney, NSW 2006, Sydney, Australia, ramin@arch.su.edu.au), and F. R. Fricke (Univ. of Sydney, NSW 2006, Sydney, Australia)

The level of intermittent transportation noise into the building is controlled by activating an intelligent control system to close the window when the noise occurs. This method is a possible solution for conflict between passive noise control and natural ventilation for places with a moderate climate. A knowledge-based control strategy has been implemented to control the position of the window using acoustic feedback. There are three major steps, training, source classification, and control, in which fuzzy logic and rule-based algorithms are implemented. If somebody plays music or speaks near the window, the intelligent window must not react even though the level of the sound may be higher than an approaching aircraft or vehicle. To identify an approaching noise source, centroid of the Wigner-Ville distribution and the peaks of STFT are used. The cross correlation between output data from indoor and outdoor microphones indicates the effect of outdoor transportation noise on the indoor noise level. A training software is considered to update the thresholds of noise classification. The experimental result of a prototype intelligent window is presented. A noise attenuation of more than 10 dB in L_{eq} and L_{max} in each single noise events is achieved. [Work supported by RTA of NSW.]

1:15

3pNS2. Active control of jet engine inlet noise. Jerome P. Smith, Ricardo A. Burdisso, and Chris R. Fuller (Vib. and Acoust. Labs., Dept. of Mech. Eng., Virginia Tech., Blacksburg, VA 24060)

This paper summarizes the recent advances in the active control of tonal inlet noise from an operational Pratt and Whitney JT15D turbofan engine. The most recent results show that multiple circumferential control source arrays reduce the radiation spillover to the sideline that is often observed with a single control array. An optimal control configuration results in near global attenuation in all directions of the directivity field, and a sound power reduction of 5.1 dB. A power reduction of 4.1 dB was obtained over a desired sector spanning the angles from 40 to 60 degrees in the far field using a properly configured control system. Results using a hybrid passive-active inlet have also been obtained showing that a passive liner in combination with an active control system will eliminate much of the sideline control spillover. An inlet-mounted sensing strategy using a wave-number observation technique was implemented and the preliminary results show the potential of the method for eliminating far-field microphones. The latest results utilizing hybrid passive-active control with far-field and inlet mounted sensors will be presented and discussed. [Work supported by NASA Langley Research Center.]

1:30

3pNS3. An alternative feedback structure for the adaptive active control of time-varying periodic or time-varying narrow-band disturbance. Martin Bouchard (Dept. of Mech. Eng., Univ. of Sherbrooke, Sherbrooke, PQ J1K 2R1, Canada) and Bruno Paillard (Univ. of Sherbrooke, Sherbrooke, PQ J1K 2R1, Canada)

For the feedback active control of noise and vibrations, there are still a few situations where today's adaptive control algorithms do not perform well (slow convergence and/or bad transient response). In this paper, a feedback structure suited for the control of time-varying periodic or time-varying narrow-band disturbances is inspired from a technique used in today's low bit-rate speech coders. The main concept is to separate the information and the processing pertaining to the plant from the information and the processing pertaining to the excitation signal (disturbance). Since the time constants of the variations of the plant are usually much longer than the time constants of the disturbance fluctuations, the adaptation effort can then be put on the prediction of the disturbance. Simulations will be presented, showing that either for low damping systems or multiharmonic disturbance signals, the algorithm presented in this paper can be compared favorably to a popular internal model control (IMC) structure that uses a stochastic gradient algorithm. [Work supported by I.R.S.S.T.]

1:45

3pNS4. Multi-channel active control of acoustic energy density. Scott D. Sommerfeldt (Dept. of Phys., Brigham Young Univ., 277 FB, Provo, UT 84602), John W. Parkins, and Jiri Tichy (Penn State Univ., State College, PA 16804)

Previous numerical work has indicated the potential advantages of minimizing the acoustic energy density for active control of sound in enclosures. Preliminary experimental results have generally confirmed this numerical work. A computerized scanning system has recently been implemented in an enclosure at Penn State Univ. University that provides greatly improved scanning capabilities to investigate the uncontrolled and controlled fields for active control experiments. Using this system, the active control of enclosed fields using energy density has been further investigated for both single-channel and multi-channel implementations. With the scanning capabilities, it is also possible to represent the experimentally measured field in terms of its modal components. Recent results that have been obtained will be presented, to show the control achieved using energy density, both in terms of the spatial dependence of the field, as well as in terms of its modal components.

3pNS5. Active noise control with multiple, correlated reference signals. Yifeng Tu and Chris R. Fuller (Vib. and Acoust. Lab., Dept. of Mech. Eng., Virginia Polytechnic Inst. and State Univ., VA 24060)

The application of active noise control has been extended from single noise source, one-dimensional acoustical fields to complicated multiple noise source, three-dimensional acoustical fields. The application to a multiple noise source environment usually requires multiple reference sensors to generate a complete vector of reference signals. A widely adopted control configuration feeds each reference signal into a different control filter. This approach suffers from the problem of ill-conditioning when the reference signals are correlated. The problem of ill-conditioning results in slow convergence rate and high sensitivity to measurement error especially when the FXLMS algorithm is applied. A practical method is presented to pre-process the reference signals based on the Wiener filter theory and Gram-Schmidt orthogonalization which overcomes these problems. Simulations and experiments have been conducted to verify the issues associ-

ated with multiple reference ANC when the signals are correlated. The results presented are consistent with the theoretical analysis. [Work supported by ONR.]

3pNS6. Effect of unlined duct reflections on coherence for active noise control applications. Patrick Marks and Chad Schweitzer (Digisonix, Inc., 8401 Murphy Dr., Middleton, WI 53562)

Because of a major concern regarding internal air quality (IAQ) issues, much of the current work at Digisonix involves unlined duct for HVAC systems. In addition to the well-known problems associated with these systems, such as an abundance of high-frequency noise, the large number of reflective surfaces in the acoustic paths of unlined duct systems can create long reflection times. This can lead to decreased measured coherence between input and error microphones, and may also result in poor cancellation performance. This paper will show that for a particular unlined duct system, a substantial increase in measured coherence can be achieved through changes in the way in which the coherence measurements are made. These changes and their practical implementation will be discussed.

WEDNESDAY AFTERNOON, 18 JUNE 1997

ROOM S, 1:00 TO 3:20 P.M.

Session 3pPP

Psychological and Physiological Acoustics: Physiological Processing, Speech Processing, and Auditory Impairment (Poster Session)

Janet D. Koehnke, Chair

Department of Speech Pathology and Audiology, University of South Alabama, UCOM 2000, Mobile, Alabama 36688-0002

Contributed Papers

All posters will be on display from 1:00 p.m. to 3:20 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:00 p.m. to 2:10 p.m. and contributors of even-numbered papers will be at their posters from 2:10 p.m. to 3:20 p.m.

3pPP1. Rate encoding of binaurally alternating low-frequency tone bursts in macaque A1. David H. Reser, Yonatan I. Fishman, Joseph C. Arezzo, and Mitchell Steinschneider (Dept. of Neurosci., Albert Einstein College of Medicine, Rm. 322 Kennedy Bldg., 1410 Pelham Pkwy. S., Bronx, NY 10461)

Integration of sequential components is an essential feature of auditory scene analysis. Humans consistently underestimate by 35–50% the frequency (F) of click trains, when the component clicks alternate between ears, and F exceeds 10 Hz [Axelrod *et al.*, J. Acoust. Soc. Am. **43**, 51–55 (1968)]. This suggests a boundary for integration of sequential components. Multiunit Activity (MUA) and current source density (CSD) in A1 of the awake monkey were used to explore neural correlates of this illusion. Phase-locked responses of cortical neuronal ensembles correlate with the perceptions of human listeners [Soc. Neurosci. Abstr. 22:637.18 (1996)]. Above 10 Hz, inhibition in thalamorecipient layers suppressed the response to each ipsilateral click in the train, in regions of high best frequency (BF). Here these observations are extended to low BF sites, which respond poorly to clicks. Responses to 25-ms low-frequency alternating tone bursts exhibit a boundary above 10 Hz, in agreement with click train data. The consistency of this physiologic boundary across the tonotopic map of A1 may represent an important neural substrate for assigning individual components to perceptual streams. [Work supported by DC00657 and the Institute for Music and Neurologic Function of Beth Abraham Hospital.]

3pPP2. Strings or springs: Insight into the nature of links in chick cochlear hair bundles. R. Keith Duncan (Dept. of Bioeng., 120 Hayden Hall, Univ. of Pennsylvania, Philadelphia, PA 19104, duncanrk@eniac.seas.upenn.edu), Marc D. Eisen, and James C. Saunders (Univ. of Pennsylvania, Philadelphia, PA 19104)

The present study considers the issue of symmetry in relative hair bundle motion. The stereocilia on *in vitro* chick cochlear hair cells were viewed in profile, stimulated by an oscillating water jet, and observed under stroboscopic illumination. The angular deflection of the tallest hair was set to either 10, 15, or 20 deg. The angular displacements for the tallest and shortest hairs were measured in the excitatory and inhibitory directions relative to the rest position. Motion of the tallest hair was approximately symmetric for each stimulus magnitude, while short hair motion was greater in the excitatory than inhibitory direction ($p < 0.01$). Moreover, short hair angular displacement with respect to the tall hair was greater in the excitatory direction ($p < 0.01$), but equal in the inhibitory direction ($p > 0.01$). These results suggested that stereocilia linkages were more analogous to strings than springs, in that they transmitted tension but collapsed under compression. Also, breakage of tip links in low calcium reduced relative motion between the tallest and shortest hairs, suggesting that motion in low calcium media was unchanged for the inhibitory direction, but dramatically reduced for excitatory displacements. [Work supported by NIDCD and NOHR.]

3pPP3. Resonant cavity dependence of auditory-nerve fibers of the frog *Rana pipiens*. Michael J. Anderson and Evangelia Micheli-Tzanakou (Dept. of Biomed. Eng., Rutgers Univ., Piscataway, NJ 08855, mikemike@eden.rutgers.edu)

Auditory electrophysiological experiments were conducted to study the hypothesis that sound transduction by frogs is dependent on the volume of a resonant cavity, which is made up of the middle ear, Eustachian tube, and mouth. Using a closed acoustic delivery system, the extracellular responses to the interaction of a 100-ms pure tone of varying frequency and a single-cycle sinusoidal pulse were recorded from the eighth-nerve, via a ventral approach. Results show that the resonant cavity system used in vocalizing is also involved in sound transduction. When the Eustachian tube is closed, by a small piece of oiled paper or Apiezon, there is a reduction in the mean spike count of the eighth-nerve fibers. Removing the occlusion restores the firing rate to normal levels. This information, together with knowledge of the fibers' best excitatory frequency (BEF) and inhibitability, is important because it provides insight into the peripheral origin (i.e., Amphibian papilla, Basilar papilla, or Sacculus) of the nerve fibers. Since some nerve fibers respond by phase locking to the stimulus, taking the FFT of the response also improves our understanding of sound transduction methods employed by the inner ear apparatus. [Work supported by NIH predoctoral fellowship GM17890 to MA.]

3pPP4. Discrimination of silent temporal gaps in sinusoidal markers by the budgerigar (*Melopsittacus undulatus*). Satoshi Amagai, Robert J. Dooling (Dept. of Psych., Univ. of Maryland, College Park, MD 20742, dooling@bss3.umd.edu), Craig Formby (Univ. of Maryland School of Medicine, Baltimore, MD 21201), and Timothy G. Forrest (Univ. of North Carolina, Asheville, NC 28804)

The tonal quality and temporal complexity of bird vocalizations have long suggested that birds may excel in auditory spectral or temporal processing. However, most psychophysical studies support the notion that birds have spectral and temporal resolving power roughly comparable to that measured in humans. Rarely have such psychophysical tasks involved both frequency and time cues. An intriguing aspect of gap-detection experiments with sinusoidal markers is the robust dependence on the frequency separation between the markers. In humans, silent gaps become more difficult to discriminate as the frequency difference between the markers increases. This frequency-dependent pattern of thresholds is well fit with a roex-filter model. Here, budgerigars were trained by operant conditioning to discriminate a change in the duration of a silent gap between two tonal markers. Frequency separation between the tonal markers had only a small effect on gap discrimination thresholds in budgerigars, suggesting an unusually broad peripheral auditory filter. This result stands in marked contrast to the results from humans and to estimates of the auditory filter in budgerigars using masking paradigms. Alternate explanations for the apparently broad tuning of the discrimination data will be considered. [Work supported by NIH grants DC-00198 and MH-00982 to RJD.]

3pPP5. Auditory filter shapes in the budgerigar (*Melopsittacus undulatus*) derived from notched-noise maskers. Jian-Yu Lin, Robert J. Dooling, and Michael L. Dent (Dept. of Psych., Univ. of Maryland, College Park, MD 20742, dooling@bss3.umd.edu)

Estimates of the filter bandwidths in the avian auditory system have traditionally come from either single-unit tuning curves or psychophysical masking studies using broad (critical ratio) or narrow (critical band) noises—procedures that are known to have certain limitations. The experiment reported here used the notched-noise method for determining the size and shape of the auditory filter at 2.86 kHz—the center of the budgerigar's range of hearing. Three budgerigars were trained by operant conditioning to detect a pure tone in a background of continuous noise and tested daily until thresholds stabilized. Masked thresholds were then obtained for noises of eleven different notch widths (five symmetric, six asymmetric) surrounding 2.86 kHz. Relative equivalent rectangular bandwidths were

evaluated using several procedures and these bandwidths were compared to other estimates of filter bandwidths in the budgerigar auditory system from critical ratios, directly measured critical bands, critical modulation frequency, psychophysical tuning curves, and the detection of gaps in pure tones. The size and shape of filter bandwidths in budgerigars will be compared to similar measures on other birds and other vertebrates. [Work supported by NIH Grant Nos. DC-00198 and MH-00982 to RJD.]

3pPP6. Detection of tones in noise by chinchillas using the probe-signal method. William P. Shofner and William A. Yost (Parmlly Hearing Inst., Loyola Univ., 6525 N. Sheridan Rd., Chicago, IL 60626, wshofne@luc.edu)

The probe-signal method was used to determine whether chinchillas listen selectively for a signal tone in the presence of continuous broadband noise. Animals ran in blocks of 40 trials: 32 trials were tone trials and 8 trials were blank trials. Of the 32 tone trials, the signal tone of 1000 Hz was presented on 24 trials; probe tones of 500 and 2000 Hz were presented on four trials each. The sound levels of the 500-, 1000-, and 2000-Hz tones were adjusted to be equally detectable, as determined from psychometric functions that were generated using the method of constant stimuli. The behavioral sensitivities as estimated by d' for the signal tone based on the probe-signal method were similar to or larger than the estimated d' based on the psychometric functions. The sensitivities for the detection of the probe tones were also similar to or larger than the sensitivities measured with the psychometric functions. The results indicate that chinchillas do not listen selectively to a single auditory filter, but rather appear to use a wide-band listening strategy. This apparent wide-band listening strategy may account for the higher detection thresholds observed in animals compared to human listeners. [Work supported by NIDCD.]

3pPP7. Acoustic reflex decay for pulse train signals. Namrata Pai, Joseph W. Hall III, and John Grose (Dept. of Otolaryngol.—Head and Neck Surgery, Univ. of North Carolina School of Medicine, 610 Burnett-Womack Bldg., CB7070, Chapel Hill, NC 27599-7070, jwh@med.unc.edu)

Currently, the acoustic reflex decay (ARD) test is useful in testing for auditory nerve tumors, but only at low frequencies, as even normal-hearing subjects show ARD to pure tones at frequencies higher than 1000 Hz. The purpose of this study is to develop a reliable high-frequency test of ARD and to achieve a better understanding of the mechanism underlying ARD. ARD was measured in eighteen normal-hearing subjects using pulse trains with rates of 50, 100, 200, 400, and 800 Hz. Each of these pulse trains was then bandpass filtered with a fixed high-pass setting of 4.5 kHz, and a variable low-pass setting of 5, 6, 8, or 10 kHz. In addition, random noise filtered to similar bandwidths was also tested. Results suggest that the acoustic reflex is resistant to decay for high-frequency pulse train stimuli with rates of 200 Hz and below, whereas rates of 400 Hz and above, as well as high-pass filtered noise, result in significant decay. There seems to be little effect of stimulus bandwidth. These data indicate a strong temporal component to acoustic reflex decay. [Work supported by the American Hearing Research Foundation.]

3pPP8. The effect of probe frequency on ipsilateral acoustic reflex thresholds. Vishakha W. Rawool (Commun. Disord. and Special Education, Bloomsburg Univ., Bloomsburg, PA 17815)

This study was designed to evaluate the effect of probe frequency on the ipsilateral acoustic reflex thresholds. Five young female subjects with normal hearing participated in the study. Ipsilateral acoustic reflex thresholds were established for each subject using 226-, 678-, and 1000-Hz probe tones. The criteria for reflex thresholds were based on the lowest hearing levels (HL) at which the following minimum changes in baseline admit-

tance were apparent on at least two trials: for 226 Hz, 0.02 ml; for 678 Hz, 0.06 mmho; and for 1000 Hz, 0.09 mmho. The reflex activator frequencies were 0.5, 1, and 2 kHz. The repeated measures multivariate analyses of variance were performed on the data. The results revealed a significant effect of probe frequency and a significant activator and probe interaction. Further analyses with the least significant difference (LSD) test revealed that the reflex thresholds obtained with the 678-Hz probe were significantly higher than those obtained with the 226- and 1000-Hz probes. In addition, for the activator frequency of 1 kHz, the lowest reflex thresholds were obtained with the 226-Hz probe, and for the activator frequency of 2000 Hz the lowest acoustic reflex thresholds were obtained with the 1000-Hz probe tone. [Work supported by a Bloomsburg University Grant for Research and Disciplinary projects.]

3pPP9. The relationship between ipsilateral acoustic reflex thresholds and loudness judgments in young normal women. Vishakha W. Rawool (Commun. Disord. and Special Education, Bloomsburg Univ., Bloomsburg, PA 17815) and Joseph P. Pillion (Kennedy Krieger Inst., Baltimore, MD 21205)

This study examined the relationship between ipsilateral acoustic reflex thresholds (ARTs) and loudness judgments. Eleven young women served as subjects. The ipsilateral ARTs were recorded for 0.5-, 1-, and 2-kHz activators using two probe frequencies: 678 and 1000 Hz. The subjects were also asked to judge the loudness of warble tones (0.5, 1, and 2 kHz) presented at various sound pressure levels in the following categories: not audible, very soft, soft, comfortable, loud, and very loud. From this data following levels were derived for each subject: very loud, average loud, maximum loud, average comfortable, and maximum comfortable. Correlation analyses were performed to examine the relationship of these levels with the ARTs. For 1 kHz, the maximum comfortable loudness levels were significantly correlated with the ARTs obtained with the 678-Hz probe tone and the average comfortable loudness levels were significantly correlated with the ARTs obtained with the 1000-Hz probe tone. For 2 kHz, the ARTs obtained with the 678-Hz probe tone were significantly correlated with the sound pressure levels that were judged as "very loud." For 500 Hz, none of the loudness judgments were predictable from the ARTs. Detailed analyses and possible implications of the findings will be presented. [Work supported by a Bloomsburg University Grant for Research and Disciplinary projects.]

3pPP10. Attention blinks in pure tone rapid auditory presentation. Kim M. Goddard (Dept. of Psych., Univ. of Calgary, Calgary, AB T2N 1N4, Canada, kgoddard@acs.ucalgary.ca), Matthew I. Isaak (Rice Univ., Houston, TX 77251), and Elzbieta B. Slawinski (Univ. of Calgary, Calgary, AB T2N 1N4, Canada)

When a target in a rapid (11 items/s) series of visual items is attended, the perception of a subsequent probe is impaired if the probe occurs within 500 ms. If this attentional blink (AB) reflects central processing limitations, then an AB should be found using auditory stimuli. Existing investigations, however, have used compressed speech stimuli, namely spoken letter or digit names. Because these stimuli are susceptible to visual recoding, the possibility that the AB reflects a vision-specific bottleneck cannot yet be dismissed. To redress this problem, pure tones were used in place of compressed speech stimuli. Participants heard rapid auditory streams comprising 20 tones ranging in frequency from 1000 to 2500 Hz. All tones were 50 dB except the target and probe, which were more intense on independent random halves of the trials. Participants detected the loud target and then detected whether the stream contained a second loud tone. Preliminary results show an auditory AB. The AB appears to reflect general, modality-independent limitations. Future work will examine the re-

lationship between the visual, compressed speech, and pure tone ABs. If the AB reflects central limitations, the magnitudes of the three ABs should correlate positively. [Work supported by NSERC.]

3pPP11. Effects of listening paradigm on speech recognition. Joan M. Basing, Janet Koehnke (Dept. of Speech Pathol. and Audiol., Univ. of South Alabama, UCOM 2000, Mobile, AL 36688-0002), Kim S. Abouchakra, and Tuyen V. Tran (U.S. Army Res. Lab., Aberdeen Proving Grounds, MD 21005-5425)

In everyday listening situations, binaural information enhances the ability to understand speech messages in multitalker environments. For applications to virtual auditory technology in such environments, the importance of preserving characteristics of the natural listening environment is not clear. The purpose of this study was to assess listeners' ability to monitor target (*T*) messages in the presence of synchronous competing (*C*) messages in an anechoic and a reverberant environment in three modes: (1) through loudspeakers at $\pm 45^\circ$ and $\pm 90^\circ$ with the subject seated in the actual room; (2) through the same loudspeakers with the subject listening remotely to the stimulus presented to a manikin in the actual room; and (3) with the environment simulated to create virtual sources at $\pm 45^\circ$ and $\pm 90^\circ$ azimuth under earphones. Fourteen subjects listened to messages selected from four lists of 2034 ten-syllable sentences. In every listening condition, the *T* message was presented 40 times to the subject in the presence of 0, 1, 2, and 3 *C* messages; subjects recorded the *T* messages. Results demonstrate that the ability to understand the *T* messages decreases as the number of *C* messages increases. As expected, speech recognition is better in the anechoic environment than the reverberant environment when *C* messages are present. Performance is degraded whenever the subjects do not listen with their own ears. [Work supported by the U. S. Army Research Lab.]

3pPP12. A functional neuroimaging study of speech recognition in noise. Amy R. Horwitz, Judy R. Dubno (Dept. of Otolaryngol., Medical Univ. of South Carolina, 171 Ashley Ave., Charleston, SC 29425, horwitar@musc.edu), Diana J. Vincent, and Mark S. George (Medical Univ. of South Carolina, Charleston, SC 29425)

Few studies have examined brain activity in response to auditory stimuli using functional magnetic resonance imaging (fMRI) due to the inherent loudness of the scanner noise. Here, the scanner noise has been incorporated into the design of a speech recognition in noise task. Normal-hearing subjects were scanned on a Picker EDGE 1.5-T MR scanner; structural and functional images were acquired for 15 coronal slices. In one active task, sentences included contextual information so that the final words were highly predictable from the other words in the sentence (PH); in the other active task, sentences had no contextual information so that the same final words were not predictable (PL). Using nonmetallic noise-reduction earphones, sentences were presented to subjects' right ears for 1-min blocks, alternating with 1-min blocks with no speech stimuli. Using a multiple subtraction technique, activation during rest was subtracted from activation in response to both low-context sentences (PL-rest) and high-context sentences (PH-rest). The difference between activation in response to low-context and high-context sentences was also determined (PL-PH). This analysis identifies areas of brain activation associated with processing acoustic/phonetic information within words with and without sentence context. [Work supported by NIH/NIDCD.]

3pPP13. Perceptual grouping of tone sequences by normally hearing and hearing-impaired listeners. Marina M. Rose and Brian C. J. Moore (Dept. of Exp. Psych., Univ. of Cambridge, Downing St., Cambridge CB2 3EB, England, mr209@cam.ac.uk)

This study examined the perceptual grouping of the rapid sinusoidal tone sequence ABA-ABA for normal and hearing-impaired listeners. Tone A was fixed at 250, 500, 1000, or 2000 Hz. Tone B started with a frequency well above or below that of tone A, and its frequency was swept

toward that of tone A. Listeners had to indicate when they could no longer hear two separate streams (fission boundary). For 12 normally hearing listeners, the fission boundary was roughly independent of the frequency of tone A when expressed as the difference in number of ERBs (ΔE) between A and B, consistent with a recent model of stream segregation [M. W. Beauvois and R. Meddis, *J. Acoust. Soc. Am.* **99**, 2270–2280 (1996)]. For four unilaterally hearing-impaired listeners, there was no consistent difference in the ΔE magnitudes across ears, even though the auditory filters were broader in the impaired ears. This is not consistent with the theory of Beauvois and Meddis. The ten bilaterally hearing-impaired listeners showed fission boundaries ranging from normal to well above normal. The results are discussed in terms of the factors that might influence perceptual stream formation in hearing-impaired listeners. [Work supported by the Hearing Research Trust.]

3pPP14. Laterality differences between children who clutter and stutter. Ingrid M. Blood, Gordon W. Blood, and Glen M. Tellis (Dept. of Commun. Disord., Penn State Univ., State College, PA 16801, i2b@psu.edu)

Cluttering is a disorder of speech and language processing resulting in rapid, dysrhythmic, sporadic, unorganized, and frequently unintelligible speech. It is an impairment in formulating language and can be accompanied by accelerated speech [Daly (1992)]. Numerous similarities have been reported between cluttering, stuttering, and central auditory processing disorders. This study examined cerebral dominance and processing using 120 trials of a consonant–vowel dichotic listening task in 15 subjects (5 children who stuttered, 5 children who cluttered, and 5 children with no disfluency problem). Results revealed that during the free recall task, significant right ear advantages (REA) were found for all control subjects. REAs were reported for 3 of 5 children who stuttered and 2 who had left ear advantages (LEA). Children who cluttered exhibited 4 no ear advantages (NEA) and 1 REA. During the directed right ear and directed left ear listening tasks, children who cluttered and control subjects showed strong ear advantages. Children who stutter showed mixed findings on these tasks. The discussion includes the use of dichotic ear advantages in differentiating children who stutter and clutter.

3pPP15. Audiologic findings in a cohort of adults with diabetes. Ingrid M. Blood, Peter R. Cavanagh, Jennifer Tufts, and Dana J. Boddorf (College of Health and Human Development, Penn State Univ., University Park, PA 16802, i2b@psu.edu)

The hearing status in 45 adults with diabetes (mean age = 59.87 years, s.d. = 9.83) and an age-matched control group of individuals was examined. The average duration of diabetes was 17.02 years (s.d. 9.66). Although hearing loss has been previously documented in patients with diabetes, the relationship between hearing impairment and diabetes is unclear. Etiology of hearing loss, whether cochlear or retrocochlear, has also been debated in the literature. Audiometric thresholds, word recognition scores in quiet and in competing noise of +6 dB signal-to-noise ratio, and transiently evoked otoacoustic emissions (TEOAE) and distortion product otoacoustic emissions (DPOAE) were obtained. Comparisons in audiometric results were made between diabetics with and without complications (30 with and 15 without neuropathy). Results revealed considerable variability in hearing sensitivity, audiometric configuration, and speech recognition in the presence of noise. Mean auditory thresholds and TEOAE and DPOAE values were calculated. Otoacoustic emissions were reduced in amplitude or absent in subjects with sensorineural hearing loss regardless of diabetic status. Relations between measures and clinical implications will be discussed. [Work supported by NIH.]

3pPP16. Modeling detectability and speech transmission quality of band-specific modulated-noise distortions. Martin Hansen and Birger Kollmeier (AG Medizinische Physik, Univ. Oldenburg, D-26111 Oldenburg, Germany)

A model is presented which was developed for the prediction of speech transmission quality of (low-bit-rate) speech codecs [Hansen and Kollmeier, ICASSP'97, paper #2056 (1997)]. The model is based on a quantitative psychoacoustical preprocessing scheme [Dau *et al.*, *J. Acoust. Soc. Am.* **99**, 3614–3622 (1996)] and was successfully applied to various speech codec test databases. This study presents measurements and modeling results of the detectability of band-specific modulated-noise distortions. Two sentences were used as stimuli. They were telephone-band-filtered and presented to subjects diotically via headphones in a sound-attenuating booth. To mimic band-specific speech code-like distortions, two types of distortion were applied to the stimuli: modulation with wide-band noise (MNRU) followed by critical-band bandpass-filtering, and vice versa. In an adaptive 2I-2AFC experiment, the modulation depth at threshold was measured as a function of the critical-band center-frequency. The measured results were compared with the objective speech quality measure and calculated for the same signals used in the two experiments. At threshold levels of modulation the computed objective quality measure was found to be *constant*. This indicates a direct monotonic relation between speech transmission *quality* and the *detectability* of the distortion.

3pPP17. Differences between binaural sound-field thresholds and monaural audiometric thresholds. William J. Murphy, John R. Franks (Bioacoust. and Occup. Vib. Sect., Natl. Inst. for Occup. Safety and Health, MS C-27, 4676 Columbia Pkwy., Cincinnati, OH 45226-1998), Stephanie L. Hall (Univ. Cincinnati, Cincinnati, OH 45221-0379), and Edward F. Krieg, Jr. (Natl. Inst. for Occup. Safety and Health, Cincinnati, OH 45226-1998)

Pure-tone audiograms under earphones and noise-band open-ear sound-field thresholds were collected from more than 150 subjects during hearing protector testing. The pure-tone audiograms were administered with 5-dB resolution to normal-hearing subjects (≤ 25 dB HTL) with an audiometer using TDH-50 headphones in standard cushions. Sound-field thresholds were determined by presenting a third-octave band noise in a reverberant room for the same subjects at resolution of 1 dB [Franks *et al.*, *Ear Hear.* **13**, 2–10 (1992)]. The subject's best pure-tone threshold for either ear at each frequency was converted to dB SPL for comparison to the sound-field thresholds. The transfer function between pure-tone, minimum audible pressure, to diffuse sound-field, minimum audible field, was calculated for each of the test frequencies 125–8000 Hz, including 3150 and 6300 Hz. A second experiment was performed with ten subjects where each received three pure-tone audiograms at 1-dB resolution, as well as three sound-field noise-band audiograms at 1-dB resolution. The ten-subject data set was used to test the robustness of the larger data set. The results may be useful for predicting the open-ear sound-field thresholds from a subject's pure-tone audiogram and in establishing guidelines about test precision of sound-field audiometry.

3pPP18. Forward masking patterns measured in cochlear implant subjects. Chandra S. Throckmorton and Leslie M. Collins (Dept. of Elec. and Comput. Eng., Duke Univ., P.O. Box 90291, Durham, NC 27708-0291, lcollins@ee.duke.edu)

Several psychophysical procedures have been proposed as measures of channel interaction in subjects implanted with a multichannel cochlear prosthesis. In order to compare these different measures of channel interaction, and to assess their impact on speech recognition, two psychophysical procedures were performed to measure the forward masking and electrode discrimination patterns elicited by electrical stimulation in subjects implanted with the Nucleus device. In this study, forward masking was measured for all combinations of the masker and probe electrodes. These results are presented and compared with the electrode discrimination data. The masking patterns are also compared to those described by Lim *et al.*

[J. Acoust. Soc. Am. **86**, 971–980 (1989)] and by Shannon [R. V. Shannon, *Hear. Res.* **12**, 1–16 (1983)]. Speech recognition data are also presented, and the relationship between channel interaction and speech recognition is explored. Finally, since the Nucleus device stimulates electrodes sequentially from base to apex, the potential loss in transmitted information that is indicated by the channel interaction data is discussed. [Work supported by NIH.]

3pPP19. Detection of amplitude changes in the presence of adjacent stimulation by cochlear implant recipients. John W. Hawks (School of Speech Pathol. and Audiol., Kent State Univ., Kent, OH 44242), Marios S. Fourakis (The Ohio State Univ., Columbus, OH 43210), and Peter N. Arkis (Warren Otologic Group, Warren, OH 44484)

Recent results [Hawks *et al.*, in revision for *Ear Hear.*] suggest that disambiguation of neighboring lax vowels by cochlear implant (CI) recipients using the spectral peak (SPEAK) speech coding strategy may be hindered by the lack of perceptually significant differences in the coding of $F1$ and $F2$ as represented in the electrode activation patterns. Since multiple adjacent electrodes are activated when coding these spectral prominences, it is important to know how much activation is necessary to elicit the percept of a spectral peak in the presence of adjacent stimulation. To investigate this issue, CI recipients and normal hearing adults were compared on a detection task. Continua of three-tone complexes were constructed with tone frequencies equal to the centers of the frequency bands encoded by three adjacent electrodes in the $F1$ region and again for the $F2$ region. Stimuli were varied in intensity of the center tone, incremented both above and below a fixed level for the bounding tones. An adaptive procedure was used to estimate subjects' thresholds for detection of a change in center tone amplitude from the pedestal level. Results will be discussed in terms of their significance to CI processor coding strategies of spectral information for vowels.

3pPP20. Central auditory processing in a patient with malignant glioma. D. M. Daly (DALY, Box 210 855, Dallas, TX 75211-0855)

In brain gliomas infiltrate white matter, compressing or destroying axons en passant. Such tumors rarely invade cortex; they alter local blood flow and cellular metabolism and produce seizures by corticopedal disruption and deafferentation. The case of a 15-yr-old girl who developed complex partial seizures (daze, mastication) at 13 yr is reported. EEG disclosed left anterior temporal and right mesial temporal foci. An unorthodox occipital approach removed a grade IV astrocytoma deep within the left temporal region, sparing auditory and visual functioning. The patient was tested with four sets of synthesized sounds [D. M. Daly *et al.*, *J. Neurophysiol.* **44**, 200–222 (1980)]. Pre-op auditory functioning was intact with vowels, BDG, and BW. Performance on GY was aberrant but better than chance with right ear (AD)/hand, right ear/left hand, and left ear (AS)/right hand, but nearly normal with left ear/hand; typically AD boundary increased >10 ms; AS decreased <15 ms. Post-operatively GY performance improved tenfold (LRCS). At one year, cumulative AD performance was diminished only by brief fixed-response episodes; at 2 yr performance was within normal limits. Results demonstrate distant electrophysiological foci resolved and performance improved over a considerable time once malignant irritative source was removed.

3pPP21. Frequency weighting functions for nonsense syllable test speech materials as estimated by a correlational method. Christopher W. Turner, Chiemi Tanaka, Jennifer Knapp, and Bomjun Kwon (Dept. of Speech Pathol. and Audiol., Univ. of Iowa, Iowa City, IA 52242, christopher-turner@uiowa.edu)

A correlational method [K. A. Doherty and C. W. Turner, *J. Acoust. Soc. Am.* **100**, 3769–3773 (1996)] was used to estimate the relative weights of four frequency bands for the 258-item UCLA nonsense syllable test in ten normal-hearing listeners. Random levels of corresponding band-

pass noise were added to the speech tokens on each trial. From the trial-by-trial data, correlations between the noise level and listener's performance yielded an estimate of the relative weight placed upon each band. The weighting functions were very similar across all listeners. This procedure appears to be a simple, fast, and reliable method for measuring speech recognition frequency weighting functions in individual subjects. The correlational weighting function showed differences from the frequency importance function derived using traditional articulation index (AI) filtering experiments. Some speech bands that were shown to contain significant speech information by the AI methods were not highly weighted when all four bands were presented to the listeners in the correlational procedure. It is hypothesized that these differences reflect the way in which listeners use speech information that is redundant across frequency regions. [Work supported by NIDCD.]

3pPP22. Consonant discrimination in noise with earplugs and muffs in combination. Sharon M. Abel and Deborah L. Spencer (Dept. of Otolaryngol., Univ. of Toronto, 600 University Ave., Toronto M5G 1X5, Canada, abel@mshri.on.ca)

The drawback for speech understanding of wearing sound attenuating earplugs and muffs in combination was investigated. Two groups of normal-hearing subjects were tested in backgrounds of low-frequency cable swager and speech spectrum noise, respectively. Consonant discrimination was assessed using the Four Alternative Auditory Feature Test [J. R. Foster and M. P. Haggard, *Proc. MRC Inst. of Hearing Res.*, UK, 9–12 (1979)], in a semi-reverberant chamber that met the specifications of ANSI 12.6-1984. Speech materials were presented free field at a level of 85 dB SPL, in quiet and mixed with noise levels of 80–95 dB SPL. Subjects were tested with the ears unoccluded and fitted with E-A-R foam plugs, Peltor muffs with ANR capability, and the plugs and muffs, with and without ANR operational, in combination. Preliminary results showed that consonant discrimination decreased with S/N ratio. There was no difference due to the protector condition for speech spectrum noise. In contrast, in the low-frequency cable swager noise, subjects showed greater variability in performance across listening conditions, with relatively poorer scores for the combinations. [Work supported by National Defence Canada.]

3pPP23. The effect of completely-in-the-canal hearing aids on localization performance: Normal hearing listeners. William R. D'Angelo, Linda J. Morris (AL/CFBA, Bldg. 441, 2610 Seventh St., Wright-Patterson AFB, OH 45433-7901), and Pamela J. Mishler (Veteran's Affairs Medical Ctr., Dayton, OH 45428)

The Completely-In-The-Canal (CIC) hearing aid, due to its placement in the ear canal, has acoustic advantages that are not offered by traditional amplification. Among these advantages are increased high-frequency gain, reduction of the occlusion effect, and subjective reports of improved localization ability over other types of aids. Quantitative data are readily available to verify the frequency response and occlusion effect; however, data on localization ability have been primarily anecdotal. A series of experiments has been designed to document the effect of CIC hearing aid use on auditory localization performance. A normal-hearing subject group, fit with low-gain CICs, was used in this phase; later studies will utilize hearing-impaired listeners and a variety of devices. The use of normal hearing subjects allowed an assessment of localization ability independent of the effects of hearing loss. Localization accuracy was measured using the God's eye localization pointing (GELP) technique [Gilkey *et al.*, *Behav. Res. Methods, Instrum. Comput.* **27**(1), 1–11 (1995)]. Results showed that subjects wearing the CICs were able to localize over the entire spatial presentation range of the stimulus ($\pm 180^\circ$ azimuth, -74° to $+90^\circ$ elevation). Localization accuracy data will be presented.

3pPP24. Insert earphones: Calibration and reference thresholds. Tom Frank (Dept. of Commun. Disord., Penn State Univ., S-A Moore Bldg., University Park, PA 16802)

Since their introduction in 1984, insert earphones (Etymotic ER-3A and EARTone 3A) have gained widespread popularity for numerous applications in audiology, hearing science, and for different hearing aid fitting

procedures. This informational poster describes several procedures that can be used for output calibration using different coupler types specified by ANSI, IEC, and ISO. This will be done by describing step-by-step calibration procedures. Further, ANSI and ISO insert earphone reference threshold levels for pure tones and speech will be presented for different coupler types.

WEDNESDAY AFTERNOON, 18 JUNE 1997

ROOM Q, 1:20 TO 3:20 P.M.

Session 3pSC

Speech Communication: Second Language Acquisition, Perception and Production (Poster Session)

Keiichi Tajima, Chair

Department of Linguistics, Indiana University, Memorial Hall 322, Bloomington, Indiana 47405

Contributed Papers

All posters will be on display from 1:20 p.m. to 3:20 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:20 p.m. to 2:20 p.m. and contributors of even-numbered papers will be at their posters from 2:20 p.m. to 3:20 p.m.

3pSC1. Cross-linguistic speech rhythm in a phrase repetition task. Keiichi Tajima (Dept. of Linguist. / Cognit. Sci. Program, Indiana Univ., Bloomington, IN 47405, ktajima@indiana.edu)

Language-specific and language-general aspects of speech rhythm were explored in a repetitive speech production task, whereby native English and Japanese speakers repeated phrases in time with an auditory metronome. Subjects were instructed to repeat each phrase (e.g., English "Give the dog a bone") once per metronome beep, aligning the beginning of the phrase with each beep. The rate of repetition was systematically manipulated by varying the metronome rate from slow to fast across trials. Rhythmic properties of the utterances were investigated by marking the vowel onset of each syllable, and expressing it in terms of phase with respect to the beginning of the current repetition and that of the following repetition. Preliminary results indicate that, while there was considerable variability across speakers and repetitions, there was a language-general tendency for prominent syllables to cluster around rhythmically motivated phases, such as 1/2 or 2/3. For a subset of the data, abrupt jumps from one stable phase to another were also found, suggesting quasi-categorical modes in rhythmic speech production. The data further address language-specific issues such as the nature of stress-timing in English, and the role of supra-moraic, "foot"-level units in Japanese rhythm.

3pSC2. Learning to identify a foreign language. Z. S. Bond and Verna Stockmal (Linguistics, Ohio Univ., Athens, OH 45701)

Listeners are able to identify spoken samples of foreign languages at better than chance levels. They are able to do this even when they do not know a language, presumably by using its general phonological properties, that is, its "acoustic signature." It is not known how much exposure to a foreign language is necessary to learn to identify it. Two experiments were conducted in order to determine the ease with which American listeners could learn to identify the target language, Japanese. The first experiment examined exposure time and materials. Listeners heard word lists or a short children's story; when tested immediately after exposure, listeners could identify Japanese interspersed with six other languages better than a control group. The materials did not affect listener performance. In the second experiment, listeners were presented the same materials but tested after a 45-min delay. This task proved much more difficult, although listener performance was marginally better than that of the control group.

Apparently, listeners are able to develop a representation of the way a foreign language sounds after a brief exposure, but the representation fades rapidly.

3pSC3. Sentence duration as an index of overall proficiency in an L2. Susan G. Guion, James E. Flege, Hua Liu (Dept. of Rehabilitation Sci., VH503, Univ. of Alabama, Birmingham, AL 35294, guionsg@biocom1.bioc.uab.edu), and Grace Yeni-Komshian (Univ. of Maryland, College Park, MD 20742)

Maximum reading rate in the native and second language has often been used to assess bilingual balance. In the present study, the duration of L2 sentences produced at a nonmaximal rate was measured. The subjects were 24 native English (NE) speakers and 240 native Korean (NK) speakers assigned to ten subgroups based on age of arrival in the US (2–22 yrs). The subjects repeated 21 aurally presented English sentences following a delay. The durations of sentences that were produced without pauses or repetitions were measured. The NK subjects who arrived in the US after the age of 9 years produced sentences that were significantly longer than did the NE subjects. Significant correlations were found to exist between the Koreans sentence durations and age of arrival in the US: (1) self-estimated use of both Korean and English; (2) the number of segmental errors produced in English sentences; and (3) knowledge of English morphosyntax (assessed by a 144-item Grammaticality Judgment Test). The correlation with amount of education in the US was nonsignificant. Taken together, these results suggest that sentence duration may serve as a useful index of overall proficiency in an L2. [Work supported by NIH.]

3pSC4. A longitudinal study of native Japanese speakers' production of English liquids. Timothy J. Riney (Biocommun. Lab., VH503, UAB, Birmingham, AL 35294-0019, rineytj@biocom1.bioc.uab.edu)

This study investigates English liquids [l] and [ɹ] produced by 11 Japanese students at two times during their first and fourth years at a bilingual university in Tokyo (T1, T2). Five native English speakers from the same university serve as the control group. Each speaker read the same word list at T1 and T2. The word-initial CV portion of 50 words that contained

liquids were subjected to auditory evaluation to determine if (1) the liquids were produced more accurately at $T2$ than at $T1$, and (2) to what extent the liquids, $[l]$ and $[r]$, were identifiable as $[l]$ and $[r]$, respectively, at $T1$ and $T2$. For question (1), paired comparison judgments made by ten naive, native speakers of American English were obtained. For question (2), forced-choice identification by three trained listeners were obtained. The presentation of the results will separate the identification and paired comparison judgments for $[l]$ and $[r]$ as singletons, and for $[l]$ and $[r]$ in clusters in contrastive and noncontrastive environments. The findings are discussed in light of current models dealing with adult acquisition of second language pronunciation.

3pSC5. Arabs' stress placement in English: Intelligibility test. Fares Mitleb (English Dept., Yarmouk Univ., Irbid, Jordan)

This study was intended to give an impressionistic description of Arabs' performance in English stress placement. Therefore, two native British English speakers were asked to evaluate six Arabs production of English words. Findings indicated that Arab subjects' "conscious knowledge" of a stressful speaking situation was less than satisfactory. Listeners reported a great percentage of confusion in stress placement. A number of explanations were attempted for the intelligibility problem which arises from the misplacement of stress in our subjects' production. It was inferred from confusions Arabs made that suprasegmentals are to be given emphasis in the teaching of foreign languages and should be taught simultaneously with segmental structure of the target tongue. [Research supported by Yarmouk University.]

3pSC6. The identification and discrimination of a non-native nasal series by Bengali and American English listeners. James D. Harnsberger (Program in Linguist., Univ. of Michigan, 1076 Frieze Bldg., Ann Arbor, MI 48109, jharns@umich.edu)

Predictions based on the perceptual assimilation model (PAM) (Best, 1995) concerning the perception of non-native contrasts rely primarily on three factors: the contrast in question, the inventory of native categories, and the degree of gestural similarity between them. To determine if these factors are sufficient for predicting listeners' difficulties, nasal contrasts from three languages, Malayalam, Marathi, and Oriya, varying in place of articulation (bilabial, interdental/dental, alveolar, and retroflex), syllabic position, and vowel context, were presented in an AXB discrimination test and an identification test with category-goodness ratings to sixteen speakers each of American English (AE) and Bengali. AE and Bengali were selected for their similar nasal series ($[m]$, $[n]$, and $[ŋ]$). However, Bengali differs from AE in its use of dental and retroflex oral stops. PAM should predict equivalent performance for both subject groups. However, a pilot of this study demonstrated poorer performance by Bengali listeners on dental-retroflex (34.6% mean correct to AE 53.6%) and interdental-alveolar (33.2% to AE 69.6%). The implications of these results for phonological and phonetic factors in the perception of non-native contrasts will be discussed.

3pSC7. Non-natives' identification of English consonants in noise. Ian R. A. MacKay (Dept. of Linguist., Univ. of Ottawa, 70 Laurier Ave. E., Ottawa, ON K1N 6N5, Canada, imackay@aix1.uottawa.ca), D. Meador, and James E. Flege (Univ. of Alabama, Birmingham, AL 35294)

This study examined the identification of English consonants presented in noise. The subjects were English native (NE) speakers and four groups of Italian/English bilinguals who differed according to age of arrival (AOA) in Canada and/or self-estimated percentage daily use of Italian (%ITAL). The naturally produced stimuli, which differed according to their initial or final consonant (e.g., "bado," "sado," "hodab," "hodas"), were presented at decreasing S/N ratios (+12, +6, 0, -6 dB). Subjects identified each consonant by selecting one of five written alternatives, including the target and four likely foils. Percent correct identification

scores decreased similarly across the four S/N ratios for all five groups. However, the bilinguals scored lower than did the NE subjects, despite their long residence in Canada ($M=35.4$ years). Their scores varied inversely with AOA and %ITAL, and were related to their performance on a test of phonological short-term memory. Which consonants were most difficult to identify varied across the five groups. Error analyses revealed that errors due to confusions of place and manner of articulation were similar across the five groups. Additional analyses will determine whether the influence of the L1 can be characterized in terms of finer-grained features.

3pSC8. Perception of English /n l/ with amplified consonants or transitions by native speakers of Chinese. Anna Marie Schmidt (School of Speech Pathol. and Audiol., Kent State Univ., P.O. Box 5190, OH 44242) and Amy C. Kaminski (Kent State Univ., Kent, OH 44242)

Previous work has shown that speakers of some southern Chinese dialects have difficulty perceptually differentiating English /n l/. In this study, an attempt was made to increase the perceptual salience of some possible acoustic cues. Seven native Chinese speakers labeled the first consonant of three combined sets of ten English /n l/ minimal pairs produced by a male and a female native English speaker. One set consisted of unaltered words from the earlier study. The consonant portion of the second set had been normalized for peak amplitude. The peak amplitude of the consonant was then equal to the peak amplitude of the vowel. In the third set, a portion of the CV transition had been normalized. Native English speakers labeled the words with high accuracy. Results and implications will be discussed.

3pSC9. Context effects on English word-final stops by American and Mandarin speakers. Yuan Yi (Dept. of Linguist., Indiana Univ., Bloomington, IN 47405)

The production of English word-final stops by Mandarin and English speakers in five contexts was examined. While studies show that the phonetic realizations of segments in native English speech are context sensitive [P. A. Luce and J. Charles-Luce, J. Acoust. Soc. Am. **78**, 1949-1957 (1985)], such contextual effects have not been investigated in non-native speech. Six Mandarin speakers of English and six American speakers read 10 carrier sentences containing the test words *back* and *'bag* in utterance-final position or followed by words beginning with /s/, /t/, /n/, and the vowel /ə/. Three repetitions of the words in each context were measured for the vowel duration and the word-final stop closure duration. The closure/vowel ratio was calculated, and the two subject groups were compared. Results show that for both groups, the vowel and closure durations varied with contexts: while utterance final position showed the strongest effect, among the other contexts, /n/ and /t/ triggered longer vowel and closure durations than /s/ and /ə/ did for both speaker groups. The non-native speakers' closure durations and the closure/vowel ratios were not nativelike and their articulatory gestures for the velar stops differed from those of native English speakers.

3pSC10. Nonnatives' perception of English sentences presented in noise. D. Meador, James E. Flege (Dept. of Rehabilitation Sci., Univ. of Alabama, VH 503, Birmingham, AL 35294), and Ian R. A. MacKay (Univ. of Ottawa, Ottawa, ON K1N 6N5, Canada)

This study built on earlier investigations of the effects of noise on speech identification by nonnative speakers. Oyama showed that the ability of native Italian subjects to identify speech in noise was inversely related to their age of arrival (AOA) in the United States [J. Psycholinguist. Res. **5**, 261-283 (1976)]. Mayberry and Fischer showed that more phonological and fewer semantic errors are made the later sign language is acquired, indicating that late learners allocate more attention to bottom-up processes [Mem. Cognit. **17**, 740-754 (1989)]. Subjects in the present study were

native English (NE) and Italian/English bilinguals differing in their AOA in Canada. The stimuli were naturally produced, semantically unpredictable English sentences. Subjects repeated as much of each sentence as possible as it was presented at successively higher S/N ratios ($-6, 0, 6, 12$ dB). NE subjects outperformed the bilinguals, despite the latter's lengthy

residence in Canada ($M=35.4$ years). The bilinguals' performance varied inversely as a function of AOA. Scores were also related to individual differences in phonological short term memory. Additional analyses will determine whether late learners allocate more attention to bottom-up processes than do NE subjects. [Work supported by NIH.]

WEDNESDAY AFTERNOON, 18 JUNE 1997

ROOM D, 1:30 TO 3:15 P.M.

Session 3pSP

Signal Processing in Acoustics: General Signal Processing

Mark Mahon, Chair

Applied Research Laboratory, Pennsylvania State University, New ARL Buidling, P.O. Box 30, State College, Pennsylvania 16804

Contributed Papers

1:30

3pSP1. Reduction of requirements and restrictions for near-field acoustic holography in cavities. Kelly Q. Kay (Virginia Polytech. Inst. and State Univ.—Virginia Consortium of Eng. and Sci. Universities, MS 463, NASA Langley Res. Ctr., Hampton, VA 23681, kay@actif.larc.nasa.gov)

Recently, Williams and Houston developed a near-field acoustic holography (NAH) method for generic cylindrical cavities [Proc. Internoise 96, 2553–2558]. Previously, cavity NAH methods required restrictions or gross assumptions, which limited their use to specific cases. The Williams–Houston method, on the other hand, is valid inside any practical cylindrical boundary. The method, though, requires prior measurement of the near field of all boundaries, whether they contain sources or not, even when boundaries are locally reacting with a known acoustic impedance. Also, because of singularities in the Green's function used, reconstructions could not be performed at some frequencies. This work is based on the Williams–Houston method, and presents several significant improvements and contributions. Most importantly, scattering off locally reacting cavity boundaries is accounted for in the Green's function, reducing the required number measurements and processing, for a given resolution. As well, Green's functions are solved for with consideration given to the uniqueness and existence of the acoustic field to allow reconstructions to be performed over a continuous range of frequencies. A set of cavity NAH Green's functions are developed in Cartesian coordinates, extending the method to parallel walled rectangular (box) cavities. Reconstruction results comparing the methods will be presented.

1:45

3pSP2. Echo components for aspect-independent classification. Charles F. Gaumond (Acoust. Div., Naval Res. Lab., 8555 Overlook Ave. SW, Washington, DC 20375)

Monostatic, backscattered data from a finite, ribbed cylinder are shown. These data from a nonspherical scatterer exhibit a great deal of angular dependence. This angular dependence complicates the design of a classifier which uses templates to test the similarity of a test signal with the known scatterer. The scattering data have previously been shown to primarily consist of components generated by specular reflection, elastic waves, and Bloch waves. The angular dependence of this data set is shown to be due to fading of each component from a physical mechanism, as well as the changing interference of the various components as their relative timing changes. The separation of the echoes into components is shown to decrease the angular dependence. Finally, it is shown that a modest number of templates can be derived from the angular-dependent components through the use of the Karhunen–Loeve transformation and the assumption of equal statistical weighting of each aspect angle. The ka range considered extends from approximately 3 to 10. [Work supported by ONR.]

2:00

3pSP3. Modeling temporal variations in the room impulse response. Nancy T. Tharakan, Anthony Kozloff, Charles Thompson, and Vineet Mehta (Ctr. for Adv. Computation and Telecommun., Univ. of Massachusetts, Lowell, MA 01854)

The objective of this study is to determine the impact of a time-varying room impulse response on the rate of adaptation and accuracy of monaural and stereo echo cancellers. The variation in the impulse response results from secondary echoes generated by moving objects in a room. The time-invariant portion of the response is modeled using the method of images. The time-varying component will be modeled using a Pade approximant-based solution of the Kirchhoff integral equation. Simulation results are given for a room having a volume on the order of 2744 cubic feet. It is shown that for strong acoustic scatterers located closer than the critical distance to the receiver can induce up to a 15-dB residual error for monaural cancellers. Adaptation time requirements for acoustic echo cancellers will also be discussed.

2:15

3pSP4. Time-frequency analysis adapted to the spectral characteristics of the signal. Salvador Cerdá and José Romero (Lab. de Acústica, Dept. de Física Aplicada, Facultad de Ciencias Físicas, c/o Doctor Moliner, Burjasot 46100, Spain)

The introduction of the wavelet transform (WT) and the analysis of multiresolution around the end of 1980, assumed an increase in the interest in signal processing and analysis. The development of the WT can be thought of as an attempt to improve the traditional analysis accomplished through STFT. However, STFT and WT outline resolution problems too. Attempts at improving the WT are the wavelet packets, and the designated M -band wavelet analysis. These analysis methods are similar in that they do not consider the characteristics of the signal for purposes of analysis. A method that takes into account these characteristics is the “matching pursuits” of Mallat and Zhang. This method selects from the dictionary of time-frequency atoms those atoms that are more representative in the signal. Introduced in this work is an analysis method that is adapted to spectral characteristics of the signals. First, the spectral characteristics are determined. The transform of adjustment of the window size (TAWS), is an “interpolation” between the WT analysis and the STFT. This interpolation is accomplished by taking into account the signal characteristics: If frequencies that are very well defined do not exist in the spectrum, the analysis is the WT.

3pSP5. Recovery of nonhomogeneous media impedance profile from transmitted signal. Roberto A. Tenenbaum (Acoust. and Vib. Lab., Federal Univ. of Rio de Janeiro, Rio de Janeiro, Brazil, P.O. Box 68503, CEP. 21945, roberto@serv.com.ufrj.br) and Marcelo B. S. Magalhes (Federal Univ. of Rio de Janeiro, Rio de Janeiro, Brazil)

A technique for recovering the acoustic impedance of layered media from its impulse transmission response is presented. A model for evaluation of the reflection response [R. A. Tenenbaum and M. Zindeluk, J. Acoust. Soc. Am. **92**, 3364–3370 (1992)] is extended to the transmission problem, leading to an algorithm where the amplitudes of the transmitted signal samples, H_i , are given by polynomials, P_i , of the reflection coefficients, R_i , as $H_i = P_i(R_1, \dots, R_N)$. For identification purposes, that relation leads to a highly nonlinear system. To solve it, extra information is needed: The inhomogeneity is considered to be inserted in a homogeneous medium, which yields $Q(R_1, \dots, R_N) = \prod_{k=0}^N (1 + R_k)/(1 - R_k) = 1$. A cost function is then defined as the quadratic difference between the measured transmitted signal and a synthetic one, given by P_i , that is, $S = \sum_{i=0}^N (P_i^2 - H_i^2) + Q^2 - 1$, which is minimized. This is a delicate process, since the Hessian of S is not always positive-definite, requiring the use of smart methods of optimization, such as Levenberg–Marquardt. S may have much local minima, to where minimization methods can converge, leading to the wrong results. This problem is solved by taking successive random initial estimates, and checking to see if the final value of S matches a global minimum criterion.

3pSP6. Sound velocity measurement in sub-wavelength thick samples. X. Yuan and W. Cao (Mater. Res. Lab., Penn State Univ., University Park, PA 16802, cao@math.psu.edu)

A conflict exists in ultrasonic measurements between the resolution, which requires higher frequency, and the penetration depth, which requires long wavelength. The traditional pulse–echo method for sound velocity measurement fails when the sample becomes too thin to allow the separation of repeated echoes. A signal processing scheme in frequency domain is described here to maximize the resolution of broadband transducers. The Wiener optimum filtering factor is used, together with the multiple sampling average technique, for power spectrum calculation. Using a broadband 2-MHz center frequency ultrasonic transducer, the sound velocity was successfully measured in brass shims as thin as 5% of the wavelength λ .

3pSP7. Ultrasonic measurement of the sound velocity and attenuation in some polymers. Aly H. A. Moustafa (Faculty of Electron. Eng., Dept. of Commun., Menoufia Univ., Menouf, Egypt)

The sound velocity and attenuation in perspex (poly methyl methacrylate) and in reson (hydroxyl terminated polybutadiene) are determined as functions of frequency and temperature using ultrasonic pulse technique. The importance of these materials comes from the need to examine their impedance match with water and to measure their attenuation coefficients over an extended range of frequency. A pair of transducers was placed in a small tank full of water and used to transmit and receive a 2.5-MHz pulse. The temperature of the water was varied using a temperature controlled heater (or cooler). The signal was captured with and without the material sample being inserted in the transmit/receive water path, using a digital storage oscilloscope (DSO), and then processed on a PC to obtain its FFT and to compute and plot the variation of the velocity of sound in the sample material as well as its absorption coefficient. The accuracy of the method was first checked and proved to be quite good by obtaining results for perspex and comparing them with the known published data. The measurements were then repeated for reson. A frequency range of ~ 1 –5 MHz was covered. The lowest considered temperature was 10 °C and the largest was 40 °C.

WEDNESDAY AFTERNOON, 18 JUNE 1997

ROOM R, 12:45 TO 2:15 P.M.

Session 3pUW

Underwater Acoustics: Chaos and Predictability in Long-Range Sound Propagation II

Frederick D. Tappert, Cochair

Applied Marine Physics, University of Miami, RSMAS, 4600 Rickenbacker Causeway, Miami, Florida 33149-1098

Michael G. Brown, Cochair

Applied Marine Physics, University of Miami, RSMAS, 4600 Rickenbacker Causeway, Miami, Florida 33149-1098

Contributed Papers

3pUW1. Validity and implications of ray-based studies for long-range propagation through internal waves. Jeffrey A. Simmen, Stanley M. Flatte, and Guang-Yu Wang (Office of Naval Res., 800 N. Quincy St., Arlington, VA 22217, simmenj@onr.navy.mil)

An analysis of the effects of ocean internal waves on long-range propagation is presented from a geometrical-optics point of view, and the validity of this ray-based approach is discussed. It is concluded that in the presence of internal waves, the observed depth diffusion of energy in the

late-arriving portion of the timefront is primarily a result of refraction (not diffraction) due to chaotic rays. It is found that internal waves can cause a single eigenray to break up into several microrays having upper turning points spread 10 km horizontally and 100 m vertically, affecting the resolution of ocean acoustic tomography. The validity of ray-based, semi-classical waveforms to represent received arrivals is evaluated by comparing waveforms generated with parabolic-equation simulations: The semi-classical waveforms reproduce the correct arrival time and temporal shape of almost all arrivals, even those that are made of dozens of microrays induced by internal waves. [Work supported by ONR.]

1:00

3pUW2. Breakdown of ray perturbation theory due to chaos.

X. Tang (Marine Studies, Univ. of Delaware, Newark, DE 19716) and F. D. Tappert (Univ. of Miami, Miami, FL 33149)

The possibility of ocean acoustic tomography in its usual form is based on ray perturbation theory and the following assumption: Eigenrays computed in an unperturbed ocean model agree one-to-one with eigenrays computed in the actual perturbed ocean. Using a ray numerical model with an environment containing idealized mesoscale structure, the perturbation strength ε is gradually increased and it is found that the multiplicity of eigenrays increases exponentially due to ray chaos. When the multiplicity changes then ray perturbation theory, even to very high order, necessarily breaks down. Bifurcation diagrams show eigenray travel times versus ε , and at critical values of ε pairs of new eigenrays suddenly appear as triplications in accordance with catastrophe theory. Such plots dramatically display the breakdown of ray perturbation theory. The methodology that has been developed should be useful for understanding the limitations of ocean acoustic tomography. [Work supported by ONR.]

1:15

3pUW3. Wave chaos and travel time effects of ocean mesoscale structure at global ranges.

F. D. Tappert (Appl. Marine Phys., Univ. of Miami, RSMAS, 4600 Rickenbacker Cswy., Miami, FL 33149) and X. Tang (Univ. of Delaware, Newark, DE 19716)

A range-dependent ray-trace model and the broadband UMPE model are used to model sound propagation at fixed bearing through a field of mesoscale baroclinic modes in order to study the effects of the mesoscale structure on travel time at long ranges. The ray model exhibits chaos at ranges beyond a few Mm as manifested by an exponentially increasing number of triplications of the wavefront. In addition, the ray model predicts a mesoscale travel time bias in the direction of later time ("cold bias") of the last axial arrival that is about 50–150 ms/Mm. At center frequency 75 Hz and bandwidth 50 Hz, the full-wave UMPE model qualitatively confirms the ray-trace predictions of chaos, that is now called "wave chaos" because it represents apparently random and unstable behavior of finite frequency wavefields. Also, the UMPE model quantitatively confirms the above mesoscale bias of the last axial arrival. Theories based on the adiabatic approximation are used to explain the modeled mesoscale bias in magnitude and sign. The steeper, early raylike arrivals are relatively stable in the presence of mesoscale structure and should be useful for long-range tomography. [Work supported by ONR.]

1:30

3pUW4. Mode coupling by internal waves for multi-megameter acoustic propagation in the ocean.

John A. Colosi (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Broadband parabolic-equation simulations have been performed with and without internal-wave sound-speed fluctuations. The simulations have a center frequency of 75 Hz, a bandwidth of 30 Hz, and a propagation range R of 1000, 2000, and 3000 km. For these cases it is found that acoustic propagation through internal waves is strongly nonadiabatic. In terms of modal travel times, low modes have a negative bias because they

couple into higher, faster modes, while the higher modes show a positive bias, indicating preferential coupling into lower, slower modes. The lowest modes show the least travel-time spread and bias, and these quantities increase rapidly with increasing mode number. Empirically and approximately it is found that bias grows like R^2 and spread grows like $R^{3/2}$. The modal power distributions over frequency are markedly different from the source distributions. Power is distributed roughly equally across the 30-Hz frequency band for each mode, with 5.6-dB scintillations consistent with an exponential probability distribution function for intensity. In spite of the dramatic spread and bias in the higher modes, it is found that a synthesis of these modes results in coherent wavefronts, whose characteristic timing fluctuations at 3000-km range are two orders of magnitude less than those of the corresponding modes.

1:45

3pUW5. Sound radiation by flat layered structures.

Leonid M. Lyamshev (N. N. Andreev Acoust. Inst., Russian Acad. of Sci., Shvernika str. 4, Moscow 117036, Russia)

Sound radiation by flat layered structures oscillating under the effect of random fractal forces has been investigated theoretically. The reciprocity theorem formulated earlier by the author has been used and the boundary problem solution is put down in an integral form. It is demonstrated that an acoustical pattern in the far wave field is determined by the product of the structure transmission coefficient and the Fourier spectrum of external forces. Random fractal forces are characterized by a correlation function of mass fractal dimension. It has been established that the fractal dimension of average intensity of acoustic field fluctuations coincides with the fractal dimension of the forces. Single-layer (a thin plate), double-layer, and triple-layer elastic structures are considered. It is noted that the singularity spectrum of an acoustic field radiated by a wall in a turbulent flow coincides with the singularity spectrum of wall pressure fluctuations. [This research is supported by the Russian Foundation for Basic Research.]

2:00

3pUW6. Mapping climatic temperature changes in the ocean with acoustic tomography: Navigational requirements.

John L. Spiesberger, Anatoly L. Fabrikant, Anisim A. Silivra (Dept. of Meteorol. and the Appl. Res. Lab., 512 Walker Bldg., Penn State Univ., University Park, PA 16802, jspies@gamot.psu.edu), and Harley E. Hurlburt (NRL, Stennis Space Center, MS 39529-5004)

In eddy-resolving hydrodynamic models, first-mode baroclinic Rossby waves linked to El Niño/Southern Oscillation are the dominant features which change basin-wide temperatures below the seasonal thermocline in the northeast Pacific at periods less than a decade. Realistic simulations are carried out in which Rossby waves are mapped using acoustic tomography. Based on the model which propagated these waves, a Kalman filter is used to map temperature signals for a year. At 300-m depth, 80% of the model variance is accounted for with tomography when the navigational errors of the sources and receivers are as poor as 1 km. Consequently, it may be unnecessary to accurately navigate actual tomographic instruments to map climate change [Spiesberger *et al.*, IEEE Ocean. Eng. **22** (1997)]. [Work supported by SERDP and ONR.]

Plenary Session, Business Meeting, and Awards Ceremony

Stanley L. Ehrlich, Chair
President, Acoustical Society of America

Business Meeting

Presentation of Certificates to New Fellows

Presentation of Awards

R. Bruce Lindsay Award to D. Keith Wilson

Helmholtz-Rayleigh Interdisciplinary Silver Medal to Gerhard M. Sessler

Gold Medal to Karl Uno Ingard

Musical Entertainment

The Nova Consort quartet will perform Early Medieval and Renaissance Music accompanied by soprano Sara Kelly singing Troubadour songs and Elizabethan love songs.

Session 4aAA

Architectural Acoustics: Absorption, Modeling and Case Studies

Christopher Brooks, Chair

Orpheus Acoustics, 925 Virginia Avenue, Lancaster, Pennsylvania 17603

Chair's Introduction—8:30

Contributed Papers

8:35

4aAA1. Sound absorption characteristics of open brick/block absorbing walls. Hideki Tachibana, Shin'ichi Sakamoto (Inst. of Industrial Sci., Univ. of Tokyo, Roppongi 7-22-1, Minato-Ku, Tokyo 106, Japan), Hikari Mukai (Miji Univ., Japan), and Tong-Jun Cho (Tokyo Univ., Tokyo, Japan)

Brick and block walls with openings and backing air space with porous materials are often used in auditoria as a sound absorbing treatment. These kinds of walls have a frequency-selective sound absorption property in low frequencies, which is usually explained as the Helmholtz resonator consisting of the air volume in the part of the opening and the backing cavity. In addition, these types of wall often have different peaks of sound absorption coefficient in higher frequencies (around 1000 Hz, in the case of 10-cm thickness of the brick/block). In this study, this sound absorption effect has been investigated by theoretical consideration, reverberation room measurement of a real construction, and a $\frac{1}{10}$ scale model experiment, and it has been clarified that the sound absorption is caused by the effect of open-pipe resonance which happens in the openings. In addition to these studies, the effects of the Helmholtz resonance and open-pipe resonance have been visualized by numerical analysis and physical experiment using the Kundt's dust tube method.

8:50

4aAA2. Sound absorption by seats in concert halls, occupied and unoccupied, and average sound absorption by the interior surfaces of halls—II. Leo L. Beranek (975 Memorial Dr., Ste. 804, Cambridge, MA 02138) and Takayuki Hidaka (Takenaka R&D Labs., Chiba 270-13, Japan)

The average sound absorption coefficients for the interior surfaces of eight halls, in which all finishes were completed before installation of seats, were determined from reverberation measurements. When the chairs were installed, their absorption coefficients, both occupied and unoccupied, were also determined from RT data. Five of these halls have lightly upholstered seats, one has medium upholstered seats, and two have heavily upholstered seats. The data are compared with the 15-hall data of Chap. 5 in Beranek [*Concert and Opera Halls: How They Sound*, Acoustical Society of America (1996)]. A survey of the chair design in those 15 halls has produced a correlation between chair construction and the words: "lightly," "medium," and "heavily" upholstered. The new data permit more accurate estimation of reverberation times as a function of frequency than previously found by comparing the details of a hall-in-planning with the constructions and seats of these eight halls.

9:05

4aAA3. Computation of impulse response for stage-house geometries with use of edge-diffraction models. Rendell Torres, Mendel Kleiner, and Peter Svensson (Chalmers Rm. Acoust. Group, Dept. of Appl. Acoust., Chalmers Univ. of Tech., S-412 96 Gothenburg, Sweden, rendell@ta.chalmers.se)

A mathematical model following diffraction studies by Medwin [J. Acoust. Soc. Am. **72**, 1005–1013 (1982)] is developed to calculate the impulse response of a three-dimensional "stage house" geometry in a baffle. The model includes multiple diffraction components in addition to specular reflection, and it forms a more complete approach than current room impulse-response predictors based on geometrical acoustics. Thus it may also be suitable for more accurate auralization, particularly at low frequencies. A comparison is made with results from a "cone-tracing" computer model [B. Dalenbäck, J. Acoust. Soc. Am. **100**, 899–909 (1996)] and results from physical scale-model measurements. [Work supported by Axel & Margaret Axson Johnson's Foundation, Sweden.]

9:20

4aAA4. A critique on modeling absorbent materials from a numerical simulation perspective. Debashis Basu (State Univ. of New York at Buffalo, 212 Ketter Hall, Amherst, NY 14260)

Absorbent materials, poroelastic in nature, have very complex acoustic wave propagation characteristics in terms of the physics and mathematics defining the phenomenon. Even for a single layer of absorbent, the mathematics governing the physics of the material under acoustic flow conditions is quite daunting. A critical review of the progress in the theories (and inherent assumptions) is being presented. Proper engineering judgment might allow a more simplified numerical simulation to get an idea of the complex acoustical behavior. More often it is the lack of availability of proper data which drives the numerical simulation to be done under certain inexact approximations. The present article tries to lay down the theories and semi-empirical relations developed over the past few decades, and assumptions allowing a simplified model, which at least help in the acoustic characterization of a complex problem. With growing emphasis on a "better, safer, quieter" acoustic environment and mandatory noise control measured in different countries, there would be an ever increasing demand for effective acoustic design and absorbent materials would play a key role in several applications. As would be clearly apparent from the current article, the available "standard commercial" data on absorbent materials gives an estimate of what is happening but does *not* contain enough information to accurately predict the physics of the problem in certain situations.

9:35

4aAA5. Rain and hail simulation and sound isolation tests on a wood and metal roof assembly. Russell Cooper (Jaffe Holden Scarbrough Acoust., Inc., 114A Washington St., Norwalk, CT 06854)

A series of acoustic tests were performed on a proposed roof construction for a new concert hall at The University of North Texas to evaluate its sound transmission loss properties. The test was conducted to determine if the wood and metal deck construction would be sufficient to isolate nearby

highway traffic noise (airborne) and heavy rain and hail noise (structure-borne) from being heard in the concert hall. Four variations of the roof system were constructed and tested to determine the range of possible solutions. Simulation of rain and hail noise test methods and results are presented along with standard ASTM E-90 results and the recommended solution.

9:50

4aAA6. Factors influencing the perception of bass. John S. Bradley (Natl. Res. Council, Ottawa, ON K1A 0R6, Canada), Gilbert A. Soulodre (Commun. Res. Ctr., Ottawa K2H 8S2, Canada), and Scott Norcross (American Univ., Washington, DC 20016-8058)

It is commonly assumed that low-frequency reverberation time determines the perception of bass in concert halls. Laboratory experiments were conducted to evaluate whether this and other factors influence subjects' perceptions of the strength of bass sounds in simulated sound fields. The levels of the early and late arriving low-frequency sounds as well as low-frequency reverberation times were systematically varied. Ten subjects rated the strength of the bass content of the music for each sound field on a five point scale relative to a reference sound field. Both the levels of the early and late arriving low-frequency sound had significant effects on the judgments of the bass content. Low-frequency reverberation time was not significantly related to subjective ratings of bass. The direction of arrival of low-frequency sound had smaller effects on the assessments of the bass content in the sounds. The results suggest that increased low-frequency reverberation time is not important for increasing the sense of bass in concert hall sounds. Further, the low-frequency attenuations caused by the grazing incidence seat dip effect can be compensated for by overhead reflections to increase the sense of bass in concert halls.

10:05–10:20 Break

10:20

4aAA7. Design considerations for Boettcher Hall, a vertical surround symphonic facility. Christopher Jaffe, Gary Madaras (Jaffe Holden Scarbrough Acoust., Inc., 114A Washington St., Norwalk, CT 06854), and John S. Bradley (IRC, Natl. Res. Council, Ottawa K1A 0R6, Canada)

In the early 1960s, Dr. Leo Beranek correlated reflection patterns in concert halls with qualitative listening experiences. Over the years, practitioners and other researchers have confirmed the validity of his work, proving that the direction, amplitude, frequency composition, and arrival time of the reflections define quality, regardless of room geometry. This paper discusses methods for achieving qualitatively preferred listening experiences in vertical surround halls such as Boettcher Concert Hall in Denver, Colorado, and the difficulties that can arise when breaking the mold of a traditional shoe-box hall.

10:35

4aAA8. Acoustical characteristics of a 360-degree surround hall. John S. Bradley (IRC, Natl. Res. Council, Ottawa, ON K1A 0R6, Canada), Gary Madaras, and Chris Jaffe (Jaffe, Holden, Scarborough, Norwalk, CT 06854)

Boettcher Hall in Denver is a large multipurpose hall with audience seating completely surrounding the stage and has been described as a 360-deg surround hall. Extensive acoustical measurements were carried out to evaluate the characteristics of this hall configuration and to compare it to other halls of more conventional design. Impulse response measurements were made for the 45 combinations of three on-stage source positions and 15 receiver positions distributed throughout the audience seating area. The setup of the hall was representative of a typical unoccupied concert configuration. Measured parameters included octave band values of: reverberation times, early decay times, relative sound levels, and early/late sound ratios. Various measures of spatial impression were also measured including: lateral energy fractions, inter-aural cross correlations, and the newly proposed measure of listener envelopment, the late lateral sound

level. Hall average values as well as the spatial variation of acoustical measures are compared to data obtained from a number of other halls. [Work supported by the Concert Hall Research Group.]

10:50

4aAA9. Effects of an audience and variable architectural features on the acoustics of Boettcher Concert Hall. Gary Madaras (Jaffe Holden Scarbrough Acoust., Inc., 114A Washington St., Norwalk, CT 06854), John S. Bradley (IRC, Natl. Res. Council, Ottawa K1A 0R6, Canada), and Christopher Jaffe (Jaffe Holden Scarbrough Acoust., Inc., Norwalk, CT 06854)

Acoustic parameters (including reverberation time, early decay time, relative sound level, early to late sound ratios, lateral energy fraction, and interaural cross correlation) have been measured in Boettcher Concert Hall. All parameters were measured using different settings of variable architectural features such as acoustic drapes, variable height overhead stage reflectors, and movable stage towers. The monaural parameters were also measured with and without an audience. The effects of the variable features on the acoustic parameters are assessed. The effects of the audience on some of the acoustic parameters have been estimated using standard calculation methods. Calculated differences are compared to those actually measured in the hall. [Work supported by the Concert Hall Research Group.]

11:05

4aAA10. A comparative database of salient architectural features: Performance spaces and houses of worship. Christopher Brooks (Orpheus Acoust., 925 Virginia Ave., Lancaster, PA 17603)

A set of salient architectural features for the interior volumes of houses of worship and performance spaces was selected. These features include volume, height, length, width, number of audience seats, type of plan, and predominant material. The buildings were chosen to illustrate a broad range of types within these two categories; well-known examples are included as benchmarks. Data were taken from the literature and entered into a database. With a descriptive statistics program, these data can be inspected from various perspectives. The purpose of this study is to allow a designer to enter features of a proposed design, and by inspection to get a sense of where the design lies in relation to existing buildings. For instance, a proposed church may have important similarities with an existing concert hall. Future plans include access to drawings and acoustical measurements (if available) for each building included, and posting on the World Wide Web with forms for collecting new data.

11:20

4aAA11. Average acoustical characteristics of ten atrium spaces. J. S. Bradley (Acoust. Lab., Natl. Res. Council, Montreal Rd., Ottawa, ON K1A 0R6, Canada) and Y. K. Oh (Mokpo Natl. Univ., Muan, Chonnam, 534-729 Korea)

Extensive acoustical measurements were made in ten different atrium spaces varying in volume from 3 300 to 45 000 m³. Impulse responses were measured at between 18 and 71 combinations of source and receiver position in each atrium depending on the size of the atrium. The computer-based measurement system (RAMSoft-3) used a maximum length sequence signal and a fast Hadamard transform procedure to obtain 5-s-long impulse responses. Octave band values of decay times, sound levels, and energy ratios were calculated for each impulse response. The average values of each quantity in each atrium were used to characterize typical acoustical characteristics of untreated atria. These allow rough estimates of the acoustical conditions in other atria as a function of the volume of the spaces.

Session 4aABa**Animal Bioacoustics: Animal Communication**

Peter Tyack, Chair

*Department of Biology, MS#34, Woods Hole Oceanographic Institution, 45 Water Street, Woods Hole, Massachusetts 02543-1049***Chair's Introduction—8:40*****Invited Papers*****8:45****4aABa1. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques.** W. Tecumseh Fitch (Harvard/MIT Speech and Hearing Sci. Prog., 33 Kirkland St., Rm. 1036, Cambridge, MA 02138)

Vocal tract length of 23 rhesus macaques (*Macaca mulatta*) was measured using radiographs and computer graphic techniques. LPC analysis of tape-recorded threat vocalizations was used to determine vocal tract resonance frequencies ("formants") for the same animals. A new acoustic variable is proposed, "formant dispersion," which theoretically depends upon vocal tract length. Formant dispersion is the averaged difference between successive formant frequencies, and was closely tied to both vocal tract length and body size. Despite the common claim that voice fundamental frequency (F_0) provides an acoustic indication of body size, repeated investigations have failed to support such a relationship in many vertebrate species including humans. Formant dispersion, unlike voice pitch, is proposed to be a reliable predictor of body size in macaques, and probably many other species.

9:10**4aABa2. Whistle matching in wild bottlenose dolphins.** Vincent M. Janik (School of Biological and Medical Sci., Bute Bldg., Univ. of St. Andrews, Fife KY16 9TS, UK, vj@st-and.ac.uk)

Bottlenose dolphins (*Tursiops truncatus*) produce individually distinctive signature whistles when they are in isolation. These whistles have been hypothesized to be important in individual recognition and group cohesion. However, captive individuals are known sometimes to imitate each other's signature whistles. Imitation could either lead to confusion over the caller's identity and act against individual recognition or could be used to address specific individuals in a communication network. In this study a hydrophone array with three transducers was used to investigate whether whistle copying also occurs in the wild. Recordings were made in a 500-m-wide channel in the Moray Firth, Scotland. Caller position was determined by comparing differences in the time of arrival of a sound at different hydrophones. Since signature whistles could not be determined by isolating individuals, exact whistle matching in vocal interactions was used to study copying. The results showed that whistle matching occurred in 17% of all whistle interactions. An analysis of the variability of matched whistles showed similar degrees of stereotypy as reported for signature whistles. Thus signature whistle matching seems to be an important feature of bottlenose dolphin communication in the wild. [This study was funded by a DAAD-Doktorandenstipendium HSPII/AUFE.]

9:35**4aABa3. First audiogram for marine mammals in the open ocean and at depth: Hearing and whistling by two white whales down to 30 atmospheres.** Sam Ridgway, Donald Carder, Rob Smith, Tricia Kamolnick, and Wesley Elsberry (NCCOSC RDTE DIV D3503, 49620 Beluga Rd., San Diego, CA 92152-6266)

In examining the potential impact of human-made sound on sea mammals, it was considered that whale hearing sensitivity might diminish with increasing ambient pressure. To test the effect of depth, two white whales made 885 dives to a platform at 5, 100, 200, or 300 m in the Pacific Ocean. At each stationing on the platform up to 12 min at a time, whales whistled when they heard a 500-ms tone from a hydrophone. With increasing depth, air density increase in the middle ear, sinuses, and nasal cavity changed each whale's whistle response, but did not attenuate hearing as it does in the aerial ear (of humans and other land mammals tested in pressure chambers) due to middle ear impedance changes. The findings support theories that sound is conducted through whale head tissues to the ear without the usual ear drum/ossicular chain amplification of the aerial middle ear. These first ever hearing tests in the open ocean demonstrate that whales hear as well at depth as near the surface, therefore, zones of influence for human-made sounds are just as great throughout the depths to which whales dive, or at least to 300 m. [Work supported by Office of Naval Research N0001496WK30349.]

Session 4aABb

Animal Bioacoustics: Bioacoustics Sensing of the Environment

Daniel R. Raichel, Chair

Department of Mechanical Engineering, Cooper Union, 51 Astor Place, New York, New York 10003

Invited Papers

10:15

4aABb1. The use of CF/FM sounds in bats. Yanling Lei (Cooper Union, 51 Astor Pl., New York, NY 10003) and Daniel R. Raichel (Cooper Union and CUNY Graduate School, New York, NY 10003)

The manner in which bats process their echolocation calls involves two main components: (1) producing and transmitting the signals; and (2) receiving and analyzing them [Fenton and Brock, *Bats*, Facts On File (1992)]. The high-frequency signals constituting these calls consist of constant frequency (CF) sounds and frequency modulated (FM) sounds. CF components were found to be used for prey detection and identification; FM components were found to be utilized to update distance information. The Doppler shift also helps in detection of targets and gauging the speed of moving targets. Thus, combining the pulses and the resultant echoes enables the bat to determine object distances, sizes, and shapes. Bats were found to have their own call frequencies and are insensitive to a wide range of potential jamming sounds, which allows them to clearly distinguish their own calls amidst environmental cacophony. Experiments have shown that only when entire CF/FM simulations were used did the bats perform discrimination tests badly [J. D. Altringham, *Bats: Biology and Behavior* (Oxford U.P., New York, 1996)]. This indicates that the echolocation process may be essentially a series processing of CF/FM signals rather than parallel processing of both types.

10:40

4aABb2. A multiple target acoustic scene representation model for bat echolocation signals. Min Liang and Mathew J. Palakal (Dept. of Comput. Sci., Indiana Univ.-Purdue Univ., 723 W. Michigan St., SL 280, Indianapolis, IN 46202)

Determination of the distance to targets from echo delay is an important aspect of echolocation in bats. Neurophysiological studies suggest that the tonotopically organized delay-sensitive neurons (DSN) found in the cortex of bat encode target distance. These studies also indicate that in the tonotopic area, the DSNs in each active ensemble share the same range acuity and the acuity systematically improves over a period of time. A biologically plausible model is suggested here to understand how the bat assembles responses to echoes from multiple targets. This model consists of three stages in cascade: (1) an auditory periphery system (APS); (2) a self-organization neural network (SONN) and; (3) a cortical representation neural network (CRNN). The inputs of APS are the pulse-echo (PE) signals and the outputs are the time-frequency spike trains of PE. SONN self-organizes to the spike trains of PE in frequency index. CRNN has the same architecture of SONN but with a specific neuron response function, $y(T, \delta, f) = I(\delta, f)\Phi[a_1 T(\delta - a_2^n \beta_1)/a_2^n]\Phi(f - \beta_2)$, to achieve the multi-resolution. In this function, $\Phi(\cdot)$ is the negative second derivative of a Gaussian function, $I(\delta, f)$ is the spike amplitude of frequency component f arriving at time δ , T is the time onset of echoes after vocalization, a_1 and $a_2 > 1$ are constants, n is a negative integer, β_1 is the minimum best delay, and β_2 is the best frequency. The response maps of CRNN at different T form the multi-resolution representation of the acoustic scene. Simulations are carried out for the proposed model using phantom targets and the results are compared to neurophysiological findings. [Work supported by NSF.]

Contributed Papers

11:05

4aABb3. Acoustic flow in echo amplitudes and spectra: a viable concept for obstacle avoidance in CF-bats? Rolf Müller and Hans-Ulrich Schnitzler (Dept. of Animal Physiol., Univ. of Tübingen, Morgenstelle 28, D-72076 Tübingen, Germany, rolf.mueller@uni-tuebingen.de)

Bats are capable of obstacle avoidance based solely on echolocation information. When doing so, species being classified as CF (constant frequency) bats allocate most signal energy to rather long, narrow-band portions of their echolocation pulses. These signal components may subserve obstacle avoidance by means of information conveyed by proportional changes in amplitude (due to geometric attenuation, absorption, compound directivity of emitter and receiver) and frequency content (due to Doppler shifts) of the echoes along a trajectory. It can be demonstrated that a combination of these two putative sensory variables allows, in principal, for a metrical reconstruction of target position relative to the animal's flight vector within a hemifield. Thus the contained information seemingly suffices for mediating appropriate avoidance maneuvers. CF bats' unique

narrow-band signal design is matched by a likewise peculiar layout of their auditory system. In particular, the existence of an auditory fovea featuring extremely high filter qualities may be hypothesized to aid estimation of proportional changes in frequency. In order to establish the extent to which the informational content of the signal parameters under consideration may be accessible in the auditory system's primal sketch, the properties of a gammatone filter bank fitted to physiologically determined filter qualities are explored.

11:20

4aABb4. Measurement of the echo location signals of the Atlantic Spotted Dolphin *Stenella frontalis* in the waters off the Grand Bahamas. Whitlow W. L. Au (Hawaii Inst. of Marine Biology, P.O. Box 1106, Kailua, HI 96734) and Denise L. Herzing (Florida Atlantic Univ., Boca Raton, FL 33431)

A line array of three hydrophones with a video camera attached to the array was used to measure the echo location signals of wild Atlantic Spotted Dolphins. The separation distance between hydrophones in the array

was 30 cm. The array was attached to a float that supported an amplifier-line driver assembly with the signals sent via a 76-m multi-conductor cable back to the support boat. The float and array assembly was oriented by a swimmer. The echo location signals from the hydrophone were digitized simultaneously at a sample rate of 500 kHz. Twenty files of echo location click trains were collected with the quality of the data varying from poor (files with lots of whistles and off-axis signals) to very good. The on-axis signals typically had a bimodal spectrum with a low-frequency peak at 45–60 Hz and a high-frequency peak at 120–140 kHz. Peak-to-peak source levels up to 210 dB *re*: 1 μ Pa were measured. The rms bandwidth varied between 32 and 46 kHz, with a cluster around 40 kHz.

11:35

4aABb5. Fish resonance absorption spectroscopy. Orest I. Diachok (Naval Res. Lab., Washington, DC 20375)

Measurements of the absorption coefficient were made at 12 km between 0.6 and 5.0 kHz at a shallow water (83-m) site in the Gulf of Lion in September, 1995. At night absorption lines due to dispersed sardines at 1.3 kHz at 25 m and sardine schools at 1.5 kHz at 60 m were evident. One hour before sunrise the resonance frequency, f , of dispersed sardines increased with time, becoming 2.7 kHz at sunrise, as dispersed sardines descended to 65 m; and in the following 15 minutes f decreased to 1.7 kHz, as dispersed sardines formed schools, as evidenced by echo sounder records. The signal loss at 2.7 kHz at sunrise, attributed to dispersed sardines at 65 m, was 35 dB; the signal loss at 1.7 kHz during daytime, attributed to sardine schools at 65 m, was 15 dB. The shift in f and the decrease in absorptivity that occurred during the transition are in accord with theory: Close proximity between fish in school at 65 m causes both f

and the scattering cross section to diminish (Feuillade, 1996). The latter is proportional to the extinction cross section which, together with n (number per unit volume), control the absorptivity. Estimation of n from absorption spectroscopy measurements, and the hypothesis that marine mammals might exploit this phenomenon, will be discussed.

11:50

4aABb6. Ultrasound-mediated transdermal delivery of tagging compounds into migratory fish. Joseph A. Clark (CDNSWC, Code 734, Bethesda, MD 20084 and COMB, Ste. 236, Columbus Ctr., 701 E. Pratt St., Baltimore, MD 21202), Jane A. Young (CDNSWC, Bethesda, MD 20084), Amrit N. Bart, and Yonathan Zohar (COMB, Baltimore, MD 21202)

There is a need for noninvasive, low-cost, efficient methods of tagging large numbers of aquatic animals. For example, improved tagging methods are needed to follow the migratory patterns of some endangered salmonoid species. This talk reports on an investigation of the feasibility of using cavitation levels of ultrasound to mediate the delivery of tagging compounds into fish. Two tagging compounds are investigated: calcein (Fluoroxon) and oxytetracycline-hydrochloride (Terramycin). Both of these compounds are presently used as tagging agents. However, injection methods of delivery are not efficient for cases where large numbers of wild fish are involved, and if they are placed directly into the aquatic environment of fish the compounds diffuse too slowly through the skin to adequately tag the fish before creating other health problems for them. A system for generating controlled doses of cavitation level ultrasound to schools of fish will be described and measurements of the enhancement in delivery of the tagging compounds to the fish will be reported.

THURSDAY MORNING, 19 JUNE 1997

ROOM O, 8:00 A.M. TO 12:00 NOON

Session 4aBV

Bioresponse to Vibration/Biomedical Ultrasound and Physical Acoustics: Medical Ultrasound

Christy K. Holland, Chair

Department of Radiology, University of Cincinnati, ML 742, 234 Goodman Street, Cincinnati, Ohio 45267-0742

Contributed Papers

8:00

4aBV1. Numerical modeling of biomedically applied transient high-energy sound pulses. Eckard Steiger (Inst. fuer Hoechstfrequenztechnik und Elektronik/Akustik, Univ. of Karlsruhe, Kaiserstrasse 12, 76131 Karlsruhe, Germany)

A finite difference in time-domain (FDTD) discretization is applied to simulate finite-amplitude sound pulse propagation in biomedical applications. A numerical model requires a method of accounting for correct representation, especially of shorter wave components. These are generated by nonlinear steepening during the propagation. Therefore, the FDTD discretization has to be of low-dispersive and low-dissipative nature. A high-order scheme suitable for the computation of pulsed sound including weak shocks is applied to model a Storz SL 10 shock-wave lithotripter. The SL 10's electromagnetic source generates cylindrically diverging pulsed pressure waves. By reflection at a hemi-paraboloidal brass reflector these waves are focused and steepening of the wave profiles causes weak shock-wave generation in the focal region. A high-order method to model curved boundaries (given by the reflector) is presented. Wave profiles predicted are compared with measured ones. Excellent agreements demonstrate the proposed method to be a flexible tool to simulate biomedically applied transient high-energy sound.

8:15

4aBV2. Acoustic and cavitation fields of a pressure release ellipsoidal reflector. Michael R. Bailey (Appl. Phys. Lab., Univ. of Washington, 1013 N.E. 40th St., Seattle, WA 98105)

In an electrohydraulic lithotripter (device to treat kidney stones), an underwater spark at the near focus of a rigid ellipsoidal reflector produces acoustic cavitation at the far focus. A reflector made of pressure release foam, however, dramatically reduces cavitation with little change in acoustic pressure amplitude ($\sim +35$ MPa, ~ -10 MPa). The pressure waveform at the focus of a rigid reflector is a positive spike followed by a negative tail. The negative tail sets bubbles into a growth-and-collapse cycle hundreds of ms long. With a pressure release reflector, a negative phase precedes a positive spike. Bubble growth induced by the negative phase is quickly ($<10\mu$ s) snuffed by the spike, and collapse is weak. Cavitation was measured by (1) listening for bubble collapse and (2) recording bubble-induced pits on aluminum foil placed *along* the reflector axis. Pit size was used as an experimental measure of cavitation intensity. Intensity was reduced 3–10 times by the pressure release reflector. Results are in agreement with calculations from the Gilmore equation. Comparing the bioeffects of the two reflectors provides a way to assess cavitation's role in lithotripsy. [Work supported by ONR, ARL:UT IR & D, and NIH.]

4aBV3. Localized detection of cavitation generated by lithotripsy shock waves. Robin O. Cleveland (Appl. Phys. Lab., Univ. of Washington, 1013 N.E. 40th St., Seattle, WA 98105) and Oleg A. Sapozhnikov (Moscow State Univ., Moscow 119899, Russia)

Extracorporeal shock wave lithotripters generate intense cavitation which plays a role in both kidney stone destruction and tissue damage. The cavitation can be detected by listening for the characteristic "double-bang" acoustic emission, radiated by the bubbles during their rapid expansion and collapse phases. Previous workers have used a single focused transducer to detect the double-bang signature. However, the focal spot of the transducers is typically 5 cm long and individual cavitation events can not be localized. A system has been developed that uses two confocal transducers (1 MHz, 20 cm radius of curvature, 10 cm diameter) to detect cavitation events in a localized volume. The system can be run in two modes: passive or active. In the passive mode both transducers listen to acoustic emissions generated by the cavitation field. In the active mode one transducer sends a tone and the second transducer listens to the sound scattered by any bubbles present in the focal region. In both cases the region to which the detectors are sensitive is approximately a sphere 5 mm in diameter. The passive system shows bubble rebound times far longer than predicted by theory for a single bubble. Scattering indicates that bubbles persist in the focal region for at least 6 ms. [Work supported by NIH.]

4aBV4. Subharmonic imaging with gas-filled microbubbles. William T. Shi, Flemming Forsberg, and Barry B. Goldberg (Dept. of Radiol., Thomas Jefferson Univ., Philadelphia, PA 19107)

The nonlinear properties of gaseous contrast agents can be used to create new ultrasound imaging methods, such as harmonic and subharmonic imaging. Those methods are capable of separating the echoes from the contrast bubbles in blood and those from surrounding tissue. Harmonic images are produced by transmitting sound pulses at one frequency, but receiving echoes at double that frequency, while subharmonic images receive echoes at half that frequency. Experiments revealed that the subharmonic component of echoes from some contrast agents, at relatively high acoustic pressure, was greater than the second-harmonic component. Since harmonic imaging suffers from a reduction in image contrast, due to second-harmonic generation during ultrasound propagation, as well as second-harmonic leakage from the transmitted signal, this finding makes subharmonic imaging an attractive alternative to harmonic imaging. Harmonic response measurements with a pulse-echo arrangement and subharmonic images obtained with a modified scanner will be presented. The acoustic pressure threshold for subharmonic generation was found to be quite different for contrast agents with different compositions. The optimal transmitted ultrasound pulses and the reception strategy for subharmonic imaging will also be discussed.

4aBV5. A new imaging science test object for performance measurements of ultrasonic imaging systems. Daniel Phillips, Stephen McAleavey, and Kevin J. Parker (Rochester Ctr. for Biomedical Ultrasound, Dept. of Elec. Eng., Univ. of Rochester, Rochester, NY 14627)

A novel test object for ultrasound imaging has been developed. The test object, or phantom, is manufactured with thin film techniques allowing precise placement of "digital scatterers" which can produce sophisticated test targets, similar to those that are widely used in imaging science to evaluate displays, printers, and electronic imaging systems. This type of test object can be used to determine parameters such as the MTF and resolution limits, and can facilitate the detection of imaging phenomena and artifacts such as spatial warping and spatial frequency aliasing. These test objects can be used in conjunction with standard performance evaluation

methods and devices to augment and enhance their capabilities and range of application. In addition, by imparting an appropriate motion to the target, it is possible to provide unique test stimuli for pulsed and continuous wave as well as color flow and color energy Doppler ultrasound systems. Results are presented which demonstrate how the ability to make precise line targets allowed the assessment of direction related resolution and spatial frequency response by inspection, thus facilitating the implementation of a simple evaluation procedure.

4aBV6. Abstract withdrawn.

4aBV7. Echosignal peculiarities in biological mediums. V. Genis (Biomed. Eng. and Sci. Inst., Drexel Univ., Philadelphia, PA 19104), E. I. Oboznenko, and I. L. Oboznenko (Kiev Polytech. Inst., Kiev 252056, Ukraine)

A modified method of Born's approximation is proposed. The essence of the method is that in real situations, the incident wave gains the object's speed, but not the medium's speed; which follows from the Fresnel formula. A solution algorithm for the weak-reflecting bio-organ's scattering problem is proposed in this task. The algorithm is based on a more precise definition of Born's approximation. Numerical and experimental research shows that the pathologies can have dimensions of about 1 mm, which can be identified according to echosignal spectrums. It is shown that norm and pathologies, like pectoral glands, are distinguished. If the pathology is less dense, there will be potent destruction in the echosignal spectrum. If pathology is denser, there will be less potent destruction. The proposed method permits one to classify the pathologies in bio-tissues more subtly.

4aBV8. Measurements of differential scattering cross section using a ring transducer. T. Douglas Mast (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804), Tomas T. Jansson (Lund Univ., 221 00 Lund, Sweden), and Robert C. Waag (Univ. of Rochester, Rochester, NY 14627)

A method for the measurement of intrinsic scattering object properties is presented. This procedure is based on the measurement of the scattered acoustic field as a function of scattering angle and frequency. Measurements are normalized using analytically determined expressions for emitter and detector beams resulting from a combination of unfocused linear elements arranged in a circular configuration. The spatial effects of finite emitter pulse length and detector gate length are represented by a convolution formula valid for narrow-band signals and long receiver gates. The method includes correction for target absorption as well as measurement of the directly transmitted acoustic power in the free field, yielding the average differential scattering cross section per unit volume. Measured results for two phantoms consisting of glass microspheres embedded in agar show good relative agreement with theory for varying angle, frequency, and

phantom properties. For the phantoms employed, scattering effects, rather than increased absorption, are shown to account for most of the difference in transmission loss between pure agar and agar with glass spheres.

10:00–10:15 Break

10:15

4aBV9. Evaluation of a three-dimensional fractal model of ultrasound scattering in liver. Daniel Phillips, Robert Cramblitt, and Kevin J. Parker (Rochester Ctr. for Biomedical Ultrasound, Dept. of Elec. Eng., Univ. of Rochester, Rochester, NY 14627)

It is postulated that a major source of scattering in liver is the collagen rich structure of the vessels that comprise the hepatic vasculature system. This collagen rich hepatic vessel structure manifests itself as the branching pattern of the vessels in three dimensions. Past study of these branching patterns in other organs such as the lung has resulted in data and models which have ascribed a fractal description to their organization. A computer based model has been developed of the portal vasculature system of the human liver based on anatomical, physiological data and precepts as well as a fractal methodology. A thin, planar slice of the model is used to represent sites of ultrasound scattering. A simple model is utilized to calculate the ultrasound scattering from the cross sections of the thin-shelled cylinders used to represent the collagen component of the vascular walls. Simulation of the radio frequency scattered waves is carried out along with envelope detection to allow analysis of the first and second order statistics. The resultant statistics are analyzed in light of the parameters utilized to define the model and their variation.

10:30

4aBV10. Ultrasonic backscattering of blood in simple oscillatory flow. Yu-Hong Lin (Graduate Program in Acoust., Penn State Univ., University Park, PA 16802) and K. Kirk Shung (Penn State Univ., University Park, PA 16802)

Ultrasonic backscattering of porcine whole blood (hematocrit at 40%) was measured by a 10-MHz catheter mounted transducer in a mock flow loop at different flow rates controlled by a roller pump and a pulsatile pump. The amplitude of the oscillatory flow was fixed at 40 cm/s and the baseline velocity was set at three different levels, 0, 20, and 40 cm/s, to change the shear rate range. As the flow rate increases from 0 to 40 cm/s to 20 to 60 cm/s, the peak of cyclic variation of the backscattering signal leads more of the peak of the mean flow velocity. At 40 to 80 cm/s, the magnitude of cyclic variation drops from 4 dB to 1.5 dB. It was found that at the two lower flow rate ranges the backscattering signal increases with shear rate if the mean shear rate is below 150 1/s. When the shear rate further increases, the backscattering signal decreases. The wave form of the cyclic variation is also altered as the baseline shear rate changes. Results show that, in simple oscillatory flow, the amplitude and the timing of the peak of ultrasonic backscattering from whole blood are affected mainly by the shear rate which mediates red cell aggregation.

10:45

4aBV11. Assessment of turbulent flow in arteries using Doppler ultrasound. John S. Stroud, Christy K. Holland (Dept. of Radiol., Univ. of Cincinnati, Cincinnati, OH 45267-0579), Peter J. Disimile, and Curt W. Fox (Univ. of Cincinnati, Cincinnati, OH 45267-0070)

The detection of atherosclerotic-related diseases in the early stages could potentially impact medical treatment in a cost-effective manner. Atherosclerosis is a common arterial disorder characterized by deposition of cholesterol, lipids, and cellular debris in the inner layers of large- and medium-sized arteries. As atherosclerotic deposition progresses, an area of

narrowing, or stenosis, forms, resulting in reduced circulation to organs and other tissues supplied by the artery. When the stenoses become severe in critical arteries (e.g., the carotids or the coronaries) the result can be a catastrophic stroke or heart attack. In the early stages of the disease, the hemodynamics downstream of a stenosis become disturbed. The resulting flow disturbances produce characteristic features in Doppler ultrasound scans which are of diagnostic value. A technique has been investigated for extracting the streamwise turbulence intensity, or normalized velocity variance, using Doppler ultrasound in an arterial flow model. The model consists of a pump and 1.3-cm i.d. polyurethane tubing which is optically and acoustically transparent. Correlation between Doppler ultrasound and laser Doppler anemometry (LDA) measurements was examined in laminar as well as turbulent flow. Continuous flow has been employed initially and the work will be extended to pulsatile flow to replicate cardiac output.

11:00

4aBV12. Ultrasonic wavefront distortion caused by human abdominal wall layers. Laura M. Hinkelman (Dept. of Meteorol., Penn State Univ., University Park, PA 16802), T. Douglas Mast (Penn State Univ., University Park, PA 16802), Michael J. Orr, and Robert C. Waag (Univ. of Rochester, Rochester, NY 14627)

The relative importance of the fat and muscle layers of the human abdominal wall in producing ultrasonic wavefront distortion has been assessed based on direct measurements. Specimens employed included 6 whole abdominal wall specimens and 12 partial specimens obtained by dividing each whole specimen into a fat and a muscle layer. In the measurements, a hemispheric transducer transmitted a 3.75-MHz ultrasonic pulse through a tissue section maintained at 37°. The received wavefront was measured by a linear array translated in the elevation direction to form a two-dimensional aperture with overall dimensions 92.16×46.08 mm² and a measurement spot size of 0.72×1.44 mm². Differences in arrival time and energy level between the measured waveforms and computed references that account for geometric delay and spreading were calculated. After correction for the effects of geometry, the received waveforms were synthetically focused at 180 mm. The characteristics of the distortion produced by each specimen and the quality of the resulting foci were analyzed and compared. The results indicate that both fat and muscle layers contribute significantly to the distortion of ultrasonic beams by the abdominal wall. However, the spatial characteristics of the distortion produced by fat and muscle layers differ substantially, and the total distortion produced by the abdominal wall is not equivalent to the sum of the distortion produced by the layers.

11:15

4aBV13. Infrasound method for bone mass measurements. Dimitri Donskoy (Davidson Lab., Stevens Inst. of Technol., 711 Hudson St., Hoboken, NJ 07030)

Studies of application of acoustic energy for noninvasive skeletal diagnostics have shown feasibility and demonstrated the advantages of utilizing acoustic techniques for bone mass measurements. Unlike conventional radiological techniques, acoustic techniques emit no radiation, are low cost, and utilize equipment which is portable and easy to operate. Although a significant number of acoustic tests have been performed, these techniques have not been used as bone diagnostic tools in clinical applications because of the difficulties in the interpretation of measurements. Thus, ultrasound velocity and attenuation depend on density as well as on certain other properties of bone. The effect of soft tissues creates additional difficulties in the interpretation and use of these techniques. The proposed infrasound method involves measurements of the "rigid body" resonance of a tibia or ulna with the use of specially designed arrangement, which greatly simplifies a biomechanical model of tibia (ulna). This makes the

measurement and its interpretation much more accurate. Numerical analysis of the developed model showed that the mass of bone and its loss can be accurately measured and these measurements do not depend on variation in bone flexibility and soft tissue parameters. An experimental setup was arranged and *in vivo* measurements were conducted and proved the basic concept of the proposed method. [Work supported by NASA.]

11:30

4aBV14. Methods of integral equations in tasks of bio-organs sound scattering. V. Genis (Biomed. Eng. and Sci. Inst., Drexel Univ., PA 19104), E. I. Oboznenko, and I. L. Oboznenko (Kiev Polytech. Inst., Kiev, 252056, Ukraine)

Ultrasound scattering on biological, arbitrary-form structures is considered. Methods of integral equations by volume and surface integration are used in the task. The methods allow one to explore scattering from the surface and volume of arbitrary pathological insertions. A method of small perturbations is used. In the first approximation the weak-reflecting pathology is accepted. It is noted that pathology form and value appreciably affect echosignal and spectrum structure. In particular, echosignal spectrums for less dense and denser pathologies vitally are notable. The proposed method allows one to classify the pathologies of bio-organs. Sufficient agreement with theoretical models is shown experimentally on the weak-reflecting object (rubber).

11:45

4aBV15. Numerical calculations of scattering characteristics of inhomogeneities in biological tissue. Vladimir I. Genis (School of Biomed. Eng. Sci. and Health Systems, Drexel Univ., Philadelphia, PA 19104, genis@cbs.ece.drexel.edu), Igor L. Oboznenko (Polytechnic Univ., Kiev 252037, Ukraine), and Stanley Zietz (Drexel Univ., Philadelphia, PA)

Identification of inhomogeneities is one of the main problems in medical diagnosis, particularly in tissue characterization. Often the backscatter signal does not permit discrimination between inhomogeneities in the tissue. Despite high effectiveness of the *B*-mode diagnostic systems, at the present time ultrasonography does not provide a sufficient degree of diagnostic confidence in identification of pathology. Diffraction tomography promises to extend these observations quantitatively and provide information about tissue properties. The main goal of this work is to present a theoretical basis of the scattering process on inhomogeneities and provide some preliminary experimental results confirming theoretical investigations and numerical simulations. The method of integral equations for solving forward and inverse scattering problems is used, and scattering diagrams are presented. Experimental setup is based on one transmitting and seven receiving elements. Presented experimental results support numerical simulations.

THURSDAY MORNING, 19 JUNE 1997

ROOM S, 8:30 TO 11:45 A.M.

Session 4aEA

Engineering Acoustics: Modeling and Predictive Capabilities for Transducers I

David Yeager, Chair

Motorola, Acoustics Technology Laboratory, Room 2319, 8000 West Sunrise Road, Fort Lauderdale, Florida 33322

Chair's Introduction—8:30

Invited Papers

8:35

4aEA1. Design simulation of composite transducers. Wenwu Cao (Whitaker Ctr. for Med. Ultrasonic Transducer Eng., Dept. of Mathematics, and Mater. Res. Lab., Penn State Univ., University Park, PA 16802)

Many advanced transducers, such as those for medical ultrasound imaging, are made of composite materials. These composites contain both piezoelectric active and inactive constituents and therefore have a nonuniform surface vibration profile. This nonuniform surface vibration will in turn influence many characteristics of the acoustic beam pattern produced by the transducer, such as the focusing ability, sidelobes, and the near-field pressure distribution. These complex transducer systems cannot be modeled properly using equivalent circuit models due to the 3-D nature of the transducer vibration. A 3-D finite-element modeling scheme will be given in this presentation. The modeling contains two stages: First, a finite-element method is used to study the nonuniform structural vibrations of the composite transducer and obtain the surface velocity and displacement distributions; second, this surface velocity and displacement information will be fed to the Helmholtz integral to derive the transducer radiation pattern. It is also possible to compute directly the near-field pressure distribution using FEM. Accurate predictions on the transducer behavior are obtained. In addition, stress and electric field concentrations inside the transducer can be calculated. Comparison with experimental results will be shown via an animation.

8:55

4aEA2. Optimal digital compensation networks for acoustic transducers. Paul J. Titterton, Jr. (SRI Int., 333 Ravenswood Ave., Menlo Park, CA 94025)

A digital compensation network is designed which (1) minimizes deviation from the user-determined optimal transducer response, (2) limits the sensitivity to physical perturbations of that response, and (3) ensures actuator authority limitations for known and unknown-but-bounded input sequences. H^∞ , H^2 , l^∞ , and l^1 transfer function norms translate the design goals and specifications into a convex constrained optimization problem, which is solved using semidefinite programming. The methodology can be simply extended to treat transducer arrays and optimal beamforming. A mathematically modeled underwater acoustic transducer provides a numerical example for exploring the optimal trade-off between competing specifications. [This work is partially supported by SRI International.]

4a THU. AM

4aEA3. Capabilities and limitations of “state of the art” modeling techniques for use in the design of electroacoustic transducers. Carlos I. Beltran (Chicago Lawn High Technol., 3321 W. 62 St., Chicago, IL 60629)

Several modeling systems exist to aid in the design and engineering of electroacoustic transducers. Finite element analysis (FEA) can be used to model magnetic circuits, electromagnetic circuits, thermal transfer capacity, mechanical and dynamic properties of structural parts, vibrational modes of diaphragms and other moving parts, piezoelectrics, acoustic impedance, and acoustic radiation. Several computer programs are available to model electroacoustic systems and components via lumped parameter methods derived from electrical-mechanical and electrical-acoustical analogies. Lastly, general purpose mathematics programs and spreadsheets can be used for a wide variety of problems. The capabilities and limitations of these modeling techniques are discussed using real world examples.

4aEA4. Finite-element and boundary-element computer modeling for loudspeaker cone design. Roger E. Shively (Harman-Motive, Inc., 1201 S. Ohio St., Martinsville, IN 46151)

The use of finite-element and boundary-element computer modeling for the design and optimization of loudspeaker cones is demonstrated. First, the specifics of the analytical model and theory are summarized. The experimental setup, which utilizes laser vibrometry is described, and the correlation of analytical FEM and BEM models to experimental models and anechoic measurements for a test case are reviewed. Three case histories are then illustrated: (1) correlation of a high-frequency model, (2) a design case for maintaining a current loudspeaker packaging size with improved frequency response and distortion over the current performance, and (3) a design case for evaluating a new material for a loudspeaker's diaphragm.

4aEA5. Measurement and analysis of loudspeaker component nonlinearities at low frequencies. Thomas P. Heed (Oxford Int., Ltd., 4237 W. 42nd Pl., Chicago, IL 60632)

In the constant displacement region, several displacement-dependent quantities control the behavior of electro-dynamic transducers. This research uses a tensile/compression loadframe to measure the displacement-dependent stiffness, B1-Product, and mechanical losses of various driver components. The individual data sets are curve-fitted by power series expansions. The power series coefficients are used to predict the relative low-frequency total harmonic distortion (THD). Then, the measured components are built into loudspeakers, and the predictions are compared to actual THD measurements. Experimentally, this method is very sensitive. Additionally, the effective stiffness of a transducer's suspension is merely the sum of the power series coefficients representing the spider's and surround's stiffnesses. By comparing various spiders and cones, the effective stiffness of a suspension can be linearized, thereby reducing low-frequency THD. Finite-element analysis (FEA) simulation of nonlinear suspension behavior can facilitate such a linearization process. Lastly, power series coefficients can determine the relative variability of components (e.g., which of two spider designs is more variable). OEM manufacturers of transducers must maintain statistical process control of Q_{Tot} , resonance, and THD. Q_{Tot} , resonance, and THD are all dependent on displacement-dependent quantities. Therefore, reducing component variability is helpful in achieving an acceptable c_{pk} for Q_{Tot} , resonance, and THD.

10:15–10:30 Break

4aEA6. A simplified approach to a nonlinear loudspeaker modeling. Earl Geddes (GedLee and Assoc., 1388 Medinah Dr., Itasca, IL 60143)

Exact solutions of the nonlinear equations for a loudspeaker are extremely difficult to solve as they involve Volterra series expansions and higher-order system functions and impulse responses [A. R. Kaiser *et al.*, J. Audio Eng. Soc. **35**, 421–433]. Indeed, these complex methods are required for exact system simulations, however, they can be greatly simplified with a reasonable assumption about the system. By assuming that the distortion (or nonlinearity) is not so large that it causes an excessive amount of distortion, say about 20% THD, a vastly simplified approach can be developed. It is not always the case that this near-linear assumption can be made, but clearly, it should be accurate for higher performance units. This presentation will show how to use an algebra processor (Maple) to develop these simplified predictions given the nonlinearities of the parameters. These equations will then be used to show the effect that various loudspeaker enclosure designs have on the distortion of a given (typical) loudspeaker using a commercially available software package written by the author.

4aEA7. Loudspeaker drive-unit measurements at excursion. David Clark (DLC Design, 47677 Avante Dr., Wixom, MI 48393)

Loudspeaker drive units are conventionally modeled as linear devices. Evidence is presented that shows that this assumption leads to unacceptable modeling errors for practical conditions of use. This paper presents new techniques for accurate, nondestructive measurement of linear and nonlinear electromechanical parameters. These parameters are used in more complete models which improve accuracy of frequency response and distortion estimates. The measurement workstation uses pneumatic pressure to displace the diaphragm in order for position-dependent parameters to be measured over the full excursion range. Suspension compliance and mechanical damping are measured under conditions which simulate normal operation, thereby capturing the influence of creep, aerodynamic drag, and other second-order characteristics which are not measured separately at present. Moving mass and force factor

measurements utilize a magnetic position servo and gravity to obtain data without displacing the suspension. Conventionally, maximum diaphragm displacement, x_{\max} is calculated from voice-coil and magnet geometry. This ignores the magnetic fringe field and the effect of suspension nonlinearities. An improved value of x_{\max} is derived from the new measurement data.

11:10

4aEA8. Acoustical simulation of active noise control telephone earpiece system. John C. Baumhauer, Jr. (Lucent Technologies, Bell Labs, 6602 E. 75th St., Indianapolis, IN 46250)

A lumped-parameter equivalent circuit representation of an active noise control (ANC) feedback system has been developed for telephone handset earpiece application. Since the feedback controller must be stable to avoid "squealing" in the user's ear, its filter must work with known and robust transducer transfer functions. The earpiece loudspeaker (receiver) is the weak link. As opposed to the feedback electret microphone located between the earpiece and pinna, the moving-coil receiver's response undergoes a large, negative phase change with increasing frequency owing to multiple, highly coupled degrees-of-freedom. Moreover, the magnitude and phase of the receiver's transfer function are shown to depend on its electroacoustic parameter tolerances, such as cone stiffness, coil $B1$, acoustic damping, and handset porting. Another variation is due to the acoustic leak that occurs naturally between the earpiece and the ear. This is not just a function of the earpiece design, but also of the varying handset-to-ear force applied. Not only is low-frequency response in the ear lost with an increased leak, but the high gain feedback magnitude is reduced and the overall phase lag in the receiver becomes greater. Results with simulated ITU-T coupler equivalent circuits demonstrate these effects. A prototype ANC system is discussed.

Contributed Paper

11:30

4aEA9. A BEM approach to effectively modeling differential microphones. David Yeager (Acoust. Technol. Lab., Radio Products Group, Motorola Corp., Rm. 2319, 8000 W. Sunrise Blvd., Ft. Lauderdale, FL 33322)

A flat frequency response, reasonably good signal-to-noise ratio, mechanical robustness, and low cost make omnidirectional electret microphones ideal for telecommunications applications. Differential microphones offer the potential additional advantage of improving the acoustic signal-to-noise ratio by rejecting unwanted background noise. Background noise rejection results from two independent acoustic mechanisms—directional selectivity and proximity. Modeling the acoustic performance of a differential microphone requires accurately representing these mecha-

nisms, which means properly accounting for diffraction and damping effects. For these reasons, differential microphones are more sensitive to mounting conditions and housing geometry, and therefore, designing a differential microphone into a product can be a daunting task. This paper discusses some of the design principles that dictate noise rejection, and suggests methods for modeling differential microphones using numerical techniques such as the boundary element method (BEM). A coupled BEM model enables an acoustical interface between the microphone diaphragm and nearby cavities and ports. This is especially important near resonance. Damping plays a key role in microphone performance, and is an overriding factor in the directional behavior of a differential microphone. Techniques for representing these effects in a BEM model using SYSNOISE will be discussed. [SYSNOISE is a registered trademark of LMS Numerical Technologies, Belgium.]

THURSDAY MORNING, 19 JUNE 1997

ROOM D, 9:00 TO 11:15 A.M.

Session 4aMU

Musical Acoustics: Sound Generation in Musical Instruments

George Bissinger, Chair

Department of Physics, East Carolina University, Greenville, North Carolina 27858

Contributed Papers

9:00

4aMU1. Effects of reed cell geometry on the vibration frequency and spectrum of a free reed. James P. Cottingham (Phys. Dept., Coe College, Cedar Rapids, IA 52402)

In a reed organ the free reed is mounted above the windchest and is typically surrounded by a reed cell cavity with small cross-sectional area. The length of this tubelike cavity varies from just slightly longer than the reed in some instruments to several centimeters longer in others. A series of measurements of the frequency and spectrum of the radiated sound have been made on reeds from American reed organs mounted on a laboratory windchest, using a simulated reed cell of varying dimensions. The presence of the reed cell cavity generally results in a lower frequency of vibration at the same blowing pressure and increased amplitudes of the higher harmonics relative to the fundamental, although some anomalous

effects are observed for longer reed cells. Some hysteresis effects are observed as the blowing pressure is continuously increased and decreased.

9:15

4aMU2. Acoustics of the khaen: The Laotian free-reed mouth organ. Casey A. Fetzer and James P. Cottingham (Phys. Dept., Coe College, Cedar Rapids, IA 52402)

The khaen is constructed with free reeds mounted in bamboo pipe walls inside a carved wooden windchest. Each reed of the khaen sounds for both directions of air flow (inhaling and exhaling). The reed vibration is strongly coupled to the pipe resonance, and the reed will sound only if a small hole near the reed is closed, causing the resonant frequency of the pipe to be near that of the reed. For some examples of khaen made in northeastern Thailand, variations in frequency and sound spectrum with

blowing pressure have been studied, with both positive and negative blowing pressures considered. The relationship between frequency of reed vibration and pipe length has been studied to determine the range of pipe length over which the reed can be made to sound, as well as the amount of frequency shift associated with changes in resonance frequency of the pipe. A comparison is made with a previously reported result that the sounding frequency of the reed-tube combination is higher than the natural resonance frequencies of either the reed or the pipe taken alone. [L. E. R. Picken, C. J. Adkins, and T. F. Page, *Musica Asiatica* **4**, 117–154 (1984).]

9:30

4aMU3. Applications of the 2 DOF network representation of violin cavity modes A0 and A1 to partial cavity volume and *f*-hole geometry changes in an Aluminum violin surrogate. George Bissinger (Dept. of Phys., E. Carolina Univ., Greenville, NC 27858, pbissin@ecuvm.cis.ecu.edu)

Lack of uniqueness in the circuit elements of the 2 DOF Network model [E. A. G. Shaw, *J. Acoust. Soc. Am.* **87**, 398–410 (1990)] for the two lowest violin cavity modes A0 and A1 is one difficulty in its application, e.g., Shaw provided three possible parameter sets with substantial variation among them. To provide constraints on the range of parameters, global fitting of a large database of A0 and A1 frequencies, and their upper bout–lower bout pressure ratios—obtained by adding measured quantities of water (up to 250 cm³) to a flat-plated A1 violin-shaped cavity (volume = 1800 cm³) and varying its orientation—produced an optimized set of parameters close to Shaw's preferred set. With this optimized set of parameters, it was possible to estimate: (1) the mutual inductance contribution to the *f*-hole inductance; (2) a volume dependence of the A0 mode frequency of $f \propto V^{-0.254}$ in good agreement with the experimental value of $V^{-0.269}$; and (3) predictions of the effect of sliding, rotating, and flipping the *f* holes.

9:45

4aMU4. Measurement of direct radiation from violin excited by force hammer impact at bridge. Machele Bailey and George Bissinger (Dept. of Phys., E. Carolina Univ., Greenville, NC 27858)

A violin suspended inside a small anechoic chamber was impacted at the bridge with a small force hammer. A rotatable, semicircular, equispaced seven-microphone array was used to measure the radiation in a 0–2 kHz range over a sphere of radius 30 cm (in the near field ($r < l$) region for $f < 1100$ Hz) at 15° intervals (except in the neck-fingerboard region). The input-force-normalized microphone spectra were then analyzed to separate the contributions from each violin normal mode. On animation, most of the higher-frequency radiation patterns clearly showed apparent “phasiness” among portions of the radiation patterns, an effect primarily related to near-field effects and not the bridge impulse propagation delay through the violin corpus. The measured 3-D radiation patterns for each mode compared well with those calculated with a boundary element program using as input normal mode vibrational data from a prior modal analysis for this violin.

10:00

4aMU5. V-R model predictions of averaged radiation from a violin compared with spatial average of bridge force hammer-excited direct radiation. George A. Bissinger and Machele Bailey (Dept. of Phys., E. Carolina Univ., Greenville, NC 27858, pbissin@ecuvm.cis.ecu.edu)

Mechanical modal analysis data for a violin furnish two essential components for the vibration-radiation (V-R) model [G. Bissinger, *J. Acoust. Soc. Am.* **97**, 3154–3164 (1995)] to reliably predict the averaged radiative properties of a bowed string instrument from its measured mechanical properties, viz., (1) how strongly a particular corpus mode is excited at the bridge and (2) normal mode frequency and shape data for boundary element acoustic radiation calculations of the radiation efficiency for each corpus normal mode. Combining both of these elements in the V-R model gives predictions of a spatially averaged acoustic radiation for a violin.

These were compared with similarly spatially averaged measurements of direct radiation from a violin excited by force-hammer impact at the bridge using a rotating semicircular ($r = 30$ cm) seven-microphone array in a small anechoic chamber covering almost an entire sphere around the violin. The V-R model calculations for each mode compared quite well in general to the measured average radiation.

10:15

4aMU6. Important violin resonance frequencies. Oliver E. Rodgers (179 Kendal Dr., Kennett Square, PA 19348, olerodj@juno.com)

Experimental determination of the natural frequencies and nodal line patterns of those modes which produce sound was made of a test violin which had previously been analyzed by George Bissinger using an experimental modal analysis system [*J. Acoust. Soc. Am.* **97**, 3154–3164 (1995)]. Excitation of the test violin was done by bowing the instrument and the analysis system was the CONQT software previously used by the author. Nodal lines were defined approximately with a small microphone which was held very closely to the plate surface to seek lines of minimum response. Comparison with the Bissinger results points out the very few vibration modes in the first 1000 Hz which contribute to the overall tone of the instrument. These modes are found among those displayed in Bissinger's work. Only two modes of the air cavity and four modes of the gross mechanical system are important in this frequency range. The CONQT data suggest that tone quality is determined strongly by the five to seven prominent higher modes, which are complex motions of small areas of the plates, and/or by the ability of the lower modes to transmit sound at off-peak frequencies.

10:30

4aMU7. Vibrational modes of two violins. Mark Roberts and Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115)

Using electronic TV holography, as well as other methods for observing vibrational motion and sound radiation, the normal modes of vibration in two violins have been studied. The principal modes observed in a Hutchins violin showed fairly good agreement with those reported by Marshall [*J. Acoust. Soc. Am.* **77**, 695–709 (1985)] for the same violin. The strongly radiating T_1 and C_3 modes appear to be doublets, and this phenomenon is discussed. Vibrational modes excited by a force applied to the bridge by internal sound pressure, and by the sound field of a loudspeaker are compared.

10:45

4aMU8. Subharmonic generation in cymbals at large amplitude. Charles Wilbur and Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115)

The aftersound in cymbals has been attributed to chaotic behavior at large amplitude, which provides a mechanism for the conversion of low-frequency energy to a rich assortment of modes of much higher frequency [N. H. Fletcher, “Nonlinear dynamics and chaos in musical instruments,” in *Complex Systems: from Biology to Computation*, edited by D. G. Green and T. Bossomaier (IOS, Amsterdam, 1993)]. When a cymbal is driven sinusoidally at an amplitude slightly lower than that leading to chaotic behavior, the cymbal sound includes harmonics of a fundamental having a fraction (often one-half or one-third) the drive frequency. Nonlinear behavior leading to subharmonic generation as well as chaotic behavior is discussed.

11:00

4aMU9. Investigation into the feasibility of a smart acoustic guitar. Steven F. Griffin (Air Force Phillips Lab., 3550 Aberdeen Ave. SE, Kirtland AFB, NM 87117-5776) and Sathya V. Hanagud (Georgia Inst. of Technol., Atlanta, GA 30332)

In 1990, the “Mendelssohn” Stradivarius violin sold for \$1,686,700. A good violin sells for around \$2000. What is it about the Stradivarius that makes it cost almost 1000 times as much? The structure and geometry of

the two instruments are similar, yet subtle differences in structural dynamics cause them to vibrate differently in response to a violinist's bow. This, in turn, causes differences in the sound produced by the two instruments, which ultimately determines quality and, to a large extent, price. If it were possible to force the less expensive violin to vibrate like the Stradivarius, the legendary sound would follow. This paper explores the potential for the use of active structural/acoustic control to obtain a desired acoustic response in an acoustic guitar. Detailed information on the desired acoustic

response of guitars is available, and experimental specimens are relatively easy to obtain. The model developed is an elastic plate with piezoceramic sensors and actuators backed by a rigid, vented cavity. The sensors and actuators are used with active feedback control to influence the vibration and acoustics of the guitar. The feasibility of favorably changing acoustic guitar vibration and acoustics is examined in detail, including analytical and experimental results.

THURSDAY MORNING, 19 JUNE 1997

ROOM G, 10:30 TO 11:30 A.M.

Session 4aNS

Noise: Progress Report and Discussion on the Continuing Activity on ASA's Role in Noise and Its Control

Bennett M. Brooks, Chair

Brooks Acoustics Corporation, P.O. Box 3322, Vernon, Connecticut 06066

The Technical Committee on Noise, through a subcommittee, is involved in outreach activities related to noise control. The details of a special session on meeting-room acoustics to be presented at the meeting of the Council of Engineering and Scientific Society Executives (CESSE) in Pittsburgh in July 1997 will be reviewed. Plans for a noise seminar to be held in San Diego will also be reviewed. The results of discussions with hotel industry engineers regarding meeting-room acoustics will be presented. Progress will be reported on a video of noise control demonstrations for use in classrooms. The committee is also involved in exploration of the possibilities of hands-on automated hearing testing in public spaces.

THURSDAY MORNING, 19 JUNE 1997

ROOM N, 9:00 TO 10:30 A.M.

Session 4aPA

Physical Acoustics: Porous Media

Craig J. Hickey, Chair

National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677

Contributed Papers

9:00

4aPA1. Modification of Biot's theory of porous materials. H. Tavossi and B. R. Tittmann (Penn State Univ., Dept. of Eng. Sci. and Mech., University Park, PA 16802)

Acoustic waves in porous material are traditionally investigated by using Biot's theory to describe the elastic wave propagation in such materials. However, experimental results obtained by models of porous media made of cohesionless particles forming loose solid matrix with fluid-filled pores, contradict the above theory at high frequencies. These results show that the acoustic wave speed as a function of frequency and particle size goes through a maximum and then decreases with frequency until a cutoff frequency is reached, whereas Biot's theory predicts an asymptotic increase of propagation speed with frequency, without a wave cutoff frequency. The theoretical analysis and the experimental investigation are presented that modify and adapt Biot's theory to the cohesionless porous materials. The equations describing these modifications to the theory take into account the following effects: the contact surfaces between the solid particles, their number concentration, the size of the solid particles, the

depth at which they are situated, the external pressure, their random arrangement, and the propagation direction of the acoustic waves.

9:15

4aPA2. Coupled phase theory for the complex density of rigid-porous materials. Keith Attenborough and Jon Evans (Faculty of Technol., The Open Univ., Milton Keynes MK7 6AA, England, k.attenborough@open.ac.uk)

Despite the fact that most porous materials, of influence in building and urban acoustics, are either fibrous or granular in nature, theoretical models used to predict their acoustical behavior at audio frequencies are derived from an assumed *pore-based* microstructure. This leads to the need for pore-related parameters that are either difficult to measure or not measurable independently. An alternative is to start from the known size and shape distribution and properties of the constituent solid particles. One possible particle-based approach is coupled phase theory. This has been used extensively to predict sound propagation in suspensions and emul-

4a THU. AM

sions. By means of an expression for the interphase inertial force appropriate to spherical particles, it has been shown that “frameless” Biot theory is equivalent to coupled phase theory modified to include hydrodynamic interactions between the forces acting on individual particles. The two models differ only in their expression of the interphase viscous force. Extensions of the coupled phase approach that are necessary when considering air-filled granular and fibrous materials are discussed.

9:30

4aPA3. Dynamics of single bubbles in laboratory sediments. D. Felipe Gaitan and James M. Sabatier (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677, gaitan@olemiss.edu)

Increased interest in the acoustical properties of sediments has been motivated by the fact that in shallow-water environments the seabed is the dominant factor controlling propagation. With this in mind, the dynamics of a gas bubble surrounded by a poro-elastic medium in a laboratory sediment have been studied. Measurements of the effect of the surrounding medium on the resonance frequency and the quality factor of the bubble will be presented. [Work supported by ONR.]

9:45

4aPA4. On the dynamic drag coefficient for wave propagation through fluid-saturated highly porous open-cell foams. Robert E. Slade, Keith Walton (School of Mathematical Sci., Univ. of Bath, Bath BA2 7AY, UK), Colin A. Mead, Alan T. Parsons (Winfrith Technol. Ctr., Dorchester, UK), Roger J. Pinnington, and Anthony R. Briscoe (Univ. of Southampton, Southampton, UK)

This paper considers the damping of waves propagating in high porosity (97%) open-cell foams which are fully saturated by a viscous fluid. The technique of homogenization is applied on two spatial scales to the fluid-foam system. The resulting equations are found to be those of Biot [M. A. Biot, J. Acoust. Soc. Am. **28**, 168–178 (1956)] but reduce to the simpler form given by Chase [D. M. Chase, J. Acoust. Soc. Am. **65**, 1–8 (1979)] when particular properties of the solid foam matrix are taken into consideration. A number of important material parameters are identified, and of particular significance is the measure of drag between the solid and fluid components. The results of numerical fluid flow experiments are used to obtain estimates of the damping coefficient. Comparisons are made with measurements of the steady state and frequency-dependent damping coefficients, which are determined using specially constructed experimental rigs. [Published with the permission of the Contoller of Her Britannic Majesty’s Stationary Office.]

10:00

4aPA5. The behavior of transient elastodynamic waves at a linear slip interface. Martin D. Verweij (Lab. of Electromagn. Res., Dept. of Elec. Eng., Ctr. for Technolog. Geosci., Delft Univ. of Technol., Mekelweg 4, 2628 CD Delft, The Netherlands, m.d.verweij@ctg.tudelft.nl) and Christopher H. Chapman (Schlumberger Cambridge Res., Cambridge CB3 0EL, England)

A method is presented for the analysis of the transmitted and reflected transient elastodynamic wavefield at a fracture that may be modeled as a linear slip interface. This interface model implies that the traction is continuous at the fracture, while the displacement may show a jump discontinuity that is proportional to this local traction. To start with, for both the *SH* and the *P/SV* wave systems, transform domain expressions are determined for the transmitted and reflected wavefields at the interface. Subsequently, an almost completely analytical transformation back to the space-time domain is carried out. The method provides the exact time domain transmitted and reflected waves (including the body wave, interface wave, and head wave contributions) that are caused by a spatially curved, incident wavefield due to a point source. Various numerical results are given. For the *SH* case, the directionally dependent filter behavior of the fracture is shown. For the *P/SV* case, first the reflection, transmission, and conversion of body waves is presented. The next set of results involves the interface wave contribution and includes pictures of the particle movement on both sides of the fracture. Finally, the presence of headwaves in case of *SV*→*P* and *SV*→*SV* reflection is shown.

10:15

4aPA6. Estimating pile bearing capacity via acoustic and seismic measurements. Chung-Ham Yang (Dept. of Civil Eng., Penn State Univ., University Park, PA 16802, cxy122@psu.edu) and H. Randolph Thomas (Pennsylvania Transportation Inst., University Park, PA 16802)

Current dynamic methods estimate pile bearing capacity through the measurements of in-pile stress waves. Typically, accelerometers and strain gauges are mounted to the pile shaft. This on-pile instrumentation is expensive and interrupts the driving operation. The high cost and interruptive nature of the on-pile instrumentation imposes practical limits on the number of piles that can be tested. The limited amount of tested piles cannot reliably represent the remaining piles that might be driven through a variety of subsurface conditions. This research is aimed at developing a method for estimating pile bearing capacity through the acoustic and seismic measurements. The pile–soil system is modeled as a single-degree-of-freedom mass-spring-slider-dashpot system. Based on principles of vibration, the pile bearing capacity is analytically derived as a function of the vibrating mass, damped natural frequency, damping ratio, and soil quake. The damped natural frequency and the damping ratio are determined from ground waves and air pressure waves, and the vibrating mass and soil quake are estimated based upon *in-situ* soil properties. The measurement of ground waves needs no on-pile instrumentation and does not interrupt the driving operation. This enables every pile at a site to be tested.

Session 4aPP

Psychological and Physiological Acoustics: Masking and Temporal Effects (Poster Session)

Lawrence L. Feth, Chair

Department of Speech and Hearing Sciences, Ohio State University, 110 Pressey, 1070 Carmack Road, Columbus, Ohio 43210

Contributed Papers

All posters will be on display from 8:30 a.m. to 11:30 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:30 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 11:30 a.m.

4aPP1. Growth of masking by harmonic complexes with high crest factors. Marjorie R. Leek and Van Summers (Army Audiol. & Speech Ctr., Walter Reed Army Med. Ctr., Washington, DC 20307-5001)

There is a strong link between the discrimination of small phase changes in highly peaked harmonic complexes and the sensation level of the stimuli. This relationship may stem from a differential ability across listeners to code dynamic amplitude changes within waveform periods. To test this hypothesis, growth of masking was measured for flat-amplitude harmonic complexes varying in phase structure and level. Maskers were constructed with odd harmonics in cosine phase and all even harmonics shifted in phase from 0 to 90 degrees. Masked thresholds were measured for a 5-ms Hanning-windowed probe centered either at the primary peak of the masker waveform or at a secondary peak near the midpoint of the period. Growth-of-masking functions when the probe is placed at the primary waveform peaks are generally linear. However, in normal-hearing listeners, masking of probes placed at the secondary peaks grows more nonlinearly. Reduced masking observed at low stimulus levels may be related to compression of the high-amplitude primary peaks which act as forward maskers. Differential forward-masking by the large primary peaks across listeners with normal hearing and those with hearing impairment will be related to the ability to discriminate complexes with differing phase structures. [Work supported by NIH.]

4aPP2. Psychometric functions for detecting an increment added to a continuous or gated pedestal. Huanping Dai and Donna L. Neff (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131)

Green *et al.* [J. Acoust. Soc. Am. **66**, 1051–1056 (1979)] reported that a psychometric function for detecting a 1000-Hz increment was two to three times steeper with a continuous rather than with a gated pedestal. Intrigued by their finding, the observations to frequencies of 250 and 4000 Hz were extended, and to signal types including tone in noise and noise in noise. Psychometric functions of three normal-hearing listeners were constructed from 2IFC adaptive tracks and fitted with a Gaussian function: $p_c = \Phi(d'/\sqrt{2})$, where $d' = (x/\alpha)^\beta$, x is the signal level, α is the threshold, and β is the slope. In all conditions, slopes obtained with continuous pedestals were steeper than those obtained with gated pedestals. For tone-in-tone conditions [$x = 10 \log(1 + \Delta I/I)$], slopes differed by a factor of 2 at 250 and 1000 Hz, and by a factor of 4 at 4000 Hz. For tone-in-noise conditions [$x = E/N_0$], slopes differed by a factor of 2 across the three frequencies. For noise-in-noise conditions [$x = 10 \log(1 + S_0/N_0)$], the slopes differed by a factor of 1.5. Implications of these data for existing models of intensity coding will be discussed. [Work supported by NIH/NIDCD.]

4aPP3. Masking of pure tones by sinusoidally amplitude-modulated tonal maskers. Melanie J. Gegan, Sid P. Bacon, and Jungmee Lee (Psychoacoust. Lab., Dept. of Speech and Hearing Sci., Arizona State Univ., Tempe, AZ 85287-1908, spb@asu.edu)

In experiment 1, masking patterns were obtained with an 80-dB SPL, 500-ms sinusoidally amplitude-modulated (SAM) masker ($m=1.0$). The 30-ms signal was centered at a masker peak or masker valley. Masker frequency (fm) was 750, 1350, or 2430 Hz; signal frequency (fs) was 0.8, 0.9, 0.95, 1.0, 1.2, 1.44, or 1.62 fm . Thresholds were generally higher for a signal at masker peak. The magnitude of this effect was governed by fs/fm , rather than by fs . In experiment 2, growth-of-masking functions ($fm=1350$ Hz, $fs=1.44 fm$) were measured for a SAM masker ($m=0.5, 0.75$, or 1.0). These thresholds were compared with those obtained with unmodulated maskers in forward or simultaneous masking. The comparisons suggest that thresholds for a signal at a peak of a SAM masker are due to simultaneous masking, while those in a valley are due primarily to forward masking when $m=1.0$ or simultaneous masking when $m=0.5$ or 0.75 . Finally, the slope of the masking functions in simultaneous masking (unmodulated masker or signal at peak of SAM masker) changed from a slope greater than 2.0 to a slope of 1.0 at the highest levels; this will be discussed in terms of basilar membrane nonlinearity. [Work supported by NIDCD.]

4aPP4. Comparisons between tone-in-noise and phase sensitivity tasks. Brian Branstetter, Bruce G. Berg, Curt Southworth, and Lisa Khuu (Dept. of Cognit. Sci., Univ. of California, Irvine, CA 92612, bgberg@uci.edu)

According to critical band theory, phase sensitivity as a function of stimulus bandwidth and detection thresholds for a tone masked by noise are both indicators of the frequency resolution of peripheral auditory filters. The relation between these two measures is examined. For the phase sensitivity task, stimuli consist of three tones with constant intensity and randomly sampled phases. Two identical sounds are played on each trial together with a third having different randomly sampled phases. Listeners identify the odd sound. The frequency separation between tones is varied according to an adaptive staircase procedure which yields an estimate of the bandwidth at which the odd sound is identified with 71% accuracy. Estimates are obtained for center frequencies of 500, 1000, 2000, and 4000 Hz. For the second task, thresholds for 500, 1000, 2000, and 4000 Hz tones masked by broadband Gaussian noise are estimated. Results show that performance levels for different conditions within each task are highly correlated, whereas across-task correlations are either zero or slightly

negative. These findings demonstrate that predictions from critical band theory do not generalize to the level of individual listeners. [Work supported by ONR.]

4aPP5. Influence of masker phase structure on tone detection by normal-hearing and hearing-impaired listeners. Van Summers and Marjorie R. Leek (Army Audiol. and Speech Ctr., Walter Reed Army Medical Ctr., Washington, DC 20307-5001)

For normal-hearing listeners, harmonic complexes with equivalent power spectra but which differ in temporal waveform shape (i.e., in phase spectra) may also differ in their effectiveness as maskers. Summing harmonic components in positive-Schroeder phase generally produces a less effective masker than a negative-Schroeder complex [Kohlrausch and Sander, *J. Acoust. Soc. Am.* **97**, 1817–1829 (1995)]. This influence of phase structure on masking effectiveness may reflect nonlinear compressive processing in a healthy cochlea [Carlyon and Datta, *J. Acoust. Soc. Am.* **99**, 2542 (1996)]. This study measured detection thresholds of normal-hearing and hearing-impaired listeners for tones masked by complexes in positive- and negative-Schroeder phase. Masker stimuli contained all harmonics between 200 and 5000 Hz of a 100-Hz fundamental with harmonics set equal in amplitude. Three probe frequencies (1000, 2000, and 4000 Hz) and three probe levels (60, 70, and 80 dB SPL) were tested. For normal-hearing listeners, positive-Schroeder stimuli were consistently less effective maskers than negative-Schroeder complexes, particularly at the 60-dB probe level. Differences in masker phase structure had less effect at high presentation levels and for hearing-impaired listeners. The results are consistent with nonlinear compressive auditory processing by normal-hearing listeners which is reduced at high presentation levels and in the presence of sensorineural hearing loss. [Work supported by NIH.]

4aPP6. Psychophysical suppression as a function of signal frequency. Jungmee Lee and Sid P. Bacon (Psychoacoustics Lab., Dept. of Speech and Hearing Sci., Arizona State Univ., Tempe, AZ 85287-1908)

Psychophysical suppression was investigated across signal frequency (250, 500, 1000, 2000, and 4000 Hz) in a forward-masking paradigm. Masker duration was 200 ms, signal duration was 20 or 40 ms, and signal delay was 0 or 20 ms. When using a noise masker (spectrum level of 40 dB), the amount of suppression was determined by subtracting threshold in the presence of a broadband masker from that in the presence of a critical band masker. When using a tonal masker (masker level of 50 dB SPL, suppressor level of 70 dB SPL, with the suppressor frequency being 1.2 times the masker/signal frequency), the amount of suppression was determined by subtracting the threshold in the presence of the masker plus suppressor from that in the presence of masker alone. For both masker types, the amount of suppression increased as signal frequency increased up to 1000 Hz, but then reached an asymptote or decreased somewhat as signal frequency increased to 4000 Hz. [Work supported by NIDCD.]

4aPP7. Modified masking period patterns as a function of frequency region and masker bandwidth. David A. Eddins (Psychoacoust. Lab., Dept. of Speech and Hear. Sci., Indiana Univ., Bloomington, IN 47405 deddins@indiana.edu)

The auditory system must encode dynamic aspects of acoustic stimuli at frequencies throughout the audible spectrum. This coding is limited by the temporal resolution of the ear. Recent measures of modulation detection and temporal gap detection thresholds with narrow-band noises of fixed bandwidth indicate that temporal resolution does not vary with noise carrier frequency from 500 to 4400 Hz. The present study investigated temporal processing using a modified masking period pattern technique. Masked thresholds were measured for pure tones of 500, 2000, and 4000 Hz that were 400 ms in duration. Maskers consisted of narrow-band noises centered on the signal frequency that were either unmodulated or sinusoi-

dally amplitude modulated (modulation depth=1.0) at frequencies from 4 to 256 Hz. The noise bandwidth ranged from 200 to 1600 Hz at each center frequency. By comparing masked thresholds for the unmodulated and amplitude modulated noises, one may assess both within-channel (narrow-band noises) and across-channel (wideband noises) temporal processing. Preliminary results indicate that thresholds for a given masker bandwidth depend strongly on modulation frequency but vary little with signal frequency over the range tested. The data will be discussed relative to measures of temporal resolution and comodulation masking release.

4aPP8. Effect of number and frequency spacing of masker components on multitone masking. Eunmi Oh and Robert A. Lutfi (Dept. of Psych. and Dept. of Commun. Disord., Univ. of Wisconsin, Madison, WI 53706)

Neff and Green [*Percept. Psychophys.* **41**, 409–415 (1987)] reported that the masking of single tones by random-frequency multitone maskers is greatest when the masker is comprised of between 10 and 50 tonal components. In this paper we present data to suggest that such results are related to the frequency spacing of masker components. Two conditions were run. In the first, the number of tones comprising the masker was varied from 20 to 906, with frequencies chosen at random over a fixed range from 0.1 to 10 kHz. In the second, the number of tones was fixed at 20, and the frequency range over which components were chosen was varied to match the frequency spacing of components in the first condition. An adaptive two-interval, forced-choice procedure was used to measure masked thresholds for a 1.0-kHz signal. Both conditions yielded large elevations in signal threshold. This was true even after compensating for the effects of energetic masking produced by components nearest the signal. Compensated thresholds showed little dependency on component number, but decreased systematically with decreasing frequency spacing of components. The results are consistent with a model of multitone masking based on component-relative entropy [R. A. Lutfi, *J. Acoust. Soc. Am.* **94**, 748–758 (1993)].

4aPP9. Range of performance for two tasks with random-frequency context tones. Donna L. Neff, Christina J. Kessler, Toktam Sadralodabai, and Traci R. Gleason (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131, neff@boystown.org)

Forty listeners were tested on two tasks known to show large individual differences in performance. In one task, termed sample discrimination, pairs of target tones were drawn from each of two overlapping frequency distributions, and listeners indicated which pair came from the higher distribution. Target tones were presented alone and in the presence of two flanking context tones that were either fixed at 500 and 4000 Hz or had Gaussian variation centered at these two frequencies. In the second task, masked thresholds were determined for a 1000-Hz tone in the presence of ten-component, random-frequency, simultaneous maskers. The maskers were drawn on each presentation from a pool of either 200 or 10 waveforms. For sample discrimination, the range of performance was large for both the no-context and fixed-frequency context conditions. For the Gaussian-context condition, few listeners exceeded chance performance. Significant effects of testing order were found, with exposure to Gaussian-context conditions degrading later performance. For masking, performance did not differ for large versus small waveform pools, and the 35-dB range of performance corroborated previous work. For both tasks, random-frequency context degraded performance for the majority of listeners. [Work supported by NIDCD.]

4aPP10. Analysis of the performance of a model-based optimal auditory processor on a simultaneous masking task. Lisa C. Gresham and Leslie M. Collins (Dept. of Elec. Eng., Box 90291, Duke Univ., Durham, NC 27708-0291, lcollins@ee.duke.edu)

Psychoacoustic experiments indicate that the human auditory system is less sensitive than predicted by classical signal detection theory on a simultaneous masking task [C. M. Reed *et al.*, J. Acoust. Soc. Am. **53**, 1039–1044 (1973)]. Traditionally, the difference between the matched filter and the experimental performance has been compensated for by assuming an additional internal additive noise. However, this noise has not been completely explained in physiological terms. The inconsistency may also partially be the result of the suboptimal application of signal detection theory. To investigate this discrepancy, signal detection theory is integrated with Patterson's model of the human auditory system [R. D. Patterson *et al.*, J. Acoust. Soc. Am. **98**, 1890–1895 (1995)]. The performance of the optimal detector for a simultaneous masking task is compared to experimental data, allowing the theoretical performance bounds to be determined and the model to be verified. The results of this work suggest that with a more appropriate application of signal detection theory it is possible to partially explain the difference in performance between the matched filter and experimental data as the result of physical processes inherent in the auditory system. Utilizing *a priori* information regarding these processes results in an approach which more accurately predicts psychophysical behavior.

4aPP11. Bandwidth of spectral resolution for the “C-O-G” effect in vowel-like complex sounds. Qiang Xu, Lawrence L. Feth (Dept. of Speech and Hear. Sci., Ohio State Univ., Columbus, OH 43210), Jayanth N. Anantharaman, and Ashok K. Krishnamurthy (Ohio State Univ., Columbus, OH 43210)

Several investigators [e.g., Chistovitch *et al.* (1979); Carson *et al.* (1975); Beddor and Hawkins (1990)] report that listeners match an adjustable single-formant to the spectral center of gravity of a two-formant signal, as long as the two formants are separated by 3.5 Bark or less. This suggests that listeners use some form of spectral averaging in processing vowel sounds. Assmann's 1991 results do not confirm the C-O-G predictions. The preliminary work [Feth *et al.* (1996)] used the double-staircase procedure to confirm the C-O-G effect for two-formant complex sounds using both periodic pulse train and white noise excitation of the filter used to generate the signals. In the work to be reported here, the relationship between the C-O-G effect and an independent measure of spectral resolution is explored. In 1977, Feth and O'Malley used two-component complex tones described by Voelcker (1966) to estimate spectral resolving power. Resolution bandwidths paralleled the growth of the critical bandwidth with center frequency, but were 2.5–3 times wider. For the same listeners, the limit of spectral averaging for the C-O-G effect and the resolution bandwidth using Voelcker tone pairs have been measured. These results will be compared with predictions of an IWAIF model. [Work supported by a grant from AFOSR.]

4aPP12. Effects of masker uncertainty on masking by familiar sounds. Eunmi Oh and Robert A. Lutfi (Dept. of Psych. and Dept. of Commun. Disord., Univ. of Wisconsin, Madison, WI 53706)

Detection thresholds for a tone in an unfamiliar tonal pattern are greatly elevated under conditions of masker uncertainty [D. L. Neff and D. M. Green, Percept. Psychophys. **41**, 409–415 (1987)]. The present experiment was undertaken to determine whether a similar result is obtained when the masker is a familiar sound. Fifty common environmental sounds served as familiar maskers (e.g., dog bark, door slam, phone ring, etc.). Fifty Gaussian noise samples served as unfamiliar maskers. The magnitude and phase spectra of each masker were analyzed into 906 spectral components. Masker uncertainty was introduced by using fewer than 906 components to synthesize each sound, and by selecting the frequencies of these components at random on each presentation. Masked thresholds for a 1.0-kHz signal were obtained using an adaptive, two-interval, forced-choice

procedure. Results with noise replicate earlier results showing largest elevations in threshold for 10–20 spectral components. Results with familiar sounds showed a similar pattern but with even larger elevations in threshold. Trial-by-trial analyses showed a small detection advantage for sounds rated as highly familiar by listeners. Overall the results fail to provide strong evidence that listeners use stimulus familiarity to overcome the effects of masker uncertainty. [Research supported by NIDCD.]

4aPP13. Factors in the detection of environmental sounds. Brian Gysi (Psych. Dept., Indiana Univ., Bloomington, IN 47405), Charles S. Watson, and Gary R. Kidd (Indiana Univ., Bloomington, IN 47405)

In separate experiments, detection and identification thresholds were obtained for a set of 25 common environmental sounds (e.g., dog barking, car starting). In the detection task, threshold values of the event-to-noise ratio (Ev/N) were established using an adaptive tracking procedure. The identification experiment was part of a larger individual differences study reported earlier [Watson *et al.*, J. Acoust. Soc. Am. **99**, 2516(A) (1996)]. A 3AFC identification task was used, with sounds presented at eight levels of Ev/N. Thresholds were estimated from the psychometric functions derived from the data. The correlation between thresholds in the two tasks indicates that the identifiability of these sounds is only weakly related to their detectability. Several models were evaluated to identify the properties that determine the detectability of environmental sounds. Leaky integrator and cross-correlation models were used to predict detection data; identification data were evaluated using a multidimensional scaling algorithm.

4aPP14. Auditory attention beyond the ear: Precues and postcues. Denis McKeown (Dept. of Psych., Univ. of Leeds, Leeds LS2 9JT, UK)

In a “precueing” experiment listeners may come to pay attention to a narrow cued frequency band and fail to monitor other frequencies. In the experiments a cueing tone was presented at some short delay before or after a target signal. In experiment 1, one of two cues at different frequencies was presented shortly before the single interval of a yes/no signal trial. Signal and cue were at the same frequency on 50% of the trials, so the cue was not informative as to signal frequency. Performance was analyzed separately for trials in which the signal occurred at the same frequency as the cue (valid trials) and when it was at a different frequency (invalid trials); the cue to signal delay varied. At short delays performance was enhanced on valid relative-to-invalid trials, but not at longer delays. In experiment 2, the cue was presented following the signal. In this post-cueing experiment subjects responded following the cue; the delay from signal to cue was varied. Performance was below chance performance on invalid trials but above chance on valid trials. These findings suggest that attended listening may modulate the processing of detected signals rather than modulate detection itself.

4aPP15. Change in envelope beats as a possible cue in comodulation masking release (CMR). Emily Buss, Joseph Hall III, and John Grose (Dept. of Surgery, Div. of Otolaryngol., Univ. of North Carolina, Chapel Hill, NC 27599-7070)

The detection advantage associated with masker envelope coherence across frequency has typically been described in terms of comparisons across auditory channels. More recently it has been suggested that analysis of the output of a wider initial filter, similar to that suggested for the TMTF, can account for the data [Berg, J. Acoust. Soc. Am. **100**, 1013–1023 (1996)]. In particular, a change in envelope beats was proposed as the cue to the addition of a pure-tone signal. Data are presented for the detection of a tone added to multiple maskers with coherent envelopes. In one condition a change in envelope beats is an accurate potential cue, and in others it is a much less reliable or unreliable indicator of the presence of the signal. All conditions employing maskers with coherent envelopes produce very similar thresholds, and all showing improved sensitivity over the case of detecting a signal added to a single masker centered on the

signal frequency. Results are interpreted as evidence that a change in envelope beats does not form the basis of detection in CMR. One version of the cued-listening hypothesis, an across-channel model, is reconsidered in light of these results. [Work supported by NIH-NIDCD.]

4aPP16. Further examination between the relationship of auditory filter bandwidth and temporal resolution. Elizabeth A. Strickland (AUS Dept., Purdue Univ., West Lafayette, IN 47907-1353)

In a previous paper, the relationship between auditory filter bandwidth and temporal resolution was examined in the context of an envelope detector model [E. A. Strickland, J. Acoust. Soc. Am. **97**, 3330(A) (1995)]. This model predicts that temporal resolution will be limited by peripheral filtering at low frequencies, and by a central limitation at high frequencies. In the previous paper, time constants as measured by temporal modulation transfer functions (TMTFs) were compared to time constants predicted from the 3-dB bandwidths of auditory filters, for TMTFs and filters measured as a function of frequency region and level. Time constants for the TMTFs covaried with those estimated from the filters for low frequencies and levels, but were constant at high frequencies, as predicted. In the present paper, the auditory filters were used in the first stage of the envelope detector model, and TMTFs predicted by the model were compared to measured TMTFs. Time constants for the TMTFs increased as frequency decreased, but not as much as was predicted by the model. The model predicts time constants well for high frequencies, where the limitation is central. [Work partially supported by a post-doctoral fellowship from NIH.]

4aPP17. Temporal gap detection thresholds measured for conditions that minimize off-frequency listening. C. Formby, L. P. Sherlock (Div. of Otolaryngol.-HNS, Dept. of Surgery, Univ. of Maryland School of Medicine, 16 S. Eutaw St., Ste. 500, Baltimore, MD 21201), and T. G. Forrest (Univ. of North Carolina, Asheville, NC 28804)

Temporal gap detection (TGD) thresholds measured for silent gaps as a function of the frequency separation between a fixed-frequency pre-gap sinusoidal marker (F_1) and a variable-frequency post-gap sinusoidal marker (F_2) may be confounded by off-frequency listening (Formby *et al.*, Aud. Neurosci. **3**, 1–20). Evidence of this confounding role comes from simulation experiments with a single-channel envelope detector model. TGD stimulation results suggest that listeners can improve the signal-to-noise ratio and, in turn, TGD performance by adjusting their optimal auditory filter so that it is centered between F_1 and F_2 (Forrest and Formby, Aud. Neurosci. **3**, 21–33). Off-frequency listening between F_1 and F_2 may result in underestimates of TGD thresholds that appear as asymmetries in detection performance for F_2 markers presented below and above the F_1 frequency. To obtain more precise estimates of TGD, unaffected by off-frequency listening, performance has been measured and will be reported for three listeners with markers spaced symmetrically and asymmetrically in linear frequency below and above the center frequencies

$F_c = 500$ and 4000 Hz. These new results will be compared with an otherwise comparable set of results that were likely confounded by off-frequency listening (Formby *et al.*, Aud. Neurosci. **3**, 1–20). [Research supported by NIH.]

4aPP18. Evidence from a simple two-channel model for asymptotic gap detection. C. Formby, M. J. Gerber, L. P. Sherlock (Div. of Otolaryngol.-HNS, Dept. of Surgery, Univ. of Maryland School of Medicine, 16 S. Eutaw St., Ste. 500, Baltimore, MD 21201), and L. S. Magder (Univ. of Maryland School of Medicine, Baltimore, MD 21201)

Detection of silent temporal gaps is characterized by two prominent features when measured as a function of frequency separation between two sinusoids that mark the silent gap. First, over a range of about an octave separation, gap detection thresholds (GDTs) routinely increase as the frequency difference is increased between the two sinusoids. Second, GDTs become asymptotic for greater sinusoidal frequency separations. This characteristic GDT pattern probably reflects two different processes. The first process almost certainly reflects gap detection measured within a single auditory filter or channel. The nature of the second process is less certain, but may reflect across-channel processing of the silent gap stimulus in two or more independent frequency channels. To evaluate the idea that asymptotic GDTs can be explained with a simple two-channel model, GDTs were measured as a function of frequency separation between a pre-gap sinusoid presented to the left ear (channel 1) and a post-gap sinusoid presented to the right ear (channel 2). The resulting GDTs for standard pre-gap sinusoids from 250 to 2000 Hz correspond closely with asymptotic GDTs measured for three listeners for the same set of stimulus conditions presented monaurally. This correspondence of the data supports a two-channel hypothesis. [Research supported by NIH.]

4aPP19. Optimal interstimulus interval for auditory discrimination of intensity, duration, and frequency for normal adults. Amitava Biswas, Charles Watson, Steven Barlow, Brian Gygi, and Ward Drennan (Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405)

Ten normal-hearing college students performed two-interval, two-alternative, forced choice discrimination tasks designed to determine the optimal range of interstimulus (ISI) intervals for auditory discrimination of intensity, duration, and frequency. Stimuli were randomized (roved) along frequency, intensity, and duration dimensions to reduce monotony and minimize the listener's reliance on a remembered standard. The center frequency was 500 Hz, the center duration was 30 ms, and the center intensity was 75 dB SPL. Differences to be discriminated were from 0.75 to 3.0 dB for intensity, 1.5 to 6 ms for duration, and 7 to 20 Hz for frequency. Interstimulus intervals were studied from 70 ms to 15 s. Maximum discrimination scores were observed in the range of ISIs from 1.0 to 3.0 s, for all three stimulus dimensions. These preliminary findings suggest the operation of a common, dimension-independent, sensory processing mechanism within the central nervous system. [Work supported by NIH and Neuro Logic, Inc.]

Session 4aSAa

Structural Acoustics and Vibration: Active Classification of Structures

Guillermo C. Gaunaud, Chair

Naval Surface Warfare Center, Carderock Division, Code 684, West Bethesda, Maryland 20817-5700

Invited Papers

8:00

4aSAa1. Mundane issues in active classification for naval purposes. Tommy G. Goldsberry (Appl. Res. Labs., Univ. of Texas, Austin, TX and Office of Naval Res., 800 N. Quincy St., Arlington, VA 22217-5660, goldsbt@onr.navy.mil)

The recent revolution in computational capability, combined with finite-element analysis techniques, has greatly expanded the ability to analyze vibrational modes of simple and complex structures and acoustic radiation from these structures. These developments have encouraged belief that one may be able to categorize or classify complex structures of interest to the U.S. Navy (e.g., submarines, mines, or swimmers) by analysis of the acoustic energy reflected and reradiated from such structures (echo structure). A detailed discussion of this topic is beyond the scope of this meeting. However, academic researchers wishing to obtain support for their research may benefit from some discussion of the importance of practical limitations such as finite transducer bandwidth, multipath propagation, and population densities on the relevance of classification methods for naval purposes. Some of these issues will be discussed in this presentation.

8:25

4aSAa2. An architecture for wave interaction with complex environments. Leopold B. Felsen (Dept. of Aerosp. and Mech. Eng. and Dept. of Elec. and Comput. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215)

Active classification of structures is an example of wave interrogation of complex environments. An algorithmic architecture is suggested that decomposes global complexity into problem-dependent clusters which interact across interfaces. A cluster is treated either analytically (preferably based on good physics) or numerically, deterministically and/or stochastically, etc., and the interconnects between clusters map representations on one side into those on the other side, subject to the continuity requirements of the relevant field equations. Projecting the field problem onto the cluster boundaries gives rise to a network formulation that impacts the analytic as well as numerical strategies. This field-based approach is coupled to, and modified by, considerations of signal processing. The challenge in the implementation is to effect an "optimal" decomposition into tractable clusters, and to treat each cluster problem in a fashion that anticipates its role as part of a larger conglomerate, taking into account the limitations imposed by the intended pre- and post-processing. Example scenarios are presented to illustrate these concepts.

8:50

4aSAa3. Active classification of submerged structures by means of wave packet decomposition. Nai-chyuan Yen and Louis R. Dragonette (Physical Acoust. Branch, Naval Res. Lab., Washington, DC 20375-5320, yen@noctilio.nrl.navy.mil)

This paper addresses the exploitation of the information contained in a signal by means of wave packet decomposition. This general technique has a wide range of signal processing applications: The application considered here is sonar systems with a particular emphasis on providing a broader view of time-frequency waveform analysis for target classification. The acoustic response of a submerged object depends on its physical shape and structure and on its elastic properties. A methodology of nonorthogonal decomposition is used to reduce the scattered echoes into wave packets associated with the important scattering mechanisms. This decomposition method is based on the energy distribution of the individual components and can be related to the structural scattering physics. In the case of a sonar target, mechanisms such as specular reflections, creeping waves, Bragg waves, and Bloch waves from the different parts of target or scattering centers can be incorporated into the target characterization. By applying pattern recognition logic, the present study can serve as a useful background for new sonar system development with advanced processing techniques and state-of-the-art computer hardware. Examples from numerical simulations and laboratory measurements are used to discuss the effectiveness of this target classification scheme. [Work supported in part by ONR.]

9:15

4aSAa4. Techniques for active classification of underwater structures. G. C. Gaunaud (Naval Surface Warfare Ctr., Carderock Div., Code 684, W. Bethesda, MD 20817-5700), H. C. Strifors (Natl. Defense Res. Est., S-17290 Stockholm, Sweden), P. W. B. Moore (NRaD, San Diego, CA 92152-6267), and H. Huang (Bowie, MD 20715)

Various techniques for the detailed classification of submerged shells insonified by short pulses from either an active sonar or a small explosion are discussed [Ultrasonics **33**, 147–153 (1995)]. The returned shell echoes in several signal domains are examined. These included the frequency, the time, and particularly the *joint* time-frequency domain. The use of *Wigner-type* distributions was

most informative in the latter case. Selected features in these echo-displays which provided information about a certain specific target characteristic were identified. The achieved “*in situ*” classification is rapid, unambiguous, and accurate. Examples dealing with short pulses simultaneously scattered by one, two, or more elastic shells will be shown. Theoretical predictions and measurements show good agreement. This novel analysis in the above domains determines the size, shape, wall thickness, material composition of the shell(s), and their possible filler substance(s), as well as the number of shells involved, and their ranges from the projector. The basis of this technique has been proposed as a possible explanation of how echolocating dolphins successfully identify submerged structures [J. Acoust. Soc. Am. **100**, 2820 (1996)].

Contributed Paper

9:40

4aSAa5. Broadband mine detection and classification—Target responses from a set of 2- to 20-kHz shallow-water experiments.

Peter J. Kaczowski, Christian J. Eggen, and James C. Luby (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105-6698, peter@apl.washington.edu)

The problem of discriminating between underwater mines and similarly sized false targets at ranges of several hundred meters remains a challenge. The Applied Physics Laboratory, University of Washington, in conjunction with Arete Engineering and Technologies Corporation, San Diego, have been developing detection and classification signal processing

schemes using data collected in several shallow-water experiments with a broadband (2–20 kHz) sonar. The use of relatively low frequencies for mine hunting permits greater penetration into bottom sediments for improved detection of buried or partially buried mines. Wide bandwidth signals mitigate the lack of spatial resolution expected from a narrow-band analysis at such frequencies, and provide a wealth of information for classification. A brief description of the sonar and of the experiments conducted in Puget Sound using mines and minelike false targets will be given. Target responses as a function of time, frequency, and aspect angle will be presented, as will results of using several feature extraction methods to reduce the dimensionality of the data.

THURSDAY MORNING, 19 JUNE 1997

ROOM Q, 10:10 TO 11:40 A.M.

Session 4aSAb

Structural Acoustics and Vibration: Scattering from Elastic Structures

Philip L. Marston, Chair

Physics Department, Washington State University, Pullman, Washington 99164

Contributed Papers

10:10

4aSAb1. Novel formalism for resonance scattering of acoustic and elastic waves. Huinam Rhee^{a)} and Youngjin Park (Dept. of Mech. Eng., KAIST Yusung-Ku, Taejon 305-701 Korea, email:hnrhee@ns.kopec.co.kr)

A novel formalism is proposed for the exact isolation of resonances from scattered waves for acoustic wave scattering from elastic or liquid bodies. The resonance scattering function (RSF) consisting of purely resonance information is proposed. Both magnitudes and phases of isolated resonances can be correctly obtained by using the proposed formalism while previous works based on the conventional resonance scattering theory (RST) can produce only magnitude correctly. The reason why previous works using RST could obtain correct magnitude information of resonances for the case of acoustic wave scattering is explained. Acoustic wave scattering from a variety of submerged bodies is analyzed by utilizing the RSF and the isolated resonances were compared with previously published studies. The exact π -phase shift through resonances and at the anti-resonances caused by the interaction between adjacent resonances, which has never been obtained by using RST, shows that the proposed formalism properly extracts the resonances from scattered waves. Due to the differences in phases, the novel formalism and the conventional RST generate different resonance spectrum. For elastic wave scattering, even magnitudes of resonances isolated by the proposed formalism will be different from those obtained by the conventional RST. ^{a)}Also at Dept. of Mech. Eng., KOPEC, 150 Duckjin-Dong, Yusung-ku, Taejon, 305-353 Korea.

10:25

4aSAb2. Scattering calculations by a hybrid GTD-boundary integral equation method. Paul E. Barbone and Ofer M. Michael (Dept. of Aersp. and Mech. Eng., Boston Univ., Boston, MA 02215)

We consider the application of a hybrid asymptotic/boundary integral equation (BIE) method to the problem of scattering from prismatically shaped objects. The hybrid method is based on patching a short wavelength asymptotic expansion of the scattered field to a BIE evaluation of the near field. In patching, the diffracted field shape functions with unknown amplitude are forced to agree smoothly with the solution in the near field along a curve at a prescribed distance from the diffraction points. This allows us to replace the original boundary value problem with an asymptotically equivalent boundary value problem, the domain of which is small and efficiently discretized. Since the domain of the numerical problem is small and may be chosen at will, we completely circumvent non-uniqueness problems associated with “forbidden frequencies.” Thus very high-frequency calculations can be performed using single layer potential equations with no problems of ill conditioning. The hybrid scattering solution shall be compared to a complete analytic field representation found using matched asymptotic expansions. [Work supported by ONR.]

10:40

4aSAb3. Backscattering of sound from ribbed plates and cylinders with oscillators attached to the ribs. Allan D. Pierce (Dept. of Aersp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02115)

A sequence of simple models is discussed to identify and quantify key theoretical issues pertaining to how such oscillators can affect sound scattering. The first such model consists of a sound wave incident obliquely on

a plate with a single rib which has a continuous smear of oscillators along it. The presence of the rib alters the apparent natural frequency of the oscillators and causes the backscattered sound to be anomalously high when the incident sound has the same frequency as the apparent natural frequency. The width of the spectral line in the plot of backscattered cross section versus frequency for the fixed angle of incidence is controlled both by the radiation of flexural waves away from the rib and by the overall loss of energy through radiated sound. Other models considered are an infinitely long cylinder with a single rib, smeared oscillators along the rib, an infinitely long cylinder with a single oscillator attached to a rib, a plate with a periodic array of ribs with attached oscillators, and an infinitely long cylinder with a similar hierarchy of configurations.

10:55

4aSab4. Meridional leaky ray amplitudes for tilted cylinders and high-frequency backscattering enhancements for finite cylinders. Philip L. Marston (Phys. Dept., Washington State Univ., Pullman, WA 99164-2814)

Consider the meridional plane which contains the axis of a circular cylinder and the incident acoustic wave vector. The meridional leaky rays which lie in this plane may be excited for certain tilt angles. Radiation resulting from the reflection of such rays off truncations on a cylinder gives strong high-frequency backscattering enhancements [Kaduchak *et al.*, J. Acoust. Soc. Am. **100**, 67–71 (1996)]. The present work gives ray approximations for such backscattering contributions and for the related contributions to scattering by tilted infinite cylinders. These are evaluated using a general convolution integral formulation [P. L. Marston, J. Acoust. Soc. Am. **100**, 2820(A) (1996)]. The infinite cylinder results for radiation by Rayleigh waves on a solid cylinder agree with the partial-wave series solution. The truncation enhancement for a finite cylinder causes a backward-directed three-dimensional wavefront possessing a vanishing Gaussian curvature. The amplitude from the far-field propagation integral associated with the caustic due to this wavefront can significantly exceed the amplitude for reflection off a rigid sphere having the same radius as the cylinder. The analysis does not rely on thin shell assumptions. [Work supported by the Office of Naval Research.]

11:10

4aSab5. Analytical study of the acoustic scattering from a submerged stiffened cylindrical shell with two endplates. Azriel Harari and Jeffrey E. Boisvert (Naval Undersea Warfare Ctr. Div. Newport, Newport, RI 02841-5047, harari@caspr5.npt.nuwc.navy.mil)

A mathematical model is described for the sound scattering from a finite cylindrical shell with two endplates submerged in a fluid, subjected to an incident wave impinging on the shell from a source in an arbitrary direction. The effect of structural damping and rib-stiffeners on the monostatic scattered response of the structure is considered for a full range of incident wave angles over the nondimensional frequency range $ka = 0.1$ – 30 .

11:25

4aSab6. Scattering and attenuation of an elastic wave by a viscoelastic cylinder. II. SV-wave incidence. Yin-Bin Liu and Chong-Fu Ying (Inst. of Acoust., Chinese Acad. of Sci., Beijing, 100080, PROC)

The scattering and attenuation of a shear wave by a circular, elastic cylinder with and without intrinsic attenuation are investigated. The analytical solutions for scattered and internal fields caused by a normally incident plane SV wave are derived. Resonance scattering, radiation pattern, scattering cross section, and synthetic seismogram are calculated. The calculated synthetic seismograms show that the creeping (diffraction) waves include the shear creeping wave $S_1\hat{S}_1S_1$ and exciting (or converting) compressional creeping wave $S_1\hat{P}_1P_1$. The creeping waves propagate on the surface outside the media and depend mainly on the outside media, and are little affected by the cylinder–absorbing property. The total fields are the superposition of the geometrically transmitted waves ($S_1P_2P_1$, $S_1S_2P_1$, and $S_1S_2S_1$) which go through the cylinder and creeping waves ($S_1\hat{S}_1S_1$ and $S_1\hat{P}_1P_1$) which propagate on the elastic side of the elastic–cylinder interface. The first arrivals within the shadow zone are a diffraction wave $S_1\hat{P}_1P_1$ for a low-velocity inclusion and a transmitted wave $S_1P_2P_1$ for a high-velocity inclusion. [Work supported by NSF of China.]

THURSDAY MORNING, 19 JUNE 1997

DEANS HALL II, 9:00 TO 11:00 A.M.

Session 4aSC

Speech Communication: Perception and Processing of Spoken Words (Poster Session)

Paul A. Luce, Chair

Department of Psychology, State University of New York at Buffalo, Park Hall, Buffalo, New York 14260

Contributed Papers

All posters will be on display from 9:00 a.m. to 11:00 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 9:00 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 11:00 a.m.

4aSC1. Temporal processing and phonotactic probability in spoken word recognition. Michael S. Vitevitch, Paul A. Luce (Dept. of Psych., Univ. at Buffalo, Buffalo, NY 14260), and Jan Charles-Luce (Univ. at Buffalo, Buffalo, NY 14260)

This study examined the influence of phonotactic information on spoken language recognition in adults. Subjects performed a lexical decision task on 120 phonotactically legal bisyllabic CVCCVC nonsense words that varied in their phonotactic probability and primary stress placement. Subjects responded faster to nonsense words containing low probability

initial syllables and slower to nonsense words containing high-probability initial syllables. On the other hand, subjects responded faster to nonsense words containing high probability final syllables and slower to nonsense words containing low-probability final syllables. Our results suggest that two phonotactically based processes may be used during word recognition. The first process is a winnowing of candidate words based on their phonotactic probability. The second process is a confirmatory process in which the lexical status of the sole candidate is confirmed based on the phonotactic probability of the item. [Work supported by NIH-NIDCD.]

4aSC2. The effect of word-final phonemes on spoken word recognition. Caroline S. Miner (Dept. of Psych., Univ. of Connecticut, 406 Babbidge Rd., U-20, Storrs, CT 06269-1020, csm93003@uconnvm.uconn.edu), Carol A. Fowler (Haskins Labs., New Haven, CT 06511-6695), and Jay G. Rueckl (Univ. of Connecticut, Storrs, CT 06269-1020)

Three experiments tested subjects' ability to recognize spoken words based on word-final phonological information. Experiment 1 replicated the finding [A. Salasoo and D. B. Pisoni, *J. Memory Lang.* **24**, 210–231 (1985)] that when presented with phonological information incrementally beginning from word-offset (backward gating), subjects correctly identified spoken words in the absence of word-initial phonemes. Additionally, experiment 1 demonstrated that acoustic neighborhood size is a significant predictor of recognition probability for both forward and backward gated words. In experiment 2, word-initial and word-final acoustic word fragments were used as primes in a crossmodal identity priming task with a naming response. Both word-initial and word-final primes significantly facilitated subsequent naming reaction times. Experiment 3 replicated experiment 2 using an associative priming paradigm. The results are interpreted from a connectionist perspective. [Work supported by NICHD.]

4aSC3. Effects of phonotactic probability on segmentation of words in continuous speech. Daniel E. Gaygen and Paul A. Luce (Lang. Percept. Lab., Dept. of Psych. and Ctr. for Cognit. Sci., SUNY, Buffalo, NY 14260-4110)

At the last meeting of the Society, evidence was presented regarding the role of phonotactic probability in the segmentation of spoken words in continuous speech. Results of further studies examining this issue will be presented. Participants made speeded word detection responses to sequences of spoken stimuli composed of target words preceded and followed by nonwords (i.e., NONWORD–TARGET WORD–NONWORD). Speed and accuracy of detection were examined as a function of the nonword contexts of the target words. In particular, probabilities of segmental transitions from the nonword contexts to the target words were manipulated: Pairs of segments composed of the last segment of the preceding nonword and the first segment of the target word, as well as pairs composed of the last segment of the target word and the first segment of the following nonword, were varied in terms of intraword transition probability (HIGH and NONE) and position specific segment probability (HIGH and LOW). Only CC transitions were used. Both spliced and coarticulated stimuli were used. The implications of our results for the use of phonotactic probabilities in the identification of words in fluent speech will be discussed.

4aSC4. The adaptive value of connotation in speech perception. Lee H. Wurm (Dept. of Psych., SUNY, Binghamton, NY 13902-6000, lwurm@binghamton.edu) and Douglas A. Vakoch (Vanderbilt Univ., Nashville, TN 37240)

In previous work it was demonstrated that reaction times (RTs) in a lexical decision paradigm were related to words' ratings on the dimensions of evaluation, potency, and activity (independent of word frequency). This was found to be true of words in general [D. A. Vakoch and L. H. Wurm, *Cognition and Emotion* (in press)], and also of words from the affective lexicon [L. H. Wurm and D. A. Vakoch, *Cognition and Emotion* **10**, 409–423 (1996)], but there were differences in the way these dimensions appeared to be used in each case. For emotion words, RTs were quickest for words that were rated as bad, strong, and fast (the evaluation \times potency \times activity interaction was significant), a finding interpreted in terms of the adaptiveness of danger avoidance. For words from the general lexicon, RTs were quickest to words that were rated as either *good* and strong, or *good* and fast, as though these represented resources that had to be seized quickly or lost. The current study clarifies the earlier findings, and expands the existing theory to include concepts such as the danger and usefulness of referents. [Work supported by AFOSR and NIMH.]

4aSC5. Word recognition skills of children with phonological disorders. Jan Edwards, Benjamin Munson, and Robert A. Fox (Dept. of Speech and Hearing Sci., Ohio State Univ., 110 Pressey, 1070 Carmack Rd., Columbus, OH 43210)

There has been little research on word recognition skills of phonologically disordered children, although recent work has shown they tend to perform more poorly than age peers on phoneme-identification tasks using synthetic speech stimuli. Six phonologically disordered preschool-aged children and six typically developing controls were asked to identify CVC words in which either the final consonant or the medial vowel had been removed in steps. The stimuli were digitized natural productions of familiar words. For the silent-center task, the phonologically disordered children were significantly less accurate than peers for the two conditions with the least acoustic information. For the gating task, the phonologically disordered children were significantly less accurate than peers for the two conditions with the most acoustic information. They were less accurate than peers at identifying digitized words with no acoustic information deleted, although they were not less accurate with live voice. This suggests that perceptual representation of final consonants for the phonologically disordered children is so fragile that even such a small degradation as a reduction in bandwidth affects their performance. Divergent patterns of group differences for the two tasks are probably related to differences in the nature of acoustic cues for medial vowels and final consonants.

4aSC6. Phonological variation in spoken word recognition. Thomas Deelman, Brian Lang, and Cynthia M. Connine (Dept. of Psych., SUNY, Binghamton, NY 13902-6000)

A common form of phonological variation in American English is flapping. For example, the word "pretty" may be pronounced as [prɪdi] or [prɪti]. A phonological priming experiment was conducted where a lexical decision was made on a flap or canonical target. The target was preceded by itself (repetition condition), by its variant (form condition), or by a phonologically unrelated word (control). The results showed that priming for a canonical variant was comparable in the repetition and form condition. In contrast, priming for flaps was reduced in the form condition compared to the repetition condition. These findings suggest that the underlying voiceless stop is recovered when processing a flapped variant. The results are discussed in terms of surface versus underlying representations as the currency of lexical activation.

4aSC7. The effects of aging on specificity of memory representations for spoken words. Mara B. Goodman, Paul A. Luce (Lang. Percept. Lab., Dept. of Psych. and Ctr. for Cognit. Sci., SUNY, Buffalo, NY 14260-4110, mara@deuro.fss.buffalo.edu), Jan Charles-Luce (SUNY, Buffalo, NY 14260), and Emily A. Lyons (SUNY, Buffalo, NY 14260)

Recent evidence suggests that lexical items are stored in memory as detailed exemplars that encode specific information about characteristics of speaker's voice. Lyons and Luce (1996) found that an incongruity in talker's voice (male/female) in an explicit recognition task led to slowed reaction times for college-aged participants. While this finding suggests that speaker information is preserved in stored exemplars, it can be asked whether lexical representations of aged individuals are equally detailed. Namely, it has been suggested that with age, phonological representations become more abstract or holistic in nature. If this is true, then a mismatch in talker voice should not affect the performance of aged participants. In order to explore this hypothesis, aged and college participants were tested in an explicit memory task. Preliminary results suggest that older adults are, in fact, more rather than less sensitive to variations in the surface characteristics of spoken words.

4aSC8. The interaction of lexical competition and semantic context in spoken word recognition by younger and older adults. Mitchell S. Sommers and Stephanie M. Danielson (Dept. of Psych., Washington Univ., Box 1125, St. Louis, MO 63130)

The present study examined the interaction of two operations, discrimination of lexical candidates and use of semantic context, that are central to the early stages of spoken word recognition. In addition, the investigation was designed to evaluate age differences in the interaction of these two processes. Identification scores were obtained for lexically hard (items phonetically similar to many high-frequency words) and easy (items phonetically similar to only a few low-frequency words) stimuli in three contexts: single words, low-predictability (LP) sentences, and high-predictability (HP) sentences. Although significant differences between easy and hard words were observed in all three contexts, the effects of lexical difficulty were attenuated in the HP, relative to the SW and LP conditions. The principal finding with respect to age was that older adults exhibited greater benefits of contextual information for hard, but not for easy words. The findings are discussed in terms of their implications for the interactive nature of spoken word recognition and for understanding age-related declines in speech processing. [This research was supported by the Brookdale Foundation.]

4aSC9. Positional neighborhood effects on spoken word recognition. Shigeaki Amano (NTT Basic Res. Labs. and Dept. of Psych., Indiana Univ., Bloomington, IN 47405, amano@harmonic.psych.indiana.edu), Gina M. Torretta (Indiana Univ., Bloomington, IN 47405), and Paul A. Luce (Univ. at Buffalo, Buffalo, NY 14260)

Previous work has shown that the composition of similarity neighborhoods has demonstrable effects on spoken word recognition. In a reanalysis of previously collected data of word identification and lexical decision, the effects of similarity neighborhoods on spoken word recognition were further explored by asking two specific questions. (1) Do neighbors at each phoneme position within a spoken word have equivalent effects on word

recognition? (2) Is average neighborhood frequency or the highest neighborhood frequency the superior predictor of recognition performance? Our results show that the number of neighbors defined on the basis of initial phoneme substitution affects word recognition more than the neighbors at other phoneme positions. And, it is found that the word with the maximum frequency in a neighborhood is as good a predictor of recognition performance as the average neighborhood frequency. The implications of these findings for theories of spoken word recognition will be discussed.

4aSC10. Hearing-impaired perceivers' encoding and retrieval speeds for auditory, visual, and audiovisual spoken words. Philip F. Seitz (Army Audiol. & Speech Ctr., Walter Reed Army Med. Ctr., Washington, DC 20307-5001, seitz@wrair-emhl.army.mil)

Perceptual encoding and memory retrieval processing speeds were assessed for spoken words in 26 subjects, mean age 66, with mild to moderate acquired sensorineural hearing loss. Subjects were trained to achieve error-free recognition of a set of ten spoken words in auditory, visual (speech reading), and audiovisual conditions. They then performed the Sternberg item recognition task in each of the modality conditions using the same set of ten words. The task involved presenting memory sets of one to four words, followed by a probe word to which subjects made a speeded "YES" or "NO" button response to indicate whether the probe matched any of the memory set items. Least-squares linear models provided good fits to subjects' memory-set size by reaction time functions (mean $r^2 > 0.90$ for all three conditions). Using the models' intercepts and slopes to represent encoding and retrieval times, respectively, Wilcoxon tests showed significant differences among the conditions with respect to both encoding and retrieval speed, with audiovisual fastest and visual slowest. These results are interpreted as evidence for: (1) audiovisual "benefit" to processing speed in hearing-impaired speech perception; (2) relative inefficiency of encoding visual speech; and (3) representation differences associated with the modalities. [Work supported by NIH-NIDCD.]

THURSDAY MORNING, 19 JUNE 1997

ROOM H, 8:00 TO 11:45 A.M.

Session 4aSP

Signal Processing in Acoustics: Model Based Signal Processing for Acoustic Propagation

David I. Havelock, Chair

National Research Council, Institute for Microstructural Sciences, M-36, Montreal Road, Ottawa, Ontario K1A 0R6, Canada

Chair's Introduction—8:00

Invited Papers

8:05

4aSP1. Model-based signal processing in acoustics: An overview. James V. Candy (Lawrence Livermore Natl. Lab., Univ. of California, P.O. Box 808, L-437, Livermore, CA 94551)

When the signal information to be extracted from a particular measurement is complex and severely contaminated with noise, then more information about the underlying physics, measurement dynamics, and noise must be incorporated into the processor. This concept defines the basic model-based approach to signal processing. Here, the development of model-based signal processing techniques applied to a diverse set of acoustic problems is discussed: (1) the extraction of sounds from a prosthetic heart valve for failure detection; (2) the enhancement of laser ultrasonic sounds in materials for nondestructive evaluation; and (3) the extraction of parameters from ocean acoustic measurements for target localization. Recall that model-based signal processing is a well-defined methodology enabling the inclusion of process, measurement, and noise models into sophisticated processing algorithms. The basic principle relies on the notion that the more information incorporated about the physics, measurements, and noise into the processor, the better (smaller error variance) the enhancement. When these parameters exactly "match" the process, then minimum variance can be achieved, yielding the best (in a mean-squared sense) possible results. Thus model-based processing finds its roots in minimum variance design where a cost function is investigated until it statistically yields the minimum error between predicted and measured values.

8:30

4aSP2. Model-based towed array processing. Edmund J. Sullivan (Code 82101, Naval Undersea Warfare Ctr., 1176 Howell St., Newport, RI 02841, sullivan@tech.npt.nuwc.navy.mil)

Model-based processing is a method of enhancing processing performance (smaller error variance) by the inclusion of physical models of the signal/medium, noise, and measurement systems into the processor. It is based on a state-space approach, which leads to Kalman type estimators. A major advantage of this approach is that it allows for the stochastic aspects of the problem to be included in a natural and self-consistent manner, thereby allowing for modeling errors to be accommodated by the processor. In this work we describe the application of model-based processing to the problem of processing towed array data. Here the focus is on four problems of interest in the field of towed array processing: (1) the bearing and source frequency estimation problem for plane waves; (2) the bearing, source frequency, and range estimation problem for the circular wavefront signal model; (3) problems (1) and (2) generalized to the case of the Rayleigh fading channel; and (4) the case of the stochastic broadband signal model. Results based on simulated data will be shown which will demonstrate the improvement in performance over conventional array processing techniques. In particular, it will be shown how this approach provides a passive synthetic aperture effect in a natural way.

8:55

4aSP3. Performance limits on underwater acoustic matched-field processing. Joseph Tabrikian and Jeffrey L. Krolik (Dept. of Elec. and Comput. Eng., Duke Univ., Box 90291, Durham, NC 27708)

Sensor array processing methods which use full-wave models of complex multipath propagation to facilitate estimation of underwater source locations and channel parameters are known to be prone to ambiguities and environmental variability. Particularly at low signal-to-noise ratios, the performance of matched-field methods may therefore deviate radically from the Cramer–Rao bound which assumes estimation errors are small. To study matched-field processing in the presence of anomalies, this paper considers the Barankin bound. The Barankin bound is particularly useful for determining the threshold signal-to-noise ratio below which any unbiased estimate of the source location or environmental parameters will be prone to large errors. In this paper, the Barankin bound is compared with a maximum likelihood (ML) estimator of source location in an uncertain shallow-water scenario. The results indicate that the Barankin bound predicts the threshold signal-to-diffuse-noise ratio of the ML method to within 3 dB. Further, in the presence of surface shipping, the impact of both vertical and horizontal spatial decorrelation of the interference on matched-field performance is studied for different array geometries. [Work supported by ONR.]

9:20

4aSP4. Linear and nonlinear sound wave propagation in turbulent and inhomogeneous media. Vladimir A. Krasilnikov (Dept. of Acoust., Moscow State Univ., Moscow 119899, Russia)

The author had the pleasure of being involved in the pioneering experiments on sound wave propagation in the turbulent atmosphere (1939–1950). The obtained results were interpreted from the point of view of the, then recently developed, Kolmogorov–Obukhov (KO) statistical theory of locally isotropic turbulence. These results were one of the first confirmations of this theory. It may be said that the KO theory and the experiments mentioned above had stimulated the broad developments in the problem of “waves and turbulence.” The resulting classical theory now provides a basis for the analysis of signals and the computer simulation of sound propagation through turbulence. This paper provides an historical overview of the development of the classical theory for sound propagation through turbulent media. Models of turbulence and experimental verification of these models will be discussed for sound propagation in the atmosphere, as well as in seismic and underwater media. Some examples of the problem of “wave turbulence and nonlinearity” also will be discussed.

9:45–10:00 Break

10:00

4aSP5. Some intrinsic differences between atmospheric and ocean acoustic parabolic equation models. Kenneth E. Gilbert and Xiao Di (Appl. Res. Lab. and the Graduate Program in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804)

It is often assumed that propagation models developed for ocean acoustics can be applied with minimum modification to atmospheric sound propagation. Experience shows that one often finds subtle but significant problems that can lead to serious numerical (prediction) errors, even in cases where the atmospheric environmental inputs seem similar to those in ocean acoustics. In this paper three sources of error and methods for reducing the errors to acceptable levels are discussed. The first error was recently discovered by Erik Salomons [E. M. Salomons, “Improved Green’s function parabolic equation (GFPE) method for atmospheric sound propagation,” J. Acoust. Soc. Am. (submitted)] and concerns the failure of standard FFT algorithms when applied to atmospheric split-step parabolic equation calculations for propagation over hard ground surfaces. The second problem concerns accurate numerical treatment of the surface wave that is excited by a point source over a locally reacting ground surface. The last issue discussed is the need for improved approaches to implementing an outgoing-wave boundary condition in calculations of long-range outdoor sound propagation. Numerical examples are presented showing the effect of the errors on the predicted signal. [Work supported by the Army Research Laboratory and the Pennsylvania State University Applied Research Laboratory.]

10:25

4aSP6. Atmospheric propagation modelling. John M. Noble (U.S. Army Res. Lab., 2800 Powder Mill Rd., Adelphi, MD 20783)

With the developments in modeling the propagation of sound through the atmosphere, it is now possible to merge these models with sensor models to determine the impact of the atmosphere on acoustic sensor performance. Early numerical models such as the Fast-Field Program (FFP) allowed users to predict the effects of refraction, geometric spreading, diffraction, and complex ground

impedance on sound propagation in the atmosphere. Coupling these numerical models with models for acoustic arrays, gave researchers the ability to predict the effects of the mean atmosphere on array performance. The newer numerical models like the Parabolic Equation (PE) and Green's Function PE (GFPE) have allowed researchers to incorporate homogeneous and isotropic turbulence and complex terrain into propagation models. These new capabilities will lead to determining the effects of turbulence and terrain on acoustic arrays. With the development of high-speed/low-power/low-cost microcomputers, it is now possible to even integrate simple propagation models into an acoustic array processor. This will allow the development of smart acoustic arrays which can determine the impact of the atmosphere on their performance and possibly modify how the array processes the data.

10:50

4aSP7. Modeling the propagation of sound through a turbulent atmosphere. Ph. Blanc-Benon and D. Juvé (Ecole Centrale de Lyon, LMFA UMR CNRS 5509, BP 163, 69131 Ecully Cedex, France, acous@ec-lyon.fr)

Recent advances in the numerical techniques and physical models used in sound propagation codes for the atmosphere have greatly improved the correspondence between calculated and measured fields. In particular, incorporating models of atmospheric turbulence have provided major advances. In this paper, an approach to turbulence modeling is described which the authors have developed in the last few years. The turbulence is represented as a superposition of a limited number of random Fourier modes and the acoustic waves are propagated through each individual realization of this field using a wide angle parabolic equation. Statistical results are obtained through ensemble averaging. Illustrations will be given for both line of sight propagation and for propagation into a refractive shadow zone. The influence of the physical model (vectorial versus scalar and Gaussian versus inertial) on the predicted sound field will be discussed. Extensions to time-domain analyses will also be considered.

Contributed Papers

11:15

4aSP8. Underwater acoustic communication over doubly spread channels. Trym H. Eggen (Dept. of Elec. Eng. and Comput. Sci., MIT, 77 Massachusetts Ave., Cambridge, MA 02139) and James C. Preisig (Woods Hole Oceanograph. Inst., Woods Hole, MA 02543-1053)

Acoustic communication in the ocean with carrier frequency in the range 5–50 kHz has been performed in both deep and shallow water, and the communication channels can be sorted according to their time and frequency dispersion. Some of the underwater communication channels are significantly spread in both frequency and time; thus they are doubly spread channels. Scattering function estimates computed from real data of different doubly spread channels are presented, and physical scenarios that would produce the different channels are identified by simulations using ray trace and time variant FIR filters. The frequency dispersion depends in particular on relative speed between the transmitter and channel scatterers, or between transmitter and receiver for direct paths, and the eigenrays of some of the physical scenarios can have significantly different Doppler spreads. Current receivers are unable to establish reliable communication over such channels. A new receiver is proposed which can successfully demodulate signals which have propagated through some of the doubly spread channels with multiple Doppler spreads. The receiver makes use of filter taps spaced in both the delay and the Doppler dimensions. Its capabilities and limitations are demonstrated on both real and simulated data.

11:30

4aSP9. Environmentally adaptive signal processing. Leon H. Sibul and Teresa L. Dixon (Appl. Res. Lab., Penn State Univ., State College, PA 16804-0030)

To achieve required system performance under difficult propagation and reverberation conditions, signal processing systems must be environmentally adaptive. This paper uses results from statistical signal processing and wavelet transform theory to develop an estimator-correlator (EC) approach to environmental adaptivity. The EC is an optimum (in the maximum likelihood sense) detector and estimator that incorporates propagation and interference (reverberation and ambient noise) models as an essential part of its formulation. It is a doubly adaptive processor because it uses propagation models and on-line adaptation to determine conditional mean estimates of the full-field replicas that it correlates with optimally/adaptively filtered received data. The EC concept is applicable to both active and passive sonar. This paper concentrates on wavelet transform domain implementation of active EC that achieves recombination of multipath and multihighlight signals. This paper reviews the underlying wavelet transform concepts, basic EC theory and medium characterization by wideband spreading and scattering functions. Some computer simulation results are presented. [This work was supported by ONR under Grant No. N000149510965, Code ONR 313.]

Session 4aUW**Underwater Acoustics: Synthetic Aperture Sonar**

Antares Parvelescu, Cochair

Naval Research Laboratory, Code 7130, Washington, DC 20375

Frank Henyey, Cochair

*Applied Physics Laboratory, University of Washington, 1013 N.E. 40th Street, Seattle, Washington 98105***Invited Papers****8:00****4aUW1. The prehistory of synthetic aperture sonar.** Antares Parvelescu (Naval Res. Lab., Code 7130, Washington, DC 20375)

It is necessary to go back to 1958 (A.D.) and the summer of 1959. The concept of using recorded data obtained from a single moving element to provide a "memorized array" instead of the multiple-element physical array, was offered to ONR (Aubrey Pryce), whose reaction will be described. The years 1960–68 saw slow (mini-funded) progress at Hudson Labs of Columbia University. Two algorithms were conceived and (mini-)tested; a 6-m towed "fish" was built, then progress slowed further at the Hawaii Inst. of Geophysics. Graduate students wrote theses and bootlegged computer time; names will be named and credit will be given. The unquenched urge to "see" acoustically the seabed in two or three dimensions is at least almost here, as the next few papers show. Some formerly classified reports may have been released by the time of this meeting.

8:20**4aUW2. Efficient SAS image-reconstruction algorithms.** David Hawkins and Peter Gough (Dept. of Elec. and Electron. Eng., Univ. of Canterbury, Private Bag 4800, Christchurch, New Zealand)

Now that several prototype strip-map synthetic aperture sonar (SAS) systems are in operation and producing images (albeit on an irregular basis), much of the focus of current research is on the efficiency of the various strip-map image reconstruction algorithms. Although sometimes still used by several sonar groups, the old "delay-and-sum" beamforming, or exact algorithm, is clearly inadequate for high throughput imaging. Newer algorithms such as the range/Doppler, wave number, and chirp-scaling algorithms are considered. Using a simulated target field of point reflectors, the results from all of these algorithms are compared and the relative computational loads given. As well, the more recently developed (em accelerated) chirp-scaling algorithm is explained.

8:40**4aUW3. SAMI: A low-frequency wideband prototype for synthetic aperture mapping and imaging.** Manell E. Zakharia and Jacques Chatillon (CPE Lyon, LISA/LASSO, BP 2077, Bat 308, 43 Blvd. du 11 Novembre 1918, 69616, Villeurbanne Cedex, France)

A low-frequency wideband (6–12 kHz) prototype has been developed in the framework of the SAMI (synthetic mapping and imaging) project. The prototype uses a subsurface towed fish. One transmitting array and two parallel receiving arrays 2 m long were installed on the tow-fish. Each array is associated to a synthetic aperture sonar channel. Thanks to the use of suitable transmitted signals (linear period modulated chirps), array synthesis could simply be achieved in the time domain; pulse compression and azimuthal matched filtering could be separated in order to reduce the processing complexity. Synthetic image resolution was measured at sea using reflectors on a frame; the results obtained were comparable to the predicted performance: 0.5×0.15 m. The cross correlation between both synthetic aperture channels output was used for measuring the elevation of the sea bottom reverberation cells. Synthetic aperture maps were then produced. Real-time processing was achieved on board the vessel for a swath of 500 m. Data were stored and post-processing was achieved off-line. Several experimental data obtained at sea in both shallow and deep (2400 m) water are presented. [Work supported by the EC-MAST (Marine Science and Technology) program.]

9:00**4aUW4. Test results for a synthetic aperture sonar (SAS).** Mark L. Neudorfer (Hughes Aircraft Co., Naval and Maritime Systems, 6500 Harbour Heights Pkwy., Mukilteo, WA 98275-4844, mark_neudorfer@atk.com), Dennis Garrood, Tony Luk (Hughes Aircraft Co., Mukilteo, WA 98275), and Matt Nelson (Dynamics Technology, Inc., Torrance, CA)

A special purpose synthetic aperture sonar (SAS) system has been developed and built by Hughes Aircraft Co. and Dynamics Technology, Inc. The main features of the system are a 3.2-m-long, 32-element linear array, and a data acquisition system capable of recording the elemental data in real time while preserving the phase and amplitude fidelity required by SAS processing. High-resolution, large (100^2 m²) and small (3^2 m²) SAS images of various objects will be presented that demonstrate resolutions of 20 cm at 1000 m. The innovation of the work is that resolution achievable only with short-range, high-frequency sonar has been demonstrated at ranges obtainable only with a low-frequency sonar. [Work supported by DARPA.]

9:20

4aUW5. Medium effects on high-frequency propagation: A recent experiment. Kevin Williams, Frank Henyey, Terry Ewart, Jim Grochocinski, Steve Reynolds, and Daniel Rouseff (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

In August 1996 an experiment was performed southeast of Long Island, NY (40 30.16 N, 70 30.58 W) to examine water column effects on synthetic aperture sonar (SAS). Both the experiment and the pre-experimental modeling that guided the experiment will be presented. Salient features of the experimental equipment and procedure include the use of bottom mounted towers for stability, transmission frequencies of 6, 20, 75, and 129 kHz, reciprocal tower-to-tower transmissions, and receivers spaced over a 10-m horizontal and 6-m vertical aperture. First look results will be presented for a 24-h period that included a solibore event. Large variations in propagation conditions over time scales of minutes are evident in the data. The SAS relevant quantities to be extracted from this data set will be discussed. [Work supported by ONR.]

9:40

4aUW6. An approximation theoretic approach to synthetic aperture processing. Raymond Fitzgerald (Vail Res. and Technol. Corp., 5285 Shawnee Rd., St. 210, Alexandria, VA 22312)

An algorithm for synthetic aperture processing derived using approximation theory is presented. Improvements offered by this approximation theoretic algorithm and the analytic relationships between the algorithm and other algorithms are described. The approximation theoretic approach is shown to provide a unified formalism for comparing and evaluating existing and proposed algorithms as well as a method for generalizing and improving them.

10:00

4aUW7. Mid-frequency synthetic aperture sonar imaging. Kenneth D. Rolt, Geoffrey S. Edelson, Bradford W. Gillespie (Sanders, A Lockheed Martin Co., P.O. Box 868, Nashua, NH 03061-0868, kdrolt@sanders.com), Stephen A. Celuzza, Philip A. Abbot, and Charles J. Gedney (Ocean Acoust. Services and Instrumentation Systems, Inc., Lexington, MA 02173)

A detailed review of many temporal stability and spatial coherence experiments is used to develop a reasonable model for the temporal and spatial limitations on mid-frequency synthetic aperture sonar (SAS) in shallow water. A comparison of deep- versus shallow-water coherence experiments is made based on an extensive review of the literature. This comparison reveals that the spatial limits for arrays often dominate over temporal limits. The upper bounds on real or synthesized array lengths in shallow water that result from this analysis are presented. The feasibility of mid-frequency SAS imaging in shallow water is shown, and the predicted resolution in range, azimuth, and depth is compared to the theoretical values for a perfectly stable ocean and precise motion compensation. As a first step to validate the spatial and temporal limitation analysis, a mid-frequency surface sonar is modeled to create artificial data for SAS imaging in shallow water. The modeled environment includes a time varying ocean surface and sound-speed profiles, bottom roughness, penetration, and scattering. Useful SAS imaging is shown. [Work supported by NUWC.]

10:20–10:35 Break

Contributed Papers

10:35

4aUW8. Design procedures for synthetic aperture sonar. Theo Kooij (Defense Adv. Res. Projects Agency (DARPA), 3701 N. Fairfax Dr., Arlington, VA 22203-1714, tkooij@darpa.mil)

An engineering approach is presented for the top down design of synthetic aperture sonar (SAS) systems. The principal system parameters and the equations governing their relations will be identified. The difference between the variables and the equations determines the number of free variables, which can be selected either arbitrarily or according to the system requirements. Some simple rules will be stated, and by example it will be shown how to apply the design steps.

10:50

4aUW9. Multistatic scattering from bottom targets in the presence of anisotropically rough interfaces. Jaiyong Lee and Henrik Schmidt (Dept. of Ocean Eng., MIT, Cambridge, MA 02139, jai@arctic.mit.edu)

The performance of active, multistatic sonar systems can be enhanced significantly, if the 3-dimensional environmental physics is incorporated into the processing. Thus fundamental understanding of the difference between target scattering and bottom reverberation could provide the basis for advanced, multistatic clutter reduction. Strong bottom interaction is the most distinct acoustic feature in shallow-water environments. Consequently, bottom roughness becomes the major reverberation mechanism limiting sonar system performance. To investigate the three-dimensional

spatial characteristics of multistatic target scattering and bottom reverberation, a unified numerical model is being developed. It combines existing theoretical models for scattering from shells of finite length, anisotropic interface scattering and seismo-acoustic wave propagation in stratified media. The model generates Monte Carlo statistics for multistatic target scattering in the presence of a strongly reverberant, rough elastic bottom. The effects of anisotropic statistics of the bottom roughness on the multistatic field are investigated, and the significance of the geometric and acoustic properties of target is addressed. [Work supported by ONR.]

11:05

4aUW10. Passive synthetic aperture sonars, revisited. Edmund J. Sullivan (Naval Undersea Warfare Ctr., Newport, RI 02841, sullivan@tech.npt.nuwc.navy.mil) and William M. Carey (MIT, Cambridge, MA 02139)

Experimental and analytical studies have clearly shown the feasibility and applicability of passive and active synthetic aperture sonars (SAS). These investigations have spanned the infrasonic to ultrasonic frequency range. However, passive SAS has not gained widespread use due to three arguments. The first is based on a single hydrophone system, and knowledge of the source frequency is required for bearing estimation. However, with multiple hydrophones, both the bearing and frequency can be estimated simultaneously. The second argument claimed that there was no "new gain," since the number of spatial and temporal degrees of freedom are constant. However, this analysis failed to account explicitly for the

4a THU. AM

movement of the array, thereby ignoring the Doppler information. The third argument claimed that temporal coherence time must be at least equal to the time of synthesization. This argument ignores the applicability of sequential coherence corrections. This paper will clarify and place these arguments in perspective. Experimental results and analyses showing the performance improvements will be presented.

11:20

4aUW11. Modal mapping in a complex shallow water environment.

George V. Frisk, Laurence N. Connor, Kyle M. Becker (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanograph. Inst., Woods Hole, MA 02543), and Barry J. Doust (Penn State Univ., State College, PA 16804)

An experimental method is described for mapping the wave-number spectrum of the normal-mode field as a function of position in a complex, shallow-water waveguide environment whose acoustic properties vary in three spatial dimensions. The experimental configuration consists of a fixed source radiating one or more pure tones to a field of freely drifting buoys, each containing a hydrophone, GPS and acoustic navigation, and radio telemetry. The precisely navigated, drifting hydrophones form a two-dimensional, synthetic aperture planar array that can be used to determine the evolution of the normal modes in range and azimuth. By describing the spatially varying spectral content of the modal field, the method provides a direct measure of the propagation characteristics of the waveguide. The resulting modal maps can also be used as input data to inversion techniques for obtaining the acoustic properties of the waveguide. [Work supported by ONR and ARL Penn State Univ.]

11:35

4aUW12. Three-dimensional modal mapping in shallow water using simulated acoustic field data.

Kyle M. Becker, George V. Frisk, Laurence N. Connor (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanograph. Inst., Woods Hole, MA 02543), and Barry J. Doust (Penn State Univ., State College, PA 16804)

It is well known that the acoustic field propagated through a shallow-water waveguide is sensitive to the properties of the waveguide and its boundaries. Through adiabatic mode theory, it is understood that the modal eigenfunctions and eigenvalues adapt to the local environment and changing boundary conditions. It has been proposed that modal tracking may be employed in a laterally varying waveguide to infer properties of the normal-mode field and the environment. In support of a three-dimensional modal mapping experiment, synthetic three-dimensional acoustic fields are generated to simulate propagation in a shallow-water region in the vicinity of the East Coast *strataform* site. Synthetic fields are computed using normal-mode and parabolic equation codes and transformed to the wave-number domain to obtain estimates of the spatially varying spectral content of the modal field. The resulting modal maps are discussed for several experimental regions. [Work supported by ONR and ARL Penn State Univ.]

11:50

4aUW13. Reduced wavenumber synthetic aperture. Ira Ekhaus (Synap Corp., 100 Wildwood Ave., Arlington, MA 02174)

Underwater acoustic imaging applications such as mine counter measures and sidescan seafloor surveys would benefit from an operational synthetic aperture sonar system. While the potential of high-resolution synthetic aperture (SA) sonar designs to provide unprecedented detection and discrimination performance is clear, the real-world issues of residual platform motion and temporal/spatial variations in sound speed (SSP) have prevented operational deployment of existing designs. High azimuthal resolution requires a synthesized aperture of lengths approaching ten thousand wavelengths. This imposes a tight constraint on residual motion error and the ranges of environmental scenarios where present designs can operate. Synap Corporation has developed the reduced wavenumber synthetic aperture (REWSA) algorithm/design based on an interferometric preprocessing stage operating on multiple element baseband receptions, and a unique inversion stage. The REWSA algorithm is shown to have a reduced wavenumber characteristic that is inherently robust to the effects of the sonar environment. Discussion of resolution, spatial sampling, and aperture requirements will be supported by simulation and experimental data results. [This work was sponsored by NOAA under Contract 50-DKNA-5-00206.]

12:05

4aUW14. Space-time signal processing for complicatedly moving antennas.

Igor I. Gorban (Inst. of Mathematical Machines and Systems, Natl. Acad. of Sci., 42, Ave. Acad. Glushkov, Kiev 252187, Ukraine)

The results of the author's last work in the field of space-time signal processing (STSP) were presented at the 132nd Meeting of the Acoustical Society of America [I. I. Gorban, "Space-time signal processing algorithms for moving antennae," J. Acoust. Soc. Am. **100**, 2638(A) (1996)]. It touched the two problems of STSP in complicated dynamic conditions, where there was antenna motion with variable velocity and angle rotations, and modifying of the form of the antenna. The first problem is STSP optimization in conjunction with the complicated antenna motion, the noises, and the medium together. The second is the problem of super-multichannel STSP with low calculating efforts. New approaches were proposed for solving these problems. It was found that complicated antenna motion, often regarded as a destabilizing factor, may play a greatly positive role for increasing space-time selectivity and noise immunity of the systems. Rather than struggle with this destabilized factor it was proposed that one should use it. Ways to essentially reduce calculations for complicated STSP in dynamic conditions were also found. At the 133rd Meeting, the new results of the research developing the proposed approaches, for towed arrays in particular, will be presented.

Meeting of Accredited Standards Committee S12 on Noise

to be held jointly with the

U. S. Technical Advisory Group for ISO/TC 43/SC1 Noise and ISO/TC 94/SC12 Hearing Protection

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 4005, Champaign, Illinois 61820

B. M. Brooks, Vice Chair, S12
Brooks Acoustics Group, 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, Chairman, 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. The interaction with ISO/TC 43/SC1 and ISO/TC 94/SC12 activities will also be discussed, with reference to the international standards under preparation. A report will be given on the activities of ISO/TC 43/SC1. The Chairs of the respective U.S. Technical Advisory Groups (P. D. Schomer and E. H. Berger) will report on current activities of these international Technical Subcommittees under ISO.

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.

THURSDAY AFTERNOON, 19 JUNE 1997

ROOM E, 3:30 TO 5:00 P.M.

Session 4pAA

Architectural Acoustics: Distinguished Lecture

William J. Cavanaugh, Chair

Cavanaugh Tocci Associates, Inc., 327F Boston Post Road, Sudbury, Massachusetts 01776

Chair's Introduction—3:30

Distinguished Lecture

3:45

4pAA1. Dynamic multiple-use concert hall-theater design and multi-form theater design for the twenty-first century or From dampened steel and reinforced concrete to mosquito netting and back again to dampened steel and reinforced concrete. George C. Izenour (George C. Izenour Assoc., Inc., 16 Flying Point Rd., Stony Creek, CT 06405)

Except for the largest and wealthiest urban centers, it is an undisputed fact that the *raison d'être* for design of multiple-use facilities for the performing arts in municipalities of smaller size and relative wealth is an economic necessity for capital as well as for operational funding. Solutions to this many-faceted, much-disputed and much-maligned problem *vis-a-vis* owners, architects, and engineering consultants have been and continues to be many and varied. Overall in the USA, two schools of thought have emerged. One is continuation of the traditional static architectural approach inherited from the 18th-century high baroque. The other originated and developed by this consultant and favoring a dynamic engineering approach is, instead, firmly rooted in modern times. This paper, accompanied by concise graphic exposition of certain seminal historical and related contemporary realized recent designs and two proposed unrealized designs is an extrapolation into the next century of the author's technologically based, as opposed to the conventional architecturally based, approach to the problem.

THIS SESSION WILL BE FOLLOWED BY A WINE AND CHEESE RECEPTION
IN SENATE ROOM I. SPONSORED BY THE NATIONAL COUNCIL
OF ACOUSTICAL CONSULTANTS

THURSDAY AFTERNOON, 19 JUNE 1997

ROOM P, 2:00 TO 4:40 P.M.

Session 4pAB

Animal Bioacoustics: Low Frequency Bioacoustics

Adam S. Frankel, Chair

Bioacoustics, Cornell Laboratory of Ornithology, 159 Sapsucker Woods Road, Ithaca, New York 14850

Invited Papers

2:00

4pAB1. A survey of research on low-frequency acoustic communication in elephants. Katharine Payne (Bioacoust. Res. Program, Cornell Lab. of Ornithology, Cornell Univ., Ithaca, NY 14850)

Research projects by K. Payne, W. Langbauer, E. Thomas, and J. Poole have shown that Asian and African elephants make powerful infrasonic calls, some of which are used in long-distance communication. Although full classification has not been achieved, some of these calls are social, others reproductive; behavioral responses show individual recognition. Playback experiments yielded a measure of the distances over which conspecifics respond during daylight hours (Langbauer *et al.*) and of the social information imbedded in certain clearly definable calls (K. McCoun). Long-distance communication appears to be implicated in the coordinated movements between separated elephant families; a field study using radio collars with implanted voice-activated microphones yielded suggestive results (Langbauer *et al.*). A meteorological study by meteorologists M. Garstang and D. Larom led them to predict

large-scale diurnal fluctuations in low-frequency sound propagation in dry season savannas. They and Payne plan to collaborate in research on the relation between these phenomena and calling behaviors of savanna animals. [Support for the presenter's projects came from World Wildlife, National Geographic, NSF, and a Guggenheim fellowship.]

2:25

4pAB2. Low-frequency signaling behavior in mysticete whales. Christopher W. Clark (Cornell Univ., Lab. of Ornithology, Ithaca, NY 14850) and William T. Ellison (Marine Acoust., Inc.)

Mysticete whales produce a variety of low-frequency sounds. These range from the highly labile song of the humpback whale (*Megaptera novaeangliae*) to the infrasonic sounds of the blue (*Balaenoptera musculus*) and fin (*Balaenoptera physalus*) whales. Biological functions of these sounds remain largely untested, the most common being that sounds are a male reproductive display for social communication. Recent access to fixed arrays of bottom-mounted hydrophones and towed arrays has provided a novel mechanism for documenting low-frequency whale sounds throughout large ocean areas in relation to season, bathymetric features, and ocean conditions. Ocean sound propagation conditions impose severe constraints and offer dramatic opportunities for acoustic transmission for whales. The bioacoustic data in combination with sound propagation models provide novel insights into the possible biological functions for low-frequency whale sounds, and offer testable predictions. Improved understanding of low-frequency whale bioacoustics is critical for knowing the potential impact of human-made low-frequency sounds on whales, and offers insights into improved signal design. Biological functions will be presented in conjunction with empirical evidence to support, for example, that blue and fin whales navigate using reverberation and bistatics, and humpback males should prefer deep rather than shallow water singing habitat.

Contributed Papers

2:50

4pAB3. Comparison of low-frequency communication by footdrumming in three species of solitary, desert rodent, kangaroo rats (*Dipodomys*). Jan A. Randall (Dept. of Biol., San Francisco State Univ., San Francisco, CA 94132, jrandall@sfsu.edu)

The ears of desert rodents are well adapted for receiving low-frequency vibrations created during footdrumming. Kangaroo rats drum species-specific patterns, ranging from single thumps to individual footdrumming signatures. The desert kangaroo rat, *D. deserti*, drums single thumps spaced 0.25 to 0.30 s apart. The giant kangaroo rat, *D. ingens*, drums long footrolls that can average over 100 drums at 18 drum/s. The banner-tailed kangaroo rat, *D. spectabilis*, drums footdrumming signatures consisting of 3–38 footdrums in a footroll combined into sequences of 2–13 footrolls. In playback tests, all three species stood alert and entered the burrow in response to footdrumming of their own and other species. The rats also responded in species-specific ways. *D. spectabilis* drummed to its own species' footdrumming, but not to playbacks of single drums of *D. deserti*. *D. deserti* did not footdrum in response to the playbacks, but approached the speaker more frequently than either of the other two species. *D. ingens* footdrummed equally to all footdrumming playbacks. The species' differences reflect differences in social tolerance and spacing. *D. deserti* chases visitors from the burrow, *D. spectabilis* engages in frequent footdrumming exchanges, and *D. ingens* tolerates close neighbors and footdrums periodically.

3:05

4pAB4. Comparison of the low-frequency response of the peripheral auditory organs in the goldfish and oscar. Corrie Derenburger, James J. Finneran, and Mardi C. Hastings (Dept. of Mech. Eng., Ohio State Univ., 206 W. 18th Ave., Columbus, OH 43210)

A noninvasive ultrasonic measurement system was used to measure the amplitude and phase response of the peripheral auditory organs of the goldfish, considered auditory specialists, and the oscar, having an auditory threshold noticeably higher than that of the goldfish. The fish were excited with an underwater acoustic source at frequencies ranging from 12.5–3000 Hz. For frequencies between 12.5–400 Hz, tests were performed in a 14-m acoustic waveguide. For tests above 400 Hz, a 2.43-m-diam pool with a depth of 0.4 m was used. The response of the goldfish swimbladders, Weberian ossicles, and otoliths were measured. In addition, the swimbladders were deflated and measurements were made of the Weberian ossicles and the otoliths. Similar measurements were obtained for the auditory organs of the oscars, which vary from the goldfish in that they contain only one swimbladder and lack the Weberian apparatus. Comparisons within

and between species are presented based on these four groups of data. [Work supported by ONR Grant No. N00014-94-1-0337.]

3:20–3:40 Break

3:40

4pAB5. Dynamic mechanical measurements of marine mammal tissues. Edwin R. Fitzgerald, John D. Strandberg (Johns Hopkins Univ., 3400 N. Charles St., Baltimore, MD 21218), and Brent R. Whitaker (Natl. Aquarium, Baltimore, MD 21202)

An automated measurement system for complex shear compliance, $J^* = J' - iJ''$, shear modulus, $G^* = G' + iG''$, and shear wave velocity and attenuation from 2 to 10 000 Hz at temperatures from -25 to 150°C [E. R. Fitzgerald, Proc. Am. Chem. Soc. **60**, 573–578 (1989); U. S. Patent 5,081,870 (1992)] has been used to determine these viscoelastic parameters for blubber and skin of several marine mammals. Excised samples from two stranded harbor seals, for which rescue efforts failed, were measured at 20°C from 2 to 1000 Hz, at short intervals, from 1.9 to 12 h after death. Values of the elastic (J') and viscous (J'') compliance components decreased with time after death, and showed sharp, "life-to-death" slope transitions at 4 to 6 h after death, as found in prior animal tissue measurements [Fitzgerald *et al.*, J. Acoust. Soc. Am. **29**, 61–64 (1957); E. R. Fitzgerald, Biorheology **12**, 397–408 (1975)]. These measurements are compared with those previously reported for pilot whale and dolphin blubber [E. R. Fitzgerald and J. W. Fitzgerald, Mater. Sci. Eng. C **2**, 209–214 (1995); Proc. Third Int. Conf. Intel. Mater. SPIE **2779**, 83–88 (1996)]. [This work was supported, in part, by the Advanced Research Projects Agency and the Office of Naval Research.]

3:55

4pAB6. Canine acoustics. II. Frequencies, transmission, and annoyance. Peter M. Scheifele (Natl. Undersea Res. Ctr., Univ. of Connecticut—Avery Point, Groton, CT 06340-6097, scheifel@uconnvm.uconn.edu), David G. Browning (Univ. of Rhode Island, Kingston, RI 02881), and Lesa M. Scheifele (The Lost Ark, Inc., Norwich, CT 06360)

In a previous paper [P. M. Scheifele and D. G. Browning, J. Acoust. Soc. Am. **100**, 2710(A) (1996)] it was found that the source level for a howling Eastern Coyote is a relatively high 93 dB. At a primary frequency of 500 Hz this howl should have a strong signal throughout a typical territory (4×4 km) under normal atmospheric transmission conditions. Based on reported [H. E. Heffner, Behav. Neurosci. **97**(2), 310–318

(1983)] canine hearing response, however, howls appear louder to humans than to canines. For low-frequency noise, such as from road traffic, the opposite is true. This brings up some interesting questions of mutual annoyance which will be discussed.

4:10

4pAB7. Potential impacts of low-frequency anthropogenic noise on the hearing of subarctic beluga whales in the Saint Lawrence estuary. Peter M. Scheifele (Marine Sci. and Technol. Ctr., Univ. of Connecticut, 1084 Shennecossett Rd., Groton, CT 06340-6097, scheifele@uconnvm.uconn.edu)

A small population of about 500 beluga whales, *Delphinapterus leucas*, resides in the Saint Lawrence estuary, isolated from other populations of its species. This population has been extensively studied over the last 12 years in terms of its seasonal distribution, size, age, group structure, pathology, and group behavior. It is now considered to be threatened by human activities including elevated noise levels due to a wide range of anthropogenic sources including merchant shipping and whale watching activities. A whale recovery plan is in development that will include hearing conservation. Ambient noise measurements were made at the three major sites of habitation of these whales to gain an understanding of the daily fluctuations in noise levels and to establish a baseline characterization of the acoustic environment at each site. The noise levels were related statistically by site, time of day, and human activity and compared to beluga hearing sensitivity curves [W. Au, *The Sonar of Dolphins* (1993)] and indicate the probability that hearing damage will occur for animals

occupying two of the three sites. Specifically, 200 Hz, 500 Hz, and 1 kHz and 40 kHz were scrutinized. [Work supported by World Wildlife Fund Canada, Humane Society Canada, and Parks Canada.]

4:25

4pAB8. Investigation of a smart trigger for an acoustic deterrent device. Dehua Huang, Steve G. Christensen, Stephen G. Boucher, and Roger H. Tancrill (Airmar Technol. Corp., Meadowbrook Dr., Milford, NH 03055)

Attacks of pinnipeds on fish pens cause considerable destruction to the fish farming industry. An acoustic deterrent device (ADD) operated at 10–16 kHz with a source level around 200 dB *re*: 1 μ Pa at 1 m is an effective way of predation deterrence. The present technique is to operate the ADD system on a continuous basis. This constant insonification could lead to predator habituation and could create serious man-made acoustic noise. In this paper, techniques for creating a device to trigger the existing ADD system are investigated. The trigger will turn on the ADD system only when the fish pens are threatened. Field tests were conducted by applying both passive hydrophone and active Doppler methods to detect the presence of marauding pinnipeds. During the test, acoustic devices were located inside the fish pen or outside the fish cage, and the observation objects were fish in the pen or predators in the open sea. Design concept for the trigger devices, related DSP technique and pattern recognition, along with data collected in the field will be presented and discussed. [The principal investigator, Dr. Huang, acknowledges the funding from USDA/NRAC (Project No. 5555506).]

THURSDAY AFTERNOON, 19 JUNE 1997

ROOM F, 2:00 TO 4:45 P.M.

Session 4pEA

Engineering Acoustics: Modeling and Predictive Capabilities for Transducers II

James M. Powers, Chair

Naval Undersea Warfare Center, Code 2131, New London Detachment, Newport, Rhode Island 02481

Chair's Introduction—2:00

Contributed Papers

2:05

4pEA1. Finite-element analysis of both flexural and Helmholtz resonance modes in a flextensional transducer. Dennis F. Jones (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada)

During the last decade, several types of compact piezoceramic-driven barrel-stave flextensional transducers were researched and developed at the Defence Research Establishment Atlantic (DREA) for low-frequency high-power sonar applications. These transducers were designed to operate at the fundamental flexural resonance where they radiate sound efficiently and omnidirectionally. However, an important limitation common to almost all underwater flextensional transducers is that they require some form of depth compensation for operation at water depths exceeding a few hundred meters. One technique that may eliminate the need for depth compensation (and the associated design complications) for the barrel-stave flextensional transducer is to free flood the cavity inside the cylindrical ring-stack piezoceramic driver with the surrounding acoustic fluid medium and exploit the resulting Helmholtz resonances. To examine the feasibility of such a design approach, the flexural and Helmholtz resonances of a theoretical barrel-stave flextensional transducer with a free-flooded driver have been studied using finite-element techniques, and the

results are presented in this paper. In addition, it is shown that the Helmholtz resonance in an experimental transducer does exist and may be practical.

2:20

4pEA2. A 3-kHz class VII flextensional transducer. Stephen C. Butler and Jan F. Lindberg (Naval Undersea Warfare Ctr., Newport, RI 02841)

A 3-kHz class VII flextensional transducer was developed at NUWC. The class VII flextensional is a modified class IV shape having concave radiating surfaces rather than convex. This transducer is commonly referred to as a "dogbone" flextensional because of its similar shape. This design, originally described by Merchant in U.S. Patent 3,258,738 (28 June 1966), has the advantages that all the surfaces of the shell radiate in-phase and a positive compressive stress is maintained on the ceramic driver as the operation depth increases. The shell can support thinner walls, since hydrostatic pressures do not have to be overcome in order to keep the piezoelectric ceramic driver under compression, as in a conventional class IV transducer. The transducer was designed using two different modeling approaches: finite element modeling (using ATILA) and an equivalent circuit model based on plane-wave and lumped parameters. Based on the modeled results, a transducer was fabricated and then tested in August 1996. The transducer volume is 5 in. (12.7 cm) high, 5 in. (12.7 cm) long,

and 2.5 in. (6.35 cm) wide, the shell material is stainless steel, and the ceramic stack is composed of 30 0.1-in. (2.54-mm)-thick PZT-8 plates, two of which are unpoled and used for insulators. The in-air weight is 10 lbs. The transducer performed exceptionally, matching the modeled data while producing an SL of 202 dB at 10 kV/in., a resonance frequency of 3050 Hz, a mechanical Q of 5, an in-water effective coupling coefficient of 30% (typical of class IV transducers), and an electro-acoustic efficiency of 80%. [Work supported by ONR.]

2:35

4pEA3. Nonlinear time-domain analysis of electrostrictive materials by the finite element method. Jocelyne F. Coutte-Dubois, Jean-Claude Debus, Bertrand Dubus, and Régis Bossut (IEMN, Dept. ISEN, UMR CNRS 9929, 90 rue de Solferino, 59046 Lille Cedex, France)

Lead magnesium niobate ceramics (PMN) are promising materials for applications in the field of underwater acoustic projectors. A finite element procedure has been developed in the ATILA code to model the static deformation of these materials [J. C. Debus *et al.*, J. Acoust. Soc. Am. **100**, 2584(A) (1996)]. Two different elements are available according to whether or not the saturation of polarization is included. An extension of this model to nonlinear transient analysis is presented in this paper. The procedure is derived for time stepping by finite difference, using a central difference scheme. This new capability is demonstrated by analyzing the transient response of an electrostrictive bar submitted to different electrical excitations (voltage step, charge step, continuous sine). The validation of the model is carried out by comparing the results obtained with the analytical results for a lumped constants model. The validity of commonly used transducer characteristics is discussed to describe electrostrictive transducers. [Work supported by ONR contract and DRET/DCN grant.]

2:50

4pEA4. Finite-element study and experimental verification on irregular composite transducer design. Wenkang Qi and Wenwu Cao (Mater. Res. Lab., Penn State Univ., University Park, PA 16802, wqk@sun01.mrl.psu.edu)

In this paper 2-2 composite transducers with both regular and irregular kerfs have been studied by using FEM and experimental methods. Undesirable lateral modes are shown in the electrical impedance curves of periodic 2-2 composites. These lateral modes are well suppressed in a 2-2 composite with irregular kerfs when the ceramic volume percentage is low. However, the beam pattern shows some degree of distortion. In order to restore the beam pattern, mirror symmetry is introduced into a 2-2 composite with irregular kerfs, and the beam pattern of these 2-2 composites has been experimentally measured and calculated using FEM. Excellent agreement has been achieved. As expected, when the system is more disordered, the lateral modes are well suppressed, but the beam pattern is poorly focused. On the other hand, when the system is more ordered, the lateral modes are poorly suppressed, but the beam pattern is more focused. A trade-off is a sequence of 3; in this case the lateral modes can be effectively suppressed, and the beam pattern is also well focused.

3:05

4pEA5. Finite-element and experimental study of ultrasonic beam pattern characterization. Mark Draheim and Wenwu Cao (Penn State Univ., University Park, PA 16802, mrd145@psu.edu)

In this paper a study on the characterization of the acoustic beam pattern for two PZT-5H, 3.5-MHz transducers using both experimental and FEA methods is reported. To obtain actual pressure-field measurements, a three-axis positioning system was constructed. The transducers were submerged in an acoustic test tank and the positioning system was used to maneuver a 200-micron PVDF hydrophone within the pressure field. At each location the output voltage of the hydrophone was measured along with the corresponding location. Using SPYGLASS®, these data were plotted to yield a two-dimensional "slice" of the actual pressure field. A finite-element model of the water-loaded ultrasonic transducer was created using ANSYS®. Material properties that were needed to model the piezo-

ceramic and polymers properly in the transducer were measured via pulse-echo technique. To validate the FEM results the electrical impedance of the transducers was calculated and compared to the actual measured values. The frequency-dependence of the impedance curves match closely. Using harmonic analysis, the pressure field in the water generated by the transducer was calculated and compared to the measured pressure-field data. Excellent agreement was obtained.

3:20-3:30 Break

3:30

4pEA6. Use of automatic modeler and small array receiver for acoustic source location. Xiaorong Lu, Wolfgang Sachse (Theoret. and Appl. Mech., Cornell Univ., Ithaca, NY 14853, xiaorong@msc.cornell.edu), and Igor Grabec (Univ. of Ljubljana, Slovenia)

A statistical, nonparametric regression analysis utilizing a conditional average and a smoothed empirical probability distribution forms the basis of an automatic modeler that resembles the operation of a neural-like network. Coupled with a small receiver array, the modeler can determine both the distance and direction to a source of acoustic emission. The use of such a small receiver array to locate sources of emission on a thick plate is demonstrated. The characteristic property of such a system is its capability for predicting the locations of various sources in a structure from the emitted waveforms. Using the modeler involves two steps. First, the modeler is trained with the computed Green's functions for the structure. In the examples shown, these correspond to a normal force and the corresponding displacement signal for a broad range of source/receiver separations. The database of these responses forms the basis of the modeler memory. After training, the modeler can be used to locate sources of emission on the plate. The operation of the system is demonstrated with both synthetic as well as measured waveforms corresponding to impact and step-loading forces on a thick plate. [Work supported by AFOSR.]

3:45

4pEA7. Universal equivalent circuit for transducers driven by a plate-shaped actuator. Jean C. Piquette (Naval Undersea Warfare Ctr., Code 21601, Bldg. 1171, Third Fl., 1176 Howell St., Newport, RI 02841-1708)

In recent years interest in the use of electrostrictive materials in transducers has increased greatly, in view of the development of materials such as lead magnesium niobate (PMN) and lead magnesium niobate/lead titanate (PMN/PT). These materials produce approximately an order of magnitude greater strain than that produced by piezoelectric materials when subjected to a comparable electric field level. Unfortunately, the behavior of these materials cannot be linearized by the usual procedure, since they will not retain a significant remanent pole for more than a short time. Hence, it is necessary to develop methods for predicting transducer performance based directly on the electrostrictive properties of these materials. This paper presents an equivalent circuit useful for predicting the first-order behavior of a transducer incorporating an electrostrictive material in the form of a plate-shaped actuator. The methods used are sufficiently general that the resulting circuit is also applicable to piezoelectric and electrostatic transducers. Thus, in this sense, the equivalent circuit presented is "universal," since it is equally applicable to electrostrictive, piezoelectric, and electrostatic transducers which utilize a plate-shaped actuator.

4:00

4pEA8. Sensitivity of a spherically focused transducer to focal length and aperture. Martin Manley (261 Congressional Ln., Rockville, MD 20852)

A method is presented to calculate the impulse response of a spherically focused transducer which generates a signal and then receives the reflection of the signal from a rigid surface. Results of calculations are shown in order to illustrate the effects of varying the focal length and the aperture.

These results suggest a method for empirical determination of effective values of the governing parameters of the transducer impulse response.

4:15

4pEA9. Comb transducer design for optimal guided wave performance. Semyen P. Pelts, Joseph L. Rose, and Michael J. Quarry (Dept. of Eng. Sci. and Mech., Penn State Univ., 114 Hallowell Bldg., University Park, PA 16802, jlr@psuvm.psu.edu)

Benefits of a comb transducer for producing guided waves in a structure include an ability to produce specific guided waves of choice in a low modulus material as well as good flexibility in mode choice and generation. As a result, excellent sensitivity and penetration power can be obtained by optimal mode choice. The comb transducer design for effective excitation of guided waves in an isotropic elastic plate is therefore considered in this study. The design involves optimizing geometrical and mechanical parameters of the transducer such as number of elements, element width, comb sizing, and frequency bandwidth. The mathematical model is necessary to provide framework for this design. The impulse excitation in an elastic plate is studied by applying a double Fourier transform in the wave-number and frequency domains. The method of residues

and numerical integration are used to evaluate the solution of the boundary value problem. The solution demonstrates how various parameters affect the transmission of energy in the elastic plate. Experimental results are obtained on the basis of this model. [Work supported by the Office of Naval Research.]

4:30

4pEA10. Composite transducer with spring's structure. B. L. Jiao and J. D. Zhang (Dept. of Electron., Peking Univ., 100871 Peking, PROC, xywang@radioms.radio.pku.edu.cn)

The objective of this paper is to introduce spring's structure to a piezoelectric beam for transducer application. The stacking sequence of the plies is along the thickness and the elastoelectric conversion is through the bending modes, in which some plies are compressed while the others are stretched. The piezoelectrical equations are given under the approximation for small deflection. As an application to the receivers, the leaf and spiral springs' structures are adapted to the numerical analysis. The acoustic impedance defined by the values in the mean is found much lower than that of the conventional transducers, while the sensitivities are significantly heightened.

THURSDAY AFTERNOON, 19 JUNE 1997

ROOM G, 1:25 TO 3:25 P.M.

Session 4pMU

Musical Acoustics: Music Cognition; General

Judith C. Brown, Chair

Department of Physics, Wellesley College, Wellesley, Massachusetts 02181

Contributed Papers

1:25

4pMU1. Sensitivity to key movement in a Beethoven excerpt: Theory and evidence. Lola L. Cuddy (Dept. of Psych., Queen's Univ., Kingston, Canada) and Nicholas A. Smith (Univ. of Toronto, Toronto, Canada)

After a brief introduction to historical and contemporary descriptions of musical key, key relationships, and key movement, the results of two experiments testing convergence of music theory with listener response will be reported. Perceptual sensitivity to key and key movement in the opening measures of the second movement of Beethoven's piano sonata Opus 53 (Waldstein) was evaluated by the probe-tone technique. Listeners were either familiar (experiment 1, $n=14$) or unfamiliar (experiment 2, $n=14$) with the excerpt. The excerpt was performed by an experienced pianist and recorded as MIDI files. On each trial, the excerpt was presented from the beginning to one of nine successive time points. Time points were the first beat, the first bar, and each successive bar. For each time point there were 12 presentations, each followed by a probe tone, and 1 of the 12 chromatic divisions of the octave randomly selected without replacement. Listeners judged the goodness-of-fit of each probe tone to the music presented; the set of 12 ratings yielded a probe-tone profile. A significant proportion of profile variance was accounted for by quantified music-theoretic and psychoacoustic predictors, the proportion increasing with familiarity. [Work supported by NSERC.]

1:40

4pMU2. Ability to detect changes in musical pieces is a function of musical training and musical context. Edward J. Crawley, Barbara E. Acker, and Richard E. Pastore (Ctr. for Cognit. and Psycholinguistic Sci., SUNY, Binghamton, NY 13902, br00437@binghamton.edu)

The ability of musicians and nonmusicians to detect a one-note change from a fixed three-voiced musical piece (the standard) was evaluated. Following previous studies, which had tested only musicians, several potentially important variables were manipulated. Musical context (polyphonic or homophonic) was manipulated between subjects. Melody location (high, middle, or low voice), and thus the standard, was varied for each subject across sessions. Within a session, changes varied in harmonic relatedness (related or unrelated), and voice location (high, middle, or low). Subjects performed a same-different task and indicated their confidence as to whether or not a change had been present. Furthermore, a signal detection analysis of the subjects' pattern of responses was performed to directly measure sensitivity to the presence of changes. As expected, musicians were more confident in their responses and revealed a higher sensitivity to changes than nonmusicians. Interestingly, the effects of musical training on sensitivity to changes did not interact with any variable except musical context. The results will be discussed in terms of general and music-specific aspects of perception and training. [Work supported by AFOSR.]

4pMU3. Effects of timbre manipulations on melody perception.

Barbara E. Acker and Richard E. Pastore (Ctr. for Cognit. and Psycholinguistic Sci., SUNY, Binghamton, NY 13902, bb31074@binghamton.edu)

The effects of timbre on melody perception were explored in three-voice homophonic (*H*) and polyphonic (*P*) pieces. Musically trained subjects learned a 12-note melody in isolation, then listened to *H* and *P* pieces that contained the melody in either the high, middle, or low spectral positions. Three timbre conditions were created: (1) homogeneous, where all three voices were in the same timbre; (2) enhanced, where the melody was presented in a timbre distinct from the other two voices; and (3) inconsistent, where each of the three timbres formed distinct patterns which were independent of the melody (*H* pieces only). In a discrimination task, subjects responded to errors only in the melody, although errors could occur in the other two voices. Several patterns of results occurred which were dependent on the type of timbre manipulation and the musical context employed. Thus the results demonstrate important aspects of how principles from music cognition (melody memory, timbre manipulations) and music theory (*H* and *P* pieces) can interact. [Work supported by AFOSR.]

2:10

4pMU4. The role of timbre, pitch, and loudness changes in determining perceived metrical structure. Punita Singh (Sound Sense Consultancy Services, 20-A Aurangzeb Rd., New Delhi 110011, India)

The role of differences in pitch, loudness, and timbre as determinants of metrical structures was investigated. A repeated sequence of 12 pure tones was systematically partitioned into 4 groups of 3 tones or 3 groups of 4 tones by introducing changes in *F0*, intensity, and spectral complexity between sounds at serial positions 1, 4, 7, 10 or 1, 5, 9, respectively. Listeners used a rating scale to indicate if a triple, quadruple, or ambiguous meter was perceived. For changes in only one parameter at a time, perception of rhythmic structure followed the physical markers. When changes in more than one parameter were made concurrently, multiple cues for triple or quadruple meter were available. Coincident changes led to reinforcement of the rhythm demarcated by the points of change. Conflicting changes led to different outcomes: Timbre and pitch changes dominated over a loudness-based accent structure. Pitch versus timbre stimuli were rated as having ambiguous meter. A combination of any two parameters versus the third dominated in determining the metrical structure. The results raise questions about the role of timbre in rhythm perception in general and as a tool for marking metrical units in particular.

2:25

4pMU5. Perceptual effects of simplifying musical instrument sound time-frequency representations. James W. Beauchamp (2136 Music Bldg., Univ. of Illinois, 1114 W. Nevada, Urbana, IL 61801), Stephen McAdams, and Suzanna Meneguzzi (Univ. Rene Descartes, F-75006 Paris, France and Inst. de Recherche et de Coordination Acoustique/Musique (IRCAM), F-75004 Paris, France)

The perceptual salience of several outstanding features of musical instrument quasi-harmonic time-variant spectra were investigated. These are: amplitude spectrum shape variation, harmonic amplitude (or frequency) micro-variations, spectral envelope irregularity, harmonic frequency incoherence, and frequency inharmonicity. Seven musical instrument sounds (clarinet, flute, oboe, trumpet, violin, harpsichord, and marimba) were spectrum-analyzed to produce time-varying harmonic amplitude and frequency data. Tones resynthesized from these data were equalized in pitch, loudness, and duration. Six basic data simplifications were applied together with five combinations of them: amplitude versus time smoothing, spectrum shape fixing, spectral envelope smoothing, harmonic frequency coherence, frequency versus time smoothing, and harmonic frequency fixing. Twenty subjects participated in a discrimination study to determine the ease of distinguishing sounds synthesized from simplified data from

sounds resynthesized from the full data. Averaged over the seven instruments, the mean discrimination scores were: spectral envelope smoothing, 96%; spectrum shape fixing, 91%; harmonic frequency fixing, 71%; frequency smoothing, 70%; frequency coherence, 69%; and amplitude smoothing, 66%. In conclusion, spectral envelope smoothing had a profound effect (discrimination >95%) for all instruments except the trumpet. Spectrum shape fixing, which eliminates any spectral centroid variation, had a very large effect except for the clarinet and the oboe.

2:40

4pMU6. Cluster-based probability model for musical instrument identification. Judith C. Brown (Phys. Dept., Wellesley College, Wellesley, MA 02181 and Media Lab., MIT, Cambridge, MA 02139, brown@media.mit.edu)

Mel based cepstral coefficients have been calculated for 30 short (1–2 s) segments of oboe sounds and 30 short saxophone sounds. These were used as features in a pattern analysis to determine for each of these sounds comprising the test set whether it belongs to the oboe or to the sax class. The training set consisted of longer sound segments of 1 min or more for each of the instruments. A *k* means algorithm was used to calculate clusters for the training data, and Gaussian probability density functions were formed from the mean and variance of each of the clusters. Each member of the test set was then analyzed to determine the probability that it belonged to each of the two classes, and a Bayes decision rule was invoked to assign it to one of the classes. Initial results classifying the oboe sounds have been very good, but are less impressive for the saxophone set. Results of a human perception experiment identifying these same sound segments will be reported.

2:55

4pMU7. Use of note partials' energy to improve the identification of polyphonic piano signals. Lucile Rossi and Gerard Girolami (Univ. of Corsica, Quartier Grosseto, 20250 Corte, France)

The identification of the fundamental frequency of musical sounds is a difficult task which has been studied for more than 20 years. Methods using spectral approaches are used to identify polyphonic sounds. A novel approach for identifying monophonic and polyphonic piano sound signals has been developed by the authors [L. Rossi and G. Girolami, J. Acoust. Soc. Am. **100**, 2842(A) (1996)]. The identification is carried out by comparing the incoming spectrum with a database (automatically built previously) containing the frequential distribution of the partials of each piano note. The rate of correct identification is in the 90% range for polyphonic sounds. One kind of error is due to the artificial construction of a note by the combination of partials belonging to other notes; another one results from the difficulty of separating the case of one note played a long time with the case of this same note played two successive times with an overlap. The aim of this paper is to show how the evolution of the partials' energy associated with the search for information about the time of the note attacks can be used to reduce the false identification.

3:10

4pMU8. Vocal accuracy in performance of well-known melodies. Clare L. Henderson (Dept. of Arts and Lang., School of Education, Private Bag 3105, Univ. of Waikato, Hamilton, New Zealand, clhend@waikato.ac.nz) and Lloyd Smith (Univ. of Waikato, Hamilton, New Zealand)

Vocalization is an important aspect of music making, regardless of the culture. Psychologists have carried out considerable research on human melody recognition and perception, which includes analysis of musical cognitive processes. However, questions that relate directly to the ways in which people perform what they recall have been left largely unanswered. In this experiment, a number of people were asked to sing, from memory,

ten well-known songs. Their performances were recorded on audio tape and analyzed and compared with notated versions of the songs. The results showed that pitch maintenance was more accurate when songs moved by steps rather than leaps and when intervals moved up rather than down. Pitch problems were more prevalent in songs using downward leaps or

changes of tonality. Wide leaps, such as fifths or sixths, were often compressed, with subjects singing the top note of an ascending interval flat, or the bottom of a descending interval sharp. The results of this experiment have useful applications for musicologists, music educators, composers, and those involved with data retrieval processes.

THURSDAY AFTERNOON, 19 JUNE 1997

ROOM N, 2:00 TO 3:30 P.M.

Session 4pPA

Physical Acoustics: Ultrasonics

L. C. Krysac, Chair

Department of Physics, Pennsylvania State University, 104 Davey Laboratory, University Park, Pennsylvania 16802

Contributed Papers

2:00

4pPA1. Experimental studies of diffuse decay curvature. John Burkhardt (Dept. of Eng., Purdue Univ., Fort Wayne, IN 46805-1499)

Reverberant, ultrasonic decay experiments were performed on an irregularly shaped aluminum block on which varying sized water drops were placed. The resulting nonproportionally damped systems were found to possess nonexponential decays in agreement with theoretically derived results. These findings support the author's proposition that the features of nonexponential decays are a possible quantitative nondestructive technique for the characterization of localized fatigue damage in metals.

2:15

4pPA2. Measuring the frequency-dependent attenuation in highly lossy material using spread spectrum ultrasound. Grant A. Gordon (Appl. Res. Lab., Penn State Univ., State College, PA 16804)

Broadband signals are commonly used in ultrasonic spectroscopy to measure the frequency-dependent attenuation characteristics of lossy solid media. As compared with narrow-band signals, broadband signals are preferred since they do not require tedious frequency scanning and extensive data reduction efforts. By using spread spectrum techniques to design signals with given bandwidth and temporal characteristics, the overall accuracy and resolution of ultrasonic spectroscopy can be improved. An expression for ultrasonic attenuation is developed and compared with the traditional analysis approach to illustrate the improved accuracy of the new technique. Experimental results obtained using a direct sequence modulation system will also be presented for the attenuation spectra of various lossy solids.

2:30

4pPA3. Temperature-dependent ultrasonic properties of aluminosilicates glass ceramics. J. H. So (Dept. of Phys. and Astronomy, CMSS Prog., Ohio Univ., Athens, OH 45701, jhso@helios.phy.ohiou.edu), D. H. Green, and S. S. Yun (Ohio Univ., Athens, OH 45701)

The longitudinal wave velocity and attenuation and shear wave velocity for machinable aluminosilicate glass-ceramic (MACOR) samples were measured as functions of frequency and temperature with the pulse-echo method. Frequencies of 5–30 MHz were used in the temperature range of 100–300 K. The velocity differences between 100 and 300 K were about 1% for longitudinal waves and 3% for shear waves. The monotonic decrease in the longitudinal attenuation coefficient between 100 and 300 K was approximately 50%. The longitudinal wave attenuation coefficient increases linearly with frequency and nonlinearly with decreasing temperature. This contrasts with the longitudinal and shear wave velocities, both of

which decrease with increasing temperature. These data were used to calculate the elastic moduli, Poisson's ratio, and Lamé parameters for the material. These results are interpreted in terms of thermal stress cracking and microstructure-related absorption.

2:45

4pPA4. Acoustic spectroscopy of deep levels in Cr-doped GaAs using light beam generated SAW. Peter Bury, Ivan Baják, and Karol Grondžák (Dept. of Phys., Žilina Univ., 010 26 Žilina, Slovakia, bury@fel.utc.sk)

A light-beam-generated space-charge inhomogeneity was used for the first time as a probing tool to study the behavior of deep levels in high-resistivity, Cr-doped GaAs using acoustic transient spectroscopy. The space-charge inhomogeneity produced by the illumination of photoresistive piezoelectric semiconductors in a high-frequency electric field can generate, in proper conditions, both the surface and transversal or longitudinal acoustic wave. The temperature dependence of the time development of such a generated surface acoustic wave (SAW) after illumination by pulsed monochromatic light was used to investigate deep levels in high-resistivity GaAs using the technique of acoustic deep-level transient spectroscopy (A-DLTS) [P. Bury and I. Jamnický, Proc. World Congress on Ultrasonics, Berlin, 1995, p. 535]. Deep-level parameters such as activation energy and capture cross-section were determined. [Work supported by a Grant of the Slovak Ministry of Education, No 1/1310/96.]

3:00

4pPA5. Ultrasonic study of vortex viscosity and phase transitions in $\text{YBa}_2\text{Cu}_3\text{O}_7$ single crystals. D. Dasgupta, C. Hucho,^{a)} J. R. Feller, B. K. Sarma, and M. Levy (Dept. of Phys., Univ. of Wisconsin, Milwaukee, WI 53211)

A novel ultrasonic technique (sampled cw method) is used to probe into the viscosity of vortex systems in untwinned and twinned YBCO single crystals. Motion of the flux lines in the superconducting sample is induced via coupling with the moving ionic lattice in the bulk of the sample. Hence, the observed dissipation has its origin in the (induced) motion of the vortices with respect to the ionic background of conduction charge carriers. In the experiments which were performed, pronounced field-dependent attenuation changes were observed which are indicative of transitions from a soft vortex system at low fields to a rigid vortex system at high fields. [Research supported by Office of Naval Research and C. Hucho was supported by Deutsche Forschungsgemeinschaft.] ^{a)}Currently at: FB Physik, Augsburg, Germany.

4pPA6. A novel method for characterizing contact-mode shear wave transducers in immersion. David K. Hsu, Brent A. Fischer (Center for NDE, Iowa State Univ., Ames, IA 50011, dhsu@cnde.iastate.edu), and Vinay Dayal (Iowa State Univ., Ames, IA 50011)

The dominating shear motion of a contact-mode shear wave transducer is often accompanied by a weak parasitic longitudinal mode due to its finite size and constrained boundary conditions. The amplitude and phase of this out-of-plane vibration may be exploited for determining the shear polarization direction and for detecting any asymmetry or distortion of the shear transducer's vibration pattern. In practice, the contact-mode shear

probe is immersed in water and energized by an rf pulser to serve as the transmitter. A C-scan is then made of the face of the shear transducer using a focused longitudinal wave transducer as the receiver. The resulting image shows the out-of-plane longitudinal vibration of the shear transducer. For commercial contact-mode shear probes, this pattern usually consists of two crescent-shaped regions of higher vibration amplitude but of opposite phase. A line bisecting the two crescents and passing through their centers defines the shear polarization direction. Finite-element modeling of the shear displacement of a circular disk with constrained boundary confirmed the experimentally observed behavior. This approach provides a simple method for characterizing commercial shear wave transducers. [Work supported by Center for NDE, Iowa State University.]

THURSDAY AFTERNOON, 19 JUNE 1997

DEANS HALL I, 12:45 TO 3:30 P.M.

Session 4pPPa

Psychological and Physiological Acoustics: Pitch and Loudness

Mary Florentine, Chair

Department of Speech-Language Pathology and Audiology, Northeastern University (133FR), 360 Huntington Avenue, Boston, Massachusetts 02115

Contributed Papers

12:45

4pPPa1. The effect of temporal fringes on the fundamental-frequency discrimination of resolved and unresolved harmonic complexes. Christophe Micheyl and Robert P. Carlyon (MRC-APU, 15 Chaucer Rd., Cambridge CB2 2EF, UK, christophe.micheyl@mrc-apu.cam.ac.uk)

Three experiments investigated the finding that F_0 discrimination for a harmonic complex can be impaired by temporal fringes [R. P. Carlyon, J. Acoust. Soc. Am. **99**, 525–533 (1996)]. The 100-ms target and the 200-ms forward and backward fringes were filtered (1375–1875 Hz) and presented in a pink noise background. Two F_0 's were used: 88 Hz, where the harmonics were unresolved by the peripheral auditory system, and 250 Hz, where the harmonics were resolved. The fringe produced a larger deterioration when it had a similar F_0 to the target than when the F_0 's differed. Experiment 2 showed that adding off-frequency components to the fringe did not affect this pattern of results, thereby ruling out interpretations based on spectral confusion between the target and the fringe. Experiment 3 further examined the dependence of the deterioration on the similarity between target and fringe F_0 's. It showed that FDLs for an 88-Hz target were not reduced when the components of an 88-Hz fringe were set to alternating phase, thereby doubling their pitch without affecting resolvability of the components. The implications for single- and dual-mechanism models of pitch will be discussed.

1:00

4pPPa2. Auditory pitch influenced dramatically by dynamic intensity change. Michael K. McBeath (Dept. of Psych., Kent State Univ., Kent, OH 44242-0001) and John G. Neuhoff (Lafayette College, Easton, PA 18042)

Historically, auditory pitch is considered to be principally a function of acoustic frequency with only a small effect due to absolute intensity. Yet, when tones are Doppler shifted, the pitch dramatically rises and falls with dynamic intensity. This study uses a matching procedure to document the magnitude of pitch drop of dynamic Doppler stimuli. Listeners heard Doppler-shifted tones with a mean frequency of either 1046 or 175 Hz, a total fall of 2 semitones, and an intensity change of 58 to 86 dB and back to 58 dB. They compared this drop to a pair of 75-dB, 0.25-s discrete tones

that dropped in frequency by intervals ranging from 0 to 24 semitones. The average match between experienced sizes of Doppler and discrete pitch change occurred at a discrete drop of 8 semitones, four times larger than the actual Doppler frequency change. This effect opposes and is an order of magnitude larger than the well known effect due to discrete intensity change. It is proposed that the interaction between dynamic pitch and loudness reflects a natural correlation between changes in frequency and intensity that is neurally encoded to facilitate processing of meaningful acoustic patterns.

1:15

4pPPa3. The interaction of pitch and loudness in dynamic stimuli: Beyond the Doppler illusion. John G. Neuhoff (Dept. of Psych., Lafayette College, Easton, PA 18042) and Michael K. McBeath (Kent State Univ., Kent, OH 44242)

Listeners tend to experience rising pitch even though frequency falls as a sound source approaches. The phenomenon, called the Doppler illusion [J. G. Neuhoff and M. K. McBeath, J. Exp. Psychol. Hum. Percept. Perform. **22**, 970–985 (1996)], shows that dynamic intensity change influences perceived pitch in a way that is qualitatively different from discrete static intensity change. Two new studies show a dynamic influence of intensity change on perceived pitch and a dynamic influence of frequency change on perceived loudness. Listeners were presented with square wave tones of either rising, falling, or constant intensity that either rose, fell, or remained constant in frequency for 6 s. Listeners responded in real time to either changes in pitch or loudness by moving a response wheel. It was found that dynamic intensity sweeps contribute to the perception of dynamic pitch change in the direction of the intensity sweep. In addition, dynamic frequency sweeps contribute to the perception of perceived loudness change in the direction of the frequency sweep. The results imply that pitch and loudness perception interact under dynamic conditions in a way that cannot be predicted by perceptual models derived from the presentation of discrete static tones.

4pPPa4. Discrimination thresholds for glides in frequency as a function of transition span, duration, and direction. John P. Madden (Dept. of Commun. Disord., Univ. of North Dakota, Grand Forks, ND 58202, madden@badlands.nodak.edu)

The study measured discrimination thresholds for glides in frequency as a function of magnitude of transition span, duration, and direction. The purpose was to determine whether the Weber fraction is constant across a range of transition spans, as suggested by an earlier study [J. P. Madden and K. M. Fire, *J. Acoust. Soc. Am.* **100**, 3754–3760 (1996)]. Subjects were asked to distinguish between a comparison glide with a fixed transition span and a target glide with a greater transition span. Comparison signal frequency transition spans were 0.5, 1, 2, 4, and 8 times the ERB of the auditory filter at the signal center frequency. The nominal center frequency of the glides was 1.8 kHz, but the actual signals were roved over a range of center frequencies. Both up- and down-glides were used; signal durations were 50 ms and 400 ms. The results indicate that the Weber fraction is essentially constant for transition spans of 1–8 ERBs, but increases below 1 ERB. The effects of duration and direction were small. [Work supported by Research Grant No. 1 R15 DC 02662-01 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.]

4pPPa5. Temporal integration of loudness for tones under partial masking. Mary Florentine (Commun. Res. Lab., Dept. of Speech-Lang. Pathol. and Audiol., 133 FR, Northeastern Univ., Boston, MA 02115), Søren Buus, and Monica Robinson (Northeastern Univ., Boston, MA 02115)

Temporal integration of loudness for 1-kHz tones presented in quiet and under white-noise masking was measured for 5- and 200-ms tones using an adaptive 2I,2AFC procedure. Levels ranged from 5 to 90 dB SL. Results for six listeners with normal hearing show that the amount of temporal integration, defined as the level difference between equally loud 5- and 200-ms stimuli, varies nonmonotonically with level and is greatest at moderate levels. The average amount of temporal integration is about 15 dB near threshold, increases to a peak of 27 dB when the 5-ms tone is about 53 dB SPL, and decreases to about 15 dB near 100 dB SPL. For masker levels of 40, 60, and 80 dB SPL, the amount of temporal integration near masked threshold remains about 15 dB. The maximum amount of temporal integration decreases as masker level increases and occurs at progressively higher levels. At high levels, the amount of temporal integration is nearly the same as in quiet for all masker levels. These results are consistent with the loudness function being steeper near masked threshold than at the same SPLs in quiet. [Work supported by NIH-NIDCD R01DC02241.]

4pPPa6. Loudness summation at low levels and the form of the loudness function near threshold. Søren Buus (Commun. and Digital Signal Processing Ctr., Dept. of Elec. and Comput. Eng., 409 DA, Northeastern Univ., Boston, MA 02115), Mary Florentine, and Hannes Müssch (Northeastern Univ., Boston, MA 02115)

This study uses measurements of loudness summation to examine the form of the loudness function for tones near threshold. An adaptive, two-interval, two-alternative, forced-choice, roving-level procedure was used to obtain loudness-balance measurements between a 1600-Hz tone and tone complexes centered at 1600 Hz. The complexes consisted of equal-SL components with levels between about -3 and 20 dB SL. Frequency separations were one, two, four, or six critical bands for four-tone complexes and one or two critical bands for ten-tone complexes. Results for six young listeners with normal hearing show that the SL difference between the pure tone and the components of the equally loud tone complex near threshold averages about 4 dB for the four-tone complexes and about 7 dB for the ten-tone complexes. When the component level of the four-tone complexes with wide frequency separations is around 20 dB SL, the

equally loud pure tone is near 47 dB SL. These results indicate that loudness may grow approximately as $I^{1.5}$ (where I is tone intensity) near threshold and as $I^{0.2}$ at moderate levels. [Work supported by NIH-NIDCD R01DC02241.]

4pPPa7. Loudness recalibration as a function of recalibration and comparison tone level. Dan Mapes-Riordan and William A. Yost (Parmly Hearing Inst., Loyola Univ. of Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, dmapes@luc.edu)

Presenting loud tones at one frequency and quiet tones at a different frequency makes the quiet tones appear relatively louder. This phenomenon, dubbed loudness recalibration [Marks, *J. Exp. Psychol.* **20**, 382–396 (1994)], was studied using an adaptive, two-track loudness comparison procedure. In this study, a baseline loudness comparison was initially established between two tones. Immediately following this baseline sequence, a sequence of trials were given in which the two comparison tones were preceded by a recalibration tone. The amount of steady-state loudness recalibration was measured as a function of the recalibration tone level and the baseline comparison tone level. The results showed that loudness recalibration is present when the recalibration tone level is much larger than the comparison tone level and that no loudness recalibration is generated when the recalibration tone level is less than or equal to the comparison tone level. In addition, it was found that a recalibration tone did not affect the threshold level of detection. These results support those of Marks [J. Acoust. Soc. Am. **100**, 473–480 (1996)] indicating that loudness recalibration is a centrally based, fatiguelike phenomenon. [Work supported by a Program Project Grant from NIDCD.]

4pPPa8. Magnitude estimation scaling of the loudness component of pure tones, narrow band noise, broadband noise, rock music, and babble speech. Donald Fucci (School of Hearing and Speech Sci., Ohio Univ., Lindley Hall 219, Athens, OH 45701, dfucci1@ohiou.edu), Linda Petrosino (Bowling Green State Univ., Bowling Green, OH 43403), Denise Wyatt, and Corry Wilcox (Ohio Univ., Athens, OH 45701)

This study was designed to extend the range of auditory stimuli used to study magnitude estimation scaling of loudness. The five stimuli chosen were: a 1000-Hz pure tone; narrow band noise (700–1300 Hz bandwidth); broadband noise (100–10 000 Hz bandwidth); rock music [Led Zeppelin, Atlantic Recording Corp., CD Recording No. 19127-2 (1969)]; and babble speech [D. Kalikow, K. Stevens, and L. Elliott, *J. Acoust. Soc. Am.* **61**, 1337–1351 (1977)]. Subjects were 30 normal young adult women. During the auditory magnitude estimation task for each of the five stimuli, a subject was instructed to assign numbers to that stimulus which was presented in a randomly ordered series of nine sensation levels (10, 20, 30, 40, 50, 60, 70, 80, and 90 dB SL). An analysis of variance repeated measured design showed no difference in the numerical responses of the subjects for the five different stimuli. Results suggest the presence of an underlying stabilizing factor (internal scaling mechanism) that allows adult subjects to consistently scale loudness irrespective of the type of auditory stimulus presented [J. Zwislöcki and D. Goodman, *Percept. Psychophys.* **28**, 28–38 (1980)].

4pPPa9. Effect of age on the comfortable loudness range in women. Vishakha W. Rawool (Commun. Disord. and Special Education, Bloomsburg Univ., Bloomsburg, PA 17815) and Joseph P. Pillion (Kennedy Krieger Inst., Baltimore, MD 21205)

This study is designed to evaluate the effect of age on the range of sound pressure levels perceived as “comfortable.” Ten young (22–26 years) and eight older (42–50 years) female subjects participated in the study. All the subjects had auditory sensitivity within 20 dB HL across 0.5 to 4 kHz. Subjects were asked to judge the loudness of warble tones in the following categories: not audible, very soft, soft, comfortable, loud and very loud. The loudness judgments were made at 0.5, 1, 2, and 4 kHz. For any loudness judgement to be considered valid, the same loudness judgement had to be assigned by the subject at least twice to the same sound

pressure level. The comfortable loudness range was calculated for each subject by subtracting the highest sound pressure levels judged as “soft” from the lowest sound pressure levels judged as “loud.” The mean “range of comfortable loudness” values for the older female subjects is lower than that obtained from the younger subjects. Additional data are being collected. Following completion of data collection, the data will be submitted to the mixed multivariate analyses of variance. The results and possible implications of the findings will be discussed. [Work supported by Research and Disciplinary grant from Bloomsburg University.]

3:00

4pPPa10. Puzzles of loudness constancy, loudness adaptation, and speech frequencies. Ernest M. Weiler, David E. Sandman, and Hongwei Dou (ML #379, Hearing Lab., CSD, Univ. of Cincinnati, Cincinnati, OH 45221)

Loudness constancy is suggested as a term for the failure to observe simple adaptation above 30 dB, noting however that Miskiewicz *et al.* [J. Acoust. Soc. Am. **94**, 1281–1286 (1993)] indicated that simple adaptation is found at higher intensities above 4 kHz. Conversely, using the ipsilateral comparison paradigm (ICP), Dange *et al.* [J. Gen. Psychol. **120**, 217–243 (1993)] found adaptation at 1000 Hz, from 45 to 75 dB. They saw this as revealing underlying neural adaptation, normally concealed by processes of loudness constancy. As a further test, Hellman [personal communication (1996)] urged that ICP adaptation be attempted without the usual designated modulus. In the present study ($N=20$), strong ICP loudness adaptation was indeed found with and without the designated modulus. Recently, Janson *et al.* [Br. J. Audiol. **29**, 288–297 (1996)] found that ICP

adaptation at 1 kHz was much less consistent for subjects with higher frequency loss than for nonimpaired listeners. Discussion of the possible value of loudness constancy in the speech frequencies will be discussed in light of current data and previous studies.

3:15

4pPPa11. The theory of mental testing and the correlation between physical noise level and annoyance. Karl Th. Kalveram (Dept. of Cybernetical Psych., Univ. of Duesseldorf, 40225 Duesseldorf, Germany)

Applying the “theory of psychological testing,” physical noise measurement procedures are regarded as “tests” variables and the related annoyance as the criterion variable. The concepts of “reliability,” “validity” and “equivalence” then can be used to assess the precision of different noise measurements. Referring to data of an investigation on aircraft noise in the vicinity of Munich Airport in 1969, in which different physical as well as annoyance data were sampled from the same subjects, it turns out, that the measurements called L_{eq1} , L_s , L_{eq3} , L_{eq4} , NNI and FB1 meet the equivalence criterions of test theory, that is, they cannot statistically be distinguished, and their coefficients of reliability are close to one. Other measurements like D10, H81, $\log N$ and L_{eq10} , equivalent to the former six. However, even the former six are only of moderate validity (about 0.5). Regarding psychological testing, therefore, physical noise is only a poor measure when used to predict individual annoyance, but it suffices when used to predict group averages. Moreover, it can be concluded from the theory that the attempt to enhance validity by modification of the highly reliable physical measuring procedures cannot be successful. It would be more effective to enlarge the reliability of the psychological measurement procedures of annoyance.

THURSDAY AFTERNOON, 19 JUNE 1997

DEANS HALL I, 3:35 TO 5:25 P.M.

Session 4pPPb

Psychological and Physiological Acoustics, Musical Acoustics and Noise: W. Dixon Ward Memorial Session

Christopher W. Turner, Chair

University of Iowa, Wendell Johnson Speech and Hearing Center, Iowa City, Iowa 52242-1012

Chair's Introduction—3:35

Invited Papers

3:40

4pPPb1. W. Dixon Ward—Professor, mentor, university citizen. Arlene Earley Carney (Dept. of Commun. Disord., Univ. of Minnesota, 164 Pillsbury Dr. SE, Minneapolis, MN 55455, carney005@gold.tc.umn.edu)

W. Dixon Ward had a colorful and active presence in the scientific community and in the Acoustical Society of America as an officer, editor, presenter, and author. Many of his ideas and much of his written work sparked debate and discussion. On the homefront at the University of Minnesota, he had an equally colorful presence as a mentor to graduate students and as a frequent and controversial letter writer to the Minnesota Daily, the university newspaper. Selections of Dix's letters on a variety of topics from word usage to university and world affairs will be presented. In this introduction, Dix's professional life outside the Acoustical Society will be reviewed and discussed through the eyes of his students and colleagues.

3:55

4pPPb2. W. Dixon Ward—A personal retrospective. Ira J. Hirsh (Central Inst. for the Deaf, 818 S. Euclid Ave., St. Louis, MO 63110, ira@cidmv1.wustl.edu)

“Dix” Ward's professional career began about 50 years ago, when he was an undergraduate majoring in physics and practicing a life-long interest in music, both as a student and as pianist and singer. Encouraged by one of his college teachers, W. A. Rosenblith, he pursued graduate work at Harvard in the Psychoacoustic Laboratory, under S. S. Stevens. A dissertation on subjective musical pitch and experimental study of the recovery of hearing after acoustic stimulation set him on a double course which continued his musical interests (first with D. W. Martin) and the study of hearing damage following noise exposure (with H. Davis and A. Glorig). Those

interests were to continue as he reached his final and long-lasting appointments at the University of Minnesota. His scientific contributions have been empirical, not theoretical. He focused on finding out what is demonstrably true. His contributions as a scientific citizen were many—in the Acoustical Society, in national and international standards, in the International Society of Audiology, in CHABA, and in related government activities related to noise and noise damage.

4:10

4pPPb3. The idiotone and spontaneous otoacoustic emissions. Robert C. Bilger (Dept. of Speech and Hearing Sci., Univ. of Illinois at Urbana—Champaign, 901 S. Sixth St., Champaign, IL 61820, r-bilger@uiuc.edu)

Although the preponderance of Dix Ward's work was in the area of the effects of noise on hearing, his paper on tonal monaural diplacusis [J. Acoust. Soc. Am. **27**, 365–372 (1955)] has always been my favorite from among all of his papers. In this paper, published several years after he had completed his work at Harvard and was being introduced to noise-induced hearing loss through the Central Institute for the Deaf, he performed a consummate study of his own tonal monaural diplacusis and that of a woman who was studying music at Harvard. Not only did he specify their diplacusis exhaustively, his conclusions from that study foretold of many future developments in psychoacoustics, hearing loss, tinnitus, and signal analysis. The import of this monumental work will be iterated.

4:25

4pPPb4. W. D. Ward as musical acoustician, journal editor, and reviewer. Daniel W. Martin (7349 Clough Pike, Cincinnati, OH 45244)

W. Dixon Ward's musical talent as an expert pianist and a singer, combined with experience and training in both psychology and physics, provided the background for his significant doctoral research on pitch at Harvard University; for subsequent research in musical acoustics at the Baldwin Piano and Organ Company; and for continuing research in musical sound throughout his career. Twice Dix served the *Journal* as an associate editor for psychological acoustics. He peer-reviewed many *Journal* articles in psychoacoustics, musical acoustics, and noise. He reviewed numerous books, and served on acoustical standards committees including acoustical terminology.

4:40

4pPPb5. Possible correction factors for intermittent noise exposure. D. L. Johnson (EG&G MSI, P.O. Box 9100, Albuquerque, NM 87119-9100, dljohnsn@rt66.com)

A standard on Occupational Hearing Loss, ANSI S3.44 (1996), has been recently published. One of its features, unlike ISO R-1999 upon which it is based, allows a trading rule other than one based solely on equal energy. To account for the fact that intermittent noise, as compared to steady noise of the same energy, may not be as harmful to hearing, equal energy with a correction factor is being investigated. ANSI Working Group S3-58 is evaluating at least seven different approaches. In order to evaluate the best approach for correcting for intermittency, W. Dixon Ward's temporary threshold shift (TTS) data, from a variety of intermittent noise exposures [W. D. Ward, *Effects of Noise on Hearing* (Raven, New York, 1976), pp. 407–420], will serve as one evaluation criterion. Actual hearing loss data from exposures to various intermittent noises will serve as the other evaluation criterion.

4:55

4pPPb6. Perception of musical pitch. Edward M. Burns (Univ. of Washington, 1417 NE 42nd St., Seattle, WA 98105, pwa@u.washington.edu)

Although most of Dix Ward's publications were concerned with the effects of noise on hearing, his first love was music perception. The publication derived from his doctoral dissertation, "Subjective musical pitch" [J. Acoust. Soc. Am. **26**, 369–380 (1954)], is one of the seminal works in the area, as is his treatise on absolute pitch [Sound **2**(3), 14–21 (1963); **2**(4), 33–41 (1963)]. In these publications, he proved that the performance of certain musicians on pitch tasks, such as ratio production and absolute identification, was an order of magnitude more precise than predicted by the then-current psychophysical "laws," derived from research on "normal-hearing" subjects. Dix Ward's research in this area will be discussed, as will recent research which extends and complements his findings. Taken together, this body of work gives empirical evidence for something of which Dix himself was long aware: It is music perception which most closely utilizes the full capabilities of the auditory system; and it is music, rather than speech, which is in fact "special."

5:10

4pPPb7. Risk assessment for occupational noise induced hearing loss: W. Dixon Ward and the "Anti-Chicken Little Movement." William W. Clark (Central Inst. for the Deaf, 818 S. Euclid Ave., St. Louis, MO 63110)

One of the hallmarks of W. Dixon Ward's illustrious career was his proclivity to criticize effectively the efforts of his colleagues. Depending upon the topic and the victim, this ability led him to be characterized as an outstanding editor, a thoughtful reviewer, an overzealous critic, or an impediment to consensus. Despite the variety of responses, Dix's critiques shared one characteristic: They were almost always correct. One of his favorite targets was the international standard, "Acoustics—Determination of occupational noise exposure and estimation of noise-induced hearing impairment" (ISO 1999, 1990). Throughout the two decades of its development Dix took to task the framers of the standard, challenging the assumptions which underly it. Dix was particularly vexed by ISO's use of the 3-dB exchange rate, the presumption of 0% risk of noise induced hearing loss for those nonoccupationally exposed, and the estimation that normal 18-year-olds have perfect hearing. In this tribute, Dix's criticisms of the ISO standard will be articulated and supported with recent data, and his role as the leader of the "Anti-Chicken Little Movement" justified.

Session 4pSA

Structural Acoustics and Vibration: Acoustic Radiation from Structures

Joel M. Garrelick, Chair

Cambridge Acoustical Associates, Inc., 200 Boston Avenue, Medford, Massachusetts 02155

Contributed Papers

2:00

4pSA1. The forward and backward projection of harmonic pressure fields from complex three-dimensional bodies. Peter R. Stepanishen (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett, RI 02882-1197)

The forward and backward projection of pressure fields from complex three-dimensional harmonically vibrating bodies is an important problem in noise control. A new generalized internal source density (GISD) method is presented to address the projection problem. The GISD approach is based on decomposing the pressure field on a closed surface of revolution in the fluid into a summation of circumferential orders where the pressure field for each order is associated with an internal linear distribution of ring sources on the axis of revolution. The axial variation of each source distribution is formulated as the solution of a mean square error problem and the resulting distributions can then be simply used to determine the entire external and/or surface pressure and velocity fields of the vibrating body. Analytical and numerical results are presented for several examples to illustrate the basic approach for bodies with different aspect ratios. Exact expressions for the source strength distributions for a sphere are developed and shown to have a vanishingly small region of support about the center of the sphere. For cylindrically shaped bodies with large aspect ratios, the spatial characteristics of the source strength distributions more closely match the normal velocity distribution.

2:15

4pSA2. An analytical solution for a rectangular piston independently driven at two edges. John MacGillivray (Graduate Program in Acoust., Penn State Univ., State College, PA 16804)

Far-field radiation of a baffled rectangular piston is described using a frequency-domain Rayleigh integral solution. The piston model is driven at a single frequency at two opposing edges, where these individual edge sources are independent of one another in both amplitude and phase. Directivity plots of the analytical solution are used to graphically observe directional effects caused by changing the phase and amplitude of these edge sources. Interesting and unexpected beamforming characteristics are shown to be produced by certain amplitude and phase combinations. [Work supported by IBM through the Shared University Research program and the Applied Research Laboratory at Penn State Univ. University.]

2:30

4pSA3. A resistance probe based on the lumped parameter model for the acoustic power output of a vibrating structure. John B. Fahnlne (Appl. Res. Lab., Penn State Univ., University Park, PA 16802), Gary H. Koopmann, and Weicheng Chen (Penn State Univ., University Park, PA 16802)

Previous papers by the authors addressed the analytical derivation and the numerical implementation of a lumped parameter model for the acoustic power output of a vibrating structure. Here the experimental measurement of the resistance is addressed. The acoustic resistance, which is related to the Green's function of the second kind, can be determined experimentally by: (1) constructing an acoustically hard physical model of

the boundary surface; (2) representing the discrete simple source with a small loudspeaker; and (3) measuring the acoustic field at a receiver location with a small microphone. The design of an acoustic resistance probe is described, along with an analysis of both the calibration procedure and the subsequent calculation of the resistance matrix. The analysis is validated through several example problems, with the accuracy of the measured acoustic resistance assessed by direct comparison to analytical or numerical solutions. For most of the cases considered, the experimental measurements and numerical predictions agree very well, but for some of the example problems discrepancies occur due to the limited dynamic range of the probe and acoustic excitation of structural vibrations.

2:45

4pSA4. Accurate determination of total sound power radiated from a rectangular panel. W. L. Li and H. J. Gibelung (United Technologies Carrier Corp., Carrier Pkwy., Syracuse, NY 13211)

Radiation of sound from planar vibrating panels has been extensively studied for years. In most of these investigations, the total sound power is approximately obtained from the modal radiation efficiencies without considering the contributions from modal cross couplings. The primary reason for doing this is perhaps that the inclusion of cross-coupling terms may result in a tremendous increase in computational load, particularly when a large number of structural modes is involved. In a recent paper [R. D. Snyder and N. Tanaka, J. Acoust. Soc. Am. **97**, 1702-1709 (1995)], an approximate estimate of the cross-coupling terms from the modal radiation efficiencies was discussed. It is, however, valid only in a low-frequency range and deteriorates rapidly as frequency increases. In this paper, the effect of modal cross couplings on the total sound power radiated from a simply supported rectangular panel is discussed. A formula is derived for an accurate and efficient calculation of the cross-coupling terms which allows for effortlessly including the effects of the modal cross couplings in the total sound power calculation. In addition, the present formulation is valid in the whole frequency range. Results are presented.

3:00

4pSA5. Acoustic radiation in light fluid of a vibrating beam with uncertain material properties. Jean-Sébastien S. Genot, François Charron, and Nouredine Atalla (Dept. of Mech. Eng., Univ. of Sherbrooke, 2500 boulevard de l'Université, Sherbrooke PQ J1K 2R1, Canada)

Engineers are aware of the need to account for uncertainties in their predictions in knowledge of the material and cross-section properties (e.g., Young's modulus, damping factor, section area). Furthermore, acousticians are often interested in the prediction of the maximum sound power that could be radiated by a given structure. But they face a numerical challenge due to the amount of input parameter combinations they must solve to compute the maximum response. In this context, a simple case of acoustic radiation in two dimensions—a simply supported beam in light fluid—is a good way to investigate the sensitivity of the acoustic power to the uncertainties of the input parameters. To study this case, an analytical method with a perturbation scheme up to the second order has been intensively used, and compared with a Monte Carlo method. In this presentation, it will be shown that the acoustic power is very sensitive to the

variations of the structural parameters in the neighborhood of the beam eigenfrequencies, and that second-order perturbation techniques are not sufficient. Consequently, a hybrid method is proposed to improve the results locally.

3:15–3:30 Break

3:30

4pSA6. Influence of partial coatings on the acoustic radiation from a fluid-loaded structure. Joseph M. Cuschieri (Dept. of Ocean Eng., Florida Atlantic Univ., Boca Raton, FL 33431, joe@jmc.oe.fau.edu) and David Feit (David Taylor Res. Ctr., Bethesda, MD 20084)

In general, an acoustic coating is applied uniformly on the surface of a fluid-loaded structure to minimize acoustic radiation and scattering. There are, however, some inherent advantages to optimize the distribution of the coating around areas from which the acoustic radiation appears to emanate. This would be analogous to the application of damping treatment in areas of a vibrating structure which have high vibration levels. In the case of the acoustic radiation the problem is more complex because of the coupling between the acoustic fluid and the structure. In this paper, the acoustic radiation from a partially coated fluid-loaded structure, of infinite extent and excited by a point or line force, is examined by developing an analytic/numeric solution. The solution follows along some of the procedures that the authors developed in previous work related to the scattering from fluid-loaded plates and shells. The coating is assumed to provide mainly a decoupling layer between the acoustic medium and the structure; that is, it does not add mass or stiffness to the base structure. The influence of added mass or stiffness of the coating can be included as an added inhomogeneity and treated separately in the solution. [Work sponsored by NSW-CD.]

3:45

4pSA7. A novel acceleration method for the variational boundary element approach based on multipole expansions. Michel Tournour and Noureddine Atalla (Dept. de Genie Mech., Univ. de Sherbrooke, Sherbrooke, PQ J1H 2R1, Canada)

The acoustic radiation of general structures with Neumann's boundary condition using the variational boundary element method (VBEM) is considered. The classical numerical implementation of the VBEM suffers from the computation cost associated with double surface integrals. To circumvent this limitation a novel acceleration method is proposed. It is based on the expansion of the cross influence matrices in terms of multipoles using the expansion of the Green's function in terms of spherical Bessel functions. Since the resulting multipoles are not dependent on the elements locations, the technique results in large computation time savings for homogeneous meshes. The theory behind the approach, its convergence, and its numerical implementation in a VBEM code will be presented. It will be shown that by accounting for the monopole, dipole, and quadrupole terms in the multipoles expansion, the classical convergence criteria usually used for quadratic element hold. Numerical applications to acoustic radiation from plates, open cylinders, and closed boxes will be presented to demonstrate the accuracy and efficiency of the method.

4:00

4pSA8. A normal decomposition technique for fully coupled problems in structural acoustics. Debashis Basu and Suresh Kumar S. (State Univ. of New York at Buffalo, 212 Ketter Hall, Amherst, NY 14260)

Problems in structural acoustics are inherently coupled. Whenever a structure vibrates, there is a noise associated to it. Similarly, any noise would induce vibrations in the structure, however small, depending on how thin it is. The indirect boundary element technique has been used for quite some time now to model the acoustic domain. Again, plate/shell formulations for thin structures using finite elements are well known. Usually, for a coupled structural acoustic problem, the associated eigenvalue problem of the structure is solved first, and then the acoustic system is coupled using a predetermined number of modal participation. For weak coupling, this technique gives accurate results considering a few modes, while a greater number of modes are necessary for problems with strong coupling. An alternate technique is presented in this paper, where the structural degrees of freedom are condensed to couple directly with the acoustic system. For the frequency under consideration, there are absolutely no assumptions involved and the coupling is exact. Several examples illustrate the validity of the method. For problems where the level of coupling (weak or strong) and the number of modes to be included are difficult to ascertain, the current technique provides an excellent alternative. With noise control playing a major role in developed and developing countries, a better simulation of the field problem clearly identifies areas for noise reduction measures.

4:15

4pSA9. Energy balance for radiated waves and propagating vibration in a fluid-loaded beam under point excitation. Svetlana I. Kovinskaya (Sci. Res. Ctr. for Ecological Safety, Russian Acad. of Sci., Korpusnaya St. 18, St. Petersburg 197042, Russia, svetlana@srcs.samson.spb.su)

Using a compensating solution method [S. Kovinskaya and I. Krasnov, *Acta Acust.* **82**, 237 (1996)] a field of flexural waves propagating along a submerged beam under a point excitation and a radiated acoustical field are represented analytically. The method is based on the Fourier integral transform of the vibration equation and is similar to the hybrid method [J. M. Cuschieri and D. Feit, *J. Acoust. Soc. Am.* **95**, 1998–2005 (1994)]. The main and compensating solutions are determined as a result of the inverse Fourier transform along a real wave number axis of nonuniform and uniform equations, correspondingly. A sum of the main and compensating solutions must supply the vibration radiation condition. Let the last equation determine an indefinite coefficient in the compensating solution. The energy balance equation based on the solution of the problem leads to an energetic equivalence between a decreasing vibration and radiated spherical wave and the same between a cylindrical acoustical wave and an energy difference in the excitation point and infinity for propagating flexural wave.

Session 4pSCa

Speech Communication: Production and Dynamic Modeling I (Poster Session)

P. F. Castellanos, Chair

Division of Otolaryngology, Head and Neck Surgery, University of Maryland, 22 South Greene Street, Baltimore, Maryland 21201

All posters will be on display from 1:15 p.m. to 3:15 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:15 p.m. to 2:15 p.m. and contributors of even-numbered papers will be at their posters from 2:15 p.m. to 3:15 p.m.

Contributed Papers

4pSCa1. Normalization for articulatory recovery. Richard S. McGowan (Sensimetrics Corp., 26 Lansdowne St., Cambridge, MA 01239)

Articulatory recovery in an analysis-by-synthesis procedure must account for individual differences in vocal tracts. In the procedure described here, all individual vocal tracts are referenced to the vocal tract of an articulatory synthesizer, known as the standard vocal tract. In order to normalize for any given talker, characteristics of the standard vocal tract are adjusted until its sound output closely matches the sound output of the talker over a range of speech sounds. After these standard vocal tract characteristics are fixed, the analysis-by-synthesis procedure is enabled to recover the articulation from the talker. Experiments are conducted with articulatory and acoustic data from the Wisconsin x-ray microbeam database. The midsagittal sections of the standard vocal tract are made to match those of the x-ray microbeam over a set of phones. Characteristics of the standard vocal tract are adjusted so that its acoustic output matches that of the talker as closely as possible over this range of phones. The normalization procedure is checked by choosing a novel utterance for recovery, and comparing the resulting recovered midsagittals with and without prior normalization. [This work was supported by Grant NIDCD 01247.]

4pSCa2. Acoustic-to-articulatory mapping of isolated and intervocalic fricatives. Edward L. Riegelsberger and Ashok K. Krishnamurthy (Dept. of Elec. Eng., Ohio State Univ., 2015 Neil Ave., Columbus, OH 43210, riegelse@er4.eng.ohio-state.edu)

While acoustic-to-articulatory mapping has been successful for purely voiced speech, a complete acoustic-to-articulatory mapping solution for all classes of speech sounds remains to be found. This work investigates extending acoustic-to-articulatory mapping techniques to fricative sounds. A normalized spectral correlation measure has been developed for an acoustic-to-articulatory mapping routine that produces good acoustic fits and reasonable articulatory configurations for isolated fricatives. The mapping routine can, by using this distance measure along with amplitude information, effectively separate fricative classes by place of articulation. Based on results with isolated fricatives, the acoustic-to-articulatory mapping of fricatives has been extended to the intervocalic case, where information in the voiced segments preceding and following a fricative is used to initialize its optimization. Confidence in the acoustic-to-articulatory mapping of fricatives in isolation influences how they are used in this dynamic case. Heuristics for controlling frication amplitude and mixed excitation are discussed.

4pSCa3. Visible speech revisited: An acoustically driven model of lip and tongue motion. Jay T. Moody (Dept. of Cognit. Sci., Univ. of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0515, jmoody@cogsci.ucsd.edu) and Maureen Stone (Univ. of Maryland Med. School, Baltimore, MD 21201)

A method is presented for converting acoustic speech data into a "speech readable" movie of a canonical talking face using a neural network. Target values for the network are created by finding the principal components of a set of video frames, with each frame consisting of side-by-side images of the face and tongue of a single speaker. Tongue contours are extracted from mid-sagittal ultrasound images. Input to the network consists of a set of cepstral parameters and their derivatives calculated over 22-ms windows (overlapping by 11 ms). Two such input frames are matched to each video frame (33 ms). Recurrent (backward) connections in the network encourage it to learn not only the acoustic-articulatory pairings, but also articulatory trajectories (expectations of the next articulatory state based on recent articulatory states). It is hypothesized that this supplemental trajectory information helps to alleviate the uncertainties inherent in the vocal tract inverse mapping. After training, the network is presented with new (untrained) tokens (audio only) of utterances from the training corpus and the network's video output is recorded. The network's output is compared to the actual recorded video. [Work supported by US Dept. of Education and NIH.]

4pSCa4. Muscle-based modeling of facial dynamics during speech. Jorge C. Lucero, Kevin G. Munhall (Dept. of Psych., Queen's Univ., Kingston, ON, Canada), Eric Vatikiotis-Bateson (ATR Human Information Processing Res. Labs., Kyoto, Japan), Vincent L. Gracco (Haskins Labs., New Haven, CT 06511-6695), and Demetri Terzopoulos (Univ. of Toronto, Toronto, Canada)

A dynamical, muscle-based model of the face is being developed to extend the facial model of Terzopoulos *et al.* [e.g., D. Terzopoulos and K. Waters, IEEE Trans. Pattern Anal. Machine Intell. **15**, 569–579 (1993)]. The purpose of this work is to characterize facial dynamics during speech with a physiologically realistic model. The model consists of a multilayered deformable mesh of lumped masses connected by springs and viscous elements which represents the layered structure of facial tissue. The spring and viscous constants approximate the stress/strain and viscous characteristics of facial skin. The mesh is deformed by forces generated by a set of modeled muscles of facial expression whose physical characteristics are being determined empirically. The shape of the facial model is individualized to subjects facial morphology with data from a laser ranger finder (Cyberware scanner). In this report, work on driving the model with intramuscular, perioral EMG signals is presented. Recordings of the sampled

three-dimensional position of a subject's facial surface using OPTOTRAK are compared to the patterns of deformation of the epidermal mesh in the model. Results will be discussed in terms of the strengths and weaknesses of this modeling approach. [Work supported by NIH-NIDCD Grant No. DC-00594 and NSERC.]

4pSCa5. The effect of task manipulation on the temporal ordering of articulatory events. Susan Shaiman (Dept. of Commun. Sci. and Disord., Univ. of Pittsburgh, 4033 Forbes Tower, Pittsburgh, PA 15260, shaiman@csd.upmc.edu)

This investigation examined the sequencing of lip and jaw movement onsets and peak velocities for bilabial closure as a function of speaking rate, phonetic context, and position of the closing gesture within the utterance. Articulatory movements were transduced using head-mounted strain gauge transducers. Five normal speakers repeated the target words /paep/, /paeps/, and /paepst/, within the carrier phrase, "Now say ____ again." Speaking rate (normal, fast, and slow) was manipulated using a magnitude production scale with no external time constraints. Most subjects demonstrated consistency in the sequencing of articulatory events at fast speaking rates, while normal and slow rates evidenced increasing variability. The position of the closing gesture within the utterance also affected the sequencing, with the closing gesture for the first /p/ usually exhibiting greater consistency than the closing gesture for the second /p/. The production of a single consonant versus consonant clusters appeared to have little impact on the sequencing of articulatory events for bilabial closure. Despite these general trends, individual subject differences were evident, with some speakers demonstrating greater consistency than others. These findings suggest that temporal ordering of articulatory events is flexibly organized based on the specific task demands. [Work supported in part by NSERC.]

4pSCa6. Kinematic characteristics of gestural onsets and the single-geminate consonant distinction in Italian and Finnish. Margaret Hall Dunn (Haskins Labs., 270 Crown St., New Haven, CT 06511, dunn@haskins.yale.edu)

Selspot traces of lip aperture change in the movement-to-closure portions (movement from preceding vowel into consonant) of single and geminate bilabial consonants in two unrelated languages, Italian and Finnish, were assessed to test the hypothesis that geminates are equivalent to sequences of identical overlapping single consonant gestures. If the closing portions of single and geminate consonants exhibited the same kinematic structure, this would be evidence that the geminate is composed of overlapping gestures. Duration of closing movement, displacement during closing, and peak velocity achieved during closing were compared for productions of single and geminate bilabial consonants for four speakers in each language. For Italian, differences in these kinematic variables during the closing portions of singles and geminates could be accounted for in terms of differences in the timing of geminates with respect to preceding vowels and the overlapping-gesture hypothesis was supported; in Finnish the kinematics of the two types of consonants was not equivalent, and it was concluded that Finnish includes a distinct inventory of gestures for the production of long, rather than geminate, consonants. [Work supported by NIH Grant No. DC-00121 to Haskins Laboratories.]

4pSCa7. Temporal relationships of the articulatory, acoustic, and perceptual effects of fricative-vowel coarticulation. Ying Xu (Dept. Speech and Hearing Sci., The Ohio State Univ., Columbus, OH 43210)

Anticipatory coarticulation has been extensively studied as part of the general phenomenon of coarticulation. However, the temporal relationships of the articulatory, acoustic, and perceptual effects of this particular type of coarticulation have not received much attention. The purpose of this study was to systematically examine lip protrusion gestures, acoustic and perceptual effects in fricative-vowel contexts, and explore the temporal relationships among these three effects. The onset of lip protrusion

and acoustic effects as obtained and compared by employing a minimal contrast technique. The results showed that the onset of lip protrusion was much more consistent in a relative timeframe than in an absolute timeframe, and it was much earlier than, but independent of, the onset of traceable changes in formant frequencies. A reliable effect of anticipatory coarticulation on perception was also observed and the onset of this perceptual effect occurs tens of milliseconds later than the onset of the acoustic effect.

4pSCa8. Modeling the acoustics of American English /r/. Carol Y. Espy-Wilson (Dept. of Elec. and Comput. Eng., Boston Univ., Boston, MA 02215), Narayanan Shrikanth, Abeer Alwan (Univ. of California, Los Angeles, CA, 90024), and Suzanne E. Boyce (Boston Univ., Boston, MA 02215)

It is often assumed that speakers use idiosyncratic combinations of constrictions at the lips, pharynx, and along the hard palate to achieve the low F_3 's—typically between 1300 and 1800 Hz—characteristic of American English /r/. This account was tested in a modeling study using attested vocal tract dimensions derived from three-dimensional magnetic resonance imaging data from four speakers [Alwan *et al.*, J. Acoust. Soc. Am. (to be published) (1997)]. These dimensions were used as input to the Maeda computer model. It was determined that (1) a constriction in the post-alveolar region of the palate and at least one other constriction are needed to bring F_3 into the 1300- to 1800-Hz range, (2) the maximum effect of the lip constriction was around 100 Hz, while that of the pharyngeal constriction was around 200 Hz, and (3) the effects of changing constrictions were not additive. Over all manipulations, the lowest F_3 achieved was 1495 Hz, which is high compared to many /r/ productions. Using "retroflex" dimensions, the lowest F_3 achieved was 2224 Hz. It may be that additional acoustic mechanisms are required to account for the full range of acoustic profiles seen for /r/. [Work supported by NSF and NIH.]

4pSCa9. Acoustic estimates of vowel centralization in persons who stutter. Michael Blomgren, Michael P. Robb, and Yang Chen (Dept. of Commun. Sci., Univ. of Connecticut, 850 Bolton Rd., U-85, Storrs, CT 06269-1085, mib95003@uconnvm.uconn.edu)

The magnitude of tongue movement for fluently produced utterances of persons who stutter and those of normally fluent controls was inferred through examination of the vocal tract vowel space. Fifteen adult males served as subjects comprising separate groups of untreated stutterers, stutterers who had completed a fluency-shaping treatment program, and non-stuttering controls. The steady-state portion of formant one (F_1) and formant two (F_2) was examined in the production of various CVC tokens embedded in a carrier phrase. Average F_1 and F_2 values for [i], [a], and [u] were plotted in an F_1 - F_2 vowel space triangle for each subject. The corresponding vowel space was measured by determining the area (Hz^2) of the triangle. Results indicated the untreated group of persons who stutter had the smallest average vowel space, while the largest vowel space was observed in the control group. Discussion focuses on the vocal tract articulation characterizing fluent speech production.

4pSCa10. Using principal component analysis of tongue surface shapes to distinguish among vowels and speakers. Maureen Stone (Univ. of Maryland, School of Medicine, Div. of Otolaryngol.-HNS, Frenkil Bldg., Ste. 500, 16 S. Eutaw St., Baltimore, MD 21201), Y. Cheng (Univ. of Maryland, College Park, MD 20742), and A. Lundberg (Johns Hopkins Univ., Baltimore, MD 21218)

The present study uses principal component analysis (PCA) to examine sagittal tongue contours for five English vowels taken from ultrasound images. The vowels are repeated three to five times each in a /pVp/ carrier utterance. Of particular interest is the use of plots of the coefficients of the eigenvectors to distinguish both subjects and vowels, and the building of a

linear model to fit their family of data. Data will be transformed to normalize surface length across contours. Short contours will be stretched in the x direction to the length of the longest curve. The x and y dimensions will be stretched equally for each curve to preserve scale. Preliminary data indicate fairly good success distinguishing among three subjects and four vowels using a linear model based on the first few components. Methods of improving this result are being explored using factor analysis or optimal fitting.

4pSCa11. A computerized ultrasound transducer scanning system.

Edward Davis (Dept. of Mech. Eng., Dept. of Elec. and Comput. Eng., Johns Hopkins Univ., Baltimore, MD 21218) and Maureen Stone (Univ. of Maryland, Baltimore, MD 21201-1595)

A computerized electromechanical ultrasound transducer holder is being developed to allow representation of three-dimensional tongue motion. Current three-dimensional ultrasound systems are comparatively slow (10 s/volume) and hence are not useful for the study of dynamic three-dimensional speech. The proposed mechanism will (automatically) both scan along and rotate about an axis, sensing transducer position and providing feedback. This motion will be motorized and controlled by a computer according to custom-developed software and parameters. One can program the computer to perform a synchronized series of approximately five sagittal or coronal slices. These slices can later be assembled and interpolated to form a three-dimensional surface model. Mechanically, this system is advantageous because it rotates the transducer about a virtual center, which is located along the axis of the center of the arc of the piezoelectric crystals within the transducer, and does not necessitate additional mechanical structures near the subject. The electrical and computer control system will allow automated imaging of a series of coronal or sagittal slices synchronized in time and located precisely in space. This system will be added to and used in conjunction with the HATS system described in Stone and Davis [J. Acoust. Soc. Am. **98**, 3107–3112 (1995)]. [Work supported by Northrup Grumman and NIH.]

4pSCa12. Acoustic, aerodynamic, and physiological characteristics of modal and vocal fry registers.

Yang Chen (Dept. of Commun. Sci., Univ. of Connecticut, Storrs, CT 06268, yac93001@uconnvm.uconn.edu), Michael Blomgren (Univ. of Connecticut, Storrs, CT 06268), Manwa L. Ng (Illinois State Univ., Normal, IL 61790), and Harvey R. Gilbert (Univ. of Connecticut, Storrs, CT 06268)

This study examined production of the modal and vocal fry registers using various acoustic and aerodynamic parameters. Airflow (AF), air pressure (AP), electroglottalgraphic (EGG), and acoustic data were obtained simultaneously for four isolated vowels (/i/, /æ/, /a/, /u/) and the stop consonant produced in a string of /pi/ syllables for both modal and vocal fry registers. Twenty normal speakers (10 males, 10 females) who were able to produce vocal fry participated in this study. Results demonstrated considerable differences among AF, AP, EGG, and acoustic signals for the modal and vocal fry registers. The fundamental frequency in vocal fry register (male: 49.1 Hz; female: 48.1 Hz) was found to be significantly lower than that in modal register (male: 117.5 Hz; female: 211.0 Hz). Also, the average air pressure values decreased in vocal fry register (male: 5.51 H₂O; female: 5.25 H₂O) compared to values obtained in the modal register (male: 7.45 H₂O; female: 7.56 H₂O). In vowel production, the airflow rate in the modal register was almost three times as high as the airflow rate in vocal fry register. For the production of /pi/, airflow rate associated with the stop consonant in the modal register was approximately 1.5 times higher than that in the vocal fry register.

4pSCa13. Two types of vowel devoicing in Japanese: Evidence from articulatory data.

Ayako Tsuchida (Dept. of Linguist., Cornell Univ., Ithaca, NY 14853, at18@cornell.edu), Shigeru Kiritani, and Seiji Niimi (Univ. of Tokyo, Bunkyo-ku, Tokyo, 113 Japan)

Japanese high vowels [i, u] become devoiced when they occur between voiceless segments: e.g., [kita] “north.” Vowel devoicing (VD) occurs systematically, except when a high vowel appears between two voiceless fricatives, where VD is less consistent and nonobligatory. The present study examined the laryngeal gestures during the production of Japanese vowels in various phonological contexts using fiberoptic and electromyographic techniques. Stimuli consisted of [CiCe] words, where the consonants were stops (S) or fricatives (F). The results revealed that the glottal gestures for typical devoiced vowels (those in [SiSe], [FiSe], or [SiFe]) and devoiced vowels in [FiFe] were distinct, although the two types were indistinguishable in the spectrograms. Consistent with observations in a previous study [H. Hirose, *Phonetica* **23**, 156–170 (1971)], [CiC] sequences with typical devoiced vowels were produced with a wide open glottis accompanied with high activities of the abductor muscle. In contrast, in [FiF] sequences, the glottis was wide open in each of the fricatives, but was almost closed during the vowel, resulting in a bimodal pattern. When the glottis reached complete closure during the vowel, a voiced vowel was observed. The present study shows two distinct mechanisms for VD, suggesting that they are two independent processes.

4pSCa14. Do children with phonological disorders use more ballistic articulatory movements?

Marios Fourakis (Dept. of Speech and Hearing Sci., Ohio State Univ., 110 Pressey, 1070 Carmack Rd., Columbus, OH 43210), Mary E. Beckman, and Jan Edwards (Ohio State Univ., Columbus, OH 43210)

This examination of speech motor control in phonologically disordered children adapted a paradigm that has been used in the dysarthria literature [Weismer *et al.*, J. Acoust. Soc. Am. **91**, 1085–1098 (1992)]. Six preschool-aged children with phonological disorders, six age controls, and six adults repeated the phrases “baby dog” and “good baby” at normal and fast rates. F_2 tracks of the fast rate tokens were processed using an in-house algorithm to calculate the transition slope in Hz/ms. For the alveolar and velar stops (but not for the labial stops) the F_2 slope distributions for the typically developing children were quite similar to those of the adults, and both differed from the children with phonological disorders in showing smaller slope values. The phonologically disordered children had faster movements than either their age controls or adults, but only in the sequences where the tongue–jaw complex was harnessed for the consonant as well as for the contiguous vowel. This pattern suggests that the phonologically disordered children are using a faster, more ballistic gesture for lowering the tongue–jaw complex for all three stop places of articulation, while the typically developing children and adults use this strategy only after a labial closure.

4pSCa15. Disordered phonology: An acoustic description of the acquisition of the anterior distinction among coronal fricatives.

Adele W. Miccio (Dept. of Commun. Disord., Penn State Univ., 105 Moore Bldg., University Park, PA 16802, awm4@psu.edu) and Karen Forrest (Indiana Univ., Bloomington, IN 47405)

This paper reports on a longitudinal analysis of the acoustic properties of speech samples of a child, age 5 years and 7 months, with disordered phonology. The child, who produced [s] for both alveolar and postalveolar fricatives, was enrolled in a remediation program that focused on the correct production of the voiceless postalveolar fricative [ʃ] “esh.” Training involved imitation of [ʃ] in isolation and spontaneous production of [s] and [ʃ] in minimally paired words. Acoustic measurements were taken prior to treatment, immediately following the 2-week treatment program, and again 2 months following termination of treatment. Temporal and spectral measurements were conducted. Results indicated no significant differences between [s] produced for the target alveolar fricative /s/ vs [s]

produced for the target postalveolar fricative /ʃ/ prior to treatment. Following treatment, the child produced the anterior–nonanterior distinction among coronal fricatives. Acoustic measurements showed that she did not simply add a postalveolar fricative to her repertoire of speech sounds but also changed her productions of target-appropriate /s/. [Work supported in part by NIH-DC 00260.]

4pSCa16. Variability in /s/ and /ʃ/ productions within and across talkers. Rochelle S. Newman (Psych. Dept., Park Hall, SUNY, Buffalo, NY 14260)

A primary issue in speech perception research is the apparent “lack of invariance” between the acoustic information in a signal and the listener’s phonemic perception. The same intended phoneme can be produced with a wide range of acoustic values, and different intended phonemes may be produced with the same values. Although the existence of this variability is not in question, the degree to which listeners have to account for it is unclear. Some studies have attempted to measure the extent of this variability, but few have done so for fricatives, making it difficult to determine the degree to which individuals need to normalize fricative input for the individual talker. The current experiment attempts to rectify this hole in the literature by examining the fricative centroids for over 100 utterances beginning with /s/ and /ʃ/ from each of 20 different speakers. Although there was variability in the centroid values, the categories showed virtually no overlap within a talker. However, there was substantial overlap across talkers. This suggests that if listeners use centroids to cue the /s/–/ʃ/ distinction, they would first need to normalize the signal for the individual talker’s frequency range. [Work supported by NIDCD Grant No. R01 DC00219 to SUNY at Buffalo.]

4pSCa17. An articulatory and acoustic analysis of English flaps: Evidence for attentional reduction in production. Kenneth J. de Jong (Dept. of Linguist., Indiana Univ., Bloomington, IN 47405)

Traditional models of American English phonology include a rule that coronal stops appear as flaps before unstressed vowels. Recent studies suggest that this flapping is better explained as the result of blending the articulatory specifications for the stops with neighboring vowels. This blending produces compromise articulatory positions for the stops, which

cross over the perceptual boundary between stops and flaps. This paper tests this explanation against an x-ray microbeam corpus of American English stops. The occurrence of flapping was determined via transcriptions. Regression analyses indicate that acoustic parameters such as VOT and voicing occurrence better predict transcriptions than articulatory parameters such as tongue position and motion during closure. In addition, articulatory analyses indicate a variation from [t] to flap which is generally consistent with a blending account. However, the contextual influence from the consonants on neighboring vowels and various aspects of articulatory postures in the consonants themselves is more consistent with a reduction of the stop, rather than a blending with neighboring vowels. These results suggest that flapping results from a linguistic convention which allows hypoarticulation of unstressed items and a secondary result of some of these items straying across an acoustic boundary. [Supported by the NSF and NIDCD.]

4pSCa18. Acoustic evidence for featural and subfeatural errors in speech production. Stefan Frisch and Richard Wright (Speech Res. Lab., Indiana Univ., Bloomington, IN 47401)

Phonological speech errors are a traditional source of psycholinguistic evidence for the representations of phonology theory. In an electromyographic (EMG) study of experimentally induced phonological speech errors, Mowrey and MacKay [J. Acoust. Soc. Am. **88**, 1299–1312] found that speech errors frequently occur at a subfeatural, gestural level, with no apparent effect on the percept of the word. Mowrey and MacKay’s study considered the activity of a single muscle, and thus was unable to determine whether single gestures acted independently of gestural constellations. This study is a preliminary report from an ongoing acoustic analysis of speech errors. The data are tape recordings of an error-inducing experiment using nonsense tongue twisters. Recordings of six speakers producing four different tongue twisters targeting /s/ and /z/, e.g., sit zap zoo sip, were digitized and analyzed. Some errors involved multiple changes in acoustic properties, including simultaneous changes in periodicity, amplitude of frication, and duration, while others involved a subset of these properties. This evidence suggests that errors can occur at both the single gesture level, affecting noncontrastive acoustic properties, and at the level of the gestural complex or segment, creating a perceptible, linguistically contrastive change. [Work supported by NIH Grant No. DC00012.]

THURSDAY AFTERNOON, 19 JUNE 1997

ROOM D, 3:30 TO 5:00 P.M.

Session 4pSCb

Speech Communication: Production and Dynamic Modeling II

P. F. Castellanos, Chair

Division of Otolaryngology, Head and Neck Surgery, University of Maryland, 22 South Greene Street, Baltimore, Maryland 21201

Contributed Papers

3:30

4pSCb1. Evidence for a second laryngeal sound source. P. F. Castellanos (Div. of Otolaryngol.—Head and Neck Surgery, Univ. of Maryland School of Medicine, 22 S. Greene St., Baltimore, MD 21201) and S. A. Elder (U.S. Naval Acad., Annapolis, MD 21402)

Sondhi reflectionless tube and strobed video have been used to investigate single glottal pulse gestures in the vocal fry range. There appear to be two sources of sound in normal VF phonation, one monopole, the other

quadrupole. It is quadrupole sound, in fact, that seems to define the shape of observed pressure trace in the single glottic pulse, or SGP. This sound pulse, which lasts 10 ms or less, resembles a single cycle of negative sine wave beginning just before the closing phase, and may be recognized even in sound emissions outside the tube, where continuous tone samples can be identified as SGP wave trains, each link beginning with a characteristic downturn. Monopole sound, emitted in shorter pulses during the abrupt zipping and sometimes during unzipping phases of the SGP, shows up along the main wave trace in the Sondhi tube as a small superposed peak, followed by a string of head echoes. The source of quadrupole sound can

be traced to fluctuating Bernoulli pressures during closure which produce opposing forces on the vocal folds. The quadrupole or q -wave forms the acoustic signature of the SGP.

3:45

4pSCb2. Control of oral closure in alveolar and velar stop consonant production. Anders Lofqvist and Vincent L. Gracco (Haskins Labs., 270 Crown St., New Haven, CT 06511, lofquist@haskins.yale.edu)

Earlier work has shown that the lips are moving at a high velocity at the instant of oral closure for bilabial stop consonants, resulting in tissue compression. The lips may thus have virtual targets that would require them to move beyond each other. The present study examines events at the oral closure for stops produced with the tongue. Tongue movements were recorded using a magnetometer system. Four subjects participated and produced ten tokens each of VCV sequences with all possible combinations of the three vowels /i, a, u/ and the consonants /t, d, k, g/. Consistent with the results for bilabial stops, the tongue was moving at a high velocity at the instant of oral closure. The tongue movement trajectories were more complex with a larger horizontal component than the lips. For velar stops, the tongue body was usually moving forward at stop closure, with the velocity influenced by vowel context. For alveolar stops, the tongue tip horizontal movement direction depended on the preceding vowel, but also differed across subjects. Between the onset and release of the oral closure, the tongue tip and tongue body moved through a trajectory of usually less than 1 cm. [Work supported by NIH.]

4:00

4pSCb3. Congruence of articulatory and acoustic variability. Alice Faber (Haskins Labs., 270 Crown St., New Haven, CT 06511) and Julie M. Brown (Univ. of Connecticut and Haskins Labs., New Haven, CT)

Johnson *et al.* [J. Acoust. Soc. Am. **94**, 701–714 (1993)] suggest, on the basis of observed inter-speaker variability in discriminant analyses of articulatory measures, that speakers utilize acoustically defined targets in speech production. The present paper compares inter-speaker variability in simultaneously recorded acoustic and articulatory data (s_t words) from five New England speakers. The articulatory data were x and y coordinates of coils on the tongue, lips, and jaw, transduced by the Haskins Laboratories EMMA system and recorded at three locations in the target vowel; acoustic data were $F1$, $F2$, and $F3$ measures at the same temporal locations. Discriminant analyses of the articulatory and acoustic data sets reveal congruent patterns of inter-speaker variability in the two domains. The inter-speaker differences do not reflect superficial dialect or idiolect differences (e.g., extent to which /a/ and /ɔ/ contrast, or tendency to glotalize syllable-final /t/). Rather, they reflect differences in the way subjects vary jaw position, especially height, suggesting an anatomical basis for the observed differences. Thus, within the limits of their oral morphology, these speakers are using comparable articulatory targets for speech sounds.

4:15

4pSCb4. Dynamic shear modulus of vocal-fold tissues and phonosurgical biomaterials. Roger W. Chan and Ingo R. Titze (Natl. Ctr. for Voice and Speech, Dept. of Speech Pathol. and Audiol., Univ. of Iowa, Iowa City, IA 52242)

Knowledge of the mechanical properties of vocal fold tissues is necessary for the constitutive modeling of vocal-fold mechanics. Assuming transverse isotropy in a linear, elastic continuum model of vocal-fold tissues, five material constants are needed to solve the constitutive equation

[D. A. Berry and I. R. Titze, J. Acoust. Soc. Am. **100**, 3345–3354 (1996)]. Among these constants, Poisson's ratios can be estimated by assuming tissue incompressibility, while the shear moduli and Young's moduli are related to one another. A parallel-plate rotational rheometer was used to examine the dynamic shear behavior of human vocal fold tissues and three commonly used phonosurgical biomaterials (bovine collagen suspension, absorbable gelatin suspension, and human subcutaneous fat). In oscillation at 0.1–10 Hz and at 37 °C, the magnitude of dynamic shear modulus of vocal fold mucosa was on the order of 100 Pa, close to that of fat. The shear modulus magnitudes of collagen and gelatin were an order of magnitude higher. These results suggest that the use of fat for vocal-fold augmentation surgery is more conducive to phonation, because of its similarity to the vocal-fold mucosa in shear stiffness. [Work supported by NIH.]

4:30

4pSCb5. Methods for acoustic-to-articulatory mapping and voice mimic systems. S. Chennoukh, D. Sinder, G. Richard, and J. Flanagan (CAIP Ctr., Rutgers Univ., Piscataway, NJ 08855)

The articulatory codebook approach to voice mimic systems is complicated by large amounts of data and expensive acoustic comparisons when searching the codebook. If, during codebook construction, the acoustic parameters are quantized and mapped back to the articulatory space, the inverse is accomplished without acoustic comparisons. Since the inverse is nonunique, articulatory trajectories must also be estimated to resolve acoustic parameters which map to multiple model shapes. This can be managed with a tracking technique that uses dynamic properties (position and velocity) of each articulatory parameter to estimate the next model shape. A recurrent algorithm results which finds an optimal path through the model shape variations. These approaches are helpful, but they do not address the problem that codebooks may not cover the entire articulatory space. Furthermore, there is no performance measure which indicates the extent of the coverage. This limitation is overcome by an analytic inverse mapping, rather than a table look-up, which relates the first two formant frequencies with articulatory parameters. This relationship, which is well established for vowels, uses a distinctive region model (DRM) to approximate three area function parameters as a function of the first two formant frequencies. [Research supported by ARPA DAST 63-93-C-0064.]

4:45

4pSCb6. Fluid dynamic studies of speech production. D. Sinder, M. Krane, S. Chennoukh, G. Richard, J. Flanagan (CAIP Ctr., Rutgers Univ., Piscataway, NJ 08855), S. Levinson (Lucent Technologies, Bell Labs, Murray Hill, NJ 07974), S. Slimon, and D. Davis (Electric Boat Corp., Groton, CT 06340)

Results are presented describing an investigation of aerodynamic sources in the vocal tract. The expected contribution of this study is a new parametrization for fricative sources based on geometric and flow conditions. The flow in an idealized fricative geometry is computed numerically using a slightly compressible, Reynolds-averaged form of the Navier–Stokes equations. In this formulation, the aerodynamic sources of sound are simulated with minimal approximation. This approach has produced encouraging results for vowel sounds. Comparisons of the computational results with physical experiments using the same geometry have also been performed. The results shed light on the source mechanisms for fricative sounds and also show the need for the development of numerical boundary conditions which can simultaneously handle convective and propagative phenomena. [Research supported by NSF/ARPA IRI-9314946 and ARPA DAST 63-93-C-0064.]

4p THU. PM

Session 4pUW

Underwater Acoustics: Propagation: General Topics

Martin A. Mazur, Chair

Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

Contributed Papers

1:30

4pUW1. The effect of an absorptive bottom on the propagation of the coherence function in a shallow-water channel. T. Barnard and M. J. Beran (Dept. of Elec. Eng., The Catholic Univ. of America, Cardinal Station, Washington, DC 20064)

When a transverse plane wave propagates through a shallow-water channel with random sound-speed fluctuations, the waveforms at different transverse separations no longer correlate perfectly. The associated coherence falls as the in-line propagation distance and the transverse separation increase. In the (lossless) rigid-bottom case, when the waveforms are represented as a summation of normal modes, the multimodal coherence vector obeys a first-order matrix differential equation with in-line propagation distance as the independent variable. The scattering matrix in this differential equation is a function of transverse separation. As the in-line propagation distance approaches infinity, the coherence vector approaches a constant vector times a scalar Dirac delta function centered at zero transverse separation. If the bottom is absorptive rather than rigid, an additional diffusion term appears in the matrix partial differential equation governing the coherence. Diffusion along the transverse separation axis then prevents the creation of a Dirac delta function as the in-line propagation distance increases without limit. This diffusion occurs whenever the imaginary part of the horizontal wave-number component for a particular mode is non-zero. Some graphical outputs depicting the coherence propagation for an absorptive bottom are presented in this paper. [Work supported by ONR Code 321.]

1:45

4pUW2. Singular value decomposition for normal-mode computations. Ronald T. Kessel (School of Earth and Ocean Sci., Univ. of Victoria, P.O. Box 3055, Victoria, BC V8W 3P6, Canada) and Trevor W. Dawson (Univ. of Victoria, Victoria, BC V8W 3P6, Canada)

Mode models of sound propagation in layered media begin with a search for modes in the complex wave-number plane, followed by computation of the vertical mode functions and mode superposition. Sound propagation modelers would like to search for modes ever more efficiently and thoroughly. The search amounts to the mathematical problem of finding the roots (zeros) of a complex-valued function—a problem that defies definitive solution. Rather than reducing the numerical test for modes (roots) to a single characteristic equation, singular value decomposition (SVD) can be used to test when the global matrix describing the propagation of plane waves through the layered media is singular, which is an equivalent condition for a mode. The advantages of this approach are that it is numerical stable, it automatically gives the vertical mode functions (although un-normalized), it can detect duct and interface modes at any depth, and it can identify and resolve close-mode pairs (double roots) and their antisymmetric mode functions occurring in weakly coupled sound channels.

2:00

4pUW3. Mode excitation for a source on a rough, elastic, sloping bottom. Arthur B. Baggeroer, Henrik Schmidt, and Brian Sperry (Dept. of Ocean Eng., MIT, Cambridge, MA 02139)

In the ATOC program, the source was sited on a slope at a depth of the SOFAR axis so that the signal would excite the axial modes efficiently for long-range propagation. One of the important issues was the elevation of the source off the seafloor. Buoying it on a mooring into the water column presented significantly more engineering challenges than simply resting it on a platform on the seafloor. Here, the impact upon the mode excitation spectrum of the source depth on a rough, elastic, sloping bottom is examined using the new range-dependent OASES code. This code matches vertical wave numbers across range-independent sectors, which may include interface roughness, as it steps out in range. It has been very successful in modeling the related problem of the effect of epicenter depth for T-phase excitation on a sloping bottom [Sperry *et al.*, J. Acoust. Soc. Am. **100**, 2641(A) (1996)]. A Munk sound-speed profile is assumed with a minimum at 1000 m and a sloping bottom out to a range of 30 km after which the bottom depth is constant. The field is projected onto the depth-dependent modes to determine their excitation. Also, the group delays are examined to determine the impact on travel times. Initial results suggest that even a modest (100 m) elevation concentrates the power in the low-order modes. In addition to the source elevation, the slope gradient and roughness and the elastic properties of the bottom are important parameters. [Work supported by ONR.]

2:15

4pUW4. Normal-mode analysis for signal fluctuations in the Yellow Sea. Renhe Zhang, Zhenge Sun, Liewei Sha, Liangying Lei, Longsheng Hao, Nan Sun, Fenghua Li (Natl. Lab. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC), Ji-Xun Zhou, Peter H. Rogers, and Gary W. Caille (Georgia Inst. of Technol., Atlanta, GA 30332)

This paper presents a normal mode approach for analyzing signal fluctuation in shallow water with a thermocline which fluctuates due to tides and internal waves. In the Yellow Sea 1996 experiment, variations of thermocline structure and internal waves were observed by CTD and thermistor chains. The cw and FM acoustic signals were measured over a 12-h period with two 16-hydrophone vertical arrays deployed 3.11 and 10.32 km from the source. The cw pulses were used to observe fluctuation at four specified frequencies. Amplitude fluctuations of more than 20 dB were observed. The FM signals were used to isolate individual normal modes by pulse compression and mode filtering. The influence of internal waves on the normal-mode structure and signal fluctuation was examined. It was found that the variation of the thermocline was significant in determining acoustic signal fluctuation. The normal-mode structure and dispersion relations in the channel depend strongly on the thermocline. Variation of the thermocline thus causes phase and group velocity to become variable. This leads to signal fluctuation, since the amplitude of the received signal is determined by interference between arriving modes. The study also showed that the amplitude fluctuation is weaker at shorter range.

4pUW5. Sound-field sensitivity to geoacoustic parameters in range-dependent media using normal modes. Ronald T. Kessel (School of Earth and Ocean Sci., Univ. of Victoria, P.O. Box 3055, Victoria, BC V8W 3P6, Canada)

Sensitivity in sound propagation modeling is a measure of the ability of environmental parameters to affect the sound field. The most sensitive parameters are often considered most important because errors or uncertainties in their values are among the most likely sources of error in the model predictions. There are many ways to define the measure of sensitivity. It generally depends on the sound field itself, on the energy flow through regions where the geoacoustic parameters under consideration lie, and on the way the field is monitored by receivers. But it is useful to define sensitivity without reference to a particular field or source–receiver configuration, as if it were a parameter of the environment alone. Then conclusions drawn from a sensitivity study apply as much as possible to all source–receiver configurations, and for both inverse and forward modeling. Such a measure has been derived using an adiabatic mode model for range-dependent media, exploiting the fact that the normal modes of vibration embody the essential properties of the field without reference to a particular source–receiver configuration. The new measure is demonstrated for range-dependent media using computer simulations.

4pUW6. Acoustic mode decoupling via environmentally adaptive coordinates. Peter C. Mignerey (Acoust. Div. 7120, Naval Res. Lab., Washington DC 20375-5350, mignerey@nrl.navy.mil)

A normal mode expansion is used to relate the coupling of energy between modes to horizontal derivatives of the eigenfunctions. A curvilinear coordinate system is constructed using a WKB approximation for a reference eigenfunction such that horizontal derivatives causing coupling vanish. The curvilinear coordinate system is defined such that constant surfaces of the vertical wave number integral coincide with surfaces of constant curvilinear depth. Thus any particular zero of an eigenfunction lies on a constant depth surface in the curvilinear system. The horizontal derivatives of all other modes depend on a ratio of vertical wave numbers with the reference mode. Nearly constant horizontal behavior of this wave number ratio supports the decoupling of all modes in the curvilinear system. Such coordinate systems naturally adapt to the environment even with realistic sound-speed profiles that include discontinuities as well as sloping bottom bathymetry. They provide an adiabatic normal-mode basis for constructing acoustic models in the fully three-dimensional environments of continental shelf regions. Environmentally adaptive coordinate systems also provide a rational basis for interpolating three-dimensional environmental data.

4pUW7. The generalized eigenvalue problem, warping matrices, and the transformation of an isovelocity environment to a variable velocity environment. M. F. Werby (Naval Res. Lab., Code 7181, Stennis Space Center, MS 39529) and N. A. Sidorovskaia (Univ. of New Orleans, New Orleans, LA 70148)

The calculation of eigenvalues for a layered isovelocity oceanic environment and associated vertical modes is very quick and easy to do. A method of transforming the isovelocity problem to a variable velocity problem, by means of a matrix transformation based on the theory of generalized eigenvalues, is described. Along with modern transformation techniques, the rapid reconstruction of variable velocity solutions is demonstrated. It is possible to determine the mode spectrum extremely rapidly by first starting out with those associated with a reference isovelocity eigenvalue set, and then by using an iteration procedure based on integral transform methods, the Householder transformation, polar decomposition theorems, shifting, and deflation methods in linear algebra to converge to

the set of eigenvalues rapidly. A first iteration may take a fraction of a second to calculate on a PC, while full iteration takes on the order of a second to calculate 100 eigenvalues. [Work sponsored by ONR, the Naval Res. Lab., and the Univ. of New Orleans.]

4pUW8. An adaptive coupled range-dependent normal-mode method. M. F. Werby (Naval Res. Lab., Code 7181, Stennis Space Center, MS 39529) and N. A. Sidorovskaia (Univ. of New Orleans, LA 70148)

Based on the warping method and the generalized eigenvalue problem the range-independent solution is extended to that of range dependence. In this method, coupling in range is analytical and may be either one way or two way. Since the range steps are not dependent on finite-difference schemes, the optimal range step at each point is determined. This leads to two advantages: (1) There is no need for input of range step; and (2) the optimal range step is predetermined by several conditions, thus avoiding overkill and reruns to determine stability. Comparisons with other range-dependent methods are presented. [Work sponsored by ONR, the Naval Res. Lab., and the Univ. of New Orleans.]

4pUW9. Analytic implementation of a parabolic equation solution. Dalcio K. Dacol and Michael D. Collins (Naval Res. Lab., Washington, DC 20375, dacol@abyss.nrl.navy.mil)

There are two basic approaches for discretizing the differential equations that are solved in propagation models: (1) Approximate the differential operator with a finite-difference formula; and (2) approximate the medium in terms of layers that may be treated analytically and solve for a set of coefficients. The analytic approach is more efficient for problems involving a small number of layers. Both approaches have been implemented for normal mode and wave-number integration models. The finite-difference approach has been implemented for parabolic equation models. This paper describes an analytic implementation of a parabolic equation model that is efficient for geoacoustic inversion problems [R. J. Cederberg and M. D. Collins, "Application of an improved self-starter to geoacoustic inversion," *IEEE J. Ocean. Eng.* (in press)]. Using the finite-difference implementation, the self-starter can solve nonlinear inversion problems in minutes on a workstation (techniques that were developed previously require hours of run time). The analytic implementation reduces run times to seconds. [Work supported by ONR.]

4pUW10. Energy-conserving spectral solutions. Michael D. Collins (Naval Res. Lab., Washington, DC 20375, collins@ram.nrl.navy.mil), Henrik Schmidt (MIT, Cambridge, MA 02139), and William L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180)

Energy conservation is an important issue in range-dependent propagation modeling. Energy-conserving coupled mode and parabolic equation solutions have been developed [J. Acoust. Soc. Am. **89**, 1058–1075 (1991)]. This presentation will describe an energy-conserving solution based on the wave-number spectrum. Range dependence is handled by dividing the medium into a sequence of range-independent regions and enforcing the complete energy-conservation condition [J. Acoust. Soc. Am. **94**, 975–982 (1993)] at the vertical interfaces between regions. This

condition involves the square root of the horizontal wave number in the forcing function of the spectral equation and in the spectral integral. This approach has been implemented and tested for both fluid and elastic media. [Work supported by ONR.]

4:15

4pUW11. Analysis of selected data from a recent Navy exercise. Stanley A. Chin-Bing, David B. King, Guy V. Norton (Naval Res. Lab., Stennis Space Center, MS 39529-5004), and Jorge C. Novarini (Planning Systems, Inc., Long Beach, MS 39560-9702)

Several ocean acoustic computer models were used to analyze shallow-water, higher-frequency (around 4-kHz) acoustic data taken during a recent Navy exercise. This effort enlisted models (EFEPE-CM, FEPE, ASTRAL, RAYMODE) and computer resources (Cray C90, SGI Challenge-8, 386 desktop PC) that spanned the RDT&E modeling spectrum, from basic research to Fleet operations. The research models were used to identify and prioritize dominant physical mechanisms. Once these were known, errors and weaknesses in the operational models were easily identified and corrected. The result was consistent with predictions from the entire suite of models that compared favorably with data. Advantages were gained in computational speed and accuracy. An acceptable shallow-water, high-frequency, acoustic prediction was obtained in 2 s using a Fleet operational model (ASTRAL) and a 386 desktop PC computer. Two important lessons were relearned: (1) Current ocean acoustic research models include the necessary physics to make accurate high-frequency predictions in shallow-water environments, even when realistic sea-surface conditions dominate; and (2) the more physics included in the models, the less knowledgeable need be the model operators. Details and examples will be presented. [Work supported by ONR/NRL, MMoDS, AUAMP, and DoD HPC.]

4:30

4pUW12. Parabolic equation models for anisotropic 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)

Because of the current interest in shallow water, propagation model capabilities have been developed for elastic, poro-elastic, and poro-acoustic sediments. These models incorporate depth and range heterogeneities in the sediments, which are assumed to be spatially isotropic. However, many shallow environments have a layered structure which often is more appropriately characterized as transversely isotropic. This feature arises from deposition and layering processes influenced by gravity, for example. Transversely isotropic sediments have elastic properties with considerable variations perpendicular to their natural plane, in which relatively little variation occurs. Efficient and accurate two-dimensional PE models are available for a variety of spatially isotropic sediments, and generalizations are needed for transversely isotropic cases. Our initial two-dimensional development assumes that the plane of the sediment is approximately horizontal. The PE formulation is presented for an elastic sediment and then extended to poro-elastic cases. An initial high-order implementation will be described. [Work supported by ONR.]

4:45

4pUW13. Boundary conditions for finite-difference PE solvers. David Yevick (Dept. of Elec. Eng., Queen's Univ., Kingston, ON K7L 3N6, Canada) and David J. Thomson (Defence Res. Establishment Atlantic, Victoria, BC V9A 7N2, Canada)

In theoretical/numerical models of underwater sound propagation, a downgoing radiation condition is usually imposed on the acoustic field as $z \rightarrow \infty$. For most implementations of parabolic equation (PE) solvers, this condition is approximated by appending an absorbing layer to the compu-

tational mesh. Such a layer acts to attenuate any energy that reaches the grid boundary before the waves undergo reflection and return to the ocean region. As shown by Papadakis [J. Acoust. Soc. Am. **92**, 2030–2038 (1992)], this approximate treatment can be replaced with a nonlocal boundary condition (NLBC) that *exactly* transforms the semi-infinite PE problem to an equivalent one in a bounded domain. Papadakis' method involves evaluating a spectral (wave number) integral of a singular kernel that is inversely proportional to the impedance of the subbottom medium. In this paper, an alternate procedure is described for obtaining NLBCs directly from the z -space Crank–Nicolson formulations of both the Tappert and Claerbout PEs. Formulas for the boundary field at range $r + \Delta r$ are derived in terms of the known field along $0 \rightarrow r$ by expanding the appropriate "vertical wave number" operator in powers of $\mathcal{R} = \exp(-\Delta r \partial_r)$ and applying $\mathcal{R}^j \psi(r, z) = \psi(r - j\Delta r, z)$. The effectiveness of these NLBCs is compared to Berenger's matched-layer technique [J. Comp. Phys. **114**, 185–200 (1994)] for several numerical examples relevant to one-way underwater sound propagation.

5:00

4pUW14. The PSTD algorithm: A time-domain method combining the pseudospectral technique and perfectly matched layers. Qing-Huo Liu (Klipsch School of Elec. and Comput. Eng., New Mexico State Univ., Las Cruces, NM 88003)

Simulations of underwater acoustic wave propagation require solutions for large-scale multidimensional problems. The traditional finite-difference time-domain (FDTD) method needs a fine discretization of 8–20 cells per wavelength in order to give accurate results. On the other hand, the pseudospectral method, even though efficient, suffers from the wraparound effect due to the use of discrete Fourier transform. This effect severely limits the applications of the pseudospectral method to large-scale problems. Hence, a popular way of simulating underwater acoustic wave propagation is to use the parabolic equation (PE) methods which neglect back-scattering. In this work, Berenger's perfectly matched layers (PML) are used in the pseudospectral method to eliminate the wraparound effect. To achieve a high accuracy, this method requires only two cells per wavelength which is dictated by the Nyquist sampling theorem. As a result, it can solve at least 64 times larger 3-D problems than the FDTD method with the same requirement in computer memory and CPU time. Hence, full-wave solutions of long-range underwater acoustic wave propagation become possible. Numerous simulations show the superiority of the PSTD method for large-scale problems. [Work supported by a Presidential Early Career Award for Scientists and Engineers through EPA and by Sandia National Laboratories.]

5:15

4pUW15. Perfectly matched layers for acoustic waves in viscous media: Applications to underwater acoustics. Qing-Huo Liu and Jianping Tao (Klipsch School of Elec. and Comput. Eng., New Mexico State Univ., Las Cruces, NM 88003)

Berenger's perfectly matched layer (PML) has recently been proved to exist for elastodynamic equations [Chew and Liu, Schlumberger-Doll-Res. Rep. (1995)]. This fictitious material absorbs all waves with an arbitrary incident angle and an arbitrary frequency without giving rise to any reflections. Therefore, when used as an absorbing boundary condition (ABC) at the computational edge in a finite-difference method, the PML provides orders of magnitude higher absorption than other existing ABCs. In this work, the PML is further extended as an ABC for a finite-difference simulation of acoustic waves in viscous media. For such attenuative media, an additional term involving the time-integrated wavefield is introduced to

account for the coupling between the attenuation from the PML and the normal viscous attenuation. This ABC is highly effective in absorbing outgoing waves at the computational edge even when a dipping interface intersects the outer boundary. This new material ABC is ideal for parallelization on multiprocessor computers. The algorithm is validated by ana-

lytical and numerical solutions. Various two- and three-dimensional numerical simulations will be shown to demonstrate the applications in underwater acoustics and other areas. [Work supported by a Presidential Early Career Award for Scientists and Engineers through EPA and by Sandia National Laboratories.]

THURSDAY AFTERNOON, 19 JUNE 1997

ROOM H, 1:30 TO 3:00 P.M.

Meeting of Accredited Standards Committee S3 on Bioacoustics

to be held jointly with the

U.S. Technical Advisory Group (TAG) Meetings for ISO/TC 43 Acoustics, IEC/TC 29 Electroacoustics, and ISO/TC 108/SC4 Human Exposure to Mechanical Vibration and Shock

T. A. Frank, Chair S3

Pennsylvania State University, Speech and Hearing Clinic, 110 Moore Building, University Park, Pennsylvania 16802

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 4005, Champaign, Illinois 61820

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, 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. The international activities in ISO/TC 43 Acoustics, and IEC/TC 29 Electroacoustics, and ISO/TC 108/SC4 Human Exposure to Mechanical Vibration and Shock, will also be discussed. The Chairs of the U.S. Technical Advisory Groups for ISO/TC 43 (P. D. Schomer), IEC/TC 29 (V. Nedzelnitsky), and ISO/TC 108/SC4 (J. Erdreich) will report on current activities of these international Technical Committees and Subcommittees.

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.

4p THU. PM

Meeting of Accredited Standards Committee S1 on Acoustics

to be held jointly with the

U.S. Technical Advisory Group for ISO/TC 43 Acoustics and IEC/TC 29 Electroacoustics

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, National Research Council, Ottawa, Ontario K1A 0R6, Canada

P. D. Schomer, Chair, U. S. Technical Advisory Group (TAG) for ISO/TC 43, Acoustics

U. S. CERL, P.O. Box 4005, Champaign, Illinois 61820

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, 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. The international activities in ISO/TC 43 Acoustics, and IEC/TC 29 Electroacoustics, will also be discussed. The chairs of the respective U.S. Technical Advisory Groups for ISO/TC 43 (P. D. Schomer) and IEC/TC 29 (V. Nedzelnitsky), will report on current activities of these international Technical Committees.

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 5aBV

Bioresponse to Vibration/Biomedical Ultrasound: Human Vibration Exposure

John Erdreich, Cochair

Ostergaard Acoustical Associates, 100 Executive Drive, West Orange, New Jersey 07052

Henning E. von Gierke, Cochair

1325 Meadow Lane, Yellow Springs, Ohio 45387

Chair's Introduction—8:55

Invited Papers

9:00

5aBV1. Human vibration standards—Do we ask the right questions? John Erdreich (Ostergaard Acoust. Assoc., 100 Executive Dr., West Orange, NJ 07052) and Linda S. Erdreich (Bailey Res. Assoc.)

Standards for vibration exposure promulgate different limits for hand–arm vibration (HAV) and for whole-body vibration (WBV). The discrepancy between the two should be explainable on some rational basis [Griffin, *Handbook of Human Vibration* (1990)], but there is no obvious difference in the mechanism of pathogenesis to the skin at each of the sites. The difference may arise because the endpoints chosen as the basis of each of the standards are different. One approach to estimating dose response [Jarabek, *Toxicol. Lett.* **79**, 171–184 (1995)], characterizes the steps in developing a model as proceeding from a protective to a predictive paradigm as data progress from qualitative to quantitative. To develop vibration standards along this continuum, certain data must be developed: Among these are identification of performance effects of HAV, identification of long-term effects of WBV, and clarification of the impedance of the affected biological structures. Some of these areas have been addressed in recent revisions of ISO TC108/SC4 standards. Other activities will be proposed to further the evolution of human biodynamic assessment.

9:30

5aBV2. A method of health hazard assessment for human exposure to repeated mechanical shocks. James Morrison, Jordan Nicol (School of Kinesiology, Simon Fraser Univ., Burnaby, BC V5A 1S6, Canada, jmorriso@sfu.ca), Daniel Robinson, George Roddan, Julie Springer, Steven Martin (B.C. Res., Inc., Vancouver, BC V6S 2L2, Canada), and John Albano (U.S. Army Aeromedical Res. Lab., Fort Rucker, AL 36362-0577)

This study describes a health hazard assessment (HHA) method for evaluating human exposures to repeated shocks in tactical ground vehicles. The method predicts the risk of injury to the occupants of a vehicle from the accelerations measured at the seat. The HHA will identify both acute and chronic health risks resulting from either a single exposure, or from multiple exposures to travel over rough terrain. The HHA is based on existing models, human injury data, and experimental data obtained from volunteers subjected to a range of shock profiles and to prolonged repeated shock exposures. The HHA consists of three components: (1) dynamic response models which predict seat-to-spine transmission of acceleration; (2) a dose model for exposure to repeated shocks based on fatigue failure theory and subjective tolerance data; and (3) an injury risk model based on the cumulative probability of failure. A biomechanical model is also described which analyzes spinal compression forces in response to shock. This model uses human experimental data to provide an independent evaluation of the HHA. The components of the HHA are outlined and some test results are presented. [Work supported by US Army Med. Res. & Dev. Command, Contract No. DAMD17-91-C-1115.]

10:00

5aBV3. Health effects in the upper extremities of persons operating low-vibration power tools. Anthony J. Brammer (Natl. Res. Council of Canada, Ottawa, Canada), P. Sutinen (Orton Rehabilitation Ctr., Helsinki, Finland), K. Koskimies (Rheumatoid Foundation Hospital, Helsinki, Finland), I. Pyykkö (Karolinska Hospital), and J. Starck (Inst. of Occupational Health)

A group of manual workers ($N=18$) who operate low-vibration power tools ($|a_{hw}|_{\max} < 2 \text{ m s}^{-2}$) has been examined for objective evidence of work-related changes in hand and arm function. The subjects were drawn from and selected to be representative of a cohort of chain saw operators ($N=124$) whose health has been followed prospectively, 30% of whom reported experiencing numbness in the upper extremities. A method for detecting a statistically sufficient change in tactile acuity from measurements of mechanoreceptor-specific vibrotactile perception thresholds conducted over a 5-yr interval has been developed, together with any shift in acuity from the value occurring in healthy persons. A progressive deterioration in acuity was found in $\frac{12}{36}$ hands, which is predicted to result in abnormal acuity in 39% of the group after a further 5 yr of work. The subjects predicted to be most severely affected currently exhibit reduced hand grip, and reduced hand and arm strength.

5aBV4. Ultra-low-frequency ride motions and kinetosis. E. D. Sussman, Robert DiSario, Peter Mengert, and John K. Pollard (Volpe Natl. Transportation Systems Ctr. DTS-45, 55 Broadway, Cambridge, MA 02142)

Proposed ultra-high-speed ground transportation systems, such as Maglev, may have motion characteristics affecting passenger comfort different from anything previously experienced in ground transportation. These motions not only include isolated vertical maneuvers but also repeated motions. For these repeated motions the power spectra is as important as the magnitude in predicting kinetosis. Segments of the New York State Thruway were simulated by flying an airliner through a series of several dozen roll maneuvers based on nine combinations of speed and curvature. The combinations were also simulated using a moving base simulator which provided a geometrically accurate "out of the window view." Analysis of the data lead to the following conclusions: (1) More than 95% of the public would accept isolated maneuvers involving bank angles up to 37° and roll rates up to 7°/s. (2) The majority found the plane simulation comfortable and felt no motion sickness. Differences between subjects' appear to have been greater than the physical effects of the differences in bank angle. (3) Cumulative dosage and duration of exposure showed significant correlation with the motion-sickness ratings procedure developed based on the work of M. J. Griffin and British Standard 6841:1987.

Contributed Paper

11:00

5aBV5. Just-noticeable differences in vertical vibration for seated subjects. William J. Pielemeier, Norman C. Otto, Raymond C. Meier, Jr., and Vadivelu Jeyabalan (Ford Motor Co., P.O. Box 2053/MD2122, Dearborn, MI 48121, w.pielemeier@ieee.org)

Just-noticeable differences in bandlimited vertical vibration were studied for subjects sitting on automotive seats. Stimuli consisted of octave band frozen Gaussian noise centered at 4, 8, and 16 Hz. Two-interval forced-choice paired comparisons were used, with a reference level of 8 mg. The level of the alternative stimulus in the pairs varied from 8.25 to 10

mg. Stimulus intensity was measured with a seat pad accelerometer. Stimulus durations of 4 s plus 1/2-s tapers were used for all frequency bands, with 1/2 s between stimuli. Sets of trials with 2-s durations at 16 Hz were done as a test of duration effects. Three subjects were trained with feedback until performance stabilized. Then 200–250 trials were performed in blocks of 25 or 50 for each of four alternatives at each frequency and each subject. Thresholds determined from the psychometric functions ranged from 0.6 to 1.8 mg, with most between 0.6 and 1.2. Little frequency dependence was evidenced over all frequency bands, and little duration effect was seen between 2- and 4-s stimuli at 16 Hz.

FRIDAY MORNING, 20 JUNE 1997

ROOM E, 8:00 A.M. TO 12:00 NOON

Session 5aEA

Engineering Acoustics and Noise: General Engineering Acoustics and Noise

Peter H. Rogers, Chair

G. W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0405

Chair's Introduction—8:00

Contributed Papers

8:05

5aEA1. Modal solution for a two-dimensional waveguide having a locally reacting wall. Thomas M. Logan, Jerry H. Ginsberg, and Peter H. Rogers (G. W. Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405)

Consider a two-dimensional waveguide consisting of a rigid wall and a parallel wall having uniform specific impedance. When the impedance is purely imaginary, the characteristic equation is real. In the springlike case, the eigenvalue of the fundamental mode (defined as the mode that varies most slowly in the transverse direction) is imaginary, whereas all other eigenvalues are real. This type of mode does not exist if the impedance is masslike. Morse and Ingard [*Theoretical Acoustics* (McGraw-Hill, New York, 1968), p. 496] fail to recognize the differing behavior of the fundamental mode for either type of lossless impedance, and subsequently no one seems to have rectified this oversight. The fundamental mode in the springlike case behaves much like the plane mode in a hard-walled waveguide in that it has no cutoff frequency, but the phase speed of this mode is subsonic at all frequencies. Introducing a resistive part to the impedance, corresponding to dissipation at the wall, moves the respective eigenvalues off the real or imaginary axis. The presentation will discuss the phase and group speeds, and the transverse mode shapes, in ideal and lossy cases.

8:20

5aEA2. Comparison of directivity patterns measured in the long-line hydrophone calibrator to predicted data. L. D. Luker and M. K. Beason (Underwater Sound Reference Div., Naval Undersea Warfare Ctr. Div. Newport, P.O. Box 568337, Orlando, FL 32856-8337)

This presentation describes the results of the first calibration of line arrays in the long-line hydrophone calibrator (LLHC). The LLHC is a water-filled traveling-wave tube that simulates a plane-wave pressure field arriving from any bearing. It provides an alternative to open-water sensitivity calibrations and measures both sensitivity and directivity patterns over the array's operational environment. This is accomplished by measuring the electroacoustic transfer matrix between hydrophones and projectors that are placed within a steel pipe with an inner diameter of approximately 30 cm. This transfer matrix is then used to compute projector drives suitable for creating the desired sound field. The current LLHC configuration is capable of calibrating line arrays up to 48-m long at pressures from 0 to 6.9 MPa and frequencies from 5 to 700 Hz. The line arrays that were calibrated were approximately 43-m long containing up to 16 sensors. The calibrations were performed at pressures up to 2.7 MPa and

frequencies ranging from 10 to 800 Hz. Measured directivity patterns for various frequencies will be presented and compared to predicted data. [Work supported by NAVSEA.]

8:35

5aEA3. Modeling perforates. N. S. Dickey and A. Selamet (Dept. of Mech. Eng. and Ctr. for Automotive Res., Ohio State Univ., 206 W. 18th Ave., Columbus, OH 43210-1107)

A number of computational and analytical tools are available to model perforated elements by treating the perforate interface as a distribution of localized reacting sources. Since the unsteady flow through an orifice exhibits a wide range of behavior, the results of such analyses depend strongly on the perforate model. Frequency-domain techniques are the most common approach for perforated element analysis and, correspondingly, nearly all available experimental data are presented in terms of frequency-domain parameters. Nonlinear behavior is generally addressed through a nonlinear orifice impedance, defined at a single frequency. While the frequency-domain approach is clearly justified for linear physics, or for nonlinear cases with a single driving frequency (the frequency is a natural scaling parameter), the use of this technique for nonlinear behavior caused by multifrequency disturbances may become questionable. The present study investigates unsteady flow through a circular orifice from a time-domain perspective. Development of an empirical perforate model based on relationships obtained from time-domain experimental data is discussed. The orifice model is incorporated into a time-domain computational scheme for perforated tube silencers and numerical predictions are compared to experimental data. The relative benefits and limitations of the present approach are addressed.

8:50

5aEA4. Acoustic frequency spectrum shifts in failing artificial hearts. Asim Choudhri, Bijan Salehizadeh, Howard Levin, and Mehmet Oz (Columbia Univ., Dept. of Surgery, New York, NY)

Left ventricular assist devices (LVAD) are implantable mechanical blood pumps providing circulatory assistance for patients in heart failure, and are currently being investigated as a long-term alternative to cardiac transplantation. While effective in short term use as a bridge to transplantation, the durability of the device for extended use has yet to be demonstrated. To prospectively follow the mechanical status of the devices, bi-weekly transabdominal acoustic readings from a standard microphone and accelerometer are obtained. Time-domain analysis of the recordings allows the timing of ejection (device systole) and passive filling (device diastole) to be measured. Ejection duration was measured at 240 ms in 11 new devices and exceeded 300 ms in two aging devices which subsequently failed within 3 weeks of measurement. Frequency-domain analysis showed the presence of a 230-Hz signal in both failing devices, possibly indicative of bearing wear within the motor. In addition, a 50-Hz signal present in all properly functioning devices was missing in one of the devices which failed. Acoustical analysis will play a vital role in the continuing care of LVAD patients and will allow a more rapid and noninvasive method of assessing device status.

9:05

5aEA5. Acoustic microscopy for characterization of high-temperature superconducting tape. Chiaki Miyasaka, Chris Cobucci, and Bernhard Tittmann (Eng. Sci. and Mech. Dept., Penn State Univ., University Park, PA 16804)

Although material scientists constantly discover superconducting compounds with higher critical temperatures (T_c 's), manufacturing of the high-temperature superconductors (HTS) remains a problem and long lengths (>1 mile) have yet to be produced. In an effort to produce long length superconductors, manufacturing steps for HTS tape production have been critically looked at to find their effects in producing tape with the desired characteristics. In support of determining superconducting tape character-

istics, acoustic microscopy offers the potential for internal microstructural material characterization. This research will ultimately support in-process monitoring of HTS manufacturing as part of an advanced sensing system to determine the presence of defects and/or the effects of process variables on the HTS tape. This presentation will overview scanning acoustic microscopy and present images of HTS tape at several frequencies ranging from 50 to 500 MHz. The results clearly demonstrate the feasibility of determining the Ag/ceramic interface location and the general integrity of the constituents.

9:20

5aEA6. Process monitoring of 2-D carbon-carbon composites with passive ultrasonics and mass spectroscopy. C. Eric Yen and Bernhard R. Tittmann (Penn State Univ., 227 Hammond Bldg., University Park, PA 16802)

2-D woven carbon-carbon composites are a class of advanced structure composite materials which can maintain good mechanical property and high stiffness constants at elevated temperatures. The high failure rate during the initial charring process has been the major obstacle in attaining increased popularity of these materials. This paper reports the use of passive ultrasonics and mass spectroscopy techniques to monitor the charring process of 2-D woven carbon-carbon composites in order to accurately and sensitively detect *in situ* the development of macro-porosity. The passive ultrasonic technique with high-temperature waveguide sensors is demonstrated to effectively detect the acoustic emission (AE) associated with the microcracking events during the charring process. The mass spectrometry (MS) technique together with a unique "sniffing" tube is shown to detect major gas evolution, such as water, methane, and hydrogen during the initial pyrolysis. The gas information is important to understand the nature of catastrophic delamination. Post-process material characterizations, such as thickness shrinkage, weight loss, and microscopic characterization, were conducted to provide information about the physical changes due to the carbonization process and proved to correlate well with the AE and MS data.

9:35

5aEA7. Excitation and detection of Lamb waves by cylindrical ultrasonic waveguides. C. Eric Yen and Bernhard R. Tittmann (Penn State Univ., 227 Hammond Bldg., University Park, PA 16803)

In this paper a novel technique to generate and detect Lamb waves by use of cylindrical ultrasonic waveguides with off-shelf ultrasound transducers attached to the ends will be investigated. Due to the impedance mismatch between the transducer and the cylindrical rod waveguide, the energy transmission from an off-shelf transducer to the rod is not very efficient. An unique technique, which makes use of solid conical waveguide as an impedance transformer between the transducer and the rod waveguide, was investigated to solve the problem. Insertion loss measurements for two different waveguide systems, i.e., transducer rod and transducer-conical waveguide rod, were conducted to verify the improvement of the acoustic energy transmission of a transducer-conical waveguide-rod combination. Transient plate waves excited by Nelson-Hsu pencil lead breaks were detected and recorded on an aluminum plate by the as-developed waveguides. Dispersion analysis was conducted to verify the plate modes excited by pencil lead breaks at different locations of the plate. Excitation of Lamb waves was also exercised by use of the waveguides. Experimental measurement of the mode velocity was compared with the theoretical predication.

9:50

5aEA8. Surface stress waves when the elastic constants are functions of depth. Dov Hazony (Dept. of Elec. Eng. and Appl. Phys., Case Western Reserve Univ., Cleveland, OH, 44106)

Of concern are distortionless surface stress waves propagating in a medium which may be nonuniform relative to depth. A similar problem was posed by Burke *et al.* [Q. Appl. Math. 183-194 (1976)] where a unidimensional electrical ladder network model, nonuniform along the

principle axis, was used in the investigation. Presently, a two-dimensional development is given based on Hooke's law and Newton's law leading to the well-known Rayleigh-waves formulation in a uniform medium. It will be seen that broadband high-pass propagation modes exist along the principle surface. As these propagate along the surface, they are accompanied by a phase change, linear with distance, in the frequency domain, and an associated symmetry change, periodic with distance, in the time domain. Applications may be found in high-frequency acousto-optics, *in situ* ultrasonic monitoring of elongated bodies, high-frequency SAW filters, and whenever surface acoustic waves are employed in an environment of high precision or relatively large distances. [Work supported by ONR.]

10:05

5aEA9. A numerical solution of acoustic scattering by an aircraft wake vortex. Chingwei M. Shieh and Philip J. Morris (Dept. of Aerosp. Eng., Penn State Univ., 233 Hammond Bldg., University Park, PA 16802)

Monostatic and bistatic acoustic scattering has been used in studies of the characteristics and motions of the atmosphere. This technique has been successfully implemented in the Monostatic Acoustic Vortex Sensing System (MAVSS) for the measurement of the characteristics of aircraft wake vortices during landing and take-off. In the present study, calculations are performed to simulate the MAVSS. The results of this study are giving information about the acoustic scattering by the vortex, as well as providing a controlled test bed for the instrument's operation. Since the Doppler shifts, characteristic of the vortex velocity, are dependent on the incoherent scattering of the acoustic waves by turbulence advected by the vortex, a statistical model is proposed that describes the turbulence imbedded in the vortex. The simulations are performed using high-order accuracy finite-difference algorithms that have been developed for computational aeroacoustic (CAA) and wave propagation problems.

10:20–10:30 Break

10:30

5aEA10. Rotorcraft noise model. Michael J. Lucas (Wyle Labs., 2001 Jefferson Davis Hwy., Ste. 701, Arlington, VA 22202) and Michael A. Marcolini (NASA Langley Res. Ctr., Hampton, VA 23665)

The rotorcraft noise model (RNM) is a multiplatform computer program being developed for NASA-Langley Research Center which calculates sound levels at receiver positions on the ground from helicopter and tiltrotor operations. RNM calculates levels at positions on a uniform grid or a specific defined locations. The basic computational model written in FORTRAN calculates maximum A-weighted sound level, sound exposure level (overall, C- and A-weighted), and effective perceived noise level (EPNL). Contained within the program are sound propagation and EPNL algorithms that are the same as, or equivalent to, corresponding algorithms in NASA's Aircraft Noise Prediction Program (ANOPP). Acoustic properties of the noise source are defined by a set of sound-pressure noise spheres. One set of noise spheres provides broadband data in the form of one-third octave band sound-pressure levels. The other set of noise spheres provides narrow-band data in the form of pure-tone sound-pressure levels and phase. RNM outputs a file that is input to NMPLLOT, a noise contouring/smoothing program used by the Department of Defense (DOD) and other government agencies. This paper will describe the basic element of the noise model and compare the noise models prediction with XV-15 noise data collected by NASA-Langley.

10:45

5aEA11. The radiated noise from realistic airfoils encountering vortical gusts. David P. Lockard and Philip J. Morris (Dept. of Aerosp. Eng., Penn State Univ., 233 Hammond Bldg., University Park, PA 16802)

The long-term objective of the current research is to use computational aeroacoustics (CAA) methodology and parallel computers to increase the understanding of broadband blade noise. In a previous paper by the authors [D. P. Lockard and P. J. Morris, AIAA-97-0286 (1997)] the effect of

thickness on the sound radiated by airfoils encountering incompressible, vortical gusts was investigated. This work extends those calculations by examining the effect of camber on the radiated noise. A two-dimensional, Navier-Stokes code, implemented using the message passing library and FORTRAN 90 on the IBM SP2, is used to perform the calculations. The code uses Runge-Kutta time integration and seven-point, central-difference, dispersion-relation-preserving spatial operators. A multigrid method is used to obtain converged steady-state solutions before the gust is introduced in a source region inside the domain. [Work supported by NASA Langley Research Center.]

11:00

5aEA12. Application of a parallel three-dimensional nonlinear method to the acoustics of flow-over cavities. Jingmei Liu and Lyle N. Long (Dept. of Aerosp. Eng., Penn State Univ., University Park, PA 16802, lnl@psu.edu)

A three-dimensional computational aeroacoustics (CAA) method combined with advanced parallel computing techniques is applied to the simulation of the acoustic response of flow-over cavities (Helmholtz resonator). Direct numerical simulation is conducted by solving a newly developed nonlinear disturbance equation [Morris *et al.*, Comp. Phys. (in press)] which consists of linear and nonlinear fluctuation terms, spatially varying coefficients that depend on the mean-flow properties, and a mean-flow source term that is independent of the fluctuations. The mean-flow component is obtained from a numerical solution of the Reynolds-averaged Navier-Stokes equations with a turbulence model by taking advantage of traditional, robust CFD algorithms with good convergence characteristics. The capturing of time-accurate characteristics of the unsteady flow field and the the nondispersive and nondissipative characteristics of the acoustic field, is conducted by using CAA methodology for the perturbations about the mean flow. In order to reduce the computational time, the calculations are performed in parallel, using a domain decomposition strategy. This attempt at understanding and overcoming some obstacles that arise in direct simulation of CAA, opens new prospects for predicting the acoustics of aircraft engine liners.

11:15

5aEA13. Source deconvolution and system identification of flow-excited acoustic systems. Jacklyn D. Bezemek and Luc Mongeau (School of Mech. Eng., Purdue Univ., West Lafayette, IN 47907)

A method for the characterization of flow-excited acoustic systems in the frequency domain was investigated. The approach, based on dimensional analysis and the hypothesis that the system may be described using a linear source-filter model, allows the source spectral distribution and the magnitude of the frequency response function to be obtained directly from sound-pressure spectra measured over a range of flow velocities. The objective as to determine whether the method could be used to predict the noise emissions of a small axial-flow cooling fan in a given installation from source and response functions evaluated using different calibration facilities. The spectral decomposition algorithms were first validated using a theoretical model. The influence of discretization errors and random errors in the input spectra was evaluated. The noise emissions of two different fans were then measured for two different computer enclosures and a calibrated plenum, following standardized procedures. The results demonstrated that the use of the spectral decomposition method as a pre-

dictive tool was accurate in this case, provided the fan inflow was reasonably uniform. In addition, the method was found to be very robust to random errors in the input data. Discretization errors, however, had a significant impact on accuracy.

11:30

5aEA14. Experimental investigation of flow-excited cavity pressure fluctuations. Jacklyn D. Bezemek and Luc Mongeau (School of Mech. Eng., Purdue Univ., West Lafayette, IN 47907)

Cavity pressure fluctuations induced by flow grazing over a long, narrow, open cavity were investigated. The apparatus stimulated a typical road vehicle door gap/door cavity/primary bulb seal assembly, tested using a wind tunnel. The cavity pressure distribution was measured over a range of flow velocities. Pressure spectra were decomposed into a source spectral distribution function and an acoustic frequency response function, using a system identification method based on dimensional analysis and a linear source-filter model. The results indicated a weak coupling between the cavity response and the flow excitation. The excitation mechanism was dominated by flow-edge interactions, i.e., the "edge tone." The characterization method was validated by varying the acoustic properties of the cavity and the excitation strength. The source function was independent of sensor location. The system response was shown to be associated with transverse standing waves. Tapering the orifice trailing edge simulta-

neously reduced the excitation strength, the excitation frequency, and the cavity response. The presence of flow aspiration through the seal raised the broadband noise floor of the system response function. This anomaly underlined one limitation of the spectral decomposition method.

11:45

5aEA15. Source modeling for the axial fan noise. Jongmin Kang and Seungchul Park (Samsung Heavy Industries Co., Ltd., Core Technologies Team, P.O. Box 43, Daeduk Science Town, Taejeon, Korea 305-600)

A program for predicting the sound power level from earth-moving equipment is under development in SHI. Fan noise is one of the major sources, whose contribution is greater than 40% of the total radiated noise power. The experiments show that both the blade passing frequency (BPF) component and the broadband component are important, and the directivity pattern is not symmetric with respect to the plane between the upstream and downstream flow. A noise source model of the fan radiator combination for predicting both the directivity and the frequency spectrum is developed from the experimental measurements. In this paper, a physical interpretation is made for modeling the types, shapes, and locations of the simple sources depending on the fan noise generating mechanisms. The minimum number of measurements for the cross-spectral approach are also considered based on the tolerable accuracy and the physically interpretable simplicity.

FRIDAY MORNING, 20 JUNE 1997

ROOM D, 9:00 TO 11:35 A.M.

Session 5aMU

Musical Acoustics: Convergence of Music Cognition and Music Theory

David Butler, Chair

School of Music, Ohio State University, 1866 North College Road, Columbus, Ohio 43210-1170

Chair's Introduction—9:00

Invited Papers

9:05

5aMU1. The perception of tonal closure by skilled musicians. Elizabeth W. Marvin and Alexander R. Brinkman (Eastman School of Music, Univ. of Rochester, 26 Gibbs St., Rochester, NY 14604)

This study arises in response to Cook (1987) and Karno and Konecni (1992), who call into question the ability of listeners to perceive tonal closure in the tonic key, a primary structural principle of tonal composition. Participants in their experiments heard unaltered compositions and variants that destroyed the compositions' tonal structure by transposing the second half to a new key or by reordering formal units. Neither study found a correlation between the listeners aesthetic preferences and the tonal closure of the original compositions. The blurred distinction between aesthetic response and music-structural judgments in these designs is misleading; the experiment reported here asks the music-structural question explicitly. Our stimuli consisted of 12 unaltered excerpts from classical CD recordings in three categories: nonmodulating, modulating to the dominant, modulating to a key other than the dominant. For each excerpt, 36 students at a professional music school answered six questions, one of which asked whether the concluding key was the same as the initial key. Participants' accuracy was significantly above chance in this task. Musicians in scholarly disciplines and performers discriminated tonal closure equally well, although there was a significant effect for number of years of theory training.

9:35

5aMU2. The evolution of single- and double-reed musical instruments: A study of spectral characteristics. James Pyne and David Butler (School of Music, Ohio State Univ., 1866 N. College Rd., Columbus, OH 43210)

Most modern orchestral instruments have evolved through many generations of development, in many cases through an evolutionary process spanning hundreds of years. As the process of refinement went on, many earlier instruments were essentially discarded and are now heard only rarely, in early-music concerts. In some cases, these nearly extinct instruments were abandoned because their descendants were much easier to control, or could produce tones spanning increased ranges. In other cases, the earlier instruments may have been discarded because of limitations of their sounds: Their dynamic ranges may have been too limited, or their timbres (tone colors) may have been simply judged unappealing. It may also be the case that these sound characteristics were related: for instance,

that the sound of an instrument might have a low “ceiling” of volume and a high “floor” of noise, with the result that the instrument may have produced a fairly weak sound that, when forced to a higher intensity level, rapidly deteriorated in tone quality. This paper will report the early results of a study involving both physical measurements and perceptual judgments of tones produced on early single- and double-reed woodwind instruments.

10:05

5aMU3. Perceptual measures of visual and auditory cues in film music. Scott D. Lipscomb (Inst. for Music Res., UTSA Div. of Music, 6900 N. Loop W., San Antonio, TX 78249)

Both musical sound and visual images have the ability to communicate. Visual images often convey explicit (verbalizable) content, while musical sound lies most often in the implicit, nonverbalizable realm. Music perceived as organized by a listener follows a culture-dependent rule system, differing from that found in verbal communication. Numerous investigations have determined specific cognitive strategies utilized in musical listening. Likewise, there have been similar studies devoted to strategies utilized in the cognition of visual images. Recently, there has been a significant amount of interest in the perceptual interaction between the auditory and visual systems in multi-modal contexts. Both psychologists and musicians are beginning to investigate the manner in which a stimulus perceived in one sensory modality may affect the cognitive processing of a stimulus in a separate modality. Beginning in the 1950s, a series of psychophysical investigations revealed cross-modal influences using simple stimuli. A number of investigators are presently attempting to incorporate more complex stimuli to determine the interrelationship of auditory and visual cues processed simultaneously in ecologically valid experimental contexts. Drawing upon research carried out by the author, this paper will highlight methods utilized successfully in such investigations, summarizing results and proposing future directions.

10:35

5aMU4. The concept of stability and expectation in tonal music. Rene Van Egmond (School of Music, The Ohio State Univ., 1866 N. College Rd., Columbus, OH 43210)

There are two important aspects when listening to Western tonal music: the recognition of structures and the way expectations in music are evoked. Although these two aspects are not independent it can be argued that the first aspect relates to the listener's sense of coherence in a piece, whereas the second aspect relates to the listener's sense of a dynamic flow. Both aspects are dependent on the underlying feature of tonality, which has its fundamentals in the acoustical properties of sounds (overtone series) and the set theoretical properties of “musical scales.” The primary focus will be on the expectations of listeners given a certain key (i.e., tonal expectations). In literature it is assumed that these expectations rely on the “mental stability” associated with tones in a certain key (represented by the so-called tonal hierarchy). It is often argued that the difference in stability relates to a tonal expectation for a listener. Problems with this approach are discussed, considering music theoretical literature and data from experimental literature, especially emphasizing the difference between the function of individual tones and the harmonic related functions. Furthermore, data will be presented from newly performed experiments that test expectations in tonal music.

11:05

5aMU5. Expressive performance of jazz ballad melodies. Richard Ashley (School of Music, Northwestern Univ., 711 Elgin Rd., Evanston, IL 20208-1200, r-ashley@nwu.edu)

This study examines the ways in which jazz musicians use rhythmic alterations of standard ballad melodies for expressive purposes. Without altering a single pitch, these performers take melodies which are so well-known as to be cliched and make them sound fresh. The primary technique is that of extensive rubato, where the melody is played at a varying tempo against the steady rhythmic background of a rhythm section (typically drums, bass, and piano). This results in a displacement of the melodic note relative to its nominal position against the background. Types and degrees of rhythmic variation are shown to depend on motivic structure (parallel motives are performed in similar ways, preserving the originals proportional rhythmic structures), metric emphasis (downbeats are treated differently than other locations), and harmonic structure (melodic tones belonging to the underlying harmonies are rhythmically displaced in a more extreme manner than tones outside the harmonies). Consideration is given to the way in which very familiar tunes are treated compared with less-familiar melodies (for example, “My Funny Valentine” compared with “Goodbye, Pork-Pie Hat”), leading to discussion of the role of long-term memory in expressive performance. Analyses are given of several melodies, taken from commercial recordings by master improvisors.

Session 5aPP**Psychological and Physiological Acoustics: Auditory System Plasticity and Perceptual Learning**

Marjorie R. Leek, Chair

*Walter Reed Army Medical Center, Audiology and Speech Center, Research Section, Washington, DC 20307-5001***Chair's Introduction—8:30*****Invited Papers*****8:40**

5aPP1. The auditory nerve in congenitally deaf white cats: Correlations between anatomy and electrophysiology. David K. Ryugo, Brian T. Rosenbaum, Tan Pongstaporn, Ahmed A. Saada, and John K. Niparko (Dept. of Otolaryngol.-HNS, Johns Hopkins Univ. School of Medicine, Baltimore, MD 21205, dryugo@bme.jhu.edu)

It is well known that cochlear pathology produces alterations in neurons of the cochlear nucleus, but there are few data that directly relate such neuronal changes to the physiology of auditory nerve fibers and structure of auditory nerve endings. These relationships are addressed in a naturally occurring model of deafness, the congenitally deaf white cat. Cats with no intact organ of Corti exhibited no measurable hearing thresholds and had severely reduced spontaneous activity in single auditory nerve fibers. Furthermore, there was structural atrophy of endbulbs of Held in the anteroventral cochlear nucleus, and hypertrophied synapses. These hypertrophied synapses were characterized by large postsynaptic densities with highly thickened membrane specializations and markedly diminished numbers of synaptic vesicles. Cats with elevated but measurable hearing thresholds had normal-appearing organs of Corti with abundant spontaneous single unit activity but infrequent sound-evoked activity. These partially deaf cats displayed modest atrophy of their endbulbs and exhibited more normal appearing synapses. The data reveal that early onset partial or total deafness that is uncorrected for 6 months or more produces unambiguous alterations in synaptic structure. These observations are relevant to selection strategies for human cochlear implant candidates. [Work supported by NIH/NIDCD.]

9:05

5aPP2. The role of the brain in the functional development of the ear. Edward J. Walsh and JoAnn McGee (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131, walsh@boystown.org)

The olivocochlear bundle (OCB) invades and organizes within the organ of Corti around the time of birth in cats, roughly coincident with the first detectable sound-elicited responses from the cochlea or auditory nerve. A model of innervation is proposed, in which the medial OCB (MOC) regulates the development of both inner (IHC) and outer hair cell (OHC) systems. In stage 1, early arriving fibers of the MOC make transient axosomatic synaptic contact with IHCs, rhythmically modulating their transduction currents. Operating through an excitatory synapse, the MOC contributes to an activity-dependent system through which exuberant auditory nerve fiber contacts with IHCs are refined. In stage 2, MOC fibers detach from IHCs, migrate across the tunnel of Corti, and contact OHCs forming their well recognized calyces with the basolateral wall of the cell, and in the process "energize" the cochlear amplifier. By removing the OCBs influence at birth, prior to the completion of its innervation of the cochlea, its role during development can be evaluated. The otopathology resulting from neonatal cochlear differentiation is consistent with the overall predictions that would be made based upon the proposed model. [Work supported by NIH/NIDCD.]

9:30

5aPP3. Auditory perception and plasticity in the avian auditory system. Robert J. Dooling (Dept. of Psych., Univ. of Maryland, College Park, MD 20742, dooling@bss3.umd.edu) and Brenda M. Ryals (James Madison Univ., Harrisonburg, VA 22807)

Auditory perception plays a crucial role in a variety of avian behaviors including individual and species recognition, mate selection, territorial defense, song learning, and predator detection. Because the premium on hearing well is high, birds provide a window on the evolution and adaptation of auditory systems to different environments. The auditory perceptual capabilities of the avian auditory system are reviewed with special emphasis on the discrimination, recognition, and learning of complex sounds including species-specific vocalizations. These processes in birds are no where more plastic than when they subserve vocal development and learning. Because birds regenerate hair cells in the inner ear following damage, they also provide a unique opportunity to examine the return of auditory perception and the precision of vocal production and vocal learning following replacement of the auditory epithelium in an adult organism. Results of these studies are relevant to the perception and learning of speech sounds in humans and to the application of auditory prosthetic devices in humans. [Work supported by NIH Grant Nos. DC-00198 and MH-00982 to RJD and DC-01372 to RJD and BMR.]

5aPP4. Is “plasticity” a useful construct in behavioral studies of auditory processing? Charles S. Watson (Depts. of Speech and Hearing Sci. and Psych., Indiana Univ., Bloomington, IN 47405, watson@indiana.edu)

Recently the term “plasticity” has been used in discussions of certain changes in human listeners’ abilities to detect, discriminate, and identify simple and complex sounds. In some cases this appears to be merely an effort to use language appropriate to the Zeitgeist. To conclude that any specific neuroanatomical or neurophysiological change is responsible for improved or changed psychophysical performance requires that physiological measurements be made on the same subjects before and after the behavioral changes occur. That is rarely done. There are, however, useful contributions to the understanding of neural plasticity that might be made by a thorough study of the time course of auditory perceptual learning for various tasks and classes of stimuli. More relevant psychophysical data can be collected when long periods of training on a single psychophysical task are followed by tests on other tasks, making it possible to estimate the range of stimuli and tasks to which the learning has generalized. Studies meeting these criteria demonstrate great variation in the time course of auditory learning and in its specificity. [Prepared with the support of NIH/NIDCD.]

10:20–10:35 Break

10:35

5aPP5. Vowel perception and neural plasticity in cochlear implant users: The effect of more-basal-than-normal stimulation. Mario A. Svirsky (Dept. of Otolaryngol.–HNS, Indiana Univ. School of Medicine, Indianapolis, IN 46202 and Elec. Eng. Dept., Purdue Univ., West Lafayette, IN)

Because formant frequencies are a crucial cue in identifying English vowels, some pulsatile multichannel cochlear implants attempt to deliver formant frequency information by stimulating different electrodes depending on instantaneous formant values. For users of these cochlear implants, it may be possible to predict performance in vowel confusion studies based on the subject’s ability to discriminate different electrodes. One possible confounding factor is that cochlear implants stimulate cochlear locations that are more basal than normal (i.e., associated with higher frequencies). To what extent is this more basal-than-normal stimulation a problem? Is it important to stimulate the “physiologically correct” places in the cochlea, or (to state a diametrically opposed hypothesis) is the place code infinitely plastic? A mathematical model of vowel perception with cochlear implants is proposed that explicitly tests both hypotheses. The model attempts to predict a postlingually deaf subject’s vowel confusions based on his electrode discrimination and on his “perceptual bias” (meaning, in this case, the extent to which he has adapted to more-basal-than-normal stimulation). The conclusion is that more-basal-than-normal stimulation of the cochlea degrades speech perception shortly after implantation, but becomes less critical as subjects adapt to the new location.

11:00

5aPP6. The role of infant speech perception capacities in discovering the sound structure of the native language. Peter W. Jusczyk (Dept. of Psych., Johns Hopkins Univ., Baltimore, MD 21218, jusczyk@jhu.edu)

The past quarter century of research on the development of speech perception has documented a number of important capacities that infants possess for recovering information from the speech signal. Such capacities provide the infant with the means to initially categorize speech input. Yet, because individual languages differ greatly in their sound structures, any language-general capacities for processing speech must eventually give way to ones that are more appropriately tailored to the specific language spoken in the infant’s native environment. Recent studies suggest that during the second half of their first year, infants are learning about the specific properties and organization of native language sound patterns. Moreover, there is evidence that infants begin to adapt and adjust their capacities to the structure of their language input. In addition to outlining some of the landmarks observed in the development of speech perception capacities during the first year, some consideration will be given to how these capacities interact with the input to facilitate the acquisition of a native language.

Contributed Papers

11:25

5aPP7. Further studies of perception of speech and music in children. Eduardo Castro-Sierra (Lab. of Psychoacoust., Hospital Infantil de México, Dr. Márquez No. 162, 06720 México, D. F., Mexico), Eliseo Paredes-Díaz (Hospital Infantil de México, México, D. F.), Yolanda Lastra, Martha Gómez-Gama, and Santiago J. Perez-Ruiz (UNAM, México, D. F.)

The present research is a continuation of an investigation carried out at C.C.R.M.A., Stanford University [J. Acoust. Soc. Am. **93**, 2403(A) (1993)], on the perception of the pitch of speech and music in children. Experiments sought to determine whether male and female subjects

[N=120 (40×3)] 6 to 14 years of age, monolingual speakers of either Spanish, a tongue with prosodic pitch contrasts, or Otomi or Zapotec, Otomanguean tongues with lexical pitch contrasts between words, and having training in singing or in diverse musical instruments would perceive the pitch of speech as they perceived the pitch of musical sounds. Data obtained provide information that Spanish-speaking subjects at all ages and younger Otomi- and Zapotec-speaking subjects perceived the pitch of samples of their language in a manner that was similar to their perception of the pitch of synthesized musical samples. These results applied irrespective of musical background. On the other hand, older subjects speaking Otomanguean tongues seemed to perceive the lexical pitch shifts in words of their language in a manner contrasting with the perception of musical pitch. Neurosurgical studies are in progress to assess whether

these differences are due to changes in the cerebral localization of pitch perception in the course of central auditory development.

11:40

5aPP8. Changes in auditory localization responses are mediated by visual input. Barbara G. Shinn-Cunningham^{a)} and John Park (Res. Lab. of Electron., MIT, Cambridge, MA 02139)

Previous studies have shown that, with training, subjects can learn a new relationship between spatial auditory stimuli and location [Shinn-Cunningham *et al.*, J. Acoust. Soc. Am. **92**, 2334(A) (1994); **95**, 2896(A) (1994)]. These earlier studies involved training in which subjects turned to face a heard and seen target source, followed by testing using a simple forced-choice identification paradigm. In the current study, subject localization was tested using a head-pointing response method. The amount of visual information available to the subject during training and testing was systematically varied. It was found that when subjects were blindfolded, they showed no changes in localization performance over time; however, if subjects were provided with a visual field, adaptation was achieved. These results are consistent with the hypothesis that the felt position of the head is not an adequate cue to allow subjects to relearn localization responses, but that the addition of the visual field provides more robust information about the instantaneous position of the head so that learning can take place. [Work supported by the AFOSR.] ^{a)}Currently at the Department of Cognitive and Neural Systems, Boston University.

11:55

5aPP9. Learning new stimulus-response relationships in a localization paradigm: Modeling bias and resolution. Barbara G. Shinn-Cunningham^{a)} and Gregory Lin (Res. Lab. of Electron., MIT, Cambridge, MA 02139)

In previous work [Shinn-Cunningham, J. Acoust. Soc. Am. **97**, 3411(A) (1995)], the context-coding model of Braid and Durlach [J. Acoust. Soc. Am. **51**, 483–502 (1970)] was extended to account for changes in subject performance that occurred in a spatial adaptation paradigm. In these earlier experiments, subjects were trained to identify the azimuthal location of a source in a way that contradicted their normal experience (e.g., they were taught to respond that a stimulus normally perceived at 30 deg was actually at the position labeled 40 deg). The extended model was able to account for observed changes in both response bias and the ability to resolve stimuli. The model assumes that changes in bias and resolution are linked, and that there are limits in the plasticity of the stimulus-response relationship. In the current work, the model was further tested by examining how performance changed when subjects were asked to learn two different stimulus-response sets over the course of an experimental session. Observed changes in bias and resolution were predicted by the model, lending further support to its basic structure and assumptions. [Work supported by the AFOSR.] ^{a)}Currently at Cognitive and Neural Systems, Boston University.

FRIDAY MORNING, 20 JUNE 1997

DEANS HALL II, 9:00 A.M. TO 12:00 NOON

Session 5aSC

Speech Communication: Prosodic and Temporal Characteristics of Speech (Poster Session)

Mary E. Beckman, Chair

Linguistics, Ohio State University, 222 Oxley Hall, 1712 Neil Avenue, Columbus, Ohio 43210-1298

Contributed Papers

All posters will be on display from 9:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 9:00 a.m. to 10:30 a.m. and contributors of even-numbered papers will be at their posters from 10:30 a.m. to 12:00 noon.

5aSC1. A neural model predicting context effects in the identification of stops and glides. Christopher W. Myers, Michael A. Cohen, and Stephen Grossberg (Dept. of Cognit. and Neural Systems, Boston Univ., 677 Beacon St., Boston, MA 02215)

One class of phonetic context effects involves durational contrasts: A signal can be perceived as a stop consonant or a glide, depending on the length of the adjacent segment. For example, the syllable [wa] is perceived as [ba] if the steady-state frequencies corresponding to the vowel are lengthened. Thus the vocalic information affects reported phonetic judgments of the prior consonantal segment. The neural model presented processes consonant-vowel transitions and steady-state vowel signals by transient and sustained channels. This model predicts the experimental effects of vowel duration, CV transition rate, and frequency extent in identifying [ba] and [wa]. The product of a leaky integrator's output with the quotient of two linear expressions in the stimulus parameters generates the model's output. With a single small set of parameters, the model predicts over 99% of the variance in two data sets [J. L. Miller and A. M. Liberman, *Percept.*

Psychophys. **25**(6), 457–465 (1979); Schwab *et al.*, *Percept. Psychophys.* **29**(2), 121–128 (1981)]. It outperforms a number of simple alternative models. Implications for other phonetic distinctions and physiological analogs to the model in monkey cortex are described. [Work supported by AFOSR and ONR.]

5aSC2. The prosody of broad and narrow focus within noun phrases. Steven R. Hoskins (Univ. of Delaware, Appl. Sci. and Eng. Labs., A. I. duPont Inst. Hospital for Children, P.O. Box 269, Wilmington DE 19899, hoskins@asel.udel.edu)

Linguistic focus is an important factor in determining nuclear and prenuclear accents, according to current phonological theories [E. O. Selkirk, *Phonology and Syntax* (MIT, Cambridge, MA, 1984); E. O. Selkirk, "Sentence prosody: Intonation, stress, and phrasing," in *Handbook of Phonological Theory*, edited by J. Goldsmith (Blackwell, Oxford, UK,

1995), pp. 550–569]. In some cases, one prosodic realization of a sentence may correspond to more than one focus [E. O. Selkirk, *Phonology and Syntax* (MIT, Cambridge, MA, 1984); E. O. Selkirk, “Sentence prosody: Intonation, stress, and phrasing,” in *Handbook of Phonological Theory*, edited by J. Goldsmith (Blackwell, Oxford, UK, 1995), pp. 550–569]. There is little experimental data that directly bear on this issue, however. This study presents a speech production experiment on broad and narrow focus within one structure predicted to allow prosodic ambiguity: noun phrases (NPs) with an adjectival modifier. Sentences were read by 15 subjects under three focus conditions: (1) broad (on NP); (2) narrow-final (on noun); (3) narrow-early (on adjective). Speech and electroglottographic (EGG) data were recorded and analyzed. Acoustic measurements (duration, f_0 , rms amplitude) and a prosodic transcription revealed the following results: (1) Accents were more likely to be present on pre-nuclear adjectives under broad focus; (2) broad and narrow focus were often distinguished by the nuclear accent type. [Work supported by NIDRR Grant No. H133E30010.]

5aSC3. A moraic duration in speech processing in Japanese. Takashi Otake (Dept. of Linguist., Ohio State Univ., Columbus, OH 43210 and Dept. of English, Dokkyo Univ., 1-1, Gakuen-cho, Soka, Saitama, 340 Japan, otake@dokkyo.ac.jp) and Kiyoko Yoneyama (Ohio State Univ., Columbus, OH 43210-1298)

Although it is often recognized that morae *per se* can be defined both phonetically and phonologically in the literature, it is not well understood how they are involved with each other in Japanese speech processing. Recent off-line studies have suggested that recognition of morae in Japanese is highly involved with duration and syllable position [Otake and Yoneyama, *ICSLP 94* 3, 1427–1430 (1994); *ICPhS 95*, 2, 686–689 (1995)]. In the present study, the duration of C1V1 and V2.C3V3 segments in C1V1.C2V2.C3V3 nonsense words where C2 is a nasal with three different speech rates was manipulated in order to see whether C2, an onset nasal, can be recognized as a moraic nasal depending upon preceding and following segmental duration. The materials were presented to 15 Japanese listeners to dictate what they heard in Roman characters. The results showed that the manipulation of C1V1 duration never caused C2 to be recognized as a moraic nasal, while V2.C3V3 duration did. The shorter the V2.C3V3 were, the more C2 were recognized as a moraic nasal. This suggests that a moraic duration is not absolute, but relative to following segments. [Work supported by the Japanese Ministry of Education.]

5aSC4. Temporal characteristics of coarticulatory vowel nasalization in English. Jill Tanowitz and Patrice Speeter Beddor (Program in Linguist., Univ. of Michigan, Ann Arbor, MI 48109, beddor@umich.edu)

Previous research has shown that, in English /VNC/ sequences, vowels and nasal consonants are shorter when the final oral consonant is voiceless than when voiced [O. Fujimura, *Symp. on Articulatory Modeling*, Grenoble (1977)]. This study investigates whether consonant voicing affects the temporal characteristics of coarticulatory vowel nasalization, as well as V and N duration. Specifically, is vowel nasalization proportionately longer in voiceless contexts, compensating for the shorter N? Speech materials are multiple instances of real-word and nonsense /bVNC/ sequences (V=/i/ e/ æ/, N=/m/ n/ ŋ/, C=/p/ t/ k/ s/ and /b/ d/ g/ z/) produced by four American English speakers. Temporal measures include duration of V, N, C and vowel nasalization. Onset of vowel nasalization was identified as onset of the nasal formant and/or widening of formant bandwidth in the low-frequency region of FFT spectra sampled across the vowel in 10-ms increments. Vowel duration was on average 20% shorter and N duration 40% shorter in voiceless than in voiced contexts. However, vowel nasalization duration was not compensatorily longer in voiceless contexts. Regardless of N duration, coarticulatory nasalization remained a constant

proportion (about 80%) of total vowel duration across all segmental contexts. Variation in this proportion across vowels and speakers is minor. [Work supported by NSF.]

5aSC5. Analysis of children's speech. Pitch and formant frequency. Sungbok Lee (Central Inst. for the Deaf, St. Louis, MO 63110, lee@cidmv1.wustl.edu), Alexandros Potamianos, and Shrikanth Narayanan (AT&T Labs, Rm. 2C-318, 600 Mountain Ave., Murray Hill, NJ 07974)

Pitch and formant measurements were performed on ten monophthongs of American English included in a speech database collected from 436 children (ages 5–18) and 56 adults [J. D. Miller *et al.*, *ICASSP*, pp. 849–852 (1996)]. After automatic segmentation and labeling of each waveform, the pitch and the first four formant frequencies of each vocalic segment were estimated using a formant-tracking program (Entropic Research Laboratory, Inc.). Median values of each pitch and formant track were computed as the representative value of the track. The raw formant data were refined based on the linear regression statistics of each formant track and manual correction. The pitch and refined formant data were analyzed as a function of vowel, age, and gender. Results indicate that the rapid change in male pitch due to puberty occurs from age 12 to age 15, lower than the range (ages 13–17) in Kent [R. D. Kent, *J. Speech Hear. Res.* 19, 421–447 (1976)]. Age-dependent formant changes are explained by a linear scaling of the vocal tract length. The intra-subject variation of formant and pitch decreases as a function of age, supporting the results of a previous study [S. Eguchi and I. J. Hirsh, *Acta Otolaryng.*, Suppl. 257 (1969)]. Implications of these and other findings on speech development will be discussed. [Work supported by AT&T.]

5aSC6. Searching for the nature of motor speech control: Vowel reduction in habitually slow versus fast talkers. Ying-Chiao Tsao (Dept. of Special Ed./Commun. Disord., Univ. of Nebraska, Kearney, NE 68849, tsaoy@platte.unk.edu) and Gary Weismer (Univ. of Wisconsin, Madison, WI 53706)

Vowel reduction as a function of speaking rate within individuals has been well established (Fourakis, 1991; van Bergem, 1995; Weismer *et al.*, 1995). This phenomenon is challenged by the finding that differences exist in the control of speech rate for both slow and fast talker groups (see Tsao, 1995). This study will therefore examine vowel reduction in the speech of individuals separated on the rate continuum (i.e., slow versus average versus fast). The hypothesis that the habitual vowel spaces and the rate-induced degree of vowel reduction is different for habitually slow versus fast talkers will be explored. Subjects will produce a reading passage, and vowel reduction will be measured based on the target frequencies for F_1 , F_2 , and F_3 for the four corner vowels (e.g., /a/, /i/, /ae/, /u/). A linear regression function will be used to demonstrate the relationship between vowel reduction of the 60 subjects (15 slow, 15 fast, and 30 average) and their habitual speech rate along the rate continuum. [Work supported by University of Nebraska—Kearney and UW—Madison Speech Acoustics Laboratory.]

5aSC7. A longitudinal investigation of duration and temporal variability in children's speech production: Additional data. Bruce L. Smith and Mary Kay Kenney (Dept. of Commun. Sci. and Disord., Northwestern Univ., 2299 N. Campus Dr., Evanston, IL 60208-3570, bsmith@casbah.acns.nwu.edu)

To follow up on an earlier, 1½-yr longitudinal study of temporal characteristics in speech production development [Smith *et al.*, *J. Acoust. Soc. Am.* 99, 2344–2349 (1996)], the same 12 children were seen again to extend the analysis interval to about 3 yr. The children ranged from ap-

proximately 7–12 years of age initially to about 10–15 years at the end of the study. For each recording, they produced 25 repetitions of each of several different words and short phrases, from which various segments and syllables were measured. Duration and temporal variability measures for up to half the subjects demonstrated nonadultlike performance at the initial and intermediate recordings, whereas average duration and variability were within an adultlike range for all 12 children by the final session. In general, the greatest decreases across the 3 yr occurred during the first $1\frac{1}{2}$ yr, particularly regarding temporal variability; however, small decreases in duration and variability occurred from the intermediate to the final recording, as well. A tendency was also observed for children whose average segment duration or variability was more adultlike at the initial recording to show more adultlike values at the intermediate and final recordings.

5aSC8. Pitch accent, phrase tones, and vowel devoicing in Tokyo Japanese. Mafuyu Kitahara (Dept. of Linguist., Indiana Univ., Bloomington, IN 47405)

In Tokyo Japanese, short high vowels devoice between voiceless consonants, which conflicts with the realization of pitch accent. The fundamental frequency around the devoiced region in a corpus of 12 speakers was measured to investigate if and how speakers compensate for the loss of information normally conveyed by the pitch accent. Previous studies revealed that elevation and abrupt pitch fall immediately after the devoiced region is characteristic of utterances containing a devoiced accented syllable. However, preliminary analyses of these data show the occurrence of the characteristic pitch pattern after a devoiced region is dependent on the pitch height before the devoiced region. If two syllables are consecutively devoiced and the first one bears an accent, the pitch elevation only occurs when the antepenultimate syllable has a low phrase tone. This suggests a nonlocal pitch adjustment exists dependent on the occurrence of devoicing and the location of phrase tones. This also implies that vowel devoicing is not just a low-level consonantal effect, but has a long distance impact on the execution of phrasal pitch contours.

5aSC9. Accent, stress, and spectral tilt. Nick Campbell (ATR Interpreting Telecommunications Res. Labs., 2-2 Hikaridai, Soraku-gun, Seika-cho, Kyoto, 619-02 Japan) and Mary E. Beckman (Ohio State Univ., Columbus, OH 43210-1298)

In English and Dutch, pitch accents occur only on lexically prominent syllables. Such syllables are not always accented in longer utterances, however, and traditional descriptions differentiated “stress” proper as a local increase in loudness, although the intonational event is the most salient cue to prominence, far outweighing any differences in overall rms amplitude. Recent work by Sluijter and colleagues [J. Acoust. Soc. Am. **101**, 503–513 (1997)] indicates that stressed syllables in Dutch are associated with differentially increased energy at frequencies well above the fundamental, and that these spectral tilt differences are a robust cue to relative syllable prominence — whether or not the word is in focal prominence. Because accents are not necessarily associated with focused words, however, their experiments do not tell us whether spectral tilt differentiates lexically stressed from unstressed syllables in the absence of an associated intonational prominence. The current study examines five acoustic measures, including spectral tilt and open quotient, for accented, unaccented but stressed, and unstressed syllables in a corpus of utterances produced by four speakers of American English. Results confirm that spectral tilt is affected by accent independent of F_0 increase, and that it is affected (to a lesser extent) by stress in the absence of accent.

5aSC10. Simultaneous effects on the duration of syllable nuclei: A covariance structure analysis approach. Mary L. Erickson (Vanderbilt Voice Ctr., 1500 21st Ave. South, Ste. 2700, Nashville, TN 37212)

Using covariance structure analysis, two linear models of vowel duration were tested. The first model, the Direct Effects Only model, tested the direct effects on vowel duration of intrinsic duration, following consonant voicing, stress, position in utterance, change in fundamental frequency, and final lengthening. The second model, the Direct and Indirect Effects model, tested all of the above direct effects on duration, plus the indirect effects of stress and phrase final position on vowel duration via their effects on fundamental frequency change. Model goodness of fit was estimated by three fit indices: the Normed Fit Index (NFI), the Nonnormed Fit Index (NNFI), and the Comparative Fit Index (CFI). The Direct Effects Only model yielded the following fit indices: NFI=0.810, NNFI=0.788, and CFI=0.821. The Direct and Indirect Effects model yielded slightly higher fit indices: NFI=0.822, NNFI=0.802, and CFI=0.833. Position in syllable, position in phrase, intrinsic duration, phrase final position, stress, and change in frequency were significant predictors of duration in both models. In the Direct and Indirect Effects model, phrase final position had an indirect effect on vowel duration via a significant effect on change in frequency; stress did not.

5aSC11. The long-time average spectra of newborn infant crying. Alexander M. Goberman and Michael P. Robb (Dept. of Commun. Sci., Univ. of Connecticut, U-85, 850 Bolton Rd., Storrs, CT 06269, amg93001@uconnvm.uconn.edu)

The majority of research pertaining to acoustic analysis of infant crying is based on the collection of a limited number of cry segments. Far less research is based on examination of whole crying episodes. Long-time average spectrum (LTAS) analysis offers a quantitative approach to inferring the vocal source as derived from whole crying episodes. The purpose of this study was to quantitatively examine a variety of LTAS characteristics (e.g., energy concentrations, spectral tilt) in the crying episodes of ten normal full-term neonates and ten at-risk preterm neonates. Results of this analysis indicated that some but not all of the LTAS features were discriminating of normal versus at-risk infants. The discussion focuses on the application of cry analysis in the evaluation of laryngeal behavior and its potential usefulness in determining an infant’s health status.

5aSC12. Prosodic phrasal structure and lexical interpretation. Shari R. Speer and Amy Schafer (Speech-Lang.-Hearing, Dole 3031, Univ. of Kansas, Lawrence, KS 66045)

In phonological theory, phonological phrases (PPhs) and intonational phrases (IPhs) are distinct levels of prosodic phrasing, with different edge tones, different constraints on well-formedness, and different acoustic properties. Although many studies have demonstrated prosodic effects on syntactic processing, few studies have examined prosodic effects on semantic processing, and the majority of studies have not compared the effects PPhs and IPhs [but see Kjellaard (1995)]. It is demonstrated that, with respect to the interpretation of lexical, syntactic, and semantic ambiguities, PPhs and IPhs have distinct processing effects. Results will be presented from several experiments testing lexical ambiguities of several types, using both cross-modal naming and end-of-sentence tasks. For example, it will be shown that reanalysis to the subordinate meaning of a polysemous word presented in a neutral context takes longer when the word is in a preceding IPh than in a preceding PPh, for lexically and syntactically identical materials. These results indicate that prosodic phrasing influences the ambiguity resolution process at lexical, syntactic, and semantic levels. It is argued that the sentence comprehension mechanism makes differential use of two levels of prosodic phrasing specified in phonological theory. Implications for current processing models will be discussed. [Work supported by NIMH.]

5aSC13. Lexical stress and metrical stress in auditory word recognition. Emily G. Soltano and Louisa M. Slowiaczek (Dept. of Psych., SUNY, Albany, NY 12222, egs572@cnisbm.albany.edu)

Three experiments examined the influence of word stress in spoken word recognition. Experiments 1 and 2 investigated the role of lexical stress on response times in lexical decision and shadowing tasks. Primes and targets shared lexical stress patterns (e.g., BIBlical—JUVenile) or did not (e.g., deFENDED—JUVenile). The results did not support an influence of lexical stress in word recognition. Experiment 3 used a shadowing task, in which metrical stress was manipulated. Metrical stress defines weak

syllables as containing schwa and strong syllables as containing all other vowel sounds. Primes and targets shared metrical stress patterns (e.g., resist—cigar) or did not (e.g., hostile—cigar). Targets with strong first syllables were shadowed faster than targets with weak first syllables. Moreover, the type of first syllable (weak or strong) interacted with whether the prime and target had the same metrical stress pattern; responses were faster when the target had a strong first syllable and the prime-target metrical stress patterns matched compared to all other conditions. These data provide converging evidence for the view that metrical stress and not lexical stress influences word recognition. [Work supported by Grant No. NS29286 from NIH.]

FRIDAY MORNING, 20 JUNE 1997

ROOM R, 8:00 A.M. TO 12:00 NOON

Session 5aUW

Underwater Acoustics: Bubbles, Scattering and Noise

Steven L. Means, Chair

Code 7121, Building 2, Room 129D, Naval Research Laboratory, Washington, DC 20375

Contributed Papers

8:00

5aUW1. Dispersion estimates for bubbly seawater. Ralph R. Goodman (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804), Jerald W. Caruthers, and Paul A. Elmore (Naval Res. Lab., Stennis Space Center, MS 39529)

In April 1995, the Naval Research Laboratory performed high-frequency direct-path experiments in shallow water, near Panama City. During that period there were two days of strong weather which gave rise to slower sound speed, larger fluctuations over a few seconds, and dispersion. It led to the conclusion that only the presence of bubbles could cause these changes. The experimental results will be published soon. In an effort to understand these changes, analytical and numerical analyses were undertaken. Straightforward dispersion theory, based on the change of compressibility gives an integral expression that is a function of frequency and bubble distribution. It was found that there is a simple, exact integration for the case where the distribution of bubbles of radius a is given by the inverse of fourth power, a^{-4} . Numerical studies have been performed that can be used to compare dispersion for other powers, from a^{-3} to a^{-5} . The analyses give insight into the role that void fraction, bubble sizes, and their distribution play in dispersion. The analyses exhibit, as well, the sensitivity to accuracy that would be required if dispersion is to be used to determine the distribution of bubble size in the sea. [Work supported by NRL.]

8:15

5aUW2. High-frequency surface bubble loss. Jun He and F. D. Tappert (Appl. Marine Phys., Univ. of Miami, RSMAS, 4600 Rickenbacker Cswy., Miami, FL 33149)

The problem of predicting the high-frequency forward propagation loss due to the near-surface bubble layer, called the surface bubble loss (SBL), is solved by use of the full-field University of Miami parabolic equation (UMPE) acoustic model. Curves of SBL versus grazing angle are presented for many wind speeds and frequencies. These results demonstrate the importance of surface-image interference at small grazing angles, an effect that ray-tracing models do not predict. It is also found that the effective surface is displaced downward by the lossy bubble layer, a novel effect that also is not predicted by ray-tracing models. [Work supported by ONR.]

8:30

5aUW3. Radiation from point sources within a bubble cloud. Steven L. Means and Richard M. Heitmeyer (Naval Res. Lab., Code 7121, Washington, DC 20375)

Bubble cloud resonances have been proposed as an explanation of the low-frequency acoustic radiation produced by breaking waves. A previous model [H. N. Oguz, J. Acoust. Soc. Am. **95**, 1895–1912 (1994)] considered excitation of the bubble cloud by a rigid piston at the base of a hemispherical bubble cloud. The present model considers excitation of the cloud by individual point sources within the cloud. A Green's function is obtained for a point source displaced from the origin of a hemispherical bubble cloud beneath a pressure release surface. The radiation pattern for multiple point sources distributed within a slice of the hemispherical cloud is obtained by superposition and the method of images. The effects of bubble cloud characteristics on the radiation field will be shown for various frequencies and ranges of point source depths. The method presented here is applicable to modeling sound generation of breaking waves having bubble clouds that are not amenable to being modeled by a hemispherical geometry. For instance, bubble clouds produced by breaking surf might be modeled more accurately using a cylindrical bubble cloud geometry. [Work supported by ONR.]

8:45

5aUW4. Multivariate Fourier and wavelet analysis of acoustic backscattering and environmental data from the acoustic surface reverberation experiment. Neil J. Williams, Charles L. Monjo, Hien B. Nguyen, and Harry A. Deferrari (Div. of Appl. Marine Phys., Rosenstiel School of Marine and Atmospheric Sci., Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149-1098)

Data from the acoustic surface reverberation experiment (ASREX) [Williams *et al.*, J. Acoust. Soc. Am. **97**, 3404 (1995)], is analyzed using multivariate Fourier and wavelet techniques to investigate relationships between environmental conditions and observed acoustic surface backscatter at 200, 400, and 800-Hz frequencies. Environmental variables observed included wind speed and direction, wave height, air and sea temperatures, current speed and direction, ambient noise, and various measurements of bubble activity characterized by depth-integrated air volume and e-folding depth. The relative contribution of the environmental parameters to ob-

served backscattering strengths at the various frequencies will be discussed. The analysis techniques will be described, compared, and contrasted. [Work supported by ONR.]

9:00

5aUW5. On the use of acoustic ambient noise to characterize backscattering strength. Christopher W. Day, Neil J. Williams, Charles L. Monjo, Hien B. Nguyen, and Harry A. Deferrari (Div. of Appl. Marine Phys., Rosenstiel School of Marine and Atmospheric Sci., Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149-1098)

Strong correlation between acoustic ambient noise and backscattering strength was observed during analysis of data from the acoustic surface reverberation experiment (ASREX) [Williams *et al.*, J. Acoust. Soc. Am. **97**, 3404 (1995)]. Traditional empirical formulas for acoustic backscattering strength, such as those of Chapman and Harris [R. P. Chapman and J. H. Harris, J. Acoust. Soc. Am. **34**, 1592–1597 (1962)] and Ogden and Erskine [P. M. Ogden and F. T. Erskine, J. Acoust. Soc. Am. **95**, 746–761 (1994)], utilize wind speed, frequency, and grazing angle as their input variables. An alternative approach which utilizes ambient noise in place of wind speed is presented here. Unlike wind speed, ambient noise measurements are easily made from subsurface platforms and this technique may therefore be advantageous for certain applications (i.e., submarines and covert operations). [Work supported by ONR.]

9:15

5aUW6. An Arctic acoustic array system: Installation, *in situ*, repair, and preliminary data analysis. Barbara Sotirin (NCCOSC RDTE D881, 49575 Gate Rd., Rm. 170, San Diego, CA 92152-6743, sotirin@nosc.mil) and Jon Thorleifson (DREA Esquimalt Defence Res. Detachment, Victoria, BC V9A 7N2, Canada)

Under Project Spinnaker a joint Canadian/U.S. large multiple aperture array was deployed and cabled back to shore under the Arctic ice pack during April 1996 to acquire long-term measurements of Arctic ambient noise. The data collected have been documented and are being distributed to investigators for analysis. As well, the system is planned as a receive station for the joint U.S./Russian Arctic Climate Observations Using Underwater Sound (ACOUS) project. Although battered for a 3-to-5-year life, in June 1996 the signal from the array terminated abruptly. After detailed analysis of the last recorded data, a repair effort was planned for the Spring of 1997. This paper will describe the analysis, methods, and outcome of the *in situ* repair and future use of the deployed acoustic array system. [Work supported by ONR, SPAWAR, and CRAD.]

9:30

5aUW7. Ship-radiated noise-field characterization. Stephen C. Wales (Code 7120, Acoust. Div., Naval Res. Lab., Washington, DC 20375-5350)

Sonobuoy measurements of merchant ships were processed for mean levels and fluctuation dynamics of the broadband and line components from a large database of ships. Direct path results are reported. The mean levels are compared with various ship source level models. The levels are shown to agree with Ross-type models with some minor differences noted. The importance of source characterization, particularly source depth, in interpreting these results is emphasized. Fluctuations of the broadband component are shown to vary by as much as 6 dB with periods of 10–20 s for a limited number of ships. The results are shown to be consistent with the hypothesis that they are due to ship motion in the water. Examples are provided of ship line temporal variation in frequency and amplitude. Preliminary work on categorizing the data according to the spectra characteristics and time behavior is discussed.

9:45

5aUW8. Modeling and simulation of marine mammal generated sound. I. Stochastic modeling of marine mammal spatial distributions, collective motions, and vocalization occurrence times. Thomas J. Hayward (Naval Res. Lab., Code 7121, Washington, DC 20375-5350)

Stochastic models and algorithms are developed for representing and simulating marine mammal spatial distributions, collective motions, and vocalization occurrence times. A general class of Markov-process models of collective motion is first identified. Within this class, particular representations are constructed that incorporate interaction mechanisms of the type developed in distributed behavior modeling [C. Reynolds, Comput. Graph. **21**, 25 (1987)]. Collective motion is represented by a Markov process with a stationary probability density defined in terms of an “energy” function that expresses the propensity for clustering and polarization within groups of individuals. Computer-generated sample functions of the Markov processes, based on variants of the Metropolis algorithm, provide for simulation of a variety of collective motions. These range from nearly independent motions of solitary individuals to motions in highly clustered and polarized configurations simulating motions of marine mammals in pods. Occurrence times of vocalizations are represented by Markov processes and associated simulation algorithms. Simulations of patterned click sequences (codas) [Watkins and Schevill, J. Acoust. Soc. Am. **62**, 1485 (1977)] produced by sperm whales (*Physeter catodon*) are presented as an example. Potential applications to sonar systems and marine mammal acoustic surveys are briefly discussed. [Work supported by ONR and NRL.]

10:00–10:15 Break

10:15

5aUW9. The effect of surface noise in low-frequency passive detection in coastal waters. William M. Sanders and Frederick D. Tappert (Appl. Marine Phys., Univ. of Miami RSMAS, Miami, FL 33149, wsanders@rsmas.miami.edu)

The effect of the environment on sensor performance is twofold. Signal level and local noise level are both affected by acoustic propagation. In performance analyses, noise level is often taken to be a spatially homogeneous and temporally constant value. This paper utilizes a three-dimensional propagation model to determine the detection properties of bottom-mounted passive low-frequency sensors. The parabolic equation (PE) model is used to propagate noise from a field of surface dipole sources to the sensor location. Summation of noise from the entire area under consideration yields the local noise level. Two environments are considered. In the Adriatic Sea, the acoustic environment emphasizes local noise sources. In the Florida Straits, noise from distant sources affects local noise levels.

10:30

5aUW10. Scattering from a fluctuating object in a fluctuating waveguide. Nicholas C. Makris (Naval Res. Lab., Washington, DC 20375)

In many practical applications of active target scattering in ocean acoustics, there is a significant amount of uncertainty in both the scattering properties of the target and the propagation characteristics of the medium. In order to consistently combine a reasonable description of this uncertainty with the established physics of scattering from submerged objects, a statistical model is developed that describes both the spatial and temporal characteristics of the scattered field. An essential aspect of this model is the multiplicative nature of the scattering process. The acoustic field propagates from the source through a random waveguide. It is then scattered from a random target that undergoes fluctuations that are uncorrelated with those of the incident field. This scattered field then propagates into azimuthal directions where the random waveguide to the receiver may be

partially or fully uncorrelated with the incident waveguide. Standard assumptions for statistical fluctuations of the waveguide channels and target lead to a non-Gaussian field at the receiver. Possible randomness of the sound source leads to a further departure from Gaussianity in the received field statistics. Expressions for the spatio-temporal mutual intensity structure of the received field will be provided, with illustrative examples for a shallow water scenario.

10:45

5aUW11. Wide bandwidth and low-frequency target response measurements with a large PVDF sheet source outside the laboratory. Scot F. Morse, B. Todd Hefner, Philip L. Marston (Phys. Dept., Washington State Univ., Pullman, WA 99164-2814), and Kevin L. Williams (Univ. of Washington, Seattle, WA 98105)

It is demonstrated that elastic response features in the few kilohertz frequency range and at higher frequencies can be observed using a large, acoustically transparent PVDF sheet source. The source was operated in both the impulse mode [Kaduchak *et al.*, J. Acoust. Soc. Am. **97**, 2699–2708 (1995)] and with tone burst excitation. It was used in Lake Union (Seattle) to measure resonances and other elastic features of a stainless steel spherical shell having a radius of 29.8 cm. Both the source and sphere were supported from the APL barge. One especially noteworthy feature of the source is its ability to generate tone bursts that turn on and off abruptly. This is because of the absence of resonances of the source. The potential applications of sheet sources of this type are not limited to laboratory experiments. [Work supported in part by the Office of Naval Research and by APL.]

11:00

5aUW12. Space and frequency dependence in low-frequency, narrow-band bistatic reverberation: Theoretical results for variable wind speed and surface directionality. Robert F. Gragg (Naval Res. Lab., Code 7140, Washington, DC 20375-5350)

Analytical methods are used to express the dependence of low-frequency bistatic reverberation on the power and direction spectra of the time-dependent sea surface. The 3-D cross-spectral density (CSD) of the reverberation is obtained, to second order in the small-waveheight approximation, as a sum of dc and sideband terms. For numerical simulations, this theoretical formulation is applied to a model ocean consisting of a homogeneous half-space. The sensitivity of the CSD to the power spectrum of the source, the power and direction spectra of the surface, and bistatic source/receiver geometry are assessed for both the zero-Doppler components and the Bragg–Doppler sidebands. [Work supported by ONR.]

11:15

5aUW13. Biomimetic sonar differentiates coin head from tail. Roman B. Kuc (Dept. of Elec. Eng., Yale Univ., 15 Prospect St., New Haven, CT 06520-8284)

An active sonar using Polaroid electrostatic transducers positioned at the end of a robot arm can recognize an object from its echoes. The sonar adaptively changes its location and configuration in response to the echoes. The sonar mimics biological echo-location systems, such as those employed by bats and dolphins, in that there is a center transmitter flanked by

two adjustable receivers, the sonar has rotational and translational mobility, and the echo processing contains elements that have been observed in the mammalian auditory system. The sonar translates in a horizontal plane and rotates in pitch and yaw to position an object at a standard location within the beam patterns. The transmitter points at the object to maximize the incident acoustic intensity and the receivers rotate to maximize the echo amplitude and bandwidth and to minimize the echo-producing region. This procedure results in a unique echo vector associated with each object at a given object pose. Recognition is accomplished by extracting 32 values from the binaural echo patterns and searching a database that is constructed during a learning phase. The system can differentiate reliably the head and tail sides of a coin. [Work supported by NSF Grant IRI-9504079.]

11:30

5aUW14. Environmental impact on active-sonar performance predictions in shallow water. James H. Leclerc, Mona J. Collins, and Donald R. Del Balzo (Naval Res. Lab., Stennis Space Center, MS 39529-5004, delbalzo@nrlssc.navy.mil)

Predictions of active-sonar system performance at 750 Hz and 3 kHz in shallow water were reported [J. H. Leclerc *et al.*, J. Acoust. Soc. Am. **100**, 2604(A) (1996)] using a simple propagation model (FAME) in reverberation-limited situations. The water sound-speed profiles and bottom geoacoustic descriptions were derived from historical data collected in the Strait of Sicily. This analysis extends the previous work with more accurate propagation calculations (RASP) and the inclusion of signal spreading in harsh bottom-limited environments. Transmission loss, reverberation level, signal spreading (echo splitting loss), and signal excess for a generic sonar system are calculated for several realistic environmental descriptions in 100- and 800-m water depths. The sensitivity of sonar system predictions to spatial environmental variations are quantified and implications about space sampling strategies are discussed. [Work supported by ONR/NRL.]

11:45

5aUW15. Echosignal temporal structure from an object of complex form in an irregular waveguide. E. I. Oboznenko (Dept. of Acoust. and Acoustoelectron., Kiev Polytech. Inst., Kiev, 252056, Ukraine)

Principal aspects of the problem of sound scattering from an object of complex form (CSO) in a free field and in an irregular waveguide are investigated. The CSOs are prototypes of real artificial submarine objects [I. L. Oboznenko, *Complex Shaped Object Detection and Classification* (EG&G WASC, Washington, 1995)]. The CSO is represented by a sum of the interacting elements of local reflection (LREs). It is shown that the LREs CSO are partly coherent scattered. In this condition, the CSO is a multipatch object containing glistening points (part of the first Fresnel zone of each LREs). Temporal structure computations of the multipatch echosignals from the four CSO's models of real constructions have been made within ray approximation in imitation of an irregular waveguide. It is shown that the real signal can vitally distort and exceed the length of the CSO's echosignal, an irregular waveguide caused by the multipath propagation and the multipatch CSO. This difference is observed by a statistical shift of the CSO along the waveguide direction at a different depth. Some classification indications of the CSO were provoked by scattering in an irregular waveguide.

Session 5pSC

Speech Communication: Potpourri (Poster Session)

Stephen A. Zahorian, Chair

Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, Virginia 23529

Contributed Papers

All posters will be on display from 1:30 p.m. to 4:30 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:30 p.m. to 3:00 p.m. and contributors of even-numbered papers will be at their posters from 3:00 p.m. to 4:30 p.m.

5pSC1. Vowel articulation training aid for the hearing impaired: An update. A. Matthew Zimmer, Stephen A. Zahorian (Dept. of Elec. and Comput. Eng., Old Dominion Univ., Norfolk, VA 23529, zahorian@ece.odu.edu), and Stefan Auberg (Syracuse Lang. Systems, Syracuse, NY 13214-2845)

A computer-based system which provides real-time feedback for vowel articulation training for the hearing impaired is described. This system is a revised version of the training aid described in previous papers [Zahorian and Correal, *J. Acoust. Soc. Am.* **95**, 3014(A) (1994); Beck and Zahorian, *ICASSP-92*, II-241-244]. Revised feature extraction and classification algorithms improve accuracy and allow processing of vowels spoken in CVC contexts. Cost has been reduced by a new Windows 95/NT-based implementation which requires no specialized DSP or audio hardware. Steady-state vowel and CVC tokens have been collected from 24 adult male, 28 adult female, and 46 child speakers to provide training data for semi-speaker-independent neural network classifiers used in the system. Displays include a two-dimensional F1/F2 style display which provides continuous feedback, and a vowel bar graph display. Each accommodates four speaker groups: adult female, adult male, child, and general. Tests with ten steady-state monophthong vowels produced by speakers outside the training set indicate that typically over 80% of vowels are correctly depicted by the bar graph display and over 75% by the 2-D display. A third display for vowels spoken in CVC contexts is under development. [Work funded by NSF, Grant No. NSF-BES-9411607.]

5pSC2. Where lipreaders look on the face: Task and individual differences. Charissa R. Lansing (Dept. of Speech and Hearing Sci., Univ. of Illinois, Champaign, IL 61820) and George W. McConkie (Univ. of Illinois, Champaign, IL 61820)

Lipreaders' eye movements were monitored in real time to determine where they direct their attention on the speaker's face to obtain linguistically relevant information for speech perception. Full-motion video sequences displaying a full-face view of a male speaker were presented without sound to proficient lipreaders who were deaf or hearing. A still frame of the speaker's face was displayed for 1 s preceding and following each video segment. The video was followed by a display of a written sentence that was identical or different from the sentence produced by the speaker. Lipreaders indicated same/different by pressing a button. Results indicate that eye gaze was most often directed at the mouth as the speaker produced the sentence but occasionally moved to other parts of the face. The number of these off-mouth gazes during speech varied among lipreaders and with sentence characteristics. Both task and lipreader characteristics influence where people look for information on the face of the speaker. [Work supported by NIDCD, NIH.]

5pSC3. Synchronic evidence for the historical development of Khmer phonology. Ratree Wayland and Allard Jongman (Cornell Univ., Cornell Phonetics Lab., Ithaca, NY 14853)

In this paper a detailed acoustic investigation of the difference between clear and breathy phonation types in the Chanthaburi dialect of Khmer is presented. The study of this dialect is of paramount importance since it may prove the only way to quantitatively evaluate the historical claim that Khmer once had contrasting phonation types. Data were collected from five speakers. The phonation types of the vowels were then investigated by means of a variety of algorithms that have been proposed in the phonetic literature, including harmonic-to-noise ratio [HNR; e.g., Wayland *et al.* (1995)]; $H1-H2$, where $H1$ and $H2$ refer to the amplitude of the first and second harmonic, respectively [e.g., Klatt and Klatt (1990)]; $H1-A1$, and $H1-A3$, where $A1$ and $A3$ refer to the amplitude of the most prominent harmonic in the region of the first and third formant, respectively [e.g., Stevens and Hanson (1994)]. A comparison of the different metrics will be presented. The results from some analysis methods show that the Chanthaburi dialect distinguishes clear and breathy vowels. The finding that there still is a dialect of Khmer with a phonation-type contrast provides crucial evidence for the historical reconstruction of Khmer phonology.

5pSC4. Categorical perception as a function of stimulus quality. A. J. van Hessen and M. E. H. Schouten (Utrecht Inst. of Linguist. OTS, Utrecht Univ., Trans 10, 3512 JK Utrecht, The Netherlands, bert.schouten@let.ruu.nl)

A number of experiments were carried out in order to test the hypothesis that the degree of categorical perception of speech stimuli is a function of synthesis quality; specifically, that the greater complexity of more natural speech stimuli makes it difficult for listeners to focus on particular stimulus parameters such as psychoacoustic cues. The results show that there is some increase in categorical perception as synthesis quality improves from a simple synthesis by rule, via LPC synthesis, to a much more complex type called sine-wave generation.

5pSC5. The relationship between speech information perceived by cochlear implantees in different spectral regions and electrode discrimination. Belinda A. Henry, Colette M. McKay, Hugh J. McDermott, and Graeme M. Clark (Dept. of Otolaryngol., Univ. of Melbourne, Parkville, Victoria, 3052, Australia); (henryb@mail.medoto.unimelb.edu.au)

This study investigated whether the relative amount of speech information contained in specific frequency regions of the speech spectrum differs between implantees and normal-hearing subjects, or between implantees having different speech perception abilities. In addition, the rela-

tionship between the amount of speech information implantees perceived in each frequency region and their ability to discriminate adjacent electrodes was investigated. Fifteen subjects using the Nucleus CI-22M implant (SPECTRA-22 processor/SPEAK strategy), and 16 normal-hearing listeners participated. Articulation index procedures were used to measure the amount of speech information perceived in different frequency bands. Results showed that, compared to the normal-hearing listeners, the implantees obtained less speech information in all frequency regions, but more information in the higher, compared to the lower, frequencies. Correlation analysis showed that electrode discrimination and perceivable speech information were correlated in the lower frequency regions, but not in the highest region, suggesting that the perception of speech information in the lower regions requires the ability to discriminate electrodes, while the perception of high-frequency information does not. These findings predict that speech perception may be improved by changing the assignment of frequencies to electrode positions to provide better discrimination of low- to mid-frequency speech information.

5pSC6. Coarticulation of disyllables by deaf and hearing speakers. Areti Okalidou and Katherine S. Harris (Ph.D. Program in Speech and Hearing Sci., Grad. School and Univ. Ctr. of the City Univ. of New York, 33 W. 42nd St., New York, NY 10036, okalidou@egnatia.ee.auth.gr)

The coarticulatory patterns of three deaf and three hearing adult speakers were assessed as a function of segment type and phonetic context, by means of acoustic analysis of the F_2 . The F_2 differences between /ə#CVC/ disyllables minimally contrasting in vocalic ([i], [u], [a]) and in constant ([b], [d]) context, respectively, were examined at five measurement locations, namely, at ə onset, ə midpoint, ə offset, and transition onset and vowel midpoint of the stressed vowel. Intervocalic coarticulation was greater for either deaf or hearing speakers, depending on the phonetic composition of disyllables (bVb or dVd). Moreover, the deaf speakers showed less marked consonant influences on the ə and stressed vowel portions of disyllables compared to the hearing controls. The above findings partly dispute the conclusions drawn from previous studies. Furthermore, they suggest that the speech of deaf speakers is guided by patterns of articulatory organization similar in certain respects to those of young children learning to speak.

5pSC7. Relations among three measures of auditory-visual integration. Ken W. Grant (Walter Reed Army Medical Ctr., Army Audiol. & Speech Ctr., Washington, DC 20307-5001)

Three measures of auditory-visual (AV) integration using both congruent and discrepant speech stimuli were evaluated in 30 hearing-impaired subjects. The first measure was based on Braidá's model of pre-labeling integration [Q. J. Exp. Psych. **43**, 647–677 (1991)] which makes predictions of AV consonant recognition from separate auditory (A) and visual (V) consonant confusion matrices. The ratio of observed-to-predicted AV score was used as an index of integration, with unity representing optimal integration. The second measure used discrepant A and V consonant-vowel stimuli to determine susceptibility to "McGurk" effects. Susceptibility was defined as the difference between A and AV consonant recognition, with greater differences reflecting greater susceptibility. The third measure used temporally misaligned AV speech stimuli in a sentence recognition task to determine individual sensitivity to AV asynchrony. Auditory onset delays between 0–320 ms were studied and individual sensitivity to AV asynchrony was defined in terms of decrements in sentence intelligibility relative to the synchronous condition (0-ms delay). Relations among these three potential measures of AV integration, and between integration measures and a separate measure of relative AV sentence benefit, will be discussed. [Work supported by NIH Grant No. DC00792 and the Department of Clinical Investigation.]

5pSC8. Improving the process of automatically extracting diphones with context-dependent boundaries for synthesized speech. Debra M. Yarrington, Steven R. Hoskins, and H. Timothy Bunnell (Appl. Sci. and Eng. Labs., A. I. duPont Institute Hospital for Children, Univ. of Delaware, Wilmington, DE 19899, yarringt@asel.udel.edu)

A method has been developed for automatically extracting diphone speech segments with context-dependent boundaries. When compared with speech synthesized from manually extracted diphone speech segments, it was found that speech synthesized from the automatically extracted segments was, overall, slightly less intelligible but slightly more natural sounding [Yarrington *et al.*, "Robust automatic extraction of diphones with variable boundaries," in *EUROSPEECH '95, 4th European Conference on Speech Communication and Technology*, Vol. 3, pp. 1845–1848, Madrid, Spain (1995)]. The lower intelligibility appeared to be due to a small number of very poor diphone segments. While it is feasible to correct this problem by manually replacing misleading diphones, several changes have been made to the extraction procedure to eliminate or at least reduce the frequency of occurrence of incorrect diphones. In particular, a different spectral measure is being used for estimates of spectral similarity, and F_0 plus a spectral robustness measure have been added as features which are used in selecting diphones for extraction. Perceptual data comparing words generated from a diphone inventory constructed using the new algorithm with words generated from the original manually constructed inventory will be presented. [Work supported by NIDRR Grant No. H133E30010, the Nemours Research Programs, and the Microsoft Corporation.]

5pSC9. Modeling perceptual confusions to dysarthric speech. H. Timothy Bunnell, Xavier Menendez-Pidal, and James B. Polikoff (duPont Hospital for Children and Univ. of Delaware, 1600 Rockland Rd., Wilmington, DE 19803)

An HMM labeler has been extended to detect poor correspondence between phonetic labels and underlying acoustic data. This paper will present work extending the labeler to model perceptual confusions of human listeners from a forced-choice word identification experiment which used dysarthric speech. The speech and perception data are from the Nemours Dysarthric Speech database [Menendez *et al.*, Proceedings of ICSLP 96, SaP2P1.19 (1996)]. The perceptual data comprise distributions of listener identification responses over sets of from four to six words (the intended word plus several phonetically similar foils). In all, 37 words were produced twice by each of 10 dysarthric talkers providing a total dataset of 740 items. Each of these items was identified at least 12 times by five naive listeners for a total of at least 60 responses per item. Half of this data set will be used to adapt parameters of the HMM labeler to reproduce the distribution of human responses to the speech. The remaining half of the data set will be used to assess the ability of the labeler to select phonetic responses in a manner reflecting patterns of human perceptual confusions among the response set items. [Work supported by the Nemours Research Programs and NIDRR.]

5pSC10. Automatic phoneme labeler in the TIMIT database. Xavier Menendez-Pidal and H. Tim Bunnell (Appl. Sci. and Eng. Labs., A. I. duPont Institute and Univ. of Delaware, P.O. Box 269, Wilmington, DE 19899)

A phonetic HMM labeler was developed for use in an automatic diphone extraction system. Using the TIMIT database for training, the best HMM labeler performance was obtained when the number of states for a given phoneme model was proportional to the average duration of that phoneme over the TIMIT database, and the HMM topology allowed both repetition of states and skipping of adjacent states. The acoustic feature set was comprised of 30 features per frame (8 Mel cepstral coefficients, the log rms amplitude, and the zero crossing rate of a frame, as well as their first and second time derivatives). Labeling accuracy was tested using all speech files in the TIMIT test set and accuracy was assessed in terms of the degree of separation between labeler-assigned phoneme boundaries and

the nominal phoneme boundaries provided in the database. 97.4% of the boundaries assigned by the labeler were within 30 ms of the nominal phoneme boundaries, and 86.8% were within 10 ms (i.e., one analysis frame) of the nominal boundaries.

5pSC11. Recent progress in the INRS speech recognition system. Douglas D. O'Shaughnessy, Clark Z. Lee, Christophe Savariaux, Azarshid Farhat, Rachida El Méliani, Rivarol Vergin, and Michel Héon (INRS-Telecommunications, 16 Pl. du Commerce, Verdun, PQ H3E 1H6, Canada)

The INRS large-vocabulary continuous-speech recognition system employs a two-pass search. First, inexpensive models prune the search space; then a powerful language model and detailed acoustic-phonetic models scrutinize the data. A new fast match with two-phone lookahead and pruning speeds up the search. In language modeling, excluding low-count statistics reduces memory (50% fewer bigrams and 92% fewer trigrams); with *Wall Street Journal* texts, excluding single-occurrence bigrams and trigrams with counts less than five yields little performance decrease. In acoustic modeling, separate male and female right-context VQ models and a bigram language model are used in the first pass, and right-context continuous models and a trigram language model are used in the second pass. A shared-distribution clustering uses a distortion measure based only on the weights of Gaussian mixtures in the HMM model. Testing the system with a 5000-word vocabulary, the word inclusion rate (i.e., correct word retained in the first pass) is about 99%; word recognition accuracy is about 92.5%. Keyword spotting with new types of fillers retains accuracy with 1.2 false alarms/hour/keyword. [Work supported by NSERC-Canada and FCAR-Quebec.]

5pSC12. Predicting speech intelligibility in noise: The role of factors other than pure-tone sensitivity. Sigfrid D. Soli and Michael J. Nilsson (Dept. of Human Commun. Sci. and Devices, House Ear Inst., 2100 W. 3rd St., Los Angeles, CA 90057)

The ability to communicate with speech in noisy environments is critical in many job settings and occupations. Pure-tone thresholds from the audiogram traditionally have been used to screen individuals who seek employment in hearing-critical jobs. This paper reports on analyses of the relationship between pure-tone thresholds and English speech intelligibility in noise, as measured with the Hearing In Noise Test (HINT) [Nilsson *et al.*, 1993], in a sample of 515 job candidates. The sample was composed of 426 males and 89 females ranging in age from 20 to 50 years, with hearing levels from 0 to 115 dB HL. Subjects were classified into five language categories ranging from native, monolingual English speakers to bilingual speakers of English as a second language. Regression analyses and distributional analyses of the relationship between pure-tone thresholds and HINT scores from four conditions (quiet, noise front, noise right, and noise left) will be reported. These analyses will be used to characterize the range of speech intelligibility in noise among individuals with normal pure-tone thresholds, the prediction of speech intelligibility in noise from pure-tone thresholds, and the relationship of experience with English to speech intelligibility in noise.

5pSC13. Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. Robert W. Peters (Div. of Speech and Hear. Sci. and Dept. of Psych., Univ. of North Carolina, Chapel Hill, NC 27599-7190, bpeters@css.unc.edu), Brian C. J. Moore, and Thomas Baer (Univ. of Cambridge, Cambridge CB2 3EB, England)

Normally hearing people have much lower speech reception thresholds (SRTs) in a background of a single talker than in the background of speech-shaped noise, whereas hearing-impaired people do not. The hearing impaired appear to be less able than normal to take advantage of temporal and spectral "dips" in interfering speech. SRTs were measured in back-

ground sounds that varied to the extent that they contained such dips. The subjects tested were: (a) young with normal hearing; (b) elderly with near-normal hearing; (c) young with moderate to severe cochlear hearing loss and (d) elderly with moderate to severe cochlear hearing loss. In a background that contained both spectral and temporal dips, the hearing-impaired and elderly with near-normal hearing performed much more poorly than normals. The signal-to-background ratio required for 50% intelligibility was about 19 dB higher for elderly hearing impaired than for young normals. Young hearing-impaired subjects showed a slightly smaller deficit, but still a substantial one. Linear amplification combined with appropriate frequency-response shaping (NAL amplification) only partially compensated for these deficits. It is proposed that noise with spectral and temporal dips provides a potentially useful way of evaluating the effects of signal processing such as frequency-selective amplification and compression.

5pSC14. Intelligibility of temporal information extracted from narrow-band sentences. Eric W. Healy and Richard M. Warren (Dept. of Psych., Univ. of Wisconsin, Milwaukee, WI 53201, ehealy@csd.uwm.edu)

Sentences can retain intelligibility when heard through narrow spectral slits even though the fluctuating amplitude envelope within the narrow frequency bands contains little spectral information [K. R. Riener *et al.*, J. Acoust. Soc. Am. **91**, 2339(A) (1992); R. M. Warren *et al.*, Percept. Psychophys. **57**, 175–182 (1995)]. Other research has demonstrated that intelligibility is retained when broadband speech is divided into contiguous spectral regions and used to modulate the amplitude of corresponding regions of noise [R. V. Shannon *et al.*, Science **270**, 303–304 (1995)]. The current report combines aspects of these previous studies and provides information concerning the use of temporal pattern information in various regions of the spectrum: Two noncontiguous narrow bands of speech having various center frequencies were used separately to modulate the amplitude of noises, which were then refiltered to corresponding center frequencies and bandwidths. Despite the absence of spectral information within each of the narrow noise bands, listeners were able to identify many of the 500 keywords in the standard Central Institute for the Deaf sentences. It is concluded that intelligibility can be retained when acoustic information in sentences consists solely of amplitude changes occurring in two noncontiguous bands. [Work supported by NIH.]

5pSC15. The mismatch negativity auditory evoked potential: A reflection of acoustic or phonetic difference detection? Susan D. Dalebout and Janet W. Stack (Commun. Disord. Prog., Univ. of Virginia, 2205 Fontaine Ave., Ste. 202, Charlottesville, VA 22903 jws8n@virginia.edu)

The mismatch negativity (MMN) auditory evoked potential reflects detection of minimal stimulus differences at a pre-attentive level. It has been reported that the MMN reflects acoustic, rather than phonetic, processing of speech stimuli [Sharma *et al.*, Electroenceph. and Clin. Neurophys. **88**, 64–71 (1993)]. If this is true, then every stimulus pair from a continuum along which the acoustic differences between each step are equivalent should elicit an equivalent MMN, even when the acoustic differences are not perceptible to the listener (e.g., within-category differences). In this study, stimuli were two-step syllable pairs taken from a nine-step synthetic continuum varying in place of articulation from /da/ to /ga/. Results of conventional labeling and discrimination tasks demonstrated that the stimuli were categorically perceived. The MMN was tested using three stimulus pairs: a pair in which syllables were not perceptibly different (items seven and nine, a within-category distinction), the pair best discriminated (behaviorally) by each individual subject (a between-category distinction), and a pair comprising the endpoints of the continuum (items one and nine). Results are discussed in terms of their implications for using the MMN as an objective measure of speech discrimination ability and as a tool for studying aspects of speech perception.

Session 5pUW

Underwater Acoustics: Pulse Propagation

Michael A. Wolfson, Chair

Earth Systems Science Center, Pennsylvania State University, University Park, Pennsylvania 16802

Contributed Papers

1:00

5pUW1. Numerical investigation of strong-pulse propagation from shallow to deep water. Andrew A. Piacsek (Dept. of Phys., Central Washington Univ., Ellensburg, WA 98926, piacsek@tahoma.cwu.edu)

Finite-difference calculations of strong pulses propagating downslope in typical littoral environments were performed using the NPE program [B. E. McDonald and W. A. Kuperman, *J. Acoust. Soc. Am.* **81**, 1406–1417 (1987)], modified to include a seabed depth that varies with range and which provides for a specification of sound-speed gradients in the bottom as well as in the water column. In these simulations, the bottom is modeled with a (compressional wave) sound-speed profile that corresponds to mud and silt overlying a sandstone bedrock; an attenuation coefficient of 0.6 dB per wavelength is applied throughout the bottom. Results indicate that a significant fraction of the acoustic energy will generally reach deep water off the continental shelf, although this is strongly influenced by the water depth at the source and the distance of shallow-water propagation leading to the shelf. In addition, the strong refraction of the pulse in the upper seabed layers and the subsequent reradiation into the water column generates head waves that are generally preserved into the deep water. This feature may potentially serve as a discriminant between seismic and explosive sources and, in the latter case, to characterize the source environment. [Work supported, in part, by DOE through the Lawrence Livermore National Laboratory.]

1:15

5pUW2. Modeling nonlinear pulses in sediments with attenuation and dispersion. Rahul S. Kulkarni, William L. Siegmann (Dept. of Math. Sci., Rensselaer Polytechnic Inst., Troy, NY 12180-3590), Michael D. Collins, and B. Edward McDonald (Naval Res. Lab., Washington, DC 20375-5320)

A hybrid treatment for wide-angle paraxial propagation that includes effects of both sediment dispersion and weak nonlinearity has been developed. A Fourier transform approach is used to combine effects of refraction, diffraction, and sediment dispersion in the frequency domain, and nonlinear effects in the time domain. A nonlinear wide-angle time-domain equation developed recently is first split into linear and nonlinear component equations. The linear equation is decomposed into its discrete frequency-domain counterparts. Sediment attenuation and dispersion are incorporated using a complex wave number along with a frequency-dependent formula for phase velocity to satisfy causality. The numerical implementation consists of first decomposing a broadband source into its frequency components and propagating them over a range step using a wide-angle PE. Fourier synthesis is used to reconstruct the signal, which is then corrected to account for nonlinear effects in the time domain. Numerical examples illustrating the effects of dispersion and nonlinearity on shallow-water wide-angle propagation will be presented and compared with available results. [Work supported by ONR.]

1:30

5pUW3. Rate of energy transport and arrival times in pulse propagation in a waveguide. N. A. Sidorovskaia (Phys. Dept., Univ. of New Orleans, New Orleans, LA 70148) and M. F. Werby (Naval Res. Lab., Stennis Space Center, MS 39529)

The concept of group velocity has one of three origins and is described in terms of a mathematical construct to predict arrival times in pulse propagation in a waveguide, particularly within the context of normal mode theory. Under certain conditions, group velocity can lead to confusion. On the other hand, the rate of energy transport, which is what one really seeks, may be derived by taking the ratio of the mean energy flux through a vertical section with the energy density between two planes separated by a wavelength. In the context of normal mode theory, this leads to an integral expression that involves the normalized vertical wavefunctions and the velocity profile leading to a physically appealing expression that lends insight to expected modal arrivals. Under certain conditions, it leads to the same results as the group velocity concept: It can be calculated with only one frequency and is valid near cutoff frequencies. This supplies one with an integral expression for a velocity that better characterizes arrival times of packets of energy. [Work sponsored by ONR, the Naval Research Laboratory, and the University of New Orleans.]

1:45

5pUW4. Analysis of pulse propagation in range-independent and range-dependent sound channels with double duct. Natalia A. Sidorovskaia (Phys. Dept., Univ. of New Orleans, New Orleans, LA 70148, nasph@uno.edu), Michael F. Werby (NRL, Stennis Space Center, MS 39529), Harry DeFerrari, and Neil J. Williams (Univ. of Miami, Miami, FL 33149-1098)

An acoustic dataset collected during the 1983 Florida Straits Experiment [H. A. DeFerrari and H. B. Nguyen, *J. Acoust. Soc. Am.* **79**, 299 (1986)] is modeled using the normal-mode acoustic model SWAMP (shallow water acoustic modal propagation by M. F. Werby and N. A. Sidorovskaia) and parabolic acoustic model RAM (range-dependent acoustic model by M. Collins). The goal of the computer simulation is to get systematic studies of the influence of range-independent and range-dependent environments on the received pulse structure observed in the experiment. In the experiment the pulse, with central frequency 460 Hz, was sent through three different shallow-water acoustic channels with surface and bottom ducts. It was found that the surface duct generated an important temporally stable effect for the range-independent environment. At a range 42 km, six surface trapped precursors were observed. The range-dependent propagation channel implied that the energy transfers between modes made the effect less pronounced when only three weak precursors could be seen. The obtained systematic studies can be important in constructing inverse tomography techniques, in predicting sea-bed characteristics, and in better understanding the influence of the ocean ducts on sound propagation.

5pUW5. Interductile energy transfer of high-frequency pulse propagation in a two-ducted ocean waveguide in the yellow sea. N. A. Sidorovskaia (Dept. of Phys., Univ. of New Orleans, LA 70148) and M. F. Werby (Naval Research Lab., Stennis Space Center, MS 39529)

In a multiducted ocean wave guide, energy may be trapped in each of the ducted regions. For sources suitably placed, energy may be stimulated initially in only one of the ducts when eigenfunctions associated with secondary ducts have negligible overlap with the primary duct. For fixed frequency, no energy is transferred, and although high-angle modes eventually traverse to the secondary ducted region, the energy is not trapped there for monochromatic sources. In reality one must take into consideration a spread of frequencies. Monjo and DeFerri have noted this effect in collecting and interpreting data. This has led to a very exciting phenomenon, the "Monjo-DeFerri effect," which is quite dramatic for the case they studied. The Monjo-DeFerri effect is a mechanism whereby energy in a lower duct, with the source in the lower duct, may in time transfer energy to an upper ducted region. The energy transport is cyclical, and at suitably high frequencies energy is trapped in the upper channel in discrete time segments. This leads to the arrival of discrete packets of energy in time which are highly localized spatially. These "precursors" arrive in advance of predominant pulse arrivals, and may determine, among other things, the source distance. Here, a study of interductile energy transfer is made for the Yellow Sea. In suitable cases, discrete packets of energy are

captured at each high-angle modal contact with the secondary duct. Expressions that predict the arrivals of the upper ducted discrete arrivals, the number in range, the lag time between arrivals, and the lowest frequencies above which this will happen, are presented. [Work sponsored by ONR, NRL, and the University of New Orleans.]

5pUW6. Full-wave simulation of the forward scattering of sound in a structured ocean: A comparison with observations. Michael A. Wolfson (Earth Systems Sci. Ctr., Penn State Univ., University Park, PA 16802, wolfson@essc.psu.edu) and John L. Spiesberger (Penn State Univ., University Park, PA 16802)

The propagation of a 133 ± 16 -Hz acoustic signal through the eastern North Pacific Ocean is simulated over basin-scale ranges. A high angle c_0 -insensitive parabolic equation model is used for the simulations. The ocean environment was constructed using CTD data or climatologically averaged data from the Levitus database, along the geodesic from Kaneohe, near Oahu to the receiver, located 3709 km to the northeast. A projection of a three-dimensional frozen internal wave field is superposed. The stable earlier arrivals measured from experimental data were used as a reference anchor, and it is demonstrated that internal wave energy in addition to mesoscale structure is responsible for vertically scattering the late "axial" acoustic energy over 1 km.